

Università degli Studi Roma Tre

Dottorato di ricerca in Filosofia e teoria delle scienze umane, XXI ciclo

**Le ragioni dell'azione: economisti, psicologi e filosofi
di fronte ai paradossi della razionalità nella teoria della scelta**

SETTORE SCIENTIFICO-DISCIPLINARE DI AFFERENZA: M-FIL/03

Tesi di dottorato di **Grazia Farina**

matr. 5816-225

Coordinatore del dottorato
prof. Paolo D'Angelo

Tutore del dottorando
prof. Mario De Caro

Indice

Indice	2
Introduzione	5
Capitolo 1 - Expected Utility Theory	13
1.0 Premessa.....	13
1.1 Dall'utilità alla preferenza	15
1.2 Teoria delle preferenze rivelate.....	21
1.3 La scelta in condizioni di certezza.....	25
1.4 La scelta in condizioni di incertezza	27
1.5 La sintesi di Savage. Subjective Expected Utility (SEU).....	32
1.6 Il paradosso di Allais.....	36
1.7 Il paradosso di Ellsberg.....	42
1.8 <i>Preference reversal</i> . Interpretazioni di economisti e filosofi a confronto	45
1.9 Conseguenze delle violazioni degli assiomi sulla teoria della scelta razionale	56
Capitolo 2 - La prospettiva della psicologia cognitiva	62
2.1 Riunificare psicologia ed economia. Verso la <i>behavioral economics</i>	62
2.2 Heuristic and bias program	67
2.3 Euristiche del giudizio	68

2.4	Gli errori delle euristiche della rappresentatività.....	70
2.5	Gli errori delle euristiche della disponibilità e dell'ancoraggio ...	74
2.6	Errori nella valutazione del rischio	76
2.7	<i>Prospect theory</i> : una spiegazione alternativa delle scelte in condizioni di rischio.....	78
2.8	Effetto framing.....	87
2.9	Un'alternativa alla razionalità come coerenza.....	93
2.10	Due processi cognitivi.....	104
2.11	<i>Affect heuristic</i>	108
2.12	Il contributo delle neuroscienze.....	110
2.13	Un limite metodologico. Le caratteristiche fondamentali dei processi neurali	114
2.14	Una prima distinzione: processi automatici e controllati.....	118
2.15	Una seconda distinzione: processi affettivi e cognitivi	123
2.16	<i>Visceral factors</i>	125
2.17	Interazione come cooperazione o competizione.....	128
Capitolo 3 - Discussioni filosofiche sulla razionalità pratica.....		134
3.0	Premessa.....	134
3.1	Donald Davidson. Credenza, desiderio, intenzionalità.....	137
3.2	Ancora Davidson: comportamenti irrazionali.....	146
3.3	Derek Parfit e Bernard Williams: ragioni esterne e interne.....	155
3.4	Bernard Williams: quale razionalità per le ragioni interne?	164

3.5 John Broome: <i>instrumental reasoning</i>	178
3.6 Broome: ragioni e requisiti normativi.....	184
3.7 <i>Ruling passions</i> : la plasticità della ragione.....	191
Capitolo 4 - Robert Sugden: dalla non expected utility theory alla behavioral economy	206
4.0 Premessa.....	206
4.1 <i>Regret Theory</i>	209
4.2 <i>Disappointment theory</i>	215
4.3 Un diverso concetto di razionalità: <i>non-expected utility theory</i> .	219
4.4 I fondamenti filosofici della scelta razionale	225
4.5 Oltre la razionalità strumentale: Hume e la <i>behavioral economics</i>	242
4.6 Hume. Il ruolo delle passioni calme	254
Conclusioni in forma di sommario	258
Bibliografia	265

Introduzione

Questo lavoro si propone di mettere a confronto il diverso modo in cui economisti, psicologi e filosofi si sono confrontati con il tema della scelta, in particolare con la possibilità di costruire una teoria della scelta razionale. Il punto di partenza per questo confronto è l'Expected Utility Theory, la teoria standard della scelta nella teoria economica, che è stata il punto di riferimento più comune anche per le altre discipline. Come si vedrà nel corso del lavoro, la psicologia cognitiva si è direttamente confrontata con gli economisti sulla validità della teoria, e sia la filosofia dell'azione sia la filosofia morale ne hanno discusso, in riferimento in particolare alle concezioni della razionalità pratica. Ma se è vero che la teoria dell'utilità attesa è oggi tema assai discusso, e in molti settori diversi, è anche vero che le discussioni al riguardo restano di norma confinate nei rispettivi ambiti. Uno degli obiettivi principali di questa tesi è dunque confrontare e raccordare tali discussioni, cercando di verificare se e in che misura idee e suggerimenti avanzati in un settore possano risultare utili o produttive anche da prospettive diverse, e suggerendo l'utilità di un approccio multidisciplinare al problema.

Una delle questioni centrali che saranno oggetto del lavoro è quella di capire qual è la definizione di razionalità a cui il modello standard – ma anche gli altri modelli – fanno riferimento; in particolare, se si tratti di una caratteristica puramente logica o se invece chiami in causa i nostri effettivi meccanismi cognitivi. Si tratterà in questo caso di mostrare come le differenze metodologiche fra le tre prospettive – quella degli economisti, quella dei

filosofi e quella degli psicologi – si riflettano nelle discussioni che si svolgono all'interno di ciascun ambito disciplinare sul modo di intendere la razionalità della scelta, e su come vadano interpretate le sfide alla teoria della scelta razionale rappresentate dalle violazioni dei suoi assiomi.

L'intento di questo confronto è di capire quali fra le tesi proposte e argomentate a partire dalle diverse prospettive si dimostri più convincente, e se (ed eventualmente in quali casi) si delinei un percorso convergente di riflessione che possa portare anche a una reciproca validazione.

Nel **primo capitolo** viene presentata la teoria della scelta razionale nella sua forma classica, e cioè l'Expected Utility Theory¹, dando conto anche dell'evoluzione storica dei suoi fondamenti concettuali. In particolare, ci soffermeremo su alcuni importanti punti di svolta, anche di carattere metodologico: dall'utilità di Jeremy Bentham al concetto di preferenza ordinale, dalla teoria delle preferenze rivelate di Paul Samuelson fino alla sistemazione di quello che sarà il modello standard della scelta, e cioè dell'Expected Utility Theory di Von Neumann-Morgenstern. Si esamineranno quindi gli assiomi previsti da questa teoria per la scelta sia in condizione di certezza sia in condizione di incertezza, spiegando in che senso il rispetto degli assiomi garantisca la razionalità della scelta e di conseguenza quali siano per questa teoria i requisiti fondamentali di una scelta razionale. A questo punto si introdurranno i paradossi classici di Maurice Allais (1952) e Daniel

¹ Nel corso del lavoro per riferirsi alla teoria mainstream si utilizzerà più spesso l'acronimo EUT. Con l'acronimo SEU ci si riferisce invece alla Subjective Expected Utility.

Ellsberg (1961) che mostrando la violazione di importanti assiomi della teoria rappresentano le prime sfide che la teoria si è trovata ad affrontare. Si prenderanno quindi in esame le reazioni degli economisti sia a questi paradossi sia ai risultati di esperimenti compiuti anche da psicologi fin dagli anni '70, che rivelano, ad esempio, che le preferenze non sono invarianti come previsto dalla teoria. Le reazioni, e le proposte di riforma, serviranno anche a comprendere meglio gli intenti della EUT. Come si vedrà, gli economisti che non si sono trincerati dietro il mito dell'infallibilità in linea di principio della teoria, hanno proposto diverse ipotesi di riforma (non-expected utility theory). E, contemporaneamente, anche diversi psicologi cognitivi hanno avanzato modelli esplicativi che potessero dar conto dei comportamenti cosiddetti irrazionali.

Il **secondo capitolo** è dedicato alla prospettiva degli psicologi cognitivi e ripercorrerà le tappe principali del percorso di Kahneman e Tversky, i cui esperimenti mettono in discussione i capisaldi della teoria *mainstream*, e aprono la strada allo studio dei meccanismi cognitivi sui cui si basano le decisioni: “judgment and economic decision making”. Come si vedrà, questi studi danno inizio a un programma di ricerca - l'*heuristics and biases approach* – che cercherà di trovare una spiegazione per i comportamenti non conformi ai criteri di razionalità della EUT. Si analizzeranno i loro studi sulle distorsioni percettive e cognitive che, come si spiegherà, hanno come presupposto teorico la concezione della “bounded rationality” di Herbert Simon. All'interno di questa prospettiva si esaminerà, in particolare, l'ipotesi che riguarda il frequente ricorso - quando si è di fronte a una decisione o a una scelta - a percorsi cognitivi alternativi al calcolo razionale, le euristiche, di cui

si valuteranno sia i vantaggi sia i limiti in termini di affidabilità. Come si vedrà, infatti, le euristiche sono all'origine di errori gravi e sistematici; il termine *bias* indica proprio questo genere di errori sistematici, che sono la conseguenza, a volte inevitabile, di giudizi non abbastanza meditati, che spesso diventano veri e propri pregiudizi.

Sulla base di queste prime ricerche, Kahneman e Tversky costruiranno la prospect theory, che ha lo scopo di descrivere quali siano le effettive procedure cognitive sottese alle scelte in condizioni di rischio. Come si vedrà, la prospect theory non si presentava, almeno inizialmente, come un'alternativa alla EUT, di cui infatti non mette in discussione la validità come modello di scelta per un agente pienamente razionale. Si proponeva però di tener presente che i comportamenti degli agenti reali, basandosi prevalentemente su procedure euristiche, si discostano inevitabilmente dalle aspettative di massimizzazione previste dalla EUT. Ad esempio, come si mostrerà attraverso le ricostruzioni proposte dai due psicologi, quando si è di fronte a un problema di scelta, la necessità di elaborarne una rappresentazione semplice ma efficace viene influenzata dal modo in cui viene presentato il problema (*framing effect*) e dalle aspettative del decisore, con il risultato che l'ordine delle preferenze non sarà invariante, come è invece richiesto dalla EUT.

Sulla base delle riflessioni di Kahneman sulla razionalità, si proverà a valutare se e fino a che punto la prospect theory, e più in generale la *behavioural economics* che da qui prende le mosse, sia davvero il risultato di un ripensamento complessivo dei fondamenti teorici della EUT, oppure se si limiti a criticarne la scarsa capacità descrittiva dei comportamenti reali e a cercare di migliorarla con l'aiuto delle indagini psicologiche. L'esame delle

implicazioni che derivano dal ritorno da parte di Kahneman a un concetto benthamiano di utilità identificata con il piacere - e che trasforma l'utilità attesa nell'*experienced utility* - porterà a propendere per la seconda ipotesi.

Si esaminerà infine la tesi dell'esistenza di due tipi di processi cognitivi, che cerca di rispondere agli interrogativi sul diverso grado di controllo che siamo in grado di esercitare sui giudizi e sulle decisioni, e che ha portato alla formulazione di diversi modelli a due sistemi con denominazioni differenti, tutti accumulati dall'alternativa fra un processo automatico e uno controllato. Come si vedrà, più di recente le tecniche di visualizzazione cerebrale hanno permesso alle neuroscienze di localizzare le aree distinte del cervello coinvolte in alcuni dei processi cognitivi e affettivi. Pertanto si prenderà in esame il modo diverso in cui viene interpretato il contributo che le neuroscienze sono in grado di dare alla teoria economica; se cioè tale contributo riguardi solo la possibilità di scoprire nuove evidenze attraverso le nuove tecniche di indagine che le neuroscienze mettono a disposizione, oppure possa anche arrivare a ridisegnare radicalmente il modello interpretativo della scelta. Verranno valutati vantaggi e limiti delle due interpretazioni che, come si vedrà, sono entrambe presenti nella cosiddetta *neuroeconomics*.

Il **terzo capitolo** proverà a cercare in alcune riflessioni filosofiche sulle ragioni dell'agire, una prospettiva che eviti l'eccessiva contrapposizione fra una metodologia positivista e deduttiva e una empirica e dunque prevalentemente descrittiva. L'idea di fondo che si intende vagliare è se dal confronto fra impostazioni teoretiche anche molto distanti, si arrivi a chiarire alcune questioni riguardanti il tema della natura della razionalità, già intraviste sullo sfondo nelle prospettive esaminate dei capitoli precedenti. Inoltre, si

cercherà di capire se e come, da una prospettiva filosofica, si riesca a restituire importanza al contesto motivazionale e alle componenti sociali della scelta.

Ci si soffermerà dunque sul tema delle ragioni delle azioni nella riflessione di Donald Davidson, e in particolare sulla differenza fra cause e ragioni e dunque fra i due modelli esplicativi dell'azione, quello causalista e quello teleologico. Ci si soffermerà in particolare sulla possibilità di spiegare la scelta di un agente a partire dalla combinazione di un desiderio e di una credenza e sulla tesi di Davidson che questo "pensiero proposizionale" sia una condizione indispensabile per poterla considerare razionale. Nella stessa prospettiva, si spiega il ruolo centrale dell'intenzionalità, dato che la possibilità di descrivere un'azione come intenzionale implica il soddisfacimento dei criteri minimi di razionalità rappresentati dalla coerenza logica. Verranno evidenziati sia i pregi della caratterizzazione del contesto di scelta sia i limiti di un approccio che privilegia la traducibilità in proposizione dei comportamenti di scelta.

La discussione sulle ragioni esterne o interne, invece, servirà a mettere in luce la contrapposizione fra le ragioni normative, che fanno riferimento a una normatività basata sulla corrispondenza con fatti e valori esterni, e le ragioni motivanti, che rispecchiano una razionalità cosiddetta procedurale o strumentale, che si limita a stabilire un legame solo psicologico fra le motivazioni di un agente e le sue scelte. In questa discussione, come si vedrà, le argomentazioni utilizzate da Derek Parfit in difesa delle ragioni esterne sono particolarmente illuminanti per capire le ragioni dell'approccio normativo alla razionalità e anche per individuarne i limiti. Di contro, l'internalismo rappresentato da Bernard Williams sembra scongiurare i rischi

di una ragione troppo svincolata dalle nostre effettive motivazioni. Particolare attenzione sarà data all'analisi del complesso motivazionale (“agent’s subjective motivational set”) e del percorso deliberativo (“sound deliberative route”) svolti da Williams, per capire se, in che misura, consentano il superamento di una concezione puramente strumentale della razionalità. Proprio sulla natura della razionalità strumentale si interroga John Broome e, come si vedrà, la sua lucida analisi volta ad accertarne il carattere normativo non sembra riuscire nell'intento di trovare una risposta soddisfacente. Una impostazione molto diversa rispetto a quella di Broome è invece quella presente nelle tesi di Simon Blackburn che, critico sulla razionalità normativa e strumentale, argomenta a favore di una concezione della ragione “Humean-friendly”. Il vantaggio di una prospettiva che considera la ragione come non autosufficiente, consente di evitare l'errore di identificare la prospettiva deliberativa con quella che prende in esame i desideri e li valuta di fronte al tribunale indipendente della ragione. Si tratta di una prospettiva molto vicina a quella di Williams, di cui vuole correggere l'eccesso di relativismo, e che, come cercherò di mostrare, si rivela particolarmente convincente per le tematiche che sono al centro di questo lavoro.

Nel **quarto capitolo** si tirano le fila delle diverse prospettive seguendo il percorso attraverso cui l'economista Robert Sugden è approdato alla behavioral economics. Come si capirà, i motivi di interesse del suo percorso intellettuale ai fini di questo lavoro sono diversi. E si troveranno, in primo luogo, esaminando la sua proposta di una *non-expected utility theory*, la *regret theory*, che promette di dar conto della complessità dei motivi che danno origine alla scelta e in particolare del ruolo svolto dalle emozioni. In secondo

luogo, riflettendo sulla sua critica alla EUT, in gran parte basata sulla ricostruzione delle tappe che hanno portato all'affermazione della teoria *mainstream*, in cui la prospettiva marginalista rappresenta a suo avviso “*the road non taken*”: e cioè una possibilità di tenere assieme psicologia ed etica, abbandonata con il prevalere agli inizi del Novecento di una impostazione positivista che aspirava a fare a meno della psicologia. Per altro verso, come si vedrà, la sua critica è anche rivolta esplicitamente alla concezione di razionalità che è alla base della EUT. Da qui prende le mosse un interessante riesame dei fondamenti filosofici della scelta, in cui l'attenzione sarà focalizzata in particolare sulla teoria della motivazione di Hume. Le complesse sfaccettature dell'interpretazione di Hume proposta da Sugden, vedono il filosofo scozzese prima come fautore di una concezione strumentale della decisione, differente però da quella di Bentham, poi come sostenitore di una teoria della decisione non strumentale e non proposizionale. Una teoria che chiamerà in causa anche le diverse posizioni sulla razionalità di Davidson e Blackburn. Tale ricostruzione permetterà di spiegare l'origine delle sue argomentazioni a favore di una spiegazione puramente naturalistica dei processi motivazionali alla base delle scelte, che prescinda da ogni riferimento a una qualsiasi forma di razionalità pratica. Nella stessa direzione va la sua convinzione che sia questa la prospettiva lungo cui deve muoversi la behavioural economics, basandosi esclusivamente su presupposti di carattere psicologico che non richiedono alcun riferimento alla razionalità.

Capitolo 1 - Expected Utility Theory

1.0 Premessa

Nella teoria economica mainstream la scelta razionale è rappresentata dal modello dell'Expected Utility Theory (EUT) di Von Neumann e Morgenstern. I paradossi classici di Allais e Ellsberg prima e gli esperimenti di economisti e psicologi dagli anni '70 in poi, hanno insinuato molti dubbi sulla validità di quella teoria e, di conseguenza, sulla definizione di razionalità applicata alla scelta che propone. Alcuni economisti hanno proposto riforme della teoria standard, altri economisti e alcuni psicologi hanno invece elaborato teorie alternative. In campo economico, le due direzioni di ricerca che sembrano oggi più promettenti sono quelle della behavioural economics e della neuroeconomics.

Lo scopo di questo lavoro è di individuare – anche attraverso le diverse interpretazioni dei risultati degli esperimenti – i punti deboli della teoria, quelli che la hanno resa poco plausibile. Passioni e norme sociali, come era già chiaro a Hume, svolgono un ruolo imprescindibile nella scelta individuale – sia essa atomistica oppure in condizione di interazione strategica con altri individui –, mentre l'assiomatizzazione della teoria si è affermata parallelamente all'estromissione delle componenti psicologiche e filosofiche nella spiegazione della scelta. Si tratta di capire in che senso queste componenti siano imprescindibili e come approcci metodologici alternativi (behavioural economics, neuroeconomics) si propongano di tenerne conto. L'ipotesi di fondo di questo lavoro è che sia necessario mettere in discussione

l'idea di una razionalità definita esclusivamente sulla base del rispetto degli assiomi della teoria dell'utilità attesa senza però rinunciare all'individuazione di requisiti di razionalità per la scelta.

A grandi linee la struttura del lavoro prevede la ricostruzione, anche storica, dei concetti basilari della EUT e cioè utilità, preferenza e probabilità, la presentazione degli assiomi di von Neumann e Morgenstern, la distinzione fra scelte in condizioni di certezza, rischio e incertezza. Dopo aver descritto i paradossi e gli esperimenti di violazione degli assiomi della EUT, si esamineranno le proposte di riforma della teoria che non ne mettono in discussione il nucleo teorico avanzate dagli economisti (Non Expected Utility Theory). Sul versante degli psicologi, invece, la Prospect Theory di Kahneman e Tversky e l'ipotesi delle neuroscienze sul ragionamento come procedimento duale danno una rappresentazione profondamente diversa della scelta. Infine, il punto di vista filosofico consentirà di riesaminare in modo critico i presupposti teorici della EUT e soprattutto di mettere a confronto i diversi paradigmi di razionalità della scelta proposti in campo economico e filosofico.

La rappresentazione che è alla base del modello di scelta individuale prevede un agente impegnato nella scelta delle azioni volte a massimizzare i suoi obiettivi. Sono quindi razionali le scelte che si rivelino adeguate al raggiungimento dell'obiettivo (o al soddisfacimento di una preferenza). Di norma, una situazione di scelta implica un percorso cognitivo che coinvolge l'insieme di desideri e credenze da cui scaturisce la deliberazione. E' appunto a tale complesso motivazionale che facciamo riferimento quando ci troviamo a giustificare un'azione. La moderna teoria dell'utilità attesa (nell'acronimo inglese EUT) si fonda invece su un procedimento deduttivo a partire da assiomi che definiscono i requisiti di razionalità relativi al nesso fra la scelta e le preferenze dell'agente. In questa visione l'agente è tutt'uno con la scelta, perché quest'ultima ci dice tutto quello che è essenziale sapere sulle preferenze dell'agente. Se tra x e y l'agente (A) sceglie x , la scelta semplicemente rivela la preferenza di A per x , o almeno che A preferisce leggermente x a y . Non c'è altro da sapere. Tutto ciò che accade a livello cognitivo prima della scelta intesa come rivelazione delle preferenze è senza importanza. Il punto di partenza del processo di scelta sono soltanto le preferenze già definite e ordinate e che sono rivelate dalle scelte stesse. Per questo il primo passo sarà l'esame del concetto di preferenza.

1.1 Dall'utilità alla preferenza

La nozione tradizionale di utilità definisce intuitivamente il grado di "piacere" che l'agente ricava dal soddisfacimento dei suoi obiettivi. Elaborato nel campo dell'etica, l'utilitarismo consequenzialista di Bentham proponeva una valutazione delle azioni umane basata sulle *conseguenze* che esse hanno

per la collettività, invece che su valori intrinseci assoluti. Le conseguenze possono essere misurate in termini di piacere, una grandezza psicologica che rappresenta il livello di soddisfazione conseguito in virtù di una data azione. Il piacere è una misura cardinale che rende possibile anche confrontare i livelli di soddisfazione degli individui. L'uguaglianza fra gli individui rispetto alla capacità di sperimentare piaceri e dolori permetterebbe di stabilire se una data azione contribuisce ad aumentare o a diminuire il benessere collettivo attraverso il "calcolo felicifico", e cioè sommando gli incrementi e sottraendo le perdite di piacere di ciascun individuo.

Nonostante la prospettiva relativistica e individualistica del concetto di piacere, per Bentham l'obiettivo da raggiungere è il benessere collettivo, ossia la massima felicità del massimo numero di individui. La sua fiducia illuministica nella ragione lo porta ad avere una visione ottimistica della relazione fra comportamenti egoistici e bene comune. A suo avviso, il compito di armonizzare gli interessi privati dei singoli individui con l'interesse pubblico della collettività spetta al legislatore. Quest'ultimo, servendosi del calcolo felicifico che gli consente di valutare le misure di utilità delle diverse scelte, potrà intervenire con incentivi e disincentivi al fine di promuovere quelle scelte che sono in grado di promuovere la situazione ottimale per la società nel suo complesso, e cioè "la massima felicità del massimo numero"².

² A. Roncaglia, *La ricchezza delle idee. Storia del pensiero economico*, Bari-Roma, Laterza, 2001, p. 306 sg.

Più di mezzo secolo dopo, la teoria economica marginalista recupera l'idea del calcolo felicifico, e adotta una misura cardinale della utilità. Alla fine dell'Ottocento molti economisti cominciarono ad essere insoddisfatti di questa impostazione. Un primo problema era dato dal fatto che l'utilità, identificata con il piacere, è comunque una grandezza psicologica non osservabile, e questo sembrava renderla inadeguata per una spiegazione scientifica. Inoltre, per effettuare confronti fra due diverse utilità sembra inevitabile il ricorso a procedure indirette e poco scientifiche come l'introspezione, che non è passibile di controllo intersoggettivo, anche se Jevons, uno dei principali protagonisti della rivoluzione marginalista, aveva comunque esplicitamente escluso la possibilità di confronti interpersonali.

Vilfredo Pareto, convinto della necessità di separare la teoria economica dalla psicologia, propose in alternativa una misura ordinale che poteva essere ricavata dalla diretta osservazione delle scelte. A quel punto le preferenze ordinali - che si presentano come dati empirici, scientificamente validi in quanto ricavati dal solo comportamento - prendono il posto del concetto di utilità cardinale misurabile. Non serviranno a misurare un valore assoluto, ma un valore relativo rispetto alle altre possibili opzioni, e non permetteranno confronti intersoggettivi di utilità, che apparivano comunque problematici. Secondo Pareto, le preferenze ordinali si limitano a riassumere le regolarità empiriche dei comportamenti di scelta, e quindi non hanno bisogno di nessuna interpretazione psicologica. L'ordinalismo diventerà l'approccio dominante della teoria del consumatore, in cui secondo alcuni rimarrebbe comunque spazio per una prospettiva psicologica sulla motivazione: le persone agiscono in modo intenzionale, perciò hanno preferenze ordinali che non sono

razionalizzazioni ex post del comportamento ma hanno un significato psicologico³. Secondo questa interpretazione, Pareto si sarebbe limitato a sostituire una forma meccanicistica di psicologia con il formalismo matematico. Ma la proposta di Pareto di separare la teoria economica dalla psicologia può invece essere vista come un significativo cambiamento di rotta rispetto al modo corrente fra gli economisti di intendere la natura della spiegazione⁴. Secondo Pareto le teorie economiche dovevano essere dedotte da proposizioni empiriche saldamente definite sulla scelta piuttosto che sulle sensazioni, e quindi senza bisogno di ricorrere a concetti psicologici come quelli di utilità o piacere. Pur non negando lo status di scienza della psicologia, Pareto ritiene che ogni scienza debba avere le sue proprie leggi fondamentali e fare deduzioni a partire da esse.

His proposal is that economics should constitute itself as a separate science – separate, in particular, from psychology and from sociology. So what is economics to be the science of? Economics – or at least *pure* economics – is to be *the science of logical action*⁵.

³ S.B. Lewin, “Economics and psychology: lessons for our own day from the early twentieth century”, *Journal of Economic Literature*, 1996, v. XXXIV, pp. 1293-1323, p.1311. Ma questa lettura è controversa, cfr. L. Bruni e R. Sugden, “The road not taken: how psychology was removed from economics and how it might be brought back”, *The Economic Journal*, 2007, 117, pp. 146-173. Corsivi nel testo.

⁴ È questa l’interpretazione di L. Bruni e R. Sugden, cit.

⁵ L. Bruni e R. Sugden, cit., p. 155.

Per Pareto un'azione è logica se e solo se è il risultato di un *ragionamento strumentale* valido, a partire da premesse oggettivamente vere. Nel capitolo sull'equilibrio economico del *Manuale di politica economica* si propone di studiare “le azioni logiche, ripetute in gran numero, che fanno gli uomini per procacciarsi le cose che soddisfano ai gusti loro⁶”.

Ossia ci occupiamo solo di certe relazioni tra fatti oggettivi e fatti soggettivi, che sono principalmente i gusti degli uomini. Inoltre semplifichiamo ancora il problema, supponendo che il fatto soggettivo si adatti perfettamente al fatto oggettivo; e ciò possiamo fare perché consideriamo solo azioni ripetute; il che pure ci concede di ritenere che il nesso tra le azioni sia quello logico⁷.

Pareto usa quindi due criteri per identificare le azioni logiche nella teoria economica. Il primo è che si tratti di azioni che vengano “ripetute in gran numero”. In questi casi è ragionevole assumere che le credenze sul mondo – fatti soggettivi – coincidano con il mondo come realmente è – fatti oggettivi. Restringendo il dominio della teoria economica alle situazioni in cui gli agenti hanno avuto adeguate possibilità di apprendere le conseguenze delle scelte alternative, Pareto anticipa la tesi delle preferenze rivelate. Il secondo criterio è che perché un'azione economica possa dirsi logica deve essere rivolta al soddisfacimento dei gusti attraverso l'acquisizione di beni. I due criteri identificano quindi una categoria di azioni che sono strumentalmente razionali rispetto ai fatti oggettivi. L'idea complessiva è che dopo un numero sufficiente

⁶ V. Pareto, *Manuale di economia politica*, Milano, Società editrice libraria, 1906, rist. Roma, Bizzarri, 1965, cap. 3, p.1.

⁷ *Ibid*

di ripetizioni ogni agente acquisisce un sistema stabile di gusti, strutturato in modo tale che i problemi di scelta strumentalmente razionali siano ben definiti.

Exactly what that structure is depends on how rational choice is formalized. For pare to, who interprets rational choice in terms of constrained maximization, the requirement is that tastes take the form of an ordering over the relevant set of end states⁸.

La razionalità dell'azione logica è strumentale, le azioni economiche – la compravendita di beni e servizi – sono i mezzi, mentre la soddisfazione dei gusti è il fine.

Sulla base dello stesso approccio, la teoria delle preferenze rivelate di Samuelson proporrà di prendere in considerazione solo le scelte, dato che queste ultime esprimono - in modo osservabile e quindi oggettivo - le preferenze dell'agente. In questo modo si elimina ogni riferimento a contenuti psicologici potenzialmente oscuri e per definizione non osservabili e quindi poco scientifici. Infine, negli anni '40 l'assiomatizzazione della teoria della scelta in condizione di rischio ad opera di von Neumann e Morgenstern stabilirà in maniera deduttiva gli assiomi che definiscono la razionalità di un comportamento.

È importante anche solo citare una posizione dissonante con questa impostazione, che vorrebbe far acquisire alla teoria economica il rigore della

⁸ L. Bruni e R. Sugden, cit., p. 158.

meccanica, quale è quella di Keynes, convinto che la teoria economica non potesse aspirare al rigore formale delle leggi della fisica: “the differentia of economic laws, as contrasted with purely physical laws, consists in the fact that the former imply voluntary human action⁹”. Sono le motivazioni – in particolare quella del “goal fulfillment” – a fare la differenza fra comportamento meccanico e agire umano.

Per riassumere, la transizione dalla cardinalità all’ordinalità consente di formalizzare un modello di scelta razionale non più vincolato a fattori psicologici e ha anche il vantaggio di attenuare il carattere teleologico della scelta, che appariva problematico nel confronto con i modelli causali delle scienze meccaniche. Una volta eliminato ogni riferimento alle motivazioni dell’agente, il criterio della coerenza interna rimane il solo indicatore della razionalità della scelta.

1.2 Teoria delle preferenze rivelate

La teoria di Samuelson delle preferenze rivelate intendeva basarsi su un concetto di preferenza libero da qualsiasi riferimento a stati mentali, e che non avesse un significato indipendente dall’atto di scelta. In definitiva, il suo obiettivo era “to develop the theory of consumer’s behaviour freed from any

⁹ J. N. Keynes, *The scope and method of political economy*, London, Macmillan, [1890], 1917, p. 86.

vestigial traces of utility concept¹⁰”. In questo modo veniva portata a compimento la svolta iniziata da Pareto che, eliminando i concetti psicologici, fonda la teoria economica sui principi della scelta razionale.

Il punto di forza della teoria delle preferenze rivelate è dato dalla possibilità di dedurre la struttura non osservabile di preferenze dall’osservazione diretta del comportamento di scelta. In questo modo si ottiene maggiore oggettività rispetto al meccanismo introspettivo, liberando il campo da ipotesi interpretative di carattere psicologico in termini di motivazione, e in particolare dalla interpretazione delle preferenze come stati mentali. Alla base di questa teoria c’è l’idea di basarsi su un dato osservativo, che in quanto tale è affidabile e oggettivo, eliminando ogni riferimento a “ciò che sta dietro” la scelta osservabile.

The outside observer notices that this person *chose* x when y is available and infers that he *preferred* x to y . From the point of view of introspection of the person in question, the process runs from his preference to his choice, but from the point of view of the scientific observe the arrow runs in the opposite direction: choices are observed first and preferences are then presumed from these observations¹¹.

¹⁰ P. Samuelson, “A note of pure theory of consumer’s behaviour”, in *Economica*, 1938. Citato in A. Sen, “Behaviour and the concept of preference”, *Economica*, 1973, pp. 241-259, p. 242.

¹¹ A. Sen, “Behaviour and the concept of preference”, *Economica*, 1973, 40, pp. 241-259, p. 241.

Di conseguenza, data la presupposta equivalenza fra razionalità e coerenza, basandosi sul solo comportamento osservabile si può accertare la razionalità del comportamento del soggetto. La condizione di coerenza che è alla base di questa teoria afferma che se un soggetto sceglie x rispetto a y significa che preferisce x a y , e quindi non sceglierà y in un contesto in cui x è disponibile (*weak axiom of revealed preference*). Non si tratta però di eliminare le preferenze a favore delle scelte. Si tratta piuttosto di non indagare sulle ragioni che hanno portato a preferire qualcosa invece di qualcos'altro, limitandosi a prendere atto del risultato finale del processo deliberativo, a prescindere dalle ragioni. È evidente del resto che il criterio di coerenza non può evitare il riferimento alle preferenze, e cioè al loro contenuto proposizionale, se vuole rilevare un'incoerenza fra due diversi comportamenti di scelta.

D'altra parte è anche vero che, come generalizzazione psicologica, l'ipotesi delle preferenze rivelate può non esser del tutto convincente. L'idea di che gli agenti rivelino le loro sottostanti preferenze attraverso le loro scelte effettive potrebbe non essere un presupposto ragionevole, se si pensa che le scelte non sono sempre il risultato di un ragionamento ponderato o di un confronto sistematico delle alternative. Inoltre, in molti casi la scelta effettiva non è espressione di una preferenza, perché condizionata da regole o norme di comportamento che inducono a compiere una scelta che non rispecchia l'effettivo ordine di preferenze.

Un altro punto controverso è costituito dalla proprietà logica di connessione che ha un ruolo importante nella teoria delle preferenze rivelate. La connessione delle preferenze – che equivale all'assioma di completezza –

stabilisce che fra due alternative si potrà preferire x a y , oppure y a x o ancora essere indifferente fra le due alternative. In base all'assioma debole delle preferenze rivelate se un agente sceglie x anziché y si presume che consideri x desiderabile almeno quanto y

non già che non avesse nessun'idea su cosa scegliere e ha scelto x perché doveva comunque scegliere qualcosa. Sen illustra questo punto servendosi di una variante della storia dell'asino di Buridano che non riesce a decidere quale dei due mucchi di fieno scegliere: gli piacciono entrambi, ma non riesce a decidere quale sia migliore. Non essendo capace di scegliere, il titubante animale finisce per morire di fame.

The dilemma of the ass is easy to understand, but it is possible that the animal would have agreed that it would have been better off by choosing either of the haystacks rather than nothing at all.[...] On the other hand, had it chosen either of the haystacks, it would have been presumed that the ass regarded that haystack to be at least as good as the other, which in this version of the story is not the case. The ass was in a real dilemma *vis-à-vis* the revealed preference approach¹².

L'interpretazione più comune del racconto è che l'asino fosse indifferente fra i due mucchi di fieno, in realtà se fosse stato davvero indifferente non avrebbe avuto ragione di esitare dal momento che una qualsiasi scelta sarebbe andata bene. Non è invece così se la relazione di preferenza non è connessa rispetto alla coppia di alternative in questione, se

¹² *Ivi*, p. 248.

cioè chi deve scegliere non può dire di preferire l'una all'altra o di essere indifferente fra le due. Quindi, per l'asino la scelta di uno dei due mucchi di fieno sarebbe apparsa come la rivelazione che il mucchio scelto non era peggiore dell'altro, mentre è proprio questa l'opinione che non può adottare poiché non riesce a decidere quale debba essere la sua preferenza. Dunque, la scelta di uno dei due mucchi avrebbe dato un segnale sbagliato al teorico delle preferenze rivelate, poiché la sua scelta avrebbe implicato che considerava il mucchio che ha scelto "at least as good as the other".

1.3 La scelta in condizioni di certezza

In caso di informazione perfetta e completa, un ordinamento di preferenze che soddisfi gli assiomi di riflessività, completezza, transitività e continuità garantirà la massimizzazione dell'utilità. L'ordinamento è univoco e stabile, dato che le relazioni fra le preferenze sono definite indipendentemente dal contesto (se x è preferito a y , è preferito in qualunque contesto siano disponibili sia x sia y oppure x, y, z).

L'assioma di *riflessività* ha lo scopo meramente formale di prendere in considerazione ciascuna preferenza in quanto tale.

L'assioma di *completezza* prescrive la possibilità di tre sole relazioni di preferenza fra due opzioni. Esse sono mutualmente esclusive e congiuntamente esaustive: x è preferito a y ; y è preferito ad x ; indifferenza fra x e y . In questo modo l'assioma certifica la validità universale della relazione di preferenza, poiché è sempre possibile confrontare e quindi ordinare qualsiasi coppia di alternative. Non è ammessa l'incommensurabilità: se A

sceglie x invece di y, significa che ha almeno una debole preferenza per x rispetto a y e riconfermerà questa preferenza in ogni altro contesto in cui sarà presente x. E' ammessa invece la possibilità di non preferire un'alternativa a un'altra, e quindi di scegliere indifferentemente ora x ora y.

L'assioma di *transitività* garantisce la coerenza logica della relazione di preferenza e impone che se A è preferito a B e B è preferito a C, allora A deve essere preferito a C. La transitività è la proprietà che meglio rappresenta la natura del vincolo imposto alla struttura delle preferenze, e cioè la coerenza, dal momento che – guardando alle proposizioni che esprimono le preferenze – il mancato rispetto della transitività costituirebbe una violazione inaccettabile delle regole logiche.¹³ Completezza e transitività sono considerati i requisiti indispensabili per avere un ordinamento di preferenze.

A questi assiomi, che riguardano proprietà implicite nella nozione stessa di ordinamento, bisogna aggiungere la condizione di *continuità*. La continuità implica che, dati due qualsiasi beni in un paniere, sarà sempre possibile – riducendo le quantità dell'uno e incrementando le quantità dell'altro – definire un altro paniere che è indifferente rispetto al primo dal punto di vista dell'utilità totale derivante dal consumo dei due beni. Ciò esclude che si possa avere un ordinamento lessicografico e cioè che qualsiasi unità del bene x abbia per il soggetto un'utilità superiore a quella conferitagli dalla prima unità del bene y. La presenza di quest'ultima condizione permette di rappresentare

¹³ In realtà nelle scelte dinamiche, cioè quelle in cui interviene il fattore tempo, e nei casi in cui le alternative non sono chiaramente specificate, la violazione della transitività appare spesso plausibile.

l'ordinamento mediante una funzione di utilità: "the individual, when acting in accordance with these preferences, can be represented as maximizing his or her utility"¹⁴.

Anticipo qui alcune considerazioni riguardo agli assiomi di completezza e di transitività su cui tornerò più avanti. L'assioma di completezza può essere considerato difficile da soddisfare in quanto sembra richiedere informazioni esaurienti per ordinare le preferenze, e comunque implica la necessità di avere sempre preferenze connesse, ma è difficile da soddisfare soprattutto nel caso di caratteristiche o beni che non sono confrontabili e non possono essere rappresentati da una struttura binaria. Anche la transitività è strettamente legata al sistema binario proprio della struttura delle preferenze, che in certi casi sembra inadeguato a descrivere la situazione di scelta. Ad esempio, nel confronto tra coppie di alternative, una stessa opzione potrebbe essere preferita o meno, proprio perché l'opzione appare sotto una luce diversa, a seconda dell'alternativa con cui è messa a confronto.

1.4 La scelta in condizioni di incertezza

L'incertezza comporta per il soggetto una più complessa determinazione della scelta. In questo caso, per la teoria della scelta, la questione più importante è trovare il modo migliore per descrivere la condizione di incertezza.

¹⁴ S. Hargreaves Heap et al., *The Theory of choice. A critical guide*, Oxford, Blackwell, 1992, p.7.

Is uncertainty about the future the same as calculable risk and mathematical probability? Is it ignorance of the facts? Or it is matter for our imagination¹⁵?

Secondo la visione neoclassica, gli agenti dispongono di tre insiemi di informazioni grazie ai quali possono ottenere una descrizione dettagliata del futuro. In primo luogo, si suppone che gli agenti conoscano con certezza tutte le possibili conseguenze delle loro azioni. Ad esempio, valutando l'acquisto di un biglietto per una partita, conosco i possibili risultati per la mia squadra, i punteggi possibili e ogni altra circostanza che potrà influenzare la visione della partita. Le combinazioni di queste possibilità sono chiamate stati di natura. In secondo luogo, gli agenti devono anche sapere l'effetto che ogni possibile stato di natura avrebbe sulla loro utilità, rispetto ad ognuna delle scelte disponibili; nell'esempio, comprare il biglietto o restare a casa. In ultimo, il terzo insieme di informazioni consiste nella distribuzione di probabilità, definita per ciascuno degli stati di natura.

Ma qual è l'origine di queste informazioni? Per il secondo insieme, quello relativo alle utilità, secondo la visione neoclassica, i gusti degli agenti sono dati e non c'è alcuna necessità di spiegarne la provenienza, dal momento che gli agenti sono i migliori giudici delle loro utilità, essendo liberi di immaginare le conseguenze delle loro azioni per ogni stato di natura. Riguardo al primo punto, è vero che, dato che gli eventi che si possono verificare in futuro sono indipendenti dall'immaginazione, possono esserci stati di natura

¹⁵ S. Hargreaves Heap et al., *The Theory of choice. A critical guide*, cit., p. 51.

che non vengono presi in considerazione; ma in realtà proprio per questo, secondo la teoria neoclassica, non influenzano la scelta.

Riguardo all'ultimo insieme, e cioè all'origine delle probabilità che vengono assegnate ai diversi stati di natura, in alcuni casi più semplici, chiaramente specificati e con cui l'agente si confronta ripetutamente, dovremmo essere in grado di attribuire probabilità oggettive attuariali. Gli esempi più comuni sono il lancio di una moneta o dei dadi. In questo caso una distribuzione di probabilità è una lista completa delle probabilità teoriche associate a ciascun evento mutualmente esclusivo. Quando però, come è più comune nelle situazioni reali, le nostre valutazioni sulla probabilità si basano su informazioni parziali, giudizi e intuizioni, assieme a criteri oggettivi, il risultato sarà una distribuzione delle probabilità *soggettive*.

Alla distinzione fra situazioni in cui è possibile un calcolo oggettivo delle probabilità e quelle invece in cui non si hanno gli elementi per un calcolo di questo tipo, fa riferimento la distinzione fra incertezza e rischio introdotta da Knight. Il primo termine descrive una condizione di forte ignoranza che rende impossibile l'attribuzione di probabilità per ogni possibile evento, o in cui comunque qualsiasi valutazione sul grado di probabilità sarebbe ritenuta troppo poco affidabile. Una condizione di rischio è invece quella in cui un soggetto è in grado di attribuire una diversa probabilità a ciascun esito, e può avere fiducia nella propria capacità valutativa. Tuttavia, la questione più rilevante per la teoria della scelta, e su cui si basa quella distinzione, riguarda l'alternativa fra una concezione oggettiva e una soggettiva di probabilità: la distribuzione di probabilità può essere ritenuta oggettiva, in quanto corrisponde all'effettiva distribuzione degli esiti, o soggettiva, se invece si

ritiene che rispecchi semplicemente le credenze individuali circa la distribuzione di quel particolare evento. Il tentativo di Keynes di definire la probabilità in termini oggettivi è particolarmente interessante perché il suo è “un approccio ‘razionalista’, centrato sul grado di fiducia che è ragionevole avere in un certo evento, dato lo stato delle conoscenze¹⁶”. Keynes critica l’approccio frequentista, poiché è scettico riguardo alla validità del metodo induttivo, e cioè sulla possibilità di inferire da una serie di osservazioni, seppure ripetute, una qualche regolarità che possa essere considerata affidabile. Pertanto definisce la probabilità come “degree of rational belief” in un’ipotesi basata sull’evidenza disponibile. La relazione logica che collega l’ipotesi all’evidenza disponibile è una relazione oggettiva di probabilità, che però dipende dallo stato delle conoscenze, che è diverso da individuo a individuo e può cambiare nel tempo. La distinzione fra incertezza e rischio è invece respinta soprattutto da chi, come il matematico ed economista Frank Ramsey, sostiene una concezione soggettiva della probabilità, che identifica le valutazioni probabilistiche con le credenze individuali dell’agente, rivelate dalle scommesse sugli esiti. Si tratta di una “belief revelation theory¹⁷”, analoga alla teoria delle preferenze rivelate di Samuelson.

Ad ogni modo, la descrizione del contesto di scelta diventa certamente più complicata quando avviene in assenza di informazioni complete, e soprattutto quando il risultato dipende da fattori che sfuggono al nostro controllo. Tali scelte possono essere espresse attribuendo dei valori di

¹⁶ A. Roncaglia, *La ricchezza delle idee* cit., p. 429.

¹⁷ P. Anand, *Foundations of rational choice under risk*, Oxford, Clarendon Press, 1993, p. 7.

probabilità agli esiti in ciascuno stato di natura. Per questo si parla in questi casi di *games against nature*. In presenza di tali forme di incertezza, il modello di von Neumann e Morgenstern mostra come, con l'aggiunta di ulteriori assiomi, si possano ordinare le preferenze individuali in modo da massimizzare la propria utilità *attesa*. In condizioni di incertezza sugli stati del mondo implicati dal procedimento di scelta, l'utilità attesa di un'azione è data dalla somma dell'utilità delle conseguenze dell'azione, moltiplicate per la loro probabilità di realizzarsi. Quindi, una volta stabilita l'utilità di un evento, quanto maggiori saranno le probabilità legate al suo verificarsi tanto maggiore sarà la sua utilità attesa. Nel caso in cui i due eventi siano ugualmente probabili, si preferirà quello che ha una maggiore utilità.

Nella rappresentazione in condizioni di incertezza, la scelta è fra prospetti o lotterie, in cui ogni opzione di scelta è associata a un possibile esito, e ciascun esito ha una data probabilità numerica. L'esempio più semplice di scelta con esito incerto è l'acquisto di un biglietto di una lotteria al costo di 1 euro con l'1% di possibilità di vincere 100 euro. La scelta di non acquistare il biglietto (100% di probabilità di non vincere né perdere nulla) è messa a confronto con un prospetto in cui si hanno 99% di probabilità di perdere 1 euro oppure 1% di probabilità di vincere 99 euro. La scelta di acquistare il biglietto può quindi essere rappresentata da un prospetto formato da una coppia di possibili esiti e dalle rispettive possibilità: (- 1 E, 99 E; 99%, 1%). Gli assiomi previsti dalla teoria Neumann Morgenstern per garantire la coerenza delle preferenze fra prospetti comprendono oltre a *completezza* e *transitività*, l'assioma di *continuità* che, una volta applicato ai prospetti, viene leggermente modificato. Immaginando che fra tre diversi prospetti x, y, z il

primo sia preferito al secondo e il secondo al terzo, esisterà una qualche probabilità p che una combinazione del primo e del terzo sia considerata indifferente (ugualmente preferita) rispetto al secondo. Un altro assioma è la **monotonicità**, che prevede la crescita della preferenza in parallelo alla crescita della probabilità, per cui, dati due prospetti che offrono identici guadagni, si preferisce quello con la probabilità più alta. Infine, l'assioma di **indipendenza**, richiede che per qualunque prospetto, una qualunque componente del prospetto o il prospetto stesso possa essere sostituito da una componente o da un prospetto che gli sia indifferente; il nuovo prospetto così ottenuto sarà indifferente rispetto al prospetto originale. Significa quindi, detto altrimenti, che sono indifferenti fra loro due lotterie che differiscono solo per il fatto che una offre x e l'altra y , quando x è indifferente rispetto a y . La presenza di una terza alternativa, z , non dovrebbe modificare la relazione di preferenza fra x e y . Le alternative rappresentano stati del mondo fra loro incompatibili. L'idea di fondo è che l'assioma di indipendenza assume che l'utilità derivata da ciascuna alternativa non è influenzata, né dipende dalle utilità derivate dalle altre alternative e viceversa.¹⁸

1.5 La sintesi di Savage. Subjective Expected Utility (SEU)

In *The Foundations of the Statistics* si trova quella che viene considerata l'esposizione più soddisfacente della teoria della scelta razionale. Anand

¹⁸ J. Bonner, *Introduction to the theory of social choice*, Baltimore, The John Hopkins University Press, 1986, p.164.

sottolinea il fatto che solo con Savage gli assiomi sono interpretati come assunzioni che nessuna persona ragionevole può rifiutare.

Axiomatic methods have curried considerable favour with mathematical economists. In so far as the approach uses logic embedded in mathematical terminology to derived consequences from simple, clearly stated premises, this can hardly be thought of as a radically new way to constructing theory. What distinguishes these theories is the particular meaning of axiom that is presupposed. Axioms are intended to be not just postulates but rather assumptions which no reasonable person could reject¹⁹.

Nella versione di Savage, il modello in quanto descrive cosa l'agente razionale deve fare, ha contenuto comportamentale, ma non può essere "behaviourally testable"²⁰. Di conseguenza, secondo Savage, le eventuali violazioni degli assiomi dovrebbero essere corrette in modo che le azioni vi si conformino. Come i principi logici possono servire a scoprire eventuali incoerenze fra le nostre credenze, così gli assiomi della teoria della scelta razionale possono essere utilizzati per controllare la coerenza delle decisioni e, laddove sia possibile, per prendere decisioni complesse sulla base di decisioni più semplici²¹. Savage infatti si propone di "estendere" i principi logici servendosi di principi ugualmente accettabili in modo da riuscire a trattare

¹⁹ P. Anand, *Foundations of rational choice under risk*, Oxford, Clarendon Press, 1993, p. 10.

²⁰ *Ivi*, p. 11.

²¹ L. Savage, *The Foundations of Statistics*, New York, Wiley, 1954, p.20.

efficacemente l'incertezza²². La sua descrizione della situazione di incertezza si basa sull'individuazione di alcuni importanti concetti primitivi, come stati del mondo, eventi, atti, conseguenze, preferenze. Con *stati del mondo* si intendono gli scenari possibili, fra loro incompatibili, che possono verificarsi in concomitanza con una determinata scelta e la cui descrizione conterrà tutti gli aspetti rilevanti per l'agente. Gli *eventi* sono insiemi di stati del mondo. Le scelte sono sempre scelte fra *atti*, in cui ciascun atto viene identificato con le conseguenze a cui può dar luogo, una per ogni stato del mondo. Una *conseguenza* è tutto ciò che può accadere a una persona che è in procinto di fare la sua scelta, e la *preferenza* tra atti, come per Samuelson, è definita attraverso la scelta realmente effettuata fra di essi. Questi concetti sono utilizzati per esporre gli assiomi, che per molti aspetti sono gli stessi individuati dal modello von Neumann -Morgenstern, come i requisiti di completezza, transitività e la struttura binaria dell'ordinamento di preferenze.

La descrizione della scelta razionale di Savage è formalizzata in quattro postulati principali. Il primo assicura che un agente razionale abbia un ordine di preferenze completo e coerente su tutte le azioni concepibili. Il secondo e il terzo permettono di definire la preferenza dell'agente fra ogni coppia di azioni, in relazione al verificarsi di ciascun evento. Tali preferenze dipendono solo dalle conseguenze di quegli atti in quell'evento e sono indipendenti dalla descrizione dell'evento stesso. Il quarto postulato permette di definire una

²²“Reasoning is commonly associated with logic [...] It has therefore often been asked whether logic cannot be extended, by principles as acceptable as logic itself, to bear more fully on uncertainty” (L. Savage, *cit.*, 1954, 1972, p. 6 e p.20).

relazione di “soggettivamente più probabile di” tra eventi. In poche parole, un evento è soggettivamente più probabile di un altro se l’agente preferisce scommettere sul primo piuttosto che sul secondo²³. L’accentuazione del punto di vista soggettivo dell’agente nella sintesi di Savage è evidente soprattutto nell’assioma che declina la tesi di Samuelson delle preferenze rivelate dalle scelte nella trattazione della probabilità. Un evento è soggettivamente più probabile di un altro se l’agente sceglie di scommettere sul primo piuttosto che sul secondo. In caso di prospetti con uguali esiti attesi, la scelta del primo dei due prospetti rivela l’attribuzione di un valore più alto di probabilità all’evento del primo prospetto. Se le preferenze dell’agente soddisfano gli assiomi, allora si può dire che sceglierà come se (*as if*) massimizasse l’utilità attesa.

Il punto decisivo della scelta in condizione di incertezza riguarda proprio il modo di affrontare l’incognita relativa allo stato del mondo che verrà a prodursi, e per Savage la probabilità è misurata attraverso la formazione delle credenze o aspettative sui possibili eventi. L’agente rivela la probabilità soggettiva di ciascun evento assegnando un numero tra 0 e 1 a ciascun evento e questi numeri misureranno la probabilità soggettiva di ciascun evento. Allo stesso modo, attribuendo un numero ad ogni conseguenza in modo da formare una graduatoria delle conseguenze che ricalchi le preferenze dell’agente, si otterranno le “utilità” delle corrispondenti conseguenze. In *Rationality in Action*, Hollis e Sugden sottolineano che il criterio di coerenza utilizzato da Savage non prevede nessun riferimento a desideri o credenze, ma si basa solo

²³ M. Hollis e R. Sugden, “Rationality in action”, *Mind*, 1993, v. 102, n. 405, pp. 1-35, p. 7.

sulle preferenze rappresentate dai numeri che misurano le utilità e le probabilità e che sono rivelate dalle scelte.

A person whose decisions are consistent in Savage's sense acts as if making complicated utilitarian calculations, using measures of utility and probability to work out the expected utility of each of the acts among which he has to choose. But this utility and probability measures do not *explain* or *justify* choices²⁴.

Come vedremo più avanti, le critiche alla Subjective Expected Utility (SEU) vanno principalmente in due direzioni: la richiesta di più razionalità, da parte di chi ritiene che una razionalità che si preoccupa solo dei mezzi non possa essere normativa, oppure di meno razionalità da parte di chi giudica possibile spiegare i meccanismi psicologici di scelta prescindendo da contenuti mentali proposizionali. Ma a mettere in discussione la teoria dell'utilità attesa sono stati in primo luogo i cosiddetti "paradossi", a cominciare dal paradosso di Allais.

1.6 Il paradosso di Allais

Durante un convegno a Parigi nel 1952 l'economista francese Maurice Allais propose ai suoi colleghi, fra i quali era presente anche Savage, di rispondere a un questionario in cui si presentava due scelte. La prima era fra due lotterie, una che prevedeva una vincita sicura di 1 milione di franchi (a) e un'altra in cui si dava la probabilità del 10% di vincere 5 milioni, l'89 di probabilità di vincere 1 milione e l'1% di probabilità di non vincere nulla (b).

²⁴ *Ibid* In corsivo nel testo.

La seconda, invece, chiedeva di scegliere fra una lotteria con il 10% di probabilità di vincere 5 milioni e 90% di non vincere nulla (c) e una lotteria con l'11% di probabilità di vincere 1 milione e l'89% di probabilità di non vincere nulla (d). La maggiori preferenze erano nel primo caso per la lotteria (a) e nel secondo caso per la (c). Anche Savage inizialmente fece questa scelta, successivamente riconobbe l'errore e la modificò, pur dichiarando di continuare a sentirsi attratto dalla prima scelta "sbagliata", un scelta che violava esplicitamente l'assioma di indipendenza.

	$p = 0,10$	$p = 0,01$	$p = 0,89$
a	1 milione	1 milione	1 milione
b	5 milioni	0	1 milione
c	5 milioni	0	0
d	1 milione	1 milione	0

Secondo l'assioma di indipendenza, infatti, la relazione di preferenza fra a e b e fra c e d non dovrebbe essere condizionata dalla componente che hanno in comune (rappresentata dall'ultima colonna). Se si sceglie (a) nel primo caso, si dovrà necessariamente scegliere (d) nel secondo caso. C'è invece la tendenza a evitare eventi molto sfavorevoli anche se hanno una probabilità molto bassa (b) e a preferire forti vincite anche se hanno a una probabilità altrettanto bassa (c). Nell'interpretazione di Kahneman e Tversky, il paradosso è il risultato di una distorsione comune nel modo di trattare l'informazione

sulle probabilità, e cioè la sopravvalutazione delle probabilità basse e la sottovalutazione di quelle alte. Non c'è distorsione invece nei casi di certezza o impossibilità a cui viene attribuito il loro effettivo valore. In questo modo si spiegherebbe perché la distanza in termini di probabilità fra le lotterie (a) e (b) è percepita in modo diverso dalla distanza fra (c) e (d). Come vedremo meglio più avanti, la *prospect theory* di Kahneman e Tversky introduce una *weighting function* che “corregge” la stima delle probabilità tenendo conto di questa distorsione. L'ottica in cui si inserisce questa spiegazione è quella di modificare la teoria per tener conto dei limiti della razionalità umana (*bounded rationality*). Ma ovviamente questo implica che il comportamento descritto dal paradosso di Allais debba essere ritenuto irrazionale, essendo il risultato di un errore di valutazione che può essere corretto.

In una diversa prospettiva, il paradosso può essere interpretato come l'effetto di un limite della teoria della scelta razionale invece che delle capacità decisionali umane. In questo caso la scelta non può essere considerata irrazionale, ma è la teoria stessa, o almeno l'assioma, ad essere messo in discussione. Secondo questo assioma, infatti, la ragione per preferire un'alternativa a un'altra deve essere ricavata da ciò che accade in uno stato del mondo considerato in se stesso, e cioè valutato separatamente da ciò che invece accade negli altri stati. La questione cruciale è se gli esiti nei diversi stati possano davvero essere valutati indipendentemente. Allais sostiene che questo non è possibile, perché possono esserci interazioni, complementarità, fra gli stati. Nel caso di una lotteria come quella che abbiamo appena descritto, l'utilità di ciascuna opzione dipende dalle altre opzioni disponibili. Allo stesso

modo, nel caso di un paniere di consumo, l'utilità di un componente di un paniere di beni non è indipendente dagli altri componenti del paniere²⁵.

Di parere contrario è Samuelson, che sostiene invece la validità dell'assioma: l'esito può essere o testa o croce, ma se esce "testa" non può uscire "croce". Di conseguenza, il valore assegnato a ciò che accade in uno stato del mondo dovrebbe dipendere soltanto da quello che succederebbe se si verificasse proprio quello stato del mondo. Dato che il verificarsi di quello stato preclude il verificarsi di qualsiasi altro stato, quello che sarebbe potuto accadere in un altro stato del mondo non dovrebbe fare nessuna differenza. Un agente razionale non dovrebbe prendere in considerazione nessun tipo di valutazione congiunta²⁶. Il paragone con il paniere di beni quindi non reggerebbe, perché mentre i beni del paniere sono consumati insieme, le opzioni della lotteria sono mutuamente esclusive.

Ci sono poi altre spiegazioni del paradosso che fanno invece riferimento al ruolo che hanno i sentimenti nella situazione di scelta. Tornando alla tabella che descrive il paradosso, la distanza fra la scelta (b) nella prima lotteria e la scelta (c) nella seconda sembra evidente. Nel primo caso, partendo dalla prospettiva di una vincita sicura nel 99% dei casi, la possibilità di non vincere nulla è fonte di forti timori e di un potenziale disappunto. Nell'altra lotteria invece le aspettative di vincita sono più contenute e la possibilità di non vincere nulla non appare altrettanto negativa. Il rilievo che hanno i sentimenti

²⁵ S. Hargreaves Heap et al., *The Theory of choice*, cit., p. 41.

²⁶ J. Broome, *Weighing Goods*, Oxford, Blackwell, 1991, p.96 sg.

di timore e di disappunto nel determinare la scelta è ignorato dall'assioma di indipendenza. Su tale linea è la spiegazione del paradosso che fa riferimento al rimpianto (*regret theory*). Se scegliendo (b) ci troviamo poi nella situazione di non vincere nulla, il confronto fra la situazione che si è verificata (0) e quella che si sarebbe potuta verificare (1 milione con certezza) può far rimpiangere (*regret*) la decisione presa. Come nel caso del disappunto, si mettono in relazione eventi mutuamente esclusivi, con la differenza che nel caso del *regret* non si tratta di una relazione fra opzioni della stessa lotteria, ma di una relazione fra la lotteria scelta e quella rifiutata.

Riassumendo, mentre secondo la *prospect theory* la violazione dell'assioma di indipendenza è dovuta a una errata elaborazione delle informazioni; secondo la *regret theory* è dovuta all'influenza esercitata dal contesto emotivo, che non consente di determinare l'utilità della lotteria prescelta prescindendo dalla valutazione delle altre lotterie²⁷.

Secondo queste interpretazioni, quindi, non è il comportamento di scelta che deve essere corretto, ma è l'assioma di indipendenza che deve essere eliminato dalla teoria dell'utilità attesa. Teorie alternative che fanno a meno dell'assioma di indipendenza mantenendo l'assioma di transitività - come la teoria dell'utilità anticipata di Quiggin, la teoria dell'utilità attesa generalizzata di Machina o la teoria di Chew dell'utilità ponderata - vanno in questa direzione.

²⁷ S. Hargreaves Heap et al., *cit.*, p. 42. Per questo le preferenze definite in funzione del rammarico possono anche non essere transitive. Della *prospect theory* si parlerà nel capitolo 2, della *regret theory* nel capitolo 4.

La proposta di Broome per evitare il paradosso descritto da Allais è invece di “affinare” la caratterizzazione degli esiti, aggiungendo il sentimento di disappunto ai payoffs monetari. Il rischio però è di indebolire troppo l’assioma della cosa sicura. A quel punto infatti sarebbe sempre possibile riscrivere gli esiti in modo da renderli compatibili con l’assioma, ma al prezzo di renderlo infalsificabile e quindi privo di valore, dal momento che non sarebbe più in grado di imporre alcun vincolo effettivo sulle preferenze²⁸. Broome pensa che sia comunque possibile introdurre un criterio per specificare gli esiti senza rendere per questo vuoto l’assioma, ogni qualvolta si riesca a dare una ragionevole giustificazione per distinguere i due esiti - come per gli esiti (b) e (c) -, se si aggiunge nel primo caso il sentimento di disappunto (*principle of individuation by justifiers*). Una descrizione degli esiti comprensiva dei sentimenti non impedirebbe quella loro valutazione congiunta che minaccia il carattere vincolante dell’assioma.

Il paradosso di Allais però contestava l’assioma di indipendenza denunciando in particolare l’impossibilità di isolare gli scenari possibili e sottolineava al contrario la loro interrelazione o complementarità. La replica di Broome è che il sentimento di disappunto, ad esempio, è la conseguenza di un processo causale che avviene all’interno di un determinato scenario, e cioè un certo tipo di riflessione su altri scenari che avrebbero potuto verificarsi: “se solo avessi deciso altrimenti, non mi troverei adesso in questa situazione”. Quindi, il suo argomento principale è che il disappunto sia comunque una “proprietà” di un determinato scenario, nonostante sia il risultato di un

²⁸ J. Broome, *Weighing Goods*, cit., p. 98 sg.

processo mentale che confronta e quindi mette in relazione scenari diversi, quello che si è effettivamente realizzato e quello che si sarebbe potuto realizzare. Tuttavia, l'argomentazione di Broome non è convincente, proprio perché quella "proprietà" discende direttamente da una valutazione congiunta degli scenari.

La tesi della *regret theory* è invece che le azioni non possano essere valutate indipendentemente l'una dall'altra. Il valore di una scelta "difficile" - cioè in condizioni di incertezza e in assenza di un predefinito ordinamento di preferenze - non può prescindere dalla capacità di prevedere i sentimenti negativi che sarebbero associati alle conseguenze di una scelta che si rivelasse "sbagliata", anche se a causa di fattori indipendenti.

1.7 Il paradosso di Ellsberg

Al centro del paradosso di Ellsberg (1961) è la reazione alla vaghezza o ambiguità definita come una scelta in cui ai payoff attesi non è possibile attribuire una valutazione probabilistica soggettiva²⁹. Ellsberg utilizzò il seguente esempio per dimostrare che l'incertezza nella scelta non è totalmente espressa dal concetto di "probabilità". Immaginate due urne, ciascuna delle quali contiene biglie rosse e nere. Nell'urna 1 ci sono 100 biglie, ma il numero di biglie rosse e nere non è noto. L'urna 2 contiene 50 biglie rosse e 50 nere. Si consideri ora una scommessa tale che, se puntate sul rosso e estraete una

²⁹ D. Ellsberg, "Risk, ambiguity and the Savage's axioms", *The Quarterly Journal of Economics*, 1961, v.75, n.4, pp. 643-669.

biglia rossa dall'urna, vincete 100 dollari di payoff; lo stesso vale per la puntata sul nero. Se, invece puntate sul colore sbagliato, il payoff è 0. Considerate per prima l'urna 1 e chiedetevi se preferite o se siete indifferenti a scommettere su una biglia rossa o su una nera (indicate rispettivamente con R1 e N1). La maggior parte delle persone è indifferente fra rosso e nero, il che implica che le probabilità soggettive dei due eventi sono uguali; cioè $p(R1) = p(N1) = .5$. Poi considerate la scelta di biglie nell'urna 2, in cui la proporzione delle biglie rosse e nere è noto. Ancora una volta la maggior parte delle persone è indifferente tra R2 e N2, il che implica che $p(R2) = p(N2) = .5$. Immaginate che ora vi si chieda di indicare se preferite estrarre una biglia rossa dall'urna 1 (dove il numero di biglie rosse non è noto) o dall'urna 2 (dove le biglie rosse sono il 50% del totale delle biglie). Di fronte a questa richiesta molte persone preferiscono l'urna 2 (piuttosto che esprimere indifferenza). Si noti che la preferenza per l'urna 2 rispetto all'urna 1 implica che $p(R2) > p(R1)$. Tuttavia, dalle precedenti scelte $p(R1) = .5$ e $p(R2) = .5$. Quindi c'è una contraddizione fra le probabilità derivate dalle scelte all'interno delle urne e quelle derivate dalle scelte fra le urne. Infine, considerate la richiesta di scegliere fra l'urna 1 e l'urna 2, se bisogna estrarre una biglia nera. Ancora una volta la maggior parte delle persone preferisce estrarre dall'urna 2, il che implica che $p(N2) > p(N1)$.

Queste scelte rivelano un' incoerenza nella valutazione delle probabilità. Anche in questo caso infatti, come nel paradosso di Allais, viene violato l'assioma della cosa sicura di Savage, oltre al principio che sia sempre possibile dare una valutazione soggettiva delle probabilità associate a qualsiasi evento.

Il paradosso segnala un'avversione all'ambiguità, e una preferenza per un'opzione basata su un'informazione precisa e oggettiva delle possibilità di vincita o di perdita. In assenza di queste informazioni non riusciamo a formarci probabilità soggettive.

Il punto in questione è l'ambiguità dell'informazione, “una qualità che dipende dalla quantità, dal tipo, dall'affidabilità e dall'“unanimità” dell'informazione e che dà luogo al grado di fiducia verso una stima di probabilità relativa”³⁰. Inoltre, l'ambiguità può essere elevata anche quando l'informazione - pur essendo sufficientemente ampia - presenti problemi di affidabilità e rilevanza. Una situazione in cui le valutazioni sulle probabilità sono ambigue o incerte, al punto che la fiducia in una particolare assegnazione di probabilità - in contrapposizione a una diversa distribuzione – sia molto bassa, è definita da Ellsberg come una situazione di alta ambiguità, in cui possono verificarsi violazioni degli assiomi di Savage. In questi casi, infatti, una regola ragionevole e “conservativa” ci spinge a dare più peso alla probabilità più bassa e a puntare sulle scelte che hanno valori attesi meno sensibili alle variazioni nella distribuzione delle probabilità. Un comportamento che può essere descritto come risultato di un'avversione all'ambiguità, ma che difficilmente può essere stigmatizzato in quanto irrazionale, dal momento che non ci sono basi evidenti per sostenere che nel lungo periodo i risultati saranno peggiori, rispetto a una scelta effettuata in conformità agli assiomi.

³⁰ D. Ellsberg, cit., p. 657.

L'influenza dei fattori emotivi è in questi casi indubitabile. Agenti avversi all'ambiguità danno importanza all'assenza di informazione, anche quando si tratta di informazioni che non modificano una decisione che avrebbero preso comunque, forse anche per mettere a tacere una forma di dissonanza cognitiva che può essere originata dal sapere che c'è qualcosa che si potrebbe conoscere ma che non si conosce.

1.8 Preference reversal. Interpretazioni di economisti e filosofi a confronto

Il fenomeno del rovesciamento delle preferenze è al centro delle discussioni fra psicologi ed economisti fin dagli anni '70 quando è stato individuato prima da Lichtenstein e Slovic e poi da Lindman³¹. La forma più comune di *preference reversal* è la seguente. Vengono proposte due scommesse, la prima presenta una probabilità relativamente alta di vincere una moderata somma di denaro (scommessa P) e la seconda prospetta invece una probabilità più bassa di vincere una somma maggiore (scommessa \$). Si chiede prima di scegliere tra le due scommesse e successivamente di dare una valutazione di ciascuna scommessa, ad esempio stabilendo il prezzo che si sarebbe disposti a pagare per partecipare a ciascuna scommessa (*reservation*

³¹ Lichtenstein and Slovic, "Reversals of preferences between bids and choices in gambling decisions", *Journal of Experimental Psychology*, 1971, vol. 89, pp. 46-55. Lindman H. "Inconsistent preferences among gambles", *Journal of Experimental Psychology*, 1971, vol. 89, pp. 390-7.

price)³². Il risultato dell'esperimento è che una parte consistente di coloro che scelgono la scommessa P nella scelta diretta danno poi una valutazione più alta alla scommessa \$, molto meno frequente è invece il rovesciamento in senso contrario.

La sfida portata alla teoria economica dell'utilità attesa da questi comportamenti è rilevante sotto due aspetti. Per un verso viene messo in discussione "il fondamentale principio che prezzi e preferenze sono sinonimi"³³, e cioè l'assunto dell'invarianza delle preferenze, sia rispetto al modo in cui le opzioni vengono descritte (probabilità di vincere, valutazione del prezzo) sia rispetto al modo in cui le preferenze vengono espresse (scelta diretta, attribuzione di valore). Per l'altro, viene violato l'assioma di transitività, data la manifesta incoerenza fra i due comportamenti, quello di scelta e quello di stima.

Si tratta di due aspetti che corrispondano a due diversi ordini di spiegazioni del fenomeno. L'interpretazione prevalente fra gli psicologi, e avanzata in primo luogo da Lichtenstein e Slovic, ritiene che le preferenze non

³² Sono stati utilizzati diversi metodi di valutazione. Chiedere ai soggetti di stabilire il prezzo minimo di vendita di ciascuna scommessa se fosse di loro proprietà; chiedere di stabilire il prezzo massimo in denaro per ciascuna scommessa a cui sarebbero disposti a comprarla; utilizzare il processo di elicitazione di Becker, DeGroot e Marschak in cui è interesse del soggetto rivelare l'equivalente effettivo in denaro che ha per lui ciascuna scommessa. Cfr. M. Machina, "Choice under uncertainty: problems solved and unsolved", *Economic Perspectives*, 1987, pp. 121- 154.

³³ S. Lichtenstein and P. Slovic, *The construction of preferences*, Cambridge, Cambridge University Press, 2006.

sono invarianti rispetto ai metodi di elicitazione - come sostenuto dalla teoria standard - e la ragione è il diverso modo in cui l'informazione viene elaborata nelle due circostanze. Si tratta sostanzialmente di due diversi processi mentali. Di fronte a una scelta fra due scommesse con prospetti è la probabilità relativa di vincita ad avere più importanza, e di conseguenza viene scelta la scommessa che ha la più alta probabilità di vincita. Più dettagliata e convincente è la spiegazione relativa alla valutazione delle due scommesse. In questo caso entra in gioco un procedimento mentale di ancoramento e aggiustamento (*anchor and adjustment*) in cui il valore di ciascuna scommessa viene stimato ancorandolo al valore della vincita più alta e poi ritoccandolo (*adjustment*) in modo da tener conto della bassa probabilità che vi è associata.

Gli economisti Loomes e Sugden³⁴ interpretano invece il fenomeno dell'inversione delle preferenze come un caso di preferenze intransitive che può essere previsto e spiegato dalla regret theory³⁵. L'argomento principale è che scegliere una scommessa che prima non si aveva è diverso dal vendere una scommessa che invece prima si aveva. Se si vende una scommessa \$ a qualcun altro c'è la possibilità che quest'ultimo si assicuri l'alta vincita, causando un rimpianto per averla venduta che è più forte di quello che si sarebbe provato se ci si fosse semplicemente limitati a non sceglierla

³⁴ Loomes G. and Sugden R., "A rationale for preference reversal", *American Economic Review*, 1983, v. 73, n. 3, pp.428-432. Cfr. Loomes, Starmer, Sugden, "Preference reversal: information processing effect or rational non-transitive choice", *The Economic Journal*, 1989, 99, pp. 140-51.

³⁵ Della regret theory si parla più diffusamente nel cap. 4.

all'inizio. Inoltre i due economisti mettono a confronto il modello standard di esperimento (scelta, stima monetaria) con un esperimento basato solo sulla scelta (*choice-only experimental design*) per dimostrare che l'inversione può essere generata anche all'interno di una situazione in cui gli agenti operano scelte dirette fra coppie di scommesse, invece di attribuire anche dei *reservation price* alle scommesse³⁶. Nel modello della scelta fra coppie di opzioni la regret theory dà conto di situazioni come queste sostituendo la funzione di utilità von Neumann - Morgenstern $U(x)$ con la funzione rimpianto/compiacimento $r(x,y)$, che rappresenta il livello di insoddisfazione o soddisfazione che l'agente sperimenta se si verifica il risultato x , quando la scelta alternativa avrebbe invece prodotto il risultato y .

Se è vero che il rammarico ha un ruolo importante nella definizione delle preferenze, ne segue che le due situazioni (cioè, nel caso dell'esperimento standard, scelta e stima) sono diverse o – in altri casi – che le preferenze possono anche non essere transitive, come consente la regret theory.

Le due interpretazioni mettono in luce la differenza di prospettiva fra psicologi ed economisti³⁷.

³⁶ Loomes, Starmer, Sugden, cit., p. 145-9.

³⁷ M. Machina, "Choice under uncertainty: problems solved and unsolved", *Journal of Economic Perspectives*, 1987, p.137. Va detto che in genere gli psicologi parlano di giustapposizione fra giudizi e scelte, mentre gli economisti parlano di comparazione di prezzi e scelte.

Va ricordato che la prima reazione fra gli economisti agli esperimenti degli psicologi è stato l'articolo di Grether e Plott³⁸ che, ammettendo esplicitamente di voler screditare il lavoro degli psicologi applicato alla teoria economica, mirava a interpretare il fenomeno dell'inversione delle preferenze come un risultato erroneo indotto dallo stesso esperimento. Dal momento che però anche negli esperimenti da loro opportunamente ridisegnati la frequenza del fenomeno restava alta, Grether e Plott finivano per riconoscerne l'incompatibilità con la teoria standard della preferenza, sottolineando con preoccupazione il rischio che venisse messa in discussione l'esistenza di un qualunque tipo di principio di ottimizzazione dietro le scelte umane e sollecitando gli economisti a trovare una spiegazione che potesse evitare questa conseguenza³⁹. Secondo la teoria economica della scelta, ogni agente possiede un ordinamento di preferenze ben definito sugli oggetti, ordinamento che è possibile portare alla luce utilizzando un qualsiasi procedimento di elicitazione affidabile, come la scelta diretta o l'attribuzione di valore. Se preferenze prestabilite e invarianza delle procedure di elicitazione sono condizioni irrinunciabili, il fenomeno del *reversal* potrà essere interpretato solo come violazione della transitività delle preferenze o dell'assioma di indipendenza⁴⁰.

³⁸D. Grether e C. Plott, "Economic theory of choice and the preference reversal phenomenon", *The American Economic Review*, 1979, pp. 623-638.

³⁹D. Grether e C. Plott, cit., p. 623.

⁴⁰ Di questo parere: C. A. Holt, "Preference reversals and the independence axiom", *American Economic Review*, 1986, 76, pp. 508-15; E. Karni, e Z. Safra, 'Preference reversals' and the observability of preferences by experimental methods, *Econometrica*,

Sul versante opposto, invece, gli psicologi negano l'esistenza di preferenze stabili e coerenti che i diversi metodi di elicitazione si limiterebbero a rivelare. Al contrario, l'idea è che le preferenze vengano costruite nel corso del processo di scelta o della definizione del giudizio valutativo e siano influenzate sia dal contesto – ad esempio dalla descrizione del problema – sia dalla stessa procedura di elicitazione. Quindi, scelta e valutazione sono processi distinti, che possono essere influenzati in modi diversi, e questo spiega le discrepanze riscontrate negli esperimenti. È evidente che il comportamento può variare anche di fronte a situazioni che gli economisti considerano identiche. Dalla prospettiva degli psicologi, il fenomeno del *preference reversal* è in realtà un esempio di un modello generale di comportamento piuttosto che una caratteristica peculiare della scelta fra scommesse⁴¹, e ogni spiegazione deve avere come premessa il fatto che il processo decisionale è un processo costruttivo, inevitabilmente esposto a diverse forme di condizionamento.

Successivamente, il caso standard considerato da Lichtenstein e Slovic è stato ulteriormente analizzato anche utilizzando situazioni sperimentali con input diversi dalle scommesse e modalità di risposte diverse da scelte e prezzi. In definitiva, le ipotesi interpretative più accreditate fra gli psicologi sono

1987, 55, 675–85 e U. Segal, “Does the preference reversal phenomenon necessarily contradict the independence axiom?”, *American Economic Review*, 1988, 78, 233–6.

⁴¹ Concordano al riguardo sia A. Tversky e R. Thaler, “Anomalies: preference reversal”, *Journal of Economic Perspectives*, 1990, pp. 201-211 sia C. Seidl “Preference reversal”, *Journal of Economic Surveys*, 2002, pp.621-655.

quella basata sulla preminenza/rilevanza (*prominence hypothesis*) e quella della compatibilità (*compatibility hypothesis*). Entrambe considerano le scommesse come scelte alternative, relative a due attributi, probabilità e vincita (*payoff*) ed entrambe muovono dall'idea che la responsabilità dell'inversione di preferenze vada attribuita al fatto che scelta e stima sono processi decisionali differenti.

Nel primo caso la causa della discrepanza è data dalla preminenza che riveste uno dei due attributi nei due diversi contesti. La probabilità nel caso della scelta diretta, la vincita nel caso della stima della scommessa. Si tratta di un'ipotesi sviluppata a partire dalla teoria lessicografica dell'eliminazione: ad ogni stadio del processo di scelta viene selezionato un aspetto con una probabilità proporzionale alla sua importanza; le alternative che non includono l'aspetto selezionato vengono eliminate⁴². Nella scelta di un'auto, ad esempio, il primo aspetto selezionato potrebbe essere l'aver ottenuto un elevato indice di sicurezza nei crash test: tutte le auto al di sotto di un certo indice vengono eliminate. Viene poi selezionato un altro requisito in base al quale sono selezionate le auto rimaste, e così di seguito fino a che non siano state eliminate tutte tranne una. La scelta fra le alternative, quindi, è fatta in base alla superiorità rispetto alle altre nel fattore considerato di volta in volta più importante.

⁴² P. Slovic, Choice between equally valued alternatives, *Journal of experimental psychology: human perception and performance*, 1975, p. 281. Anche A. Tversky inizialmente sostiene questa ipotesi in "Elimination by aspect: a theory of choice", *Psychological Review*, 1979, pp. 281-99.

Nella seconda ipotesi, invece, la discrepanza fra scelta e valutazione è indotta dalla compatibilità di scala⁴³. L'idea di compatibilità fra stimolo e risposta è al centro degli studi sulle prestazioni percettive e motorie, e viene utilizzato in questo caso per spiegare la più importante fonte di *preference reversal*, e cioè la sopravvalutazione delle scommesse con probabilità bassa e vincita alta. In questo caso infatti l'attenzione si focalizza sulle componenti dell'input (*stimulus components*) che sono più compatibili con la modalità di risposta (*response mode*). Se stimolo e risposta non fossero confrontabili sarebbero necessarie ulteriori operazioni mentali per stabilire una corrispondenza fra l'uno e l'altro, con il rischio di accrescere la probabilità di errore e di ridurre l'impatto dello stimolo. La controprova è data da esperimenti in cui gli esiti non sono monetari e l'incidenza del fenomeno di inversione diminuisce di quasi il 50%.

In realtà le due ipotesi avanzate dagli psicologi non sono necessariamente alternative. Per entrambe la tesi di fondo è che il peso degli attributi sia influenzato dal metodo di elicitazione. Procedure diverse mettono in luce aspetti diversi delle opzioni e così inducono a attribuire pesi diversi, ma il peso di una componente dell'input viene rafforzato dalla sua compatibilità - in termini di scala di misure - con l'output. Oppure, l'effetto compatibilità causa la differenza fra confronto di probabilità e confronto di

⁴³ P. Slovic, D. Griffin, A. Tversky, Compatibility effects in judgment and choice, 1990, in R. M. Hogarth (ed.) *Insights in Decision Making: theory and applications*, Chicago, The University of Chicago Press, 1990, pp. 5-27.

vincite, mentre l'effetto preminenza contribuisce alla maggiore attrattiva della scommessa P (alta probabilità vincita modesta) nella scelta diretta.

Un modo per ricomprendere la prima ipotesi all'interno della seconda è offerto da una nozione diversa di compatibilità, riferita alle strategie (*strategy compatibility*⁴⁴), e basata sulla distinzione fra strategie di scelta qualitative e quantitative. Le prime sarebbero basate su criteri puramente ordinali, le seconde su confronti o valutazioni di dimensioni. È più probabile che la prima strategia sia utilizzata in un contesto qualitativo come la scelta, e perciò l'effetto preminenza può essere attribuito alla compatibilità fra la natura qualitativa del compito e quella della strategia che si richiede. Un compito qualitativo come la scelta richiede quindi l'uso di una strategia qualitativa come la regola lessicografica, la preminenza non è altro che la strategia lessicografica utilizzata nelle scelte in cui un attributo è giudicato più importante di un altro. Nel caso di scommesse senza perdite, la probabilità è più importante della vincita e perciò dominerà la scelta.

La distanza fra economisti e psicologi si riduce con la *behavioral economics* che sostituisce il paradigma neoclassico della scoperta del valore (*value discovery*) con il nuovo paradigma della costruzione di valore (*value construction*)⁴⁵. Seidl sottolinea come la *behavioral economics* metta al centro

⁴⁴A. Tversky, S. Sattah, P. Slovic, "Contingent weighting in judgement and choice", *Psychological review*, 1988, pp. 371-384.

⁴⁵ "The notion of constructive preferences means as well that preferences are not necessarily generated by some consistent and invariant algorithm such as expected value calculation" in

i problemi generati dalla violazione della procedura di invarianza e dalla descrizione di invarianza, due capisaldi della teoria neoclassica, dal momento che sia le modalità di elicitazione sia la struttura (*framing*) del problema si sono dimostrati importanti per la costruzione delle preferenze individuali⁴⁶.

Il *preference reversal* non è il solo esempio di violazione di procedure di invarianza. Tra i più studiati Seidl cita ad esempio, l'*anchoring* che spinge valori e preferenze degli agenti nella direzione indicata dall'ancora, cioè dal punto di riferimento; il *background contrast effect* che rende attraente o meno un'alternativa a seconda se sia messa a confronto con alternative meno o più attraenti, il *tradeoff contrast effect* che fa sì che la relativa scarsità di attributi di alternative di scelta influenzi, nelle alternative presentate successivamente, il peso degli attributi di una opzione. Ancora, l'*asymmetric dominance effect* rileva che la disponibilità di una scelta alternativa Z, che è dominata da X ma non da Y, sposta le preferenze su X; a causa dell'effetto dotazione (*endowment effect*) le persone, per un oggetto in loro possesso, domandano più di quanto siano disposti invece a pagare per acquistarlo; il pregiudizio della disponibilità (*availability bias*) è invece la tendenza a giudicare la probabilità di un evento in funzione della facilità con cui quell'evento è disponibile alla mente, il che il più delle volte dipende dalla diffusione che ha sui media⁴⁷. Dal punto di vista degli psicologi anche l'intransitività è una forma di *preference*

J.W. Payne et al., Behavioral decision research: a constructive processing perspective, *Annual Review of Psychology*, 1992, 43, p. 89, citato da C. Siedl, *cit.*, p. 646.

⁴⁶ C. Siedl, *cit.*, p. 646.

⁴⁷ *Ivi*, p.647.

reversal. Come vedremo più avanti, il *framing effect* studiato da Kahneman e Tversky⁴⁸ si riferisce a una situazione in cui le preferenze vengono rovesciate - come nel caso più noto di violazione di *description invariance* - semplicemente cambiando le parole per riferirsi al destino di 600 persone, ad esempio utilizzando una prima volta l'espressione "400 persone moriranno" e una seconda volta "200 persone si salveranno"⁴⁹.

Come è stato osservato⁵⁰, sulla natura delle preferenze e dei valori sono possibili tre diverse concezioni. In base alla prima, i valori esistono e le persone li percepiscono e li descrivono nel modo migliore possibile, verosimilmente con qualche errore. La seconda concezione, invece, sostiene che le persone conoscano valori e preferenze direttamente. Nell'ultimo caso, valori e preferenze vengono costruiti durante il processo di elicitazione. Le interpretazioni del fenomeno di *preference reversal* date dagli psicologi sono in linea con quest'ultima concezione che - come abbiamo visto - vede le preferenze come risultato di un processo costruttivo, dipendente dal contesto.

⁴⁸ Cap. 2 di questo lavoro.

⁴⁹ A. Tversky, D. Kahneman, *The framing of decisions and the psychology of choice*, Science, 1981, v. 211, pp.453-8, p. 453.

⁵⁰ A. Tversky e R. Thaler, *cit.*, 1990, p. 210.

1.9 Conseguenze delle violazioni degli assiomi sulla teoria della scelta razionale

Per tentare un primo bilancio dei risultati di queste tre situazioni sperimentali “paradossali” (Allais, Ellsberg, preference reversal) e delle ricadute che hanno sulla teoria dell'utilità attesa è importante fare chiarezza sulla natura della teoria stessa che si propone di servire sia come modello normativo di come un agente pienamente razionale sceglierebbe, sia come fondamento per teorie predittive dei comportamenti effettivi di scelta di persone reali in contesti economici, politici e sociali⁵¹. Questa caratterizzazione propone sia un'interpretazione normativa sia un'interpretazione descrittiva della teoria dell'utilità attesa. Il punto però è che queste due prospettive indicano compiti diversi che difficilmente possono essere svolti da una stessa teoria, dato che si riferiscono a due diversi soggetti: rispettivamente all'agente ideale e all'agente reale.

Nel primo caso la teoria della scelta razionale intende definire, a partire da criteri a priori di razionalità formalizzati dagli assiomi, le modalità di comportamento di un'agente ideale, astraendo sia dalle caratteristiche specifiche e dai limiti cognitivi della razionalità umana sia dalle diverse componenti emotive che influenzano i comportamenti effettivi di scelta. Di conseguenza, la teoria non avrà un diretto potere predittivo sul comportamento degli agenti reali, ma fornirà piuttosto un modello di

⁵¹ R. Sugden, *How people choose*, in Shaun Hargreaves Heap, *The theory of choice: A critical guide*, cit., pp. 36-49

comportamento a cui conformarsi per quanto possibile, e cioè compatibilmente con i limiti umani strutturali e con le conoscenze di cui si può disporre. Un noto esempio di Friedman e Savage⁵² fa notare che le complesse equazioni della meccanica dei corpi rigidi assieme alla geometria piana sono in grado di prevedere perfettamente il modo in cui dei giocatori esperti di biliardo vanno in buca, anche se i giocatori sono del tutto a digiuno di meccanica e geometria. La ragione è che il modello geometrico riesce a cogliere e a rappresentare efficacemente quello che i giocatori tentano di fare, dal momento che si tratta di giocatori esperti, che hanno alle spalle anni di allenamento e di riscontri che li hanno messi in condizione di arrivare molto vicino al comportamento ottimale. Questa analogia però è stata giudicata di scarso valore⁵³, dal momento che la maggior parte delle persone non è esperta di economia e inoltre imparare dai riscontri non è affatto semplice né automatico nel caso delle decisioni quotidiane, dove l'incertezza e la variabilità dei contesti frappongono troppi ostacoli alla possibilità di imparare dall'esperienza. L'idea era comunque che il comportamento di scelta degli agenti potesse essere rappresentato come conforme agli assiomi, e cioè ai requisiti di coerenza e transitività, quasi che – “as if” – gli agenti li utilizzassero consapevolmente per le proprie decisioni.

⁵² M. Friedman e L. Savage, “The utility analysis of choices involving risk”, *The Journal of political economy*, 1948, 56, pp. 279-304, p. 298.

⁵³ Fra gli altri R. Thaler, “Toward a positive theory of consumer choice”, *Journal of Economic Behavior & Organization*, 1980, 1, pp. 39-60, p. 22. “The orthodox economic model of consumer behaviour is, in essence, a model of robot-like experts”.

Se invece, come nel secondo caso, il riferimento è l'agente reale, una teoria che voglia essere "fondamento" per previsioni di scelte reali dovrà probabilmente rinunciare a una forte assiomatizzazione, e far posto ad altri fattori che hanno un ruolo rilevante nei comportamenti effettivi. La differenza fra le due impostazioni si riflette anche nel modo di interpretare i "fallimenti" della razionalità mostrati dalle violazioni degli assiomi.

Le situazioni sperimentali di scelta come quelle appena esaminate potrebbero non essere paradossali se gli errori possono essere interpretati in termini di *bounded rationality*⁵⁴, e cioè provocati dai vincoli, di tempo, conoscenze, capacità computazionali, che sono propri della natura umana. Se le cose stanno così, gli agenti, dopo avere ascoltato le spiegazioni dello sperimentatore e analizzato con maggiore attenzione la situazione, saranno disposti a riconoscere l'errore compiuto e a modificare le proprie scelte. In questo caso saremo di fronte a errori di prestazione (performance) che non metterebbero in discussione la competenza. La stessa competenza dimostrata dal giocatore professionista di biliardo nei confronti delle leggi della fisica (ancorché non esplicita) non può comunque impedirgli di incorrere in qualche errore. Se però, come accade nella maggior parte dei casi, l'agente reale non è disposto a modificare la sua scelta, si può ipotizzare che i paradossi rivelino delle insufficienze della stessa teoria della scelta. Ad esempio, quando la scelta viene fatta in condizione di incertezza, come nel paradosso di Ellsberg, la teoria dell'utilità attesa non sembra in grado di operare un calcolo razionale che individui l'esito ottimale.

⁵⁴ Se ne parla più estesamente nel capitolo successivo.

In definitiva, sembra che una teoria che voglia proporsi come fondamento per teorie predittive dei comportamenti effettivi di scelta, oltre a definire a priori i parametri di razionalità attraverso gli assiomi, debba dar conto dei diversi meccanismi psicologici (nonché dei limiti cognitivi, informativi, ecc.) che intervengono sia nella rappresentazione del contesto della scelta, sia nella prefigurazione degli esiti possibili condizionando le decisioni dei soggetti reali. Un esempio è il *regret*, che sarebbe responsabile della violazione dell'assioma di indipendenza, ma la cui influenza sulla scelta può comunque essere prevista e quantificata⁵⁵. In generale, le violazioni degli assiomi possono offrire uno strumento importante per correggere e arricchire il modello di comportamento previsto dalla teoria.

La definizione di “razionalità” della scelta costituisce una differenza significativa fra queste due diverse prospettive. La teoria dell'utilità attesa fa riferimento a un agente ideale a cui è attribuita una razionalità, in linea di principio illimitata, definita dalla coerenza delle sue scelte effettuate in base a calcoli che hanno come obiettivo la massimizzazione dell'utilità attesa. Rinunziando ad avere un contenuto empirico, questa teoria può essere descritta come un insieme di principi formali che riguardano la relazione fra proposizioni sulle preferenze o scelte⁵⁶. Da questa prospettiva la razionalità della scelta non è altro che la conformità con gli assiomi. Ma per una teoria incentrata sull'agente reale questa definizione di razionalità può apparire insoddisfacente, non solo perché, come si è già detto, non tiene conto delle

⁵⁵ Della regret theory si parla nel capitolo 4.

⁵⁶ R. Sugden, *cit.*, p. 47

limitate capacità umane, ma anche perché lascia fuori molti altri fattori che permetterebbero di rappresentare – e prevedere – meglio le scelte effettive. Come vedremo nel secondo capitolo, la behavioral economics si propone appunto di delineare un modello in un certo senso più realistico del comportamento di scelta, e questo comporta per un verso un indebolimento degli assiomi e in generale dei requisiti di razionalità, e per altro verso la volontà di fare posto a quei meccanismi psicologici non necessariamente razionali che, come gli esperimenti mettono in luce, hanno un'influenza determinante nelle scelte.

I due diversi compiti - quello di servire sia come modello normativo di come un agente pienamente razionale sia come fondamento per teorie predittive dei comportamenti effettivi di scelta di persone reali - andrebbero tenuti distinti pur essendo per molti versi complementari. Distinguere le due prospettive aiuta anche a comprendere la differenza di prospettive fra le interpretazioni delle situazioni sperimentali di scelta proposte dagli economisti che difendono la validità del nucleo teorico della EUT e quelle degli psicologi che al contrario la ritengono inadeguata.

La tendenza dei primi è quella di proporre modifiche o riforme della teoria⁵⁷, facendo a meno ad esempio dell'assioma di indipendenza (Machina) o

⁵⁷ M. Machina's, "Expected utility" analysis without the independence axiom", *Econometrica*, 1982, 50, pp. 277-323;

S. Chew and K.R. MacCrimmon, "Alpha Utility Theory, lottery composition, and the Allais paradox", *Working paper*, University of British Columbia, Vancouver, 1979; J. Quiggin, "A theory of anticipated utility", *Journal of Economic Behavior and Organization*, 1982, 3(4),

indebolendo l'assioma di transitività (Broome). Ma la caratteristica principale del modello dell'utilità attesa è comunque data dal fatto che la razionalità dell'agente ideale è considerata indipendente dai diversi fattori che compongono il contesto della scelta (cioè dalle scelte alternative, o dalla rappresentazione o framing del problema, dai processi psicologici, emotivi). Anche l'ipotesi di Broome⁵⁸ di darne conto, ad esempio aggiungendo il sentimento di disappunto ai payoffs monetari, sembra insoddisfacente – lo stesso Broome cerca di difendersi dall'accusa che si tratti di un'ipotesi ad hoc - dato che l'incidenza del disappunto o del rammarico porta a considerare anche le possibilità controfattuali, modificando in modo sostanziale il modo in cui interpretiamo la scelta. La prospettiva degli psicologi – integrata nella behavioral o cognitive economics – cerca di ridefinire le basi della teoria attraverso una più approfondita comprensione dell'intelligenza umana. Si tratta di riconsiderare l'incidenza di questi fattori non in quanto elementi di disturbo, ma per il ruolo sicuramente rilevante che rivestono nel processo decisionale. Le divergenze all'interno di questa prospettiva, come vedremo, riguardano la relazione fra il piano della razionalità e quello dei fattori psicologici o emotivi, e cioè se vanno considerati alternativi o invece complementari.

pp. 323–43 ; M. Yaary, “The dual theory of choice under risk”, *Econometrica*, 1987, 55, pp.95–115.

⁵⁸ J. Broome, *Weighing Goods*, cit., pp. 95-115.

Capitolo 2 - La prospettiva della psicologia cognitiva

2.1 Riunificare psicologia ed economia. Verso la *behavioral economics*

La presa di distanze dalla psicologia che ha caratterizzato la teoria economica del secolo scorso può essere spiegata da due opposte tendenze. Da un lato, prendendo la fisica come modello, la teoria economica si muoveva verso la formalizzazione matematica; dall'altro, la psicologia si ricollegava alla tradizione sperimentale delle scienze naturali. La divergenza nasce sul piano metodologico: tra una teoria costituita da teoremi e strumenti matematici e una intesa come una struttura per organizzare regolarità sperimentali.

Con la svolta positivista degli anni '40, la teoria economica sarà spinta a interpretare la scelta come l'esito logico di un calcolo razionale che rispetta in modo coerente vincoli e preferenze, e che può prescindere da ogni riferimento ai processi mentali degli individui. Il carattere fortemente normativo che viene ad assumere in questo modo la razionalità viene sancito dall'interpretazione data da Friedman della teoria della decisione come teoria "positiva". La nuova svolta metodologica di Friedman – la cosiddetta "F twist" – sancisce l'inutilità di qualsiasi forma di verifica sperimentale della validità della teoria dell'utilità attesa, assumendo che le preferenze individuali non siano osservabili, e ritenendo invece il comportamento di aggregati di individui come più ragionevolmente rappresentato e formalizzato dalla curva dell'utilità attesa. La posizione di Friedman viene così descritta da Camerer:

Because theories with patently false assumptions can make surprisingly accurate predictions, economic theories that assume that individual agents are highly rational and willful, judge probability accurately, and maximize their own wealth might prove useful, even though psychology show that those assumptions are systematically false⁵⁹.

Negli stessi anni, gli esperimenti sulle preferenze individuali di Maurice Allais, che dimostravano la fallacia degli assiomi della EUT, non vennero inizialmente valutati in modo adeguato proprio perché assolutamente in controtendenza con l'impostazione dominante. In seguito, la reazione alla violazione degli assiomi porterà a modificare o ad attenuare qualche assioma della teoria in modo da ottenere previsioni più vicine ai comportamenti reali.

Sempre negli anni '50, nel diverso contesto dell'analisi del comportamento amministrativo, anche le osservazioni empiriche sui processi decisionali nelle imprese mostrano i limiti della razionalità individuale e inducono Herbert Simon a teorizzare la "bounded rationality". Per Simon non si tratta soltanto di riconoscere i limiti della capacità computazionale umana, in termini di tempo e di energie mentali, ma anche di ripensare il problema della decisione, focalizzando l'attenzione sul contesto della decisione stessa. La nozione di razionalità limitata serve infatti a prendere in considerazione il complesso processo che precede l'atto finale di scelta e che consiste nella raccolta delle informazioni rilevanti, ma soprattutto nella costruzione di un

⁵⁹ C. Camerer, "Behavioral economics: reunifying psychology and economics", *Proceeding of the National Academy of Sciences of the United States of America*, 1999, v.96, n.19, pp. 10575-10577, p. 10575.

modello mentale che permetta un'efficace rappresentazione del problema. I limiti della razionalità si manifestano proprio nella difficoltà di riuscire a organizzare una rappresentazione mentale appropriata e completa di tutte le alternative. Dato che l'aspetto cruciale del processo decisionale è individuato nell'abilità di costruire nuove rappresentazioni dei problemi e di individuare strategie alternative, il nuovo campo di indagine delle ricerche di Simon dopo la teoria dell'organizzazione sarà la psicologia cognitiva. Di conseguenza, anche nello studio sulla razionalità l'attenzione si sposta dalla coerenza/incoerenza delle scelte alla rappresentazione e all'elaborazione dei problemi⁶⁰. Significativo è il titolo di un saggio del 1979 di Simon: "From substantive to procedural rationality", in cui spiega di volersi riferire con il primo termine al concetto di razionalità che si è affermato all'interno della teoria economica e con il secondo al concetto che si è sviluppato all'interno della psicologia. Le due diverse concezioni di razionalità si giustificano con la diversa tipologia di problemi che interessano i due programmi di ricerca.

Behaviour is substantively rational when it is appropriate to the achievement of given goals within the limits imposed by given conditions and constraints. Notice that, by definition, the rationality of behavior depends upon the actor in only a single respect – his goal⁶¹.

⁶⁰ M. Egidi, L. Marengo, "Cognition, institutions, near decomposability: rethinking Herbert Simon's contribution", *CEEL, Working Paper*, 2002, n.6, pp. 1-13, p. 5.

⁶¹ H. Simon, "From substantive to procedural rationality", in *Philosophy and economic theory* edited by F. Hahn and M. Hollis, New York, Oxford University Press, 1979.

I due presupposti della teoria economica classica sono “that the actor has a particular goal”, ad esempio la massimizzazione del profitto, e “that the economic actor is substantively rational”. Sono questi due assunti a rendere la teoria economica completamente indipendente dalla psicologia, e finché non saranno messi in discussione gli economisti non avranno nessuna ragione di interessarsi agli studi psicologici sui processi cognitivi della scelta.

Behavioural is procedural rational when it is the outcome of appropriate deliberation. Its procedural rationality depends upon the process that generated it. When psychologists use the term “rational” it is usually procedural rationality they have in mind⁶².

Simon cita William James, che nei *Principles of Psychology* usa il termine “rationality” come sinonimo di “the peculiar thinking process called reasoning”. Forse è proprio perché gli psicologi sono più interessati al processo che al risultato che preferiscono parlare di “processo cognitivo” quando parlano della razionalità in riferimento ai comportamenti. Secondo Simon questa variazione terminologica può aver dato un ulteriore contributo alla separazione reciproca dei concetti di razionalità sostanziale e procedurale, ma i nuovi programmi di ricerca dell’economia comportamentale riportano l’attenzione sulla comprensione dei processi cognitivi su cui si basa la razionalità umana.

Negli anni ’70 infatti gli psicologi cognitivi cominciano a studiare “judgment and economic decision making”; prendono la massimizzazione

⁶² *Ibid*

dell'utilità e le regole logiche dei giudizi sulla probabilità come punti di riferimento e utilizzano la conformità o la deviazione da questi punti di riferimento per elaborare teorie sui meccanismi cognitivi. I risultati delle loro ricerche sono principi o costrutti psicologici che possono essere espressi in termini formali, in modo da costruire un modello di razionalità limitata più accettabile anche per gli economisti.

La behavioral economics cerca di incorporare questo tipo di ricerche nella teoria economica, provando a ridefinire i fondamenti della decisione attraverso un esame più approfondito dell'intelligenza umana, e ripensando la relazione fra psicologia ed economia lungo le linee della razionalità limitata di Simon.

"Behavioral economics" replaces strong rationality assumptions used in economic modelling with assumptions that are consistent with evidence from psychology, while maintaining an emphasis on mathematical structure and explanation of naturally-occurring (field) data. Of course, for decades social scientists have criticized economic models for assuming too much rationality, and economists defend the models as useful approximations. In behavioral economics we believe more realistic assumptions will make for better approximations. The only interesting question is how to incorporate the psychology into economics.

2.2 Heuristic and bias program

Nel dicembre del 2002 Daniel Kahneman inizia la sua Nobel⁶³ *lecture* descrivendo il lavoro per cui è stato premiato, assieme ad Amos Tversky: “we explored the psychology of intuitive beliefs and choices and examined their bounded rationality”⁶⁴. Subito dopo elenca quelli che sono stati gli argomenti più importanti con cui si sono confrontati: “heuristics of judgement, risky choice and framing effect”. A questi argomenti sono dedicati due importanti articoli pubblicati su *Science*, nel 1974 *Judgement under uncertainty: Heuristics and biases*, e nel 1981 *The framing of decisions and the psychology of choice*. Sicuramente però l’articolo che ha avuto maggiore risonanza – anche al di fuori della comunità degli psicologi – è stato *Prospect Theory. An analysis of decision under risk*, del 1979, proprio perché pubblicato su una prestigiosa rivista economica, *Econometrica*. In questi articoli, cui ne seguiranno molti altri, vengono delineate le tesi principali di quello che verrà chiamato l’*heuristics and biases approach* allo studio dei giudizi in condizioni di incertezza, quelli su cui si basano molte delle nostre decisioni. Queste tesi sono alla base di un vero e proprio programma di ricerca

⁶³ La menzione ufficiale specifica che il premio Nobel è stato assegnato a Daniel Kahneman “per aver integrato intuizioni della ricerca psicologica nella scienza economica, specialmente nel campo del giudizio e della decisione in condizioni di incertezza”; cit. in Introduzione di M. Motterlini e M. Piattelli Palmarini a *Critica della ragione economica*, Milano, Il Saggiatore, 2005.

⁶⁴ D. Kahneman, *Maps of bounded rationality: a perspective on intuitive judgement and choice*. Prize lecture, December 8, 2002.

i cui risultati più importanti, in termini sia di esperimenti sia di teorie esplicative, sono confluiti in volumi collettanei di articoli che hanno via via aggiornato i dati e ampliato lo spettro delle teorie esplicative⁶⁵.

2.3 Euristiche del giudizio

L'idea centrale di tale programma di ricerca è che i giudizi che riguardano la probabilità di eventi incerti si basano spesso su un numero limitato di euristiche – cioè di procedure che semplificano il percorso ben più impegnativo di valutazione delle probabilità attraverso complessi calcoli formali. In genere queste euristiche si rivelano risorse molto utili al raggiungimento di giudizi sufficientemente accurati, ma talvolta possono condurre a errori gravi e sistematici. Il termine *biases* indica proprio questo genere di errori sistematici, che sono la conseguenza, a volte inevitabile, di giudizi non abbastanza meditati che diventano dei pregiudizi.

Il presupposto di questa linea di ricerca è la concezione della razionalità limitata di Simon, anche se per Kahneman e Tversky i ragionamenti intuitivi che utilizzano euristiche non rispondono solo a un'esigenza di semplificazione dei compiti, quindi a un inevitabile ridimensionamento delle pretese dei modelli di razionalità, ma sono processi mentali di un genere completamente

⁶⁵ D. Kahneman, P. Slovic e A. Tversky (eds.), *Judgement under Uncertainty: heuristics and biases*, Cambridge (Ma.), Cambridge University Press, 1982. D. Kahneman, A. Tversky (eds.), *Choices, Values, and Frames*, New York, Cambridge University, 2000, T. Gilovich, D. Griffin, D. Kahneman (eds.), *Heuristics and Biases: the psychology of intuitive judgement*, New York, Cambridge University Press, 2002.

diverso. Le prime tre euristiche individuate dai due ricercatori sono *availability*, *representativeness*, e *anchoring and adjustment*. Sono procedure semplici ed efficaci perché basate su calcoli elementari che si attivano quasi automaticamente. Così ad esempio, per valutare la percentuale di attori che fanno uso di cocaina potremmo trovarci a basare la nostra stima sul numero di casi che ci vengono in mente, quindi la stima dipenderà dalla facilità con cui quell'informazione è disponibile alla mente (euristica della disponibilità). La possibilità di errore in questo caso è dovuta alla maggiore facilità con cui ricordiamo i casi più eclatanti, che ci induce, erroneamente, a considerarli indicatori di un alto indice di frequenza. D'altro canto, dovendo rispondere a una domanda riguardante la probabilità che un dato attore sia un consumatore di cocaina, ci potremmo basare sulla somiglianza fra le caratteristiche dell'attore e quelle del "consumatore tipo" di eroina (euristica della rappresentatività), tendendo a ignorare i dati percentuali di base dei consumatori di eroina. Di conseguenza, il nostro giudizio finale risulterà distorto dal cosiddetto "effetto tipicità". Un'altra possibilità di fronte a uno dei due quesiti precedenti è quella di partire da una stima iniziale - ad esempio del 50% - e poi correggerla al ribasso (euristica dell'ancoraggio e aggiustamento). L'effetto ancoraggio può indurre a errori di valutazione, perché la stima di partenza che funge da àncora viene spesso stabilita – come risulta dagli esperimenti – in maniera del tutto arbitraria.

Sulle euristiche ci sono comunque alcune importanti precisazioni da fare. Anzitutto, nonostante i giudizi erronei che possono indurre, le euristiche in sé sono procedimenti valutativi ragionevoli e niente affatto irrazionali. Inoltre, se è vero che le soluzioni cui conducono sono veloci ma approssimative (*"quick*

and dirty”), i processi sottostanti sono invece altamente sofisticati e in generale i principi euristici non servono solo a trovare una risposta a problemi molto complessi, ma anche a fornire risposte intuitive ai più semplici problemi di probabilità e di previsione. L’aspetto sicuramente positivo è quindi la loro capacità di proporre un insieme di scorciatoie mentali altamente efficienti, che forniscono soluzioni soggettivamente convincenti e spesso assolutamente funzionali. L’aspetto negativo è che la funzionalità è comunque altra cosa rispetto all’esattezza e dall’accuratezza, e di conseguenza i giudizi intuitivi tendono spesso a ignorare le leggi della probabilità. Le tre euristiche spiegano molti degli errori in cui incorriamo in modo sistematico: quando non prevediamo la regressione ai valori medi, non diamo il peso necessario alle dimensioni del campione o non prendiamo in considerazione le percentuali di base nel fare una previsione. L’individuazione di questo tipo di errori serve anche a far luce sulle caratteristiche dei processi di ragionamento che ne sono il presupposto.

2.4 Gli errori delle euristiche della rappresentatività

Nell’articolo pubblicato nel 1974 su *Science*, Kahneman e Tversky sostengono che i giudizi sulla valutazioni di probabilità assomigliano ai giudizi in cui valutiamo quantità fisiche come la distanza o le dimensioni. Si tratta in entrambi i casi di giudizi basati su dati non ottimali, che elaboriamo in base a regole euristiche. Così, ad esempio, determiniamo la distanza di un oggetto in base alla sua “chiarezza”: quanto più mi apparirà nitido e definito, tanto più mi sembrerà vicino. Questa regola ha una certa validità, ma può condurci a errori sistematici nella valutazione della distanza.

Specifically, distances are often overestimated when visibility is poor because the contours of objects are blurred. On the other hand, distances are often underestimated when visibility is good because the objects are seen sharply. Then, the reliance on clarity as an indication of distance leads to common biases. Such biases are also found in the intuitive judgment of probability⁶⁶.

Se la distinzione con cui vediamo i contorni di un oggetto non è sempre un criterio affidabile per giudicare la distanza dell'oggetto, allo stesso modo anche i giudizi intuitivi possono portarci all'errore quando focalizziamo l'attenzione su un solo aspetto. In realtà, come si dirà più avanti, il confronto con la percezione è meno significativo di quanto possa sembrare.

Nell'articolo sopra citato lo spazio maggiore è dedicato alla euristica della rappresentatività. Si è già accennato all'esempio più noto, e cioè al modo in cui cerchiamo di capire con quale probabilità A faccia parte di una certa categoria B. Immaginiamo che di un individuo scelto a caso fra la popolazione - descritto come una persona che "porta gli occhiali, è poco socievole, è grande lettore di saggi storici" - ci sia chiesto di dire se appartiene alla categoria dei commercianti o dei bibliotecari. Probabilmente la risposta ricadrebbe sulla categoria dei bibliotecari, proprio per l'attrazione esercitata da una descrizione che soddisfa lo stereotipo del bibliotecario, trascurando il fatto che all'interno della popolazione la percentuale di commercianti è certamente

⁶⁶ D. Kahneman, A. Tversky. "Judgement under Uncertainty: heuristics and biases", in *Science*, 185, n. 415 (1974), p. 1124-1131. Poi in D. Kahneman, P. Slovic e A. Tversky (eds.), *Judgement under Uncertainty: heuristics and biases*, cit., p. 585.

maggiore di quella dei bibliotecari. La forza di questa euristica è tale che problemi analoghi utilizzati in diversi esperimenti hanno dato gli stessi risultati, anche quando veniva fornita l'informazione circostanziata sulle reali proporzioni delle due categorie lavorative nella popolazione. La presenza della descrizione catalizza tutta l'attenzione, persino quando è assai poco accurata e non fornisce informazioni rilevanti; solo in sua assenza si dà la dovuta importanza alle probabilità a priori⁶⁷. Altri esempi sono l'insensibilità alla dimensione del campione statistico - la probabilità di un campione statistico valutata sulla base della rappresentatività non tiene conto della dimensione del campione - e la legge dei piccoli numeri, ovvero la tendenza a credere statisticamente vero per le piccole serie ciò che è vero - e solo approssimativamente - per serie molto lunghe. Se si tratta di prevedere, ad esempio, i futuri profitti di un'azienda, di fronte a una descrizione lusinghiera della sua attività, una previsione di profitto molto alta sembrerà più rappresentativa di quella descrizione⁶⁸. Gli esperimenti ci dicono anche che il valore attribuito a una descrizione positiva sembra non tener conto del grado di affidabilità delle informazioni che vi sono contenute o della loro rilevanza rispetto alla valutazione dei profitti futuri.

⁶⁷ La pragmatica della comunicazione suggerisce che quelle informazioni possono essere erroneamente ritenute assolutamente determinanti per la soluzione del problema, e ciò può indurre a sottostimare le probabilità a priori. Cfr. C. Bianchi, *La pragmatica del linguaggio*, Roma-Bari, Laterza, 2003.

⁶⁸ *Ivi*, p. 589.

In questi casi la previsione avviene individuando il risultato che rappresenta meglio l'input iniziale (lo stereotipo del bibliotecario, gli alti profitti dell'azienda). La fiducia riposta nella previsione sembra quindi dipendere prevalentemente dal grado di rappresentatività, cioè dalla qualità della corrispondenza fra la risposta e l'input, senza tener conto di quei fattori che potrebbero limitare l'attendibilità delle informazioni contenute dell'input.

The unwarranted confidence which is produced by a good fit between the predicted outcome and the input information may be called the illusion of validity. This illusion persists even when the judge is aware of the factors that limit the accuracy of the prediction⁶⁹.

Anche la coerenza interna del modello di input può incrementare la fiducia nella validità delle nostre previsioni, perciò siamo più sicuri nel prevedere la media scolastica finale di uno studente il cui curriculum del primo anno consiste interamente di B, di quanto lo siamo nel caso di uno studente che ha nel suo curriculum molti A e C, che pure porterebbero alla stessa media finale. Ma può essere una fiducia ingiustificata, perché la statistica della correlazione insegna che la ridondanza degli input può dar luogo a una previsione meno affidabile di quella basata invece su variabili di input indipendenti. In definitiva, i risultati costanti ottenuti dal primo studente potrebbero essere meno informativi rispetto a quelli ottenuti dal secondo, e la previsione sul punteggio finale potrebbe essere più accurata nel secondo caso.

⁶⁹ *Ivi*, p. 590.

Un altro errore sistematico che può conseguire da questo tipo di euristica è dato dall'incapacità di renderci conto della regressione ai valori medi: gli istruttori di volo notano che gli elogi dopo un atterraggio eccezionale sono seguiti da un atterraggio mediocre, mentre a un'aspra reprimenda dopo un atterraggio brusco generalmente seguirà un atterraggio migliore nel tentativo successivo. Di conseguenza, si ritiene che i rimproveri siano efficaci nell'indurre un miglioramento delle prestazioni, mentre gli elogi sembrano indurre un abbassamento del livello delle prestazioni. In realtà, molti esperimenti hanno dimostrato che le critiche e gli elogi hanno scarsa responsabilità, un atterraggio particolarmente eccellente o al contrario particolarmente insoddisfacente sono risultati estremi e dunque improbabili, che statisticamente saranno seguiti da risultati più vicini ai valori medi. Il fenomeno resta comunque elusivo perché è incompatibile con la forte tendenza a credere che il risultato previsto debba essere massimamente rappresentativo dell'input e che quindi il valore della variabile risultato debba essere altrettanto estremo del valore della variabile input.

2.5 Gli errori delle euristiche della disponibilità e dell'ancoraggio

La ricorrenza di una classe o la probabilità di un evento può essere valutata dalla facilità con cui ci vengono in mente esempi o occorrenze. La "disponibilità mentale" è un indizio molto utile per valutare frequenze o probabilità, dato che riusciamo meglio e più rapidamente a trovare esempi quando si tratta di classi o di eventi che ricorrono con maggiore frequenza. Tuttavia, ci sono anche altri fattori che possono influenzare l'euristica della disponibilità. Gli esperimenti mostrano che si tende a sovrastimare la

probabilità di eventi familiari o salienti, sia perché la diffusione sui media di un data categoria di eventi li rende più facilmente richiamabili alla mente rispetto ad altri che non hanno la stessa risonanza, sia perché ci hanno colpito di più e si sono impressi con maggiore forza nella memoria. In definitiva, si tende a considerare più probabile ciò di cui si parla di più e ciò che suscita emozioni molto forti. Talvolta valutiamo la probabilità di una classe di eventi sulla base della quantità di esempi che riusciamo noi stessi a generare: la facilità con cui immaginiamo i diversi pericoli che potrebbero verificarsi durante un viaggio un po' avventuroso, non li rende ovviamente più probabili. Un'altra causa di errore è il giudizio sulla frequenza di due eventi correlati fra loro. La forza del legame associativo è la base per giudicare la probabilità con cui i due eventi correlati si verificano, ma può anche indurre a sovrastimare la frequenza con cui i due eventi sono effettivamente correlati, sia forzando i dati che abbiamo a disposizione, sia ignorando o manipolando i dati contrari alla loro correlazione. L'illusione della correlazione laddove non abbiamo elementi sufficienti per dimostrarla è un altro degli errori legati più all'immaginazione che alla memoria.

In molte situazioni le valutazioni vengono effettuate a partire da un valore iniziale che viene corretto fino a ottenere un dato valore finale. Il valore iniziale – che rappresenta l'ancora – può essere suggerito dalla formulazione del problema, oppure può trattarsi del risultato di un calcolo parziale. In entrambi i casi le correzioni successive si rivelano inadeguate. Lo stesso pregiudizio si verifica anche nella determinazione della probabilità del primo di due eventi congiunti o disgiunti che funge poi da ancora per la valutazione della combinazione dei due eventi. Dato che l'aggiustamento rispetto alla

prima valutazione è in genere insufficiente, la stima finale resta in entrambi i casi troppo vicina alla probabilità attribuita all'evento iniziale. Rispetto al singolo evento la probabilità complessiva di due eventi congiunti è più bassa, mentre al contrario la probabilità complessiva di due eventi disgiunti è più alta. Come conseguenza dell'ancoraggio si tenderà quindi a sovrastimare la probabilità che si verifichino eventi congiunti, e a sottostimare la probabilità di eventi disgiunti.

Diversi di questi pregiudizi cognitivi sono pervasivi, nel senso che non riguardano solo l'uomo comune ma anche persone che hanno una buona preparazione statistica. Regole statistiche come la regressione ai valori medi dovrebbero essere apprese induttivamente sulla base dell'esperienza quotidiana: se questo non accade, è perché gli esempi significativi in tal senso non vengono adeguatamente codificati. Non prendiamo nota del numero di volte in cui le nostre previsioni si sono rivelate errate e questo ci impedisce di renderci conto di incorrere ripetutamente nello stesso errore, e di imparare la lezione.

2.6 Errori nella valutazione del rischio

Gli errori dei giudizi riguardanti la probabilità entrano in conflitto con la teoria soggettiva della probabilità. Quest'ultima determina la probabilità soggettiva attribuita a un evento – misurata dal grado di affidabilità che attribuiamo alla nostra credenza in merito - dall'insieme di scommesse riguardanti quell'evento che siamo disposti ad accettare. In base a questa teoria, la probabilità è soggettiva perché individui diversi possono avere

credenze diverse riguardo alla probabilità dell'evento, tuttavia le scelte sulle scommesse dovranno essere internamente coerenti. Dal punto di vista della teoria della probabilità soggettiva, quindi, sembra che qualsiasi insieme di giudizi di probabilità purché internamente coerente, sia valido. Nel caso della fallacia dello scommettitore, i giudizi che portano a ritenere che dopo un certo numero di lanci consecutivi che hanno dato come risultato "testa", le probabilità che il prossimo lancio della moneta dia come risultato "croce" aumentino, non possono ritenersi razionali, anche se sono internamente coerenti. In quel caso le stime sulla probabilità sono incompatibili con altre credenze sicuramente in possesso dell'individuo, come l'idea che una moneta non ha memoria e che perciò ogni lancio è indipendente da quello precedente. In definitiva, il criterio della coerenza richiesto dalla teoria non è sufficiente: i giudizi di probabilità dovrebbero in realtà essere compatibili con l'intera rete di credenze dell'individuo.

Gli errori o pregiudizi che spesso l'uso di euristiche comporta fanno sorgere dei dubbi sull'effettiva utilità e validità dei principi euristici stessi. Ma se ci chiediamo se valga la pena usare strategie euristiche, il punto diventa: possiamo davvero fare a meno di usarle?

La risposta a questa domanda chiama in causa un aspetto controverso delle euristiche, e cioè la possibilità di presentarle come strategie deliberatamente utilizzate dalle persone, oppure come procedure che si attivano automaticamente. Nel primo caso, la scelta di utilizzarle al fine di semplificare compiti giudicati troppo difficili per la mente umana comporta il rischio di incorrere in errori sistematici e spesso gravi. Per questo motivo l'uso delle euristiche può essere considerato un esempio di povertà cognitiva.

Tversky e Kahneman le considerano invece una sorta di valutazione istintiva indotta dalla natura stessa del problema, che può influenzare il giudizio ma non è il frutto di una deliberazione di carattere strategico. In questa prospettiva, le euristiche si presentano come una procedura alternativa a tutti gli effetti, che fa parte del nostro patrimonio cognitivo e quindi non implica una sottoutilizzazione delle nostre capacità di giudizio. In effetti, queste due concezioni possono risultare entrambe convincenti, e del resto non sono neppure necessariamente incompatibili, dal momento che è possibile pensare che in alcune circostanze le euristiche funzionino come automatismi valutativi, e in altre invece come procedure di cui ci serviamo in maniera consapevole. Tuttavia, la concezione che sottolinea la povertà cognitiva delle euristiche imputata a una pigrizia mentale ha avuto sicuramente maggiore impatto, almeno fino a quando la tesi di Tversky e Kahneman non ha avuto a sostegno la teoria dei due sistemi di ragionamento. Teoria che, come vedremo più avanti, ha permesso di conciliare le due prospettive sulle euristiche, proponendo una spiegazione del loro funzionamento.

2.7 Prospect theory: una spiegazione alternativa delle scelte in condizioni di rischio

Le tesi relative alle euristiche implicavano già una critica alla teoria della scelta razionale, dato che l'uso di queste scorciatoie mentali va evidentemente a scapito dei procedimenti rigorosi eseguiti dall'agente pienamente razionale quando, ad esempio, deve calcolare le probabilità di un evento. L'articolo su *Econometrica* inizia mettendo esplicitamente in discussione la validità della teoria *mainstream*.

Expected utility theory has dominated the analysis of decision making under risk. It has been generally accepted as a normative model of rational choice, and widely applied as a descriptive model of economic behavior. Thus, it is assumed that all reasonable people would wish to obey the axioms of the theory, and that most people actually do, most of the time⁷⁰.

La teoria dell'utilità attesa è stata proposta sia come modello normativo della scelta⁷¹, sia come modello descrittivo del comportamento economico⁷². Si presuppone che tutte le persone ragionevoli intendano conformarsi agli assiomi della teoria⁷³ – che ha quindi anche valore prescrittivo – e che, in buona misura, ci riescano.

The present paper describes several classes of choice problems in which preferences systematically violate the axioms of expected utility theory. In the light of this observations we argue that utility theory, as it is commonly interpreted and applied, is not an adequate descriptive model and we propose an alternative account of choice under risk⁷⁴.

La critica è dunque rivolta alla teoria come modello descrittivo del comportamento, a causa delle violazioni degli assiomi riscontrate negli

⁷⁰ D. Kahneman and A. Tversky, "Prospect Theory. An analysis of decision under risk", *Econometrica*, 1979 p. 263

⁷¹ Il riferimento è a Keeney e Raiffa, *Decision with Multiple Objectives: Preferences and Value Tradeoffs*, New York, Wiley, 1978

⁷² In questo caso vengono il riferimento è a Friedman e Savage e ad Arrow.

⁷³ Sul valore prescrittivo della teoria il riferimento è a von Neuman e Morgenstern e a Savage.

⁷⁴ D. Kahneman e A. Tversky, *ibid*

esperimenti sugli effettivi comportamenti di scelta. Come vedremo, i presupposti della EUT che vengono messi in discussione sono: che la massimizzazione dell'utilità attesa sia rappresentata dalle posizioni patrimoniali finali; che l'utilità sia marginalmente decrescente e gli agenti siano generalmente avversi al rischio (concavità della funzione di utilità); che l'utilità di ogni possibile esito sia ponderata con la sua probabilità.

Le violazioni degli assiomi della EUT riscontrate nelle scelte reali sono associate a fattori psicologici che possono spiegarle. Così, nel caso di una scelta fra prospetti positivi, l'“effetto certezza” contribuisce a una preferenza avversa al rischio, e quindi alla preferenza per un guadagno sicuro rispetto a uno solo probabile (come mostrato dal paradosso di Allais); mentre nel caso si tratti di prospetti negativi, in modo speculare rispetto al primo caso, l'“effetto riflesso” mostrerà una preferenza per un rischio solo probabile rispetto ad una perdita più piccola ma certa. “The same psychological principle – the overweighting of certainty – favors risk aversion in the domain of gains and risk seeking in the domain of losses”⁷⁵. Viene così eliminata la generale avversione al rischio come spiegazione dell'effetto certezza.

Il fenomeno denominato “effetto isolamento” si verifica invece quando una variazione nel modo in cui sono rappresentati gli esiti o le probabilità finisce per alterare le preferenze. E' una violazione dell'assioma dell'invarianza, secondo il quale le scelte tra prospetti sarebbero determinate unicamente dalle utilità - ponderata dalla probabilità - degli stati finali e non

⁷⁵ *Ivi*, p. 269.

dal modo in cui i prospetti sono presentati o rappresentati. Gli esperimenti invece mostrano sia che la scelta è influenzata dalla rappresentazione, sia che l'utilità non è misurata in termini di stati finali: "the carriers of value or utility are changes of wealth, rather than final asset positions that include current wealth"⁷⁶.

Dopo aver descritto i fallimenti sistematici cui va incontro la EUT come modello descrittivo, la *pars costruens* della teoria di Kahneman è rappresentata dalla *Prospect theory*, che descrive il processo effettivo di scelta come costituito da due fasi, una di organizzazione delle informazioni disponibili e l'altra di valutazione. La prima fase riguarda le operazioni di *editing* che determinano il modo in cui vengono percepiti i prospetti, e consiste nell'analisi dei prospetti e nell'elaborazione di una loro rappresentazione semplificata ma efficace. La seconda fase prende in esame i principi che sono alla base della valutazione dei prospetti e degli esiti incerti ai fini della scelta del prospetto che è risultato di valore più alto.

Il compito della prima fase – *editing phase* - è quello di riformulare e organizzare le opzioni in modo da rendere più semplice possibile la valutazione e agevolare la scelta. Una delle operazioni più importanti a questo scopo è la codificazione, e cioè la rappresentazione degli esiti possibili dei prospetti in termini di guadagni e perdite, a partire da un punto di riferimento che generalmente è dato dalla attuale situazione patrimoniale. Va tuttavia considerato che questa operazione può essere influenzata dal modo in cui sono

⁷⁶ *Ivi*, p. 273. Cfr. D. Kahneman, A. Tversky (eds.), *Choices...* cit, p. 3.

presentati i prospetti e dalle aspettative del decisore. Oltre alla codificazione, altre operazioni permettono di semplificare i prospetti, eliminando le componenti comuni, arrotondando le probabilità o gli esiti espressi in termini monetari ed escludendo i risultati estremamente improbabili. Infine, si individuano le cosiddette alternative dominate che vengono direttamente scartate senza ulteriore valutazione.

Un punto importante da sottolineare è che uno stesso prospetto può essere elaborato in modi diversi a seconda del contesto (*framing*) in cui appare. Di conseguenza, l'ordine di preferenza fra i prospetti non è necessariamente invariante, come invece imporrebbe la EUT⁷⁷. Nella descrizione di questa prima fase, la principale caratteristica della teoria alternativa alla EUT è il fatto che l'agente non guardi agli esiti in sé, ma li consideri come cambiamenti di stato – guadagni o perdite – relativi a un dato punto di riferimento, che può essere rappresentato dalla suo *status quo*, ma anche dalla condizione a cui aspira.

Inoltre, il cambiamento del punto di riferimento a partire dal quale valutare guadagni e perdite può alterare l'ordine di preferenze fra prospetti, come nelle situazioni in cui guadagni e perdite sono codificati in relazione a un livello di aspettative che non coincide con lo *status quo*. In certi casi, ad esempio, la propensione al rischio può essere indotta dall'incapacità ad adattarsi alle perdite o alla rinuncia di un guadagno atteso.

⁷⁷ L'Expected Utility Theory viene comunemente considerata la teoria mainstream all'interno delle teorie della scelta razionale.

The analysis suggests that a person who has not made peace with his losses is likely to accept gambles that would be unacceptable to him otherwise. The well known observation that the tendency to bet on long shot increases in the course of the betting day provide some support the hypothesis that a failure to adapt to losses or to attain an expected gain induces risk seeking.⁷⁸

Nella successiva fase di valutazione, il valore complessivo di un prospetto elaborato, denotato con V , è espresso nei termini di due diverse scale, π e ν . La prima associa a ogni probabilità p un peso decisionale $\pi(p)$, che riflette l'impatto di p sul valore totale del prospetto. La seconda scala, ν , assegna ad ogni risultato x un numero $\nu(x)$ che riflette il valore soggettivo del risultato, dal momento che i risultati sono definiti relativamente a un punto di riferimento. In questo modo ν misura il valore delle deviazioni rispetto a quel punto di riferimento, cioè i guadagni e le perdite⁷⁹.

In order to accommodate the effects described in the first part of the paper, we are compelled to assume that the values are attached to changes rather than to final states, and that decision weights do not coincide with stated probabilities⁸⁰.

⁷⁸ *Ivi*, p. 287.

⁷⁹ Nell'articolo in *Econometrica* a p. 276 gli autori sottolineano che l'idea di definire l'utilità in base ai guadagni e alle perdite piuttosto che in base alle posizioni patrimoniali finali era stata implicitamente accettata nella maggior parte delle misurazioni sperimentali dell'utilità. A questo riguardo viene citato il testo di Davidson, Suppes e Siegel, *Decision-making: An experimental approach*, Stanford, Stanford University Press del 1957.

⁸⁰ *Ibid*, p. 277.

Un punto centrale della *prospect theory* è la tesi che il valore sia associato alle variazioni di ricchezza o di benessere piuttosto che agli stati finali. Il valore, quindi, non è indipendente dal riferimento ed è errata la tesi - che risale alla prima versione della teoria dell'utilità attesa di Bernouilli - che i decisori valutino i risultati dell'utilità in base alle posizioni finali dell'asset. L'ipotesi implicita nella EUT è che anche i risultati descritti come guadagni o perdite siano prima trasformati in stati finali degli asset, e poi valutati in base a questa rappresentazione.

La *prospect theory* ipotizza invece una preliminare operazione di elaborazione, al fine di reincorniciare le prospettive in termini più semplici, prima della valutazione. Alla base vi è sempre l'idea che i meccanismi valutativi funzionino come quelli percettivi, che sono sintonizzati sulle valutazioni delle variazioni piuttosto che sulla valutazione di grandezze assolute. Nel caso della percezione, infatti, la reazione a uno stimolo dipenderà dal punto di riferimento, e cioè dal livello di adattamento rispetto al contesto passato dell'attuale esperienza percettiva.

Thus, an object at a given temperature may be experienced as hot or cold to the touch depending of the temperature to which one has adapted. The same principle applies to non-sensory attributes such as health, prestige, and wealth. The same level of wealth, for example, may imply abject poverty for one person and great riches for another – depending on their current assets⁸¹.

⁸¹ Ibid, p. 277.

Sono molte le dimensioni sensoriali che condividono la proprietà di produrre una risposta psicologica che varia in modo inversamente proporzionale al variare della grandezza fisica, e cioè diminuisce con il crescere della grandezza. Ad esempio, la differenza fra una variazione di temperatura da 3° a 6° è più facile da percepire rispetto a quella fra 13° e 16°; così come la differenza fra un guadagno di 100 e uno di 200 sembra più grande della differenza fra un guadagno di 1.100 e uno di 1.200. Dato che la stessa cosa vale anche per le perdite, si può dire che il valore marginale sia dei guadagni sia delle perdite diminuisce generalmente con il crescere della loro dimensione. Sia in situazioni prive di rischio sia in condizioni di rischio, la funzione di valore è concava per i guadagni e convessa per le perdite. Inoltre, le circostanze particolari in cui avvengono le scelte influenzano le funzioni di utilità, e quindi un agente non sempre avrà un atteggiamento neutro nei confronti del denaro, specialmente quando può essere in discussione il suo stile di vita. In generale, l'atteggiamento psicologico prevalente riguardo ai cambiamenti è che il peggioramento della propria condizione di vita causato da una perdita ha un maggior peso del piacere associato a un guadagno. Di conseguenza, la variazione della funzione di valore diventa più ripida nel caso delle perdite che dei guadagni.

Nella *Prospect theory* il valore di ogni esito è moltiplicato per un peso decisionale. Nonostante siano ricavati dalle scelte in modo simile a quello in cui le probabilità soggettive di Savage sono ricavate dalle preferenze, i pesi decisionali non sono assimilabili alle probabilità.

Decision weight are not probabilities: they do not obey to probability axioms and they should not be interpreted as measures of degree or belief. [...]

Decisions weight measure the impact of events on the desirability of prospects, and not merely the perceived likelihood of these events⁸².

La funzione di ponderazione π mette quindi in relazione pesi decisionali e probabilità dichiarate; $\pi(p)$ indica il rapporto del peso associato con la probabilità p con il peso associato all'evento certo. L'esame delle proprietà di questa funzione di ponderazione rivela la sistematica sovrastima delle probabilità molto piccole, la sottostima di quelle medie o molto grandi e, al contempo, il riconoscimento della certezza e dell'impossibilità per i valori estremi. Più in particolare, la riduzione della probabilità di un dato risultato ha un impatto maggiore se il risultato iniziale era certo, ma minore se il valore di probabilità era inferiore a 1. L'atteggiamento verso il rischio è evidentemente influenzato da questo modo di percepire la probabilità: la differenza di impatto di cui si è appena detto può spiegare la violazione dell'assioma di indipendenza messo in evidenza dal paradosso di Allais. Il peso decisionale associato alla probabilità di un evento dipenderà dalla percezione della probabilità, che può essere soggetta a distorsioni o influenzata dall'ambiguità come mostrato dal paradosso di Ellsberg.

The major qualitative properties of decision weights can be extended to cases in which the probabilities of outcomes are subjectively assessed rather than explicitly given. In these situations, however, decision weights may also be affected by other characteristics of an event, such as ambiguity or vagueness⁸³.

⁸² *Ivi*, p.280.

⁸³ A. Tversky, D. Kahneman, *The framing of decisions and the psychology of choice*, *Science*, v. 211, 1981, pp. 453-8, p. 254.

L'esame del processo di scelta spiega quindi le scelte incoerenti e i rovesciamenti di preferenze facendoli risalire all'interazione di due insiemi di fattori evidenziati rispettivamente nella prima e nella seconda fase: le variazioni nel modo di rappresentare le scelte, le probabilità, gli esiti e la caratteristica non linearità di valori e pesi decisionali (p e π). Entrambi risentono dell'effetto *framing*, che è un caso particolare del fenomeno più generale della dipendenza dalla rappresentazione: il modo in cui ci viene descritto o in cui noi stessi ci rappresentiamo il problema – la cornice appunto – condiziona la nostra decisione.

2.8 Effetto framing

La dipendenza delle preferenze dal *decision frame* è il tema del secondo articolo su *Science*. In tale articolo si trovano precisazioni importanti sul criterio di razionalità e sugli obiettivi della *Prospect theory*. Per la teoria della scelta razionale, la razionalità è il presupposto per poter spiegare e prevedere le scelte, che dovrebbero quindi soddisfare “some elementary requirements of consistency and coherence”⁸⁴; molto spesso questi requisiti vengono invece sistematicamente violati.

We trace these violations to the psychological principles that govern the perception of decision problems and the evaluation of options [...]. We use the term “decision frame” to refer to the decision-maker’s conception of the acts, outcomes and contingencies associated with a particular choice. The frame that

⁸⁴ *Ivi*, p. 453.

a decision-maker adopts is controlled partly by the formulation of the problem and partly by the norms, habits, and personal characteristics of the decision-maker⁸⁵.

Secondo Tversky e Kahneman, quindi, le violazioni degli assiomi della teoria sono dovuti a un'impossibile richiesta di oggettività, che appare evidentemente inesigibile in base all'effettivo funzionamento dei nostri meccanismi percettivi, alle nostre modalità rappresentative e alle nostre procedure di valutazione. La dipendenza da un punto di riferimento inevitabilmente soggettivo sembra implicare non solo la relatività dei nostri giudizi – percettivi e valutativi – ma anche la loro volatilità, in conseguenza di ogni potenziale cambiamento di prospettiva.

In realtà, il confronto con la percezione – più volte sottolineato dagli autori della *Prospect theory* – non può essere molto più che una suggestiva metafora, dal momento che nel caso della percezione abbiamo la possibilità di ricorrere a misurazioni oggettive, e che quindi con l'esperienza possiamo più facilmente correggere eventuali illusioni percettive e ammettere la parzialità del nostro punto di vista. L'assenza di standard oggettivi per i processi decisionali è del resto riconosciuta da Tversky e Kahneman⁸⁶, che tuttavia insistono sul valore euristico della somiglianza fra meccanismi percettivi e decisionali. La logica dei procedimenti decisionali è invece, a mio avviso, profondamente diversa e molto più complessa anche soltanto per la presenza

⁸⁵ Ibid

⁸⁶ *Ivi*, p. 457. Lo standard oggettivo non sempre serve nel caso delle decisioni.

di strutture argomentative condizionali che comportano un continuo slittamento fra realtà, possibilità e desideri. Per questo, mentre riusciamo molto più facilmente a imparare dall'esperienza nel caso degli errori percettivi, continuiamo – spesso senza neppure esserne consapevoli – a fare gli stessi errori in materia di scelta.

Alternative frames for a decision problem may be compared to alternative perspectives on a visual scene. Veridical perception requires that the perceived relative height of two neighboring mountain, say, should not reverse with changes of vantage point. Similarly, rational choice requires that the preference between options should not reverse with changes of frame. Because of imperfection of human perception and decision, however, changes of perspective often reverse the relative apparent size of objects and the relative desirability of options⁸⁷.

Il riferimento all'imperfezione della percezione e della decisione umane serve a rammentare che la relatività delle prospettive è una caratteristica strutturale del nostro apparato cognitivo. Ma se la dipendenza dalla rappresentazione è la sola cosa che hanno in comune, il parallelo fra l'altezza relativa di due cime montuose contigue e la desiderabilità relativa di due opzioni di scelta appare poco significativo. Nelle scelte in situazione di incertezza, lo standard oggettivo – anche qualora fosse determinabile – non sarebbe comunque sufficiente, proprio perché non potrà mai accordarsi del tutto con la nostra peculiare percezione del contesto e di conseguenza non

⁸⁷ *Ivi*, p. 453.

potrà mai darci tutte le garanzie di cui sentiamo il bisogno e metterci al riparo da ogni genere di costi impliciti nella scelta.

Possiamo metterci al riparo da situazioni potenzialmente conflittuali, come i casi di dissonanza cognitiva, provando a considerare una situazione dall'esterno, quasi fingendo che non ci riguardi direttamente, oppure mettendo a confronto diversi punti di vista in modo da riuscire a trovare un equilibrio fra la prospettiva del bicchiere mezzo vuoto e quella del bicchiere mezzo pieno. Un'altra possibilità ancora è quella di utilizzare in modo consapevole la duttilità del *framing*, ad esempio assumendo in anticipo un impegno assolutamente vincolante nei confronti di una decisione futura, che ci impedisca poi di trovarci in balia delle circostanze o degli umori del momento e di non riuscire a resistere alla tentazione di ignorare un semplice proponimento. Un caso emblematico è rappresentato dalla richiesta di Ulisse di essere legato alla nave per poter ascoltare il canto delle Sirene senza rischiare di esserne sedotto. Questi esempi mettono in evidenza i risvolti potenzialmente virtuosi della manipolazione del *framing*.

Per spiegare il fatto che da diverse rappresentazioni dello stesso problema scaturiscono decisioni differenti andrebbe indagato più a fondo il ruolo svolto dalla rappresentazione, in particolare la tendenza a servirsi della rappresentazione più familiare⁸⁸. Tuttavia, come si è detto, richiamando alla mente la rappresentazione che reinterpreta il problema in forma più semplice,

⁸⁸ Anche Davidson, come si vedrà nel capitolo successivo, parla di razionalizzazione come tendenza a reinterpretare comportamenti insoliti per ricondurli a contesti più familiari.

si corre il rischio dell'errore legato all'accessibilità: l'esempio che viene in mente con maggiore facilità spesso non è affatto quello più appropriato⁸⁹.

Gli stessi autori sottolineano che la *Prospect Theory* deve essere considerata come una descrizione approssimativa e semplificata della valutazione di scelte in condizione di rischio, dal momento che le proprietà di p e π - pur rappresentando caratteristiche strutturali di un modello comune di scelta - non sono universali. Data la loro intrinseca non linearità, infatti, l'ordine di preferenza fra opzioni di scelta, come si è già detto, sarà sempre dipendente dal *framing* con cui ci rappresentiamo azioni, possibilità, esiti. Una dipendenza che espone al rischio di molti tipi di distorsioni di cui la PT riesce almeno in parte a dar conto.

Si è già detto del diverso atteggiamento - avverso o propenso al rischio - a seconda che ci troviamo di fronte rispettivamente a un guadagno sicuro o alla probabilità di una perdita. Quando decidiamo di stipulare un'assicurazione per ridurre o eliminare la probabilità di un rischio a fronte di un costo, dovremmo tener conto che la nostra percezione del rischio è soggetta a distorsioni e può quindi essere manipolata. Possiamo ad esempio non accorgerci che è solo apparente la tranquillità che ci dà una polizza assicurativa sulla casa che promette una copertura completa rispetto a un determinato pericolo – ad esempio l'incendio – dato il carattere probabilistico

⁸⁹ D. Kahneman, *Maps of bounded rationality: a perspective on intuitive judgement and choice*. Prize lecture, December 8, 2002.

di ogni assicurazione, che riduce ma non elimina mai del tutto il rischio, in questo caso di pericoli diversi dall'incendio.

Infine, anche gli esiti sono percepiti come positivi o negativi in relazione a un quadro di riferimento costitutivamente instabile poiché determinato da diversi fattori: il nostro status effettivo, le nostre aspettative, i condizionamenti sociali. Di fronte alla complessità di una decisione non stupisce che il comportamento più diffuso consista nell'accontentarsi di descrizioni stringate della situazione, che semplificano la valutazione riducendo lo sforzo cognitivo, soddisfano l'intuizione di un forte legame causale fra azioni e conseguenze, e sono in sintonia con le caratteristiche dell'esperienza edonica, più sensibile ai cambiamenti, desiderabili o meno, che a condizioni di assoluta stabilità.

La tendenza ad agire sulla base della rappresentazione del problema più a portata di mano può essere quindi giustificata dall'esigenza di evitare lo sforzo mentale che richiederebbe il confronto fra prospettive diverse o il tentativo di evitare possibili incoerenze, come la dissonanza cognitiva di cui si è detto poco sopra. Tuttavia, per Kahneman e Tversky questa spiegazione in termini di costi e di consapevolezza dei propri limiti che si richiama alla *bounded rationality* di Simon non è che una ipotesi ad hoc. La *Prospect theory* e l'analisi del *framing* riuscirebbero invece a spiegare meglio le violazioni dei criteri di razionalità evidenziate dagli esperimenti. In questo contesto, la distinzione fra livello descrittivo e livello normativo non è più rivolta a ribadire il valore della teoria *mainstream* sul piano normativo e ad attribuire alla *Prospect theory* un compito solo descrittivo, quanto piuttosto a sottolineare la rilevanza della PT anche sul piano normativo.

The present work has been concerned primarily with the descriptive question of how decisions are made, but the psychology of choice is also relevant to the normative question of how decisions ought to be made⁹⁰.

2.9 Un'alternativa alla razionalità come coerenza

Adottando la coerenza come unico criterio di razionalità, la teoria della scelta razionale ha evitato di porsi il problema del valore della scelta, precludendosi la possibilità di dare indicazioni sul modo di risolvere eventuali incoerenze. L'inevitabile dipendenza delle preferenze dalle diverse e mutevoli contestualizzazioni della scelta rende però poco credibile l'efficacia della coerenza come criterio di razionalità. Inoltre, il criterio comune di razionalità secondo i due autori richiede che la preferenza comprenda anche l'aspettativa di un'esperienza positiva associata all'esito.

L'analisi del *framing* può essere determinante proprio per avere garanzie riguardo a quest'ultimo requisito, e la *Prospect theory* può proporsi come alternativa alla teoria *mainstream*, proprio perché è in grado di soddisfare questa legittima aspettativa riguardo agli esiti delle decisioni prese.

Thus, a man could be judged irrational either because his preferences are contradictory or because his desires and aversions do not reflect his pleasures and pains. The predictive criterion of rationality can be applied to resolve inconsistent preferences and to improve the quality of decisions.⁹¹.

⁹⁰ *Ivi*, p. 458.

⁹¹ *Ibid*

Se la coerenza è un criterio comunque insufficiente, l'attenzione andrà rivolta alla qualità della decisione, il cui indicatore è il grado di soddisfazione che ricaviamo dall'esito. L'attenzione al contesto, quindi al *framing*, serve anche a immaginare in prospettiva quale sarà l'effettivo vantaggio futuro di una scelta. Per questo di fronte a decisioni difficili per le ricadute che possono avere sul nostro futuro, la domanda giusta da porsi non è “cosa voglio veramente adesso”, ma “cosa proverò allora”. Assumendo questa prospettiva si potrà selezionare il *decision frame* più lungimirante, che rappresenti ciò che effettivamente proverò di fronte ai risultati della scelta.

La possibilità di alterare deliberatamente il *framing* può servire anche ad attribuirsi o sollevarsi dalla responsabilità per la decisione presa o a reinterpretare il significato di un risultato negativo, in modo da attenuarne l'impatto emotivo. Scegliere la prospettiva giusta da cui guardare il problema o gli esiti serve ad accrescere il grado di consapevolezza e a migliorare la qualità della decisione.

A sense of confidence in a particular choice does not ensure that the same choice would be made in another frame. It is therefore good practice to test the robustness of preferences by deliberate attempts to frame a decision problem in more than one way⁹².

L'insufficienza del criterio di razionalità come coerenza non riguarda quindi solo le violazioni per molti aspetti inevitabili data l'instabilità del *framing* e l'inaffidabilità delle euristiche, ma è anche legato al disinteresse per

⁹² D. Kahneman, A. Tversky (eds.), *Choices, Values, and Frames*, cit., p. 7.

la qualità della decisione, rappresentata dall'utilità misurata in termini di piacere.

La relazione fra preferenza e utilità è naturalmente un assunto fondamentale della EUT, ma Kahneman e Tversky interpretano l'utilità in termini di esperienza effettiva di piacere, che a loro avviso - se correttamente anticipata dalla preferenza - contribuisce in maniera rilevante alla razionalità della scelta. Su questa base poggia la distinzione fra *experience value*, che esprime il grado di soddisfazione o rimpianto nell'esperienza effettiva di un dato esito, e *decision value*, che indica invece il contributo dell'aspettativa di un determinato esito alla complessiva forza di attrazione o repulsione di un'opzione di scelta.

Nell'ottica idealizzata dell'agente *fully rational* la distinzione manca del tutto, dal momento che si presume che le previsioni dell'agente siano completamente affidabili. Nelle situazioni reali però la corrispondenza non è affatto scontata: non è sempre facile prevedere cosa proveremo, anche perché non sempre i fattori che sono stati importanti per la decisione hanno poi un impatto equivalente al momento dell'esito. Inoltre, se il punto di riferimento per l'*experience value* è determinato prevalentemente dalla condizione oggettiva di partenza, sappiamo però che è influenzato anche dalle aspettative e dal confronto sociale, e che di fronte ai cambiamenti diventa fortemente dipendente dalle dinamiche dell'adattamento.

Riassumendo, inizialmente l'obiettivo del programma di ricerca di Kahneman e Tversky era di cogliere le caratteristiche rilevanti della scelta individuale in contesti reali: un compito dichiaratamente descrittivo, che non

metteva in discussione il ruolo della teoria dell'agente razionale, il cui modello normativo di razionalità costituiva un riferimento indispensabile per misurare le violazioni e proporre delle spiegazioni dei comportamenti effettivi "devianti". In realtà, come appare dagli ultimi brani citati, Kahneman⁹³ sottolinea la differenza fra la sua posizione e quella di Simon proprio perché, mentre il modello *bounded rationality* di Simon si limita a ritenere indispensabile una semplificazione degli assiomi di razionalità a causa dei limiti cognitivi e computazionali dell'agente reale, la *Prospect theory* mira a introdurre un nuovo tipo di giudizi, che costituiscono un procedimento euristico alternativo alla teoria *mainstream*.

Se la razionalità non può esaurirsi nella coerenza, che del resto si è dimostrata spesso impossibile da ottenere data la variabilità del *framing*, la qualità della decisione identificata con il piacere ricavato dal risultato diventa il criterio più importante per giudicare la razionalità della scelta. In un articolo del 1997, significativamente intitolato *Back to Bentham*⁹⁴, Kahneman sostituisce *experience value* con *experienced utility*, e propone "a formal normative theory of the total experienced utility of temporally extended outcomes"⁹⁵.

⁹³ Amos Tversky muore nel 1996.

⁹⁴ D. Kahneman, P. Wakker, R. Sarin, "Back to Bentham. Exploration of experienced utility", *The Quarterly Journal of Economics*, v. 112, n. 2, 1997, pp. 375-405.

⁹⁵ *Ivi*, p. 375.

Si tratta di una riproposizione dei due criteri di razionalità individuati da Sen⁹⁶: coerenza interna e massimizzazione, quest'ultimo riformulato in termini benthamiani. Quello di Kahneman può essere considerato un tentativo di reintrodurre un'interpretazione psicologica della massimizzazione per ritornare a Bentham, pur con qualche significativo distinguo.

An important reservation should be started early. Our normative treatment of the utility of temporally extended outcomes adopts a hedonic interpretation of utility, but no endorsement of Bentham's view of pleasures and pain as sovereign masters of human action is intended. Our analysis applies to situations in which a separate value judgment designates experienced utility as a relevant criterion for evaluating outcomes⁹⁷.

Kahneman indaga sulla possibilità di un criterio non meramente logico di razionalità, come è la coerenza interna alle scelte in quanto preferenze rivelate, ma di un *substantive criterion*, un criterio esterno al sistema delle preferenze, e che possa rispondere a domande sulla qualità delle decisioni: se le nostre credenze siano palesemente in conflitto con le evidenze disponibili, se le decisioni perseguano o meno gli interessi dell'agente.

Ad essere messa in discussione è quindi proprio la razionalità dell'agente, la sua effettiva competenza nell'individuazione dei suoi veri interessi, che è una condizione irrinunciabile per lo spirito liberista – *laissez-faire spirit* – della teoria economica della scelta. Riprendendo la distinzione di

⁹⁶ A. Sen, *On Ethics and Economics*, Oxford, Basil Blackwell, 1987, p. 19-22.

⁹⁷ *Ivi*, p. 377. Alla fine del brano vengono citati in nota Dasgupta, Nussbaum e Sen.

Sen fra la razionalità come coerenza interna e come massimizzazione dell'utilità, Kahneman dà un significato diverso alla critica di Sen. Kahneman considera poco vincolante il primo criterio e troppo il secondo. Il primo sarebbe poco vincolante perché accetta l'avversione al rischio nelle piccole scommesse e non ha niente da obiettare nei confronti di scelte che ignorano il futuro prossimo; il secondo invece sarebbe troppo vincolante, dal momento che presume una indiscussa capacità di valutare i propri interessi.

La questione essenziale quindi è se le scelte massimizzino davvero l'utilità attesa delle conseguenze, rispecchiando quella che sarà l'effettiva esperienza. Gli elementi critici nella capacità individuale di valutare la qualità dei risultati riguardano la corretta previsione delle preferenze future e l'adeguata valutazione delle esperienze passate. Esaminando quello che viene comunque visto come uno dei possibili *substantive criterion* - "the experienced hedonic utility of outcomes"⁹⁸- si scopre la sostanziale miopia dell'agente, e cioè la sua incompetenza nella valutazione sia delle sue preferenze future sia delle sue esperienze passate.

Demonstrated deficiencies in the ability to predict future experiences and to learn from the past emerge as new challenges to the assumption of rationality⁹⁹

⁹⁸ D. Kahneman, "New challenges to the rationality assumption" in *Journal of Institutional and Theoretical Economics*, 2000, pp. 18-36. Poi in *Choices ,Values and Frames*, cit. pp. 758-774.

⁹⁹ *Ivi*, p. 760.

Riguardo alla prima ci sono esperienze diverse e diverse spiegazioni possibili. Nell'ottica della TRC la tendenza a non tener conto del futuro e a privilegiare il presente è alla base di molte decisioni poco lungimiranti che tentiamo di contenere, come si è detto, assumendo impegni vincolanti, o con altri espedienti di autocontrollo basati sulla rappresentazione e il confronto di scenari alternativi.

L'analisi di Kahneman rivela anche altri meccanismi psicologici per cui, di fronte a una decisione che implica un cambiamento, si può essere portati a concentrarsi sulla situazione di transizione da una condizione a un'altra invece che sulla situazione nuova, una volta che sia stabilizzata. In questi casi, le distorsioni del framing portano, come probabile conseguenza, a sovrastimare i fattori negativi o quelli positivi del cambiamento.

L'insegnamento che ricaviamo dalle esperienze passate si rivela altrettanto poco affidabile, dal momento che le valutazioni retrospettive degli eventi che ci riguardano si riferiscono a ricostruzioni cognitive basate su due fattori salienti, il momento di maggiore impatto emotivo e lo stato emotivo dei momenti finali – “the peak & end rule”¹⁰⁰.

Fra i volontari che si sottopongono a un esperimento che consiste in due prove successive - nella prima dovranno tenere una mano in acqua a 14° per 60 secondi, nella seconda immergere l'altra mano in acqua per i primi 60 secondi ancora a 14° e per successivi 30 secondi a 15° - la maggioranza preferirà ripetere la seconda prova, violando un principio *substantive* di

¹⁰⁰ *Ivi*, p. 769.

razionalità e cioè l'assioma di monotonicità temporale. I requisiti della razionalità sostanziale sembrano quindi incompatibili con la psicologia della memoria e della scelta. Se di episodi del passato che abbiamo vissuto in prima persona conserviamo una memoria così poco attendibile, non può sorprendere la scarsa abilità nel prevedere quali saranno le effettive preferenze dei nostri futuri sé.

The history of an individual through time can be described as a succession of separate selves, which may have incompatible preferences, and may make decisions that affect subsequent selves¹⁰¹.

Fra il soggetto che registra le emozioni momentanee che vive in prima persona e il soggetto che ricordando quelle esperienze ne dà una valutazione retrospettiva sembrano esserci pareri discordi.

Which of these selves should be granted authority over outcomes that will be experienced in the future¹⁰²?

La risposta di Kahneman è che il principio della monotonicità temporale attribuisce la priorità al soggetto che vive l'esperienza, ma nella realtà è il soggetto che ricorda ad assumersi il compito di impostare le linee guida per le azioni future, esponendosi a nuove sfide della razionalità¹⁰³. Se mantenere la

¹⁰¹ *Ivi*, p. 771. Si fa riferimento al saggio di Thomas Schelling, *Choice and consequences. Perspectives of an errant economist*, Harvard University Press, Cambridge, 1984.

¹⁰² *Ibid*

¹⁰³ Kahneman fa riferimento anche alle implicazioni etiche di questo possibile conflitto fra i sé. Cfr. M. Motterlini, "Un dilemma per il consenso informato", *Kos*, luglio 2006.

coerenza di credenze e preferenze sembrava un compito troppo impegnativo, massimizzare l'*experienced utility* del corso delle conseguenze future delle nostre scelte può rivelarsi ancor più difficile. Secondo Kahneman, anche i criteri di razionalità sostanziale sembrano destinati al fallimento.

La distinzione fra il sé che sperimenta e quello che ricorda non sembra però molto convincente: il primo sé non può vantare maggiori diritti in quanto depositario di come sarebbero effettivamente andate le cose, giacché l'eccesso di coinvolgimento spesso comporta una minore lucidità su quanto ci sta effettivamente accadendo. Perciò tendiamo ad essere ragionevolmente scettici riguardo alle nostre impressioni immediate, specie se siamo di fronte a una decisione importante. Riguardo poi alle conseguenze, proprio la necessità di rassicurarci sull'affidabilità delle nostre scelte può portarci a “farci piacere” un risultato per cui abbiamo strenuamente lottato anche quando si rivela inferiore alle aspettative.

Nella *Nobel lecture* Kahneman mette in evidenza come l'alternativa fra la EUT e la *prospect theory* sia proprio nelle interpretazioni contrastanti che le due teorie danno dell'utilità, declinando i risultati in termini di stati o di variazioni. E ciò dipende dal fatto che per la prima contano solo le conseguenze a lungo termine, mentre al contrario la *prospect theory* si preoccupa solo dei risultati a breve termine, e la funzione del valore riflette una previsione sulla valenza e l'intensità delle emozioni che si proveranno al momento della transizione da uno stato all'altro¹⁰⁴. Ma l'idea che la visione

¹⁰⁴ D. Kahneman, *Nobel Lecture*, December 8 2002, p. 464 (http://nobelprize.org/nobel_prizes/economics/laureates/2002/kahnemann-lecture.pdf).

lungimirante sia favorita dalle norme prescrittive del processo decisionale ragionevole e che la visione miope possa essere ritenuta superiore solo per scopi descrittivi non è convincente, e mette di nuovo in discussione la validità della dicotomia descrittivo - prescrittivo.

It is worth noting that an exclusive concern with the long term may be prescriptively sterile, because the long term is not where life is lived. Utility cannot be divorced from emotion, and emotion is triggered by changes. A theory of choice that completely ignored feelings such as the pain of losses and the regret of mistakes is not only descriptively unrealistic. It also lead to prescriptions that do not maximize the utility outcomes as they are actually experienced – that is utility as Bentham conceived it¹⁰⁵.

Se la forza emotiva della prospettiva del breve periodo è indubbia, la concezione benthamiana sembra essere la sola in grado di garantire un'autentica massimizzazione dell'utilità esperita.

Un'altra obiezione alle argomentazioni di Kahneman riguarda proprio la scelta del criterio edonistico di *experienced utility*, che per sua stessa ammissione non è né universale né esaustivo. Sembra che Kahneman, anche quando indica altri obiettivi possibili, come la buona reputazione o l'incremento delle capacità di cui parla Sen, tenda a considerarli come alternative al fine del piacere, invece che come una pluralità di criteri che l'agente ha a disposizione nel momento della valutazione e della scelta. Il criterio che risulterà predominante non potrà mettere completamente a tacere

¹⁰⁵ *Ivi*, p. 464-5.

gli altri, né esasperare il conflitto o la dissonanza fra i molteplici sé, che in realtà non sono solo quelli che si confrontano nel corso del tempo, ma sono rappresentati anche dalle varie identità che assumiamo nei diversi contesti di vita.

Per riassumere, nell'indagine di Kahneman e Tversky sul comportamento di scelta un punto centrale è la dipendenza dalla rappresentazione di ogni attività cognitiva coinvolta, il che comporta il rischio di alterazioni come quelle esemplificate dal *framing effect*, che può essere considerato un caso particolare della generale dipendenza dalla rappresentazione¹⁰⁶. Gli esperimenti mostrano che la rappresentazione del contesto di scelta condiziona l'interpretazione della situazione e delle opzioni disponibili e che, di conseguenza, differenti rappresentazioni portano a decisioni differenti. La semplificazione del problema della rappresentazione è rappresentata dalle euristiche che, pur fornendo delle scorciatoie potenzialmente efficaci, possono essere un'importante fonte di pregiudizi – *biases* – e indurci in errori. Pertanto l'attenzione si concentra sulle caratteristiche dei processi mentali per spiegare il ruolo svolto dalle rappresentazioni mentali nelle scelte.

¹⁰⁶ M. Egidi, "The dual process account of reasoning: historical roots, problems, perspectives", *CEEL Working paper*, n. 6, 2007, p. 9.

2.10 Due processi cognitivi

Nella sua *Nobel Lecture* Kahneman distingue due generiche modalità di funzione cognitiva: una modalità intuitiva in cui giudizi e decisioni sono formulati automaticamente e rapidamente e una modalità controllata, che è deliberata e più lenta¹⁰⁷.

There is considerable agreement on the characteristics that distinguish the two types of cognitive processes, which Stanovich and West (2000) labeled System 1 and System 2. [...]The operations of System 1 are fast, automatic, effortless, associative, and difficult to control or modify. The operations of System 2 are slower, serial, effortful and deliberately controlled; they are also relatively flexible and potentially rule-governed¹⁰⁸.

La distinzione fra due tipi di processi cognitivi nasce dagli interrogativi sull'effettivo controllo che siamo in grado di esercitare sui giudizi e sulle decisioni e ha portato a diverse ipotesi di modelli a due sistemi con denominazioni differenti¹⁰⁹, tutti accumulati dall'alternativa fra un processo automatico e uno controllato. Kahneman – che chiama i due modi di pensare rispettivamente intuizione e ragionamento – sottolinea la forte somiglianza fra

¹⁰⁷ D. Kahneman, *Nobel Lecture*, cit.

¹⁰⁸ Ibid, p. 450

¹⁰⁹ Sistemi basati su regole e sistemi associativi (Sloman); razionali e esperienziali (Kirkpatrick e Epstein), *reflective* e *reflexive* (Lieberman). Cfr. Egidi, cit., Posner e Synder; Schneider e Shrifin; Camerer; Stanovich e West

percezione e intuizione, entrambi automatici e quasi impossibili da controllare, nonché la conseguente somiglianza dei *biases* nei rispettivi giudizi.

From its earliest days, the research that Tversky and I conducted was guided by the idea that intuitive judgments occupy a position – perhaps corresponding to evolutionary history – between the automatic operations of perception and the deliberate operations of reasoning¹¹⁰.

Nel modello presentato da Kahneman sia il sistema percettivo sia le operazioni intuitive del sistema 1 generano impressioni non intenzionali degli attributi di oggetti di percezione e pensiero, e i giudizi intuitivi sono quelli che riflettono direttamente queste impressioni. Questo spiega la somiglianza fra giudizi percettivi e valutativi anche riguardo alle tipologie di errori. Gli errori nei giudizi intuitivi sono tuttavia una conseguenza del fallimento di entrambi i sistemi. Infatti, anche se è il processo intuitivo la fonte principale di errori di ragionamento, il sistema 2 - che ha anche la funzione di controllare la qualità delle operazioni mentali di entrambi i sistemi - ha la sua parte di responsabilità, non avendo effettuato un efficace intervento di controllo e correzione.

La caratterizzazione dei due sistemi permette di comprendere meglio il funzionamento delle euristiche e le ragioni dei *biases* che spesso le accompagnano. La proprietà dei pensieri intuitivi è la spontaneità, la facilità con cui vengono in mente, cioè la loro accessibilità. Kahneman sottolinea che l'accessibilità è un *continuum*, non una alternativa dicotomica, per cui alcune

¹¹⁰ *Ibid*

operazioni possono richiedere uno sforzo maggiore; tuttavia l'acquisizione di competenze aumenta selettivamente l'accessibilità di soluzioni utili e di modalità più produttive di organizzazione delle informazioni¹¹¹. Un esempio è la diversa capacità di visualizzazione della stessa scacchiera che ha un maestro di scacchi rispetto a un principiante. Inoltre, l'uso delle euristiche riguarda come si è visto in primo luogo i giudizi su eventi incerti; ciononostante, i giudizi intuitivi della probabilità – come evidenziato dalle euristiche del giudizio – sono mediati da attributi come la somiglianza e la facilità di associazione, che pure non hanno nessuna relazione intrinseca con l'incertezza.

Rispetto alla rassegna su *heuristics and biases* del 1974, per spiegare il funzionamento delle euristiche del giudizio viene proposto un processo comune di sostituzione di attributo che serve anche a delineare una nuova definizione di euristica che ne estende il campo d'azione al di là dei giudizi su eventi incerti. Nella nuova definizione, il termine “euristica” è usato in due sensi: come sostantivo, per riferirsi al processo cognitivo, e come aggettivo – “attributo euristico” – per indicare la sostituzione che avviene in un particolare giudizio.

A judgment is said to be mediated by a heuristic when an individual assesses a specified *target attribute* of a judgment object by substituting a related *heuristic attribute* that comes more readily to mind¹¹².

¹¹¹ *Ivi*, p. 453.

¹¹² *Ivi*, p. 466. Corsivi nel testo.

Talvolta, poste di fronte a un quesito difficile, le persone rispondono a un quesito più facile: dato che l'attributo obiettivo è di difficile accessibilità, un altro attributo - che è correlato al primo e che è invece facilmente accessibile - viene sostituito al primo, il più delle volte in modo inconsapevole. Gli esperimenti mostrano che, in tutte le situazioni in cui la relazione associativa dell'attributo euristico con l'attributo obiettivo sia sufficientemente stretta da superare il controllo da parte del sistema 2, gli intervistati non si accorgono di rispondere a una domanda diversa, ma quando se ne accorgono modificano il giudizio o lo sostituiscono con uno diverso. La sostituzione di attributo è descritta come un'operazione del sistema 1 che avviene automaticamente e senza sforzo, mentre l'efficacia del sistema 2, che è deliberato e richiede un qualche sforzo, può essere in diversi casi compromessa, anche solo dalla pressione temporale o da un fattore umorale, e non sempre ci mette in grado di evitare alcuni *bias*.

L'accessibilità svolge un ruolo centrale nella descrizione del funzionamento dei due sistemi e nel loro coinvolgimento nella spiegazione dell'effetto framing e delle euristiche del giudizio.

Framing effects were attributed to the fact that alternative formulations of the same situation make different aspects of it *accessible*. The core idea of prospect theory, that the normal carriers of utility are gains and losses, invoked a general principle that changes are relatively more *accessible* than absolute values. Judgment heuristics were explained as the substitution of a highly *accessible* heuristic attribute for a less *accessible* target attribute¹¹³.

¹¹³ *Ivi*, p. 481. Corsivi miei

Anche se non è disponibile un elenco complessivo dei fattori che influenzano l'accessibilità, si può sicuramente dire che la somiglianza è più accessibile della probabilità, che le variazioni di ricchezza sono più accessibili dei valori assoluti, che le medie lo sono più delle somme. Un'analisi più approfondita dei processi duali può essere d'aiuto per spiegare il funzionamento dell'accessibilità e soprattutto per chiarire la relazione fra i due processi cognitivi.

2.11 *Affect heuristic*

Il modello esplicativo di sostituzione di attributo è adatto anche per comprendere il funzionamento di una nuova euristica, l'euristica dell'affetto, che rivendica l'importanza della dimensione affettiva nei giudizi e nelle decisioni.

As used here, *affect* means the specific quality of “goodness” or “badness” (1) experienced as a feeling state (with or without consciousness) and (2) demarcating a positive or negative quality of a stimulus. [...] The reliance on such feelings can be characterized as the *affect heuristic*¹¹⁴.

L'influenza della componente affettiva nei processi valutativi e decisionali è stata a lungo misconosciuta, sia dagli economisti sia dagli psicologi, probabilmente a causa della consolidata tendenza ad associare i processi deliberativi alla razionalità e quindi a processi puramente cognitivi.

¹¹⁴ P. Slovic et al., “The Affect Heuristic”, in *Heuristics and Biases: the psychology of intuitive judgement*, cit, p. 397.

Dagli inizi degli anni Ottanta ne è stata invece sottolineata l'importanza, dal momento che le reazioni affettive agli stimoli sono spesso le prime reazioni che intervengono automaticamente e guidano l'elaborazione delle informazioni e dei giudizi.

Lo psicologo americano Zajonc in un influente articolo del 1980 sosteneva che i sistemi affettivi e cognitivi erano largamente indipendenti e che le risposte del sistema affettivo sono più forti e più immediate. In effetti, le percezioni hanno sempre una caratterizzazione affettiva: “non vediamo solo una casa: vediamo una casa elegante, una casa orribile oppure una casa pretenziosa¹¹⁵”. Allo stesso modo, molto spesso scegliamo una casa perché la troviamo bella e poi giustifichiamo la scelta con ragioni diverse. Le emozioni hanno un ruolo centrale nelle *dual-processes theories*, dato che una delle modalità espressive del sistema intuitivo o esperienziale è proprio quella affettiva.

Although analysis is certainly important in some decision-making circumstances, reliance on affect and emotion is a quicker, easier, and more efficient way to navigate in a complex, uncertain, and sometimes dangerous world¹¹⁶.

L'ipotesi di base di Slovic è che immagini contrassegnate da sentimenti positivi e negativi guidano i giudizi e le decisioni. Più in particolare, l'idea è

¹¹⁵ R.B. Zajonc, “Feeling and thinking: preferences need no inferences”, *American Psychologist*, 1980, 35, pp. 151-75, p. 154.

¹¹⁶ P. Slovic, cit., p. 398.

che le persone usano un euristica affettiva nella formazione dei giudizi che consiste nella consultazione di una sorta di “affect pool” mentale che contiene i marcatori positivi e negativi associati in esperienze precedenti, in modo consapevole o meno, alle rappresentazioni di oggetti ed eventi. Proprio come la somiglianza e la memorabilità servono come suggerimenti per i giudizi sulla probabilità – euristiche della rappresentatività e della disponibilità – l’affetto entra come fattore in molti giudizi importanti. Un’impressione affettiva globale e immediatamente disponibile può rivelarsi più semplice e più efficace rispetto al calcolo dei pro e contro, o alla ricerca nella memoria di esempi significativi, soprattutto quando si tratti di una decisione complessa o per cui non abbiamo risorse mentali sufficienti. In questo senso la valenza affettiva è una specie di valutazione naturale e quindi un candidato ideale per essere usata come attributo euristico in sostituzione di valutazioni più complesse¹¹⁷.

2.12 Il contributo delle neuroscienze

Le tecniche di visualizzazione cerebrale – in particolare quella più recente, la risonanza magnetica funzionale (fMRI) - hanno permesso alle neuroscienze di localizzare alcuni processi cognitivi e affettivi in aree distinte del cervello. Anche le ricerche mediche sulle patologie o sui danni cerebrali causati da un incidente hanno offerto suggerimenti importanti in tal senso. L’attestazione di prestazioni inferiori da parte di persone che avevano subito danni in determinate aree del cervello, rispetto a persone “sane”, nello

¹¹⁷ D. Kahneman, *Nobel Lecture*, cit., p. 470.

svolgimento di un determinato compito, fornisce un ulteriore indizio sull'importanza di quell'area per lo svolgimento di quel compito.

Automatic and controlled processes can be roughly distinguished by where they occur in the brain. Regions that support cognitive automatic activity concentrated in the occipital, parietal and temporal parts of the brain. The amygdala, buried below the cortex, is responsible for many important automatic affective responses, especially fear. Controlled processes occur mainly in the front parts of the brain. The prefrontal cortex is sometimes called the “executive” region, because it draws inputs from almost all other regions, integrates them to form near and long-term goals, and plans actions that take these goals into account¹¹⁸.

Colin Camerer è un convinto sostenitore dell'importanza del contributo che le neuroscienze possono dare alla teoria economica nella ricerca di un approccio diverso al problema della scelta. A suo avviso, esse possono contribuire aggiungendo nuove variabili a quelle previste dalla teoria standard della decisione, oppure modificando radicalmente l'insieme di strutture concettuali utilizzate per spiegare il processo decisionale.

L'alternativa rappresenta due modi diversi di interpretare il contributo delle neuroscienze: il primo ne sottolinea il carattere puramente “incrementale” ed in un certo senso ne circoscrive la portata innovativa, il secondo invece attribuisce alle neuroscienze il compito ben più radicale di disegnare una diversa struttura interpretativa della presa di decisione. In

¹¹⁸ C. Camerer, G. Loewenstein and D. Prelec, “Neuroeconomics: how neurosciences can inform economics”, *Journal of Economic Literature*, 2005, 9-64, p. 17.

entrambi i casi, la premessa è il limite metodologico della teoria della decisione che – assumendo l'impossibilità di conoscere il contenuto del cervello – aveva trovato nell'espedito del “come se” il modo per porsi al riparo da un insufficiente sostegno empirico. Rispetto agli imperscrutabili stati mentali, il vantaggio dei concetti di utilità ordinale, di preferenza rivelata, di utilità attesa soggettiva era dato anche dall'escludere la possibilità di misurarli, mentre gli sviluppi più recenti compiuti dalle neuroscienze sembrano permettere di misurare direttamente sentimenti e pensieri. Ora, riprendendo le parole di Camerer, “the brain evidence has the *potential* to suggest better theory¹¹⁹”, e può contribuire a sostenere l'idea che il nostro comportamento sia il risultato dell'interazione fra sistemi automatici e controllati da una parte e sistemi cognitivi ed emotivi dall'altra.

Ci sono anche alcuni dati relativi all'evoluzione del cervello che possono rivelarsi significativi per comprendere le relazioni fra la dimensione emotiva e quella razionale.

A crucial fact is that the human brain is basically a mammalian brain with a larger cortex. This means human behavior will generally be a compromise between highly evolved animal emotions and instincts, and more recently evolved human deliberation and foresight¹²⁰.

¹¹⁹ C. Camerer, G. Loewenstein, D. Prelec, “Neuroeconomics: why economics needs brains”, *Scandinavian Journal of Economics*, 2004, 106 (3), pp. 555-79, p. 556.

¹²⁰ *Ivi*, p. 559.

La corteccia prefrontale è la sede della razionalità, il sistema limbico è invece la sede dell'affettività e dell'emotività e costituisce la struttura arcaica del cervello: si può dire che, nella storia dell'evoluzione, le emozioni hanno preceduto la razionalità¹²¹. La nostra vita mentale può essere considerata come il frutto di un equilibrio instabile fra lobo pre-frontale e lobo limbico, i contributi provenienti dalle due diverse aree cerebrali vanno quindi considerati complementari.

Complessivamente, il contributo delle neuroscienze, pur senza negare la parte che i processi deliberativi hanno nelle scelte, porta a riconsiderare il ruolo dei processi automatici e quello dei processi emotivi, che la teoria economica standard della massimizzazione dell'utilità ha invece trascurato. Prima di esaminare più in dettaglio questa che Camerer definisce una struttura teorica a due dimensioni¹²², mi sembra importante dar conto delle aspettative – più o meno ottimistiche – sollevate da questo tipo di indagini sul funzionamento del cervello, mettendo in seguito in evidenza le caratteristiche più significative dei processi neurali rispetto al tema della scelta.

¹²¹ A. Benini, *Che cosa sono io. Il cervello alla ricerca di sé stesso*, Milano, Garzanti, 2009, p. 43.

¹²² C. Camerer, G. Loewenstein and D. Prelec, "Neuroeconomics: how neurosciences can inform economics", cit., p.15

2.13 Un limite metodologico. Le caratteristiche fondamentali dei processi neurali

L'atteggiamento verso l'approccio basato sulle tecniche di visualizzazione cerebrale – come si è detto, uno dei più importanti metodi di indagine delle neuroscienze – può essere significativamente diverso, a seconda che ne vengano sottolineate le potenzialità oppure evidenziati i limiti. Va premesso che il concetto di base di questa tecnica è che le aree cerebrali attive durante lo svolgimento di un determinato compito richiedono un maggiore apporto di sangue, e questo maggiore afflusso viene rivelato con risonanze magnetiche particolari (fMRI).

Because most of these techniques involve localization of brain activity, this can easily foster a misperception that neuroscience is merely developing “a geography of brain”, a map of which brain bits do what part of the job. If that were indeed so, then there would be little reason for economists too pay attention. In reality, however, is beginning to use regional activity differences and other clues to elucidates the principles of brain organization and functioning, which in turn, is radically changing our understanding of how the brain works¹²³.

Rispetto a questo di Camerer, un parere più cauto è espresso dal neurologo Arnaldo Benini.

¹²³ *Ibid*

Il limite della tecnica, di cui si ha ragione di temere una sopravvalutazione, è di non rivelare ciò che nelle aree attive avviene¹²⁴.

Il limite della visualizzazione consiste in ciò, che aree del cervello possono manifestarsi attive nelle risonanze funzionali senza in realtà esserlo, per una variazione stocastica del flusso sanguigno e del metabolismo. La domanda è se l'area attiva nel momento dell'indagine lo sia veramente o se sia solo aumentato casualmente il suo metabolismo¹²⁵.

L'aspetto a mio avviso più interessante della posizione di Benini è il suo mettere l'accento sull'inevitabile limite metodologico insito nella natura autoreferenziale di un'indagine sul funzionamento del cervello. Ciò premesso, i dati sperimentali raccolti dalle neuroscienze sono sicuramente di grande interesse, ma essendo ancora molto parziali, non permettono analisi sufficientemente approfondite. Di conseguenza, anche le ipotesi interpretative non possono non essere ancora lontane dal proporre una teoria compiuta sull'attività neurale.

Per il tema della scelta, i tre aspetti più importanti del funzionamento del cervello umano sono l'automaticità, la modularità e l'attribuzione di senso¹²⁶. Molte delle attività del cervello sono processi rapidi, automatici, che operano in parallelo e in assenza di consapevolezza, perciò possiamo davvero non

¹²⁴ A. Benini, cit., pp. 51-2.

¹²⁵ A. Benini, "Intervista", *Humana.Mente*, v. 9, 2009, pp. 159-64.

¹²⁶ C. Camerer, G. Loewenstein and D. Prelec, "Neuroeconomics: why economics needs brains", cit., p. 559.

conoscere le cause del nostro comportamento. In positivo, l'automaticità implica la capacità del cervello di automatizzare procedure complesse, che all'inizio richiedono un grosso sforzo – come la guida di un'auto – e che invece dopo un po' di pratica possono essere eseguite “in automatico”, diventando persino difficili da insegnare ad altri. In negativo, l'automaticità implica che superare la dipendenza o qualsiasi comportamento abitudinario richieda un forte impegno cognitivo. La modularità, invece, comporta l'esistenza di moduli, spesso anatomicamente distinti, preposti allo svolgimento di compiti specifici. Tuttavia, i comportamenti più complessi, come sono quelli che interessano alla teoria economica, richiedono la comunicazione fra moduli e funzioni più specializzate, comunicazione che diventa particolarmente frenetica quando bisogna prendere una decisione importante.

Attention in neuroeconomics is therefore focused not just on specific regions, but also on finding “circuits” or collaborative systems of specialized regions which create choice and judgement¹²⁷.

L'importanza dei circuiti implica anche che i modelli esplicativi più adeguati siano quelli computazionali, dal momento che il comportamento è il risultato della collaborazione fra diverse componenti che andranno correttamente identificate. Accanto alla modularità, un'altra caratteristica importante del cervello è la sua plasticità, cioè il suo essere sensibile e reattivo

¹²⁷ *Ivi*, p. 561.

all'ambiente circostante, una caratteristica che è particolarmente evidente nell'età infantile, ma che è rilevante anche nello sviluppo degli adolescenti.

Infine, la forte propensione del cervello all'attribuzione di senso ci porta a essere sempre pronti a trovare spiegazioni per il nostro comportamento, e a preferire spiegazioni più sofisticate, che comportano un esplicito e consapevole atto deliberativo, invece di spiegazioni più banali, che coinvolgono una semplice azione meccanica. Un aspetto importante della tendenza all'attribuzione di senso è di essere fortemente dipendente dalle aspettative: in alcune situazioni l'impulso del cervello a imporre un ordine può indurci a vederlo anche dove non ce n'è affatto¹²⁸. Inoltre, nell'assimilare nuovi dati informativi il cervello procede in modo così rapido ed efficiente da scrivere sopra ciò che aveva precedentemente creduto, con la conseguenza di creare la forte sensazione di aver potuto prevedere un evento che in realtà non avremmo potuto prevedere – quella che chiamiamo “il senso di poi”.

Queste tre caratteristiche dell'attività del cervello hanno in comune, oltre alla rapidità e all'inconsapevolezza, il fatto di essere apparentemente lontanissime da quelle utilizzate per descrivere una decisione assunta con piena cognizione di causa. Oltre al carattere rapido e inconsapevole dei procedimenti neurali, in particolare l'ultima caratteristica, il *sense-making*, dà atto in un certo senso dell'impegno teoretico del modello di scelta razionale

¹²⁸ T. Gilovich, *How we know what isn't so: the fallibility of human reason in everyday life*, New York, Free Press, 1991.

della EUT, incentrato come è sul principio della coerenza, premessa indispensabile di qualunque interpretazione sensata.

2.14 Una prima distinzione: processi automatici e controllati

The rapid emergence of various dual-self or dual-process approaches testifies to how well economic theory can be adapted to study the brain as an organisation of interacting components¹²⁹.

Come si è detto, il primo contributo delle tecniche di visualizzazione alle teorie dualiste è la possibilità di localizzare l'attività dei differenti processi in determinate aree del cervello, senza sottovalutare l'importanza delle interconnessioni fra le diverse regioni neurali.

Sulla base anche di esperimenti che coinvolgono soggetti con lesioni in determinate aree del cervello, Camerer propone una caratterizzazione bidimensionale dei processi neurali. In questo modo al dualismo dei processi automatici e controllati – i sistemi 1 e 2 di cui parla Kahneman – si sovrappone il dualismo fra processi cognitivi e affettivi. L'idea di fondo è che i processi automatici – sia cognitivi sia affettivi – rappresentino la modalità di default del funzionamento del cervello, quindi costituiscano la maggior parte dell'attività elettro-chimica del cervello e siano sempre in funzione, anche quando sogniamo. Tuttavia, i processi automatici possono subire un'interruzione, come accade in momenti particolari – di fronte a un evento

¹²⁹ C. Camerer, "Neuroeconomics: using neuroscience to make economic predictions", *The Economic Journal*, 117, 2007, pp. 26-42, p. 28.

inatteso, provando un'emozione molto intensa o affrontando una nuova sfida – ed è allora che entrano in scena i processi controllati, una volta che le procedure automatiche sono state sospese e possono essere bypassate.

To the degree that controlled processes are well described by economic calculation but parallel processes are not, one could say that economics is about the “interrupt” or “override”¹³⁰.

L'intenzione non è quella di mettere da parte o di sottovalutare i procedimenti controllati, quanto di restituire spazio a quelli che qui Camerer chiama procedimenti paralleli – sottolineandone così la capacità di operare in modalità *multitasking* – che invece risultavano intrattabili per la teoria della scelta razionale, proprio a causa della loro natura irriflessa, dunque non volontaria.

Al di là della loro localizzazione in diverse aree del cervello, sono state avanzate anche ipotesi sulla loro relazione e sul loro funzionamento. Il processo automatico, secondo Schneider and Shrifin¹³¹, è l'attivazione – in conseguenza di appropriati input – di una sequenza di elementi che è stata appresa e memorizzata, e che procede poi automaticamente, senza richiedere un'attenzione impegnativa da parte del soggetto. Le sequenze possono anche essere il risultato dell'individuazione, tramite un ragionamento, di strategie

¹³⁰ C. Camerer, G. Loewenstein and D. Prelec, “Neuroeconomics: how neurosciences can inform economics”, cit., p. 18.

¹³¹ W. Schneider e R. Shiffrin, “Controlled and automatic human information processing: I. Detection, search, and attention”, *Psychological Review*, 1977, 84, pp. 127-190.

razionali per risolvere problemi semplici. La loro successiva memorizzazione permette di recuperarli in modo automatico, per riutilizzarli anche in contesti non necessariamente identici a quelli che li hanno originati.

Numerosi esperimenti hanno mostrato come di fronte a problemi risolvibili che possono prevedere soluzioni di diverso livello di efficienza, si tenda ad utilizzare uno stesso procedimento, applicandolo anche quando non raggiunge risultati ottimali. L'espressione "meccanizzazione del pensiero" si riferisce a processi mentali di questo tipo. La meccanizzazione consente il passaggio da un tipo di attività mentale deliberata e impegnativa a un tipo di operazioni mentali quasi automatiche, inconsapevoli e agevoli.

The experimental data on puzzle solving show that most individuals, once they have been able to identify one strategy and use it repetitively until it becomes familiar, do not abandon it even in new contexts where better strategies are available¹³².

Gli esperimenti mostrano quindi come un processo cognitivo controllato, costituito da operazioni mentali lente, seriali e faticose, venga sostituito da un più semplice e facile processo automatico. Una volta che siano state memorizzate, le sequenze diventano le componenti del processo intuitivo. Gli esperimenti di Luchins¹³³ mostrano anche come, in conseguenza di ripetute

¹³² M. Egidi, The dual process account of reasoning: historical roots, problems, perspectives, *CEEL, Working paper*, 2007, n. 6, p. 11.

¹³³ A.S. Luchins and E.H. Luchins, "New experimental attempts in preventing mechanization in problem-solving", *The Journal of General Psychology*, 1950, 42, pp. 279-291.

esperienze, vengano a formarsi dei moduli specializzati, dedicati all'adempimento di compiti specifici, che diventano a loro volta dei meccanismi automatici, testimoniando l'acquisizione di abilità specifiche, quelle che fanno la differenza fra un giocatore esperto ed un principiante. Un esempio, a cui si è già accennato, è quello riportato da Kahneman.

A master chess player does not see the same board as the novice, and visualizing the tower in an array of blocks would also become virtually effortless with prolonged practice¹³⁴.

La specificità del dominio e l'attivazione automatica sono i due requisiti che rendono fondamentale il ruolo dei moduli cognitivi specializzati. Egidi¹³⁵ sottolinea che questa nozione di moduli specializzati derivati dall'esperienza non va confusa né con quella proposta da Fodor né con quella della psicologia evoluzionistica. Per Fodor¹³⁶ sono meccanismi mentali innati, limitati però ai soli sistemi di input – quelli che presiedono all'elaborazione di percezione e linguaggio – e di output – responsabili dell'azione – mentre, a suo avviso, i cosiddetti sistemi centrali che sovrintendono ai procedimenti cognitivi più elaborati non sarebbero modulari. Questa tesi, che prevede una cognizione periferica modulare e una centrale non modulare, è respinta dagli psicologici evoluzionisti, favorevoli all'ipotesi della modularità massiva, secondo cui la mente umana è largamente se non interamente composta da moduli darwiniani

¹³⁴ D. Kahneman, *Maps of bounded rationality*, cit, p. 453.

¹³⁵ M. Egidi, cit., 2007, p. 12

¹³⁶ J. Fodor, *The modularity of mind*, Cambridge (MA), MIT Press, 1983.

innati. Va detto che quest'ultima ipotesi è però molto controversa, anche perché non individua una procedura sperimentale che possa provare l'esistenza di moduli innati e il loro ruolo nello sviluppo delle specifiche abilità.

Le più recenti tecniche di visualizzazione cerebrale sembrano confermare l'emergenza di moduli intesi come capacità cognitive specializzate, che si attivano automaticamente ma che si sviluppano con l'esperienza¹³⁷.

In a process that is not well understood, the brain figures out how to do the tasks it is assigned efficiently, using the specialized systems it has at his disposal. When the brain is confronted with a new problem it initially draws heavily on diverse regions, including, often, the prefrontal cortex (where controlled processes are concentrated). But over time, activity becomes more streamlined, concentrating in regions that specialized in processing relevant to the task¹³⁸.

Quando si diventa esperti il cervello sembra tendere gradualmente a trasferire l'elaborazione a sistemi specializzati, che possono risolvere i problemi con poco sforzo e maggiore efficienza. Le conoscenze di cui disponiamo non ci permettono di attribuire la stabilità di questi moduli alla presenza di una proprietà genetica o innata, come vorrebbero gli psicologi evolucionisti.

¹³⁷ M. Egidì, cit., p. 13.

¹³⁸ C. Camerer, G. Loewenstein and D. Prelec, "Neuroeconomics: how neurosciences can inform economics", cit., p. 23-24.

L'approccio tradizionale alla razionalità assumeva l'esistenza di capacità cognitive generali applicabili alle diverse tipologie di problemi. In tal modo era plausibile ritenere che di fronte a problemi simili, le prestazioni sarebbero state equivalenti. L'approccio dei due sistemi prevede invece che le prestazioni siano fortemente dipendenti dalla corrispondenza fra il problema che si ha di fronte e le capacità acquisite con le esperienze precedenti.

L'attivazione automatica di moduli specializzati è però anche una causa potenziale di *biases*. Gli errori possono dipendere dall'uso di moduli sbagliati, quando vengono utilizzati in contesti anche solo lievemente differenti da quelli originali. Oppure si possono commettere errori nella fase preliminare alla creazione di nuovi moduli, quella che prevede di rappresentarsi un problema scomponendolo nei suoi elementi costitutivi. In quest'ultimo caso, gli errori nella rappresentazione mentale di un problema possono essere la conseguenza di una categorizzazione inappropriata dei suoi elementi costitutivi, che consente di applicarli anche al di là di quello che è il loro ambito specifico.

2.15 Una seconda distinzione: processi affettivi e cognitivi

La seconda distinzione, quella fra processi affettivi e cognitivi, già ampiamente diffusa nella psicologia contemporanea, ha uno spazio rilevante nelle neuroscienze. Sul piano della geografia della mente, la sede dell'affettività è il cosiddetto sistema limbico costituito dall'amigdala, dall'ipotalamo e dall'ippocampo. L'ippocampo, definito da Benini "governatore del cervello emotivo perché connette le emozioni alla memoria

di cui è il guardiano”¹³⁹, ha proprio per questo un ruolo particolarmente importante. L’ippocampo non è l’archivio della memoria, che è invece dislocato in diverse aree cerebrali, ma sembra essere l’organo che consente di attingere ai ricordi depositati nell’archivio e di farli “rivivere”. Il duplice collegamento con la memoria e l’affettività spiega perché eventi con una forte caratterizzazione emotiva rimangono più a lungo nella memoria: un risultato coerente con quello degli esperimenti condotti da Kahneman, che evidenziavano l’importanza della “peak & end rule” come criterio guida nelle ricostruzioni delle esperienze passate di cui siamo stati protagonisti. La stessa capacità progettuale della mente che ci permette di immaginare e simulare scenari alternativi, è resa possibile grazie alla memoria del passato mediata dall’ippocampo¹⁴⁰. Il ruolo centrale che le passioni hanno nella nostra vita sembra quindi trovare conferma anche sul piano anatomico.

Il fatto che le fibre che vanno dal cervello emotivo a quello razionale siano il doppio di quelle che vanno nel senso contrario corroborerebbe la convinzione che l’esistenza è regolata soprattutto da affetti e passioni, con i quali la razionalità avrebbe, di solito, partita persa. Decine di ricerche dimostrano che la spiegazione del comportamento è una razionalizzazione a fatti avvenuti¹⁴¹.

Per la maggior parte degli psicologici, la principale caratteristica dell’affetto non è data dagli stati sentimentali con cui viene associato, ma dal

¹³⁹ A. Benini, *Che cosa sono io*, cit., p. 43.

¹⁴⁰ *Ivi*, p. 45.

¹⁴¹ *Ivi*, p. 46.

ruolo che ha per la motivazione umana, in pieno accordo con il primato delle passioni per l'azione sostenuto da Hume. Ancora sulla scia di Hume è in effetti la distinzione di Zajonc fra i processi affettivi, a cui ci si rivolge per le questioni che hanno la forma "go/not go" e che motivano una reazione di avvicinamento o di evitamento, e i processi cognitivi che invece rispondono alle questioni della tipologia "true/false"¹⁴². Anche se il più delle volte le passioni agiscono a livello inconscio, hanno tutte una valenza positiva o negativa, e molte comportano una tendenza all'azione o effetti di altro tipo sulla percezione, sulla memoria o sulle preferenze.

2.16 *Visceral factors*

Il termine affetto o passione è comprensivo anche di quelli che vengono chiamati "visceral factors". E' indiscutibile che, rispetto ai processi di più alto livello cognitivo, stimoli come la fame o sentimenti come la paura svolgano un ruolo fondamentale per la sopravvivenza quotidiana. In realtà fattori emotivi di questo tipo hanno un ruolo essenziale, e in molti casi dominante, nel motivare il comportamento umano in generale, anche se si tende a sottostimarne l'influenza e a sovrastimare invece l'importanza di processi di più alto livello cognitivo.

Come si è detto, si tende a interpretare il comportamento come il risultato di un atto decisionale consapevole anche quando non è affatto così.

¹⁴² R.B. Zajonc, Emotions, in *The Handbook of social psychology*, New York, Oxford University Press, 1998, citato in C. Camerer, G. Loewenstein and D. Prelec, cit., p. 17.

Messi a confronto con la coerenza e la stabilità, quanto meno a breve termine, che caratterizzano le preferenze - comunemente considerate i moventi delle scelte – questi fattori viscerali sono altamente mutevoli, e pertanto la loro influenza sul comportamento viene erroneamente ritenuta del tutto imprevedibile. In realtà l’influenza che essi esercitano sul comportamento – come pure le cause che li determinano – sono altamente sistematiche e sono invece i processi deliberativi cognitivi ad essere fonte di imprevedibilità¹⁴³. L’intervento della consapevolezza svolgerebbe così la funzione diametralmente opposta di prevalere su quelli che sono prevedibili modelli di comportamento istintivo.

La cautela degli economisti nei confronti di questi fattori istintivi può essere comunque giustificata sia del fatto che spesso questi impulsi portano le persone a comportarsi in modo contrario ai loro interessi – basti pensare alle reazioni inconsulte di cui si è capaci a causa dello stress da traffico automobilistico – sia per la diffusa tendenza a sottostimare l’impatto dei fattori viscerali sul nostro comportamento presente o futuro, una tendenza che diventa la responsabile principale dei frequenti fallimenti cui vanno incontro tanti buoni propositi.

Nelle situazioni di scelta in cui questi impulsi istintivi si trovano a confrontarsi con più meditati livelli cognitivi che difendono un parere opposto al loro, si potranno verificare dei conflitti interpersonali la cui risoluzione

¹⁴³ G. Loewenstein, “Emotions in economic theory and economic behavior”, *The American Economic Review*, 2000, v. 90, n. 2, pp. 426-432.

dipenderà in parte dall'intensità degli impulsi, e in parte dalla nostra capacità di resistervi, anzitutto facendo attenzione a non sottovalutarli, prima che la situazione ci sfugga completamente di mano. In effetti, il problema che questi impulsi pongono per una teoria della decisione è che, se non si può ignorarne l'influenza sulle preferenze, la valutazione negativa che il più delle volte si dà di quest'influenza ci impedisce di mettere sullo stesso piano le preferenze influenzate da pulsioni viscerali e quelle che non lo sono.

Clearly, welfare maximization lies somewhere between the two extremes of making decisions that ignore visceral factors and treating visceral influences as no different from any other influence on tastes¹⁴⁴.

Le due circostanze in cui il ruolo svolto da questi fattori è particolarmente importante sono le scelte intratemporali, quando – il più delle volte di fronte a problemi da autocontrollo – la forza di un *visceral factor* ci fa compiere scelte che non tengono conto delle conseguenze future, e le scelte in condizioni di rischio. In quest'ultimo caso, le valutazioni cognitive del rischio possono divergere dalle reazioni emotive indotte da questo stesso rischio perché i fattori determinanti sono diversi. Infatti, di fronte a una situazione a rischio, anche se la valutazione cognitiva del rischio in assenza di nuove informazioni resta invariata, la paura può aumentare con l'approssimarsi di un pericolo e provocare il noto fenomeno del “*chickening out*”, ossia del darsela a gambe proprio nel momento cruciale. Il sistema affettivo e il sistema cognitivo rispondono a informazioni differenti, e in modo differente alle stesse

¹⁴⁴ *Ivi*, p. 429.

informazioni. Così, la vista o l'odore di un biscotto può attivare la motivazione a mangiarlo da parte del sistema affettivo, ma potrebbe anche rammentare al sistema cognitivo che si è a dieta.

Un esempio invece delle distorsioni che le emozioni possono causare nei giudizi cognitivi è il *wishful thinking*. In questo caso le emozioni creano una sorta di “motivated cognition”, come nelle situazioni in cui riusciamo a convincerci che ciò che ci piacerebbe accadesse, accadrà davvero, solo perché lo desideriamo. E' anche vero però che, se non ne sottostimiamo la forza, siamo in grado di neutralizzare gli effetti negativi di questi impulsi utilizzando semplici strategie, come quella di evitare le situazioni che possono indurci in tentazioni oppure quella di utilizzare l'immaginazione per evocare, al momento opportuno, tutte le conseguenze negative che ne seguirebbero.

2.17 Interazione come cooperazione o competizione

Al di là dell'utilità euristica di distinguere fra processi controllati e automatici e fra processi cognitivi e affettivi, la maggior parte dei giudizi e dei comportamenti è il risultato della loro interazione.

Behavior emerges from a continuous interplay between neural systems [...]
Three aspects of this interaction bear special emphasis, which we labeled
“collaboration”, “competition” and “sense-making”¹⁴⁵.

¹⁴⁵ C. Camerer, G. Loewenstein and D. Prelec, “Neuroeconomics: how neurosciences can inform economics”, cit., p. 28.

Nella visione di Camerer, la cooperazione contiene l'idea che per prendere una decisione, “which is to say “rationality” in the broad non technical sense of the word¹⁴⁶”, non occorra trasferire il potere dal livello automatico, affettivo a quello deliberativo, quanto piuttosto assicurare un appropriato livello di cooperazione fra i diversi sistemi. La competizione, invece, riflette il fatto che i diversi processi – in particolare, come si è detto, quello affettivo e quello cognitivo – possano andare in direzioni opposte e trovarsi a competere per il controllo del comportamento. Infine, l'interpretazione fa riferimento al modo in cui diamo un senso ai processi di cooperazione e competizione, e in generale alla tendenza a cercare spiegazioni soddisfacenti per il nostro comportamento. In questo caso, trattandosi di un processo cognitivo impegnato, attraverso una consapevole riflessione, a individuare le ragioni del mio agire, l'interpretazione tenderà a trovare spiegazioni della sua stessa natura, e cioè intenzionali.

Riguardo ancora alla cooperazione, va sottolineato che il bilanciamento fra i sistemi sembra essere una premessa indispensabile per giudizi e scelte “equilibrate”, mentre fra gli stili di decisione estremi si situano da un lato i comportamenti impulsivi, che sono vulnerabili e si lasciano influenzare troppo facilmente, e dall'altro quelli ossessivamente prudenti, che non si risparmiano calcoli interminabili anche per le scelte più semplici.

¹⁴⁶ *Ibid*

L'importanza della dimensione affettiva nelle decisioni non è comunque in discussione, sia in funzione collaborativa che competitiva. Damasio¹⁴⁷ ha ipotizzato l'esistenza di input provenienti dal sistema affettivo sotto forma di valutazioni affettive delle opzioni di scelta, e ha denominato tali input "marcatori somatici". I suoi studi hanno mostrato che individui senza difetti cognitivi ma con gravi deficit affettivi incontrano grandi difficoltà nel prendere decisioni, e quando ci riescono si tratta in genere di cattive decisioni. Non basta *sapere* quel che si dovrebbe fare, bisogna anche *sentirlo*. Decisioni insoddisfacenti possono anche scaturire da un impedimento esercitato dal pensiero deliberativo, che impedisce di esprimere le reazioni emotive scatenate dalla questione in gioco o dalle distorsioni di giudizi cognitivi che possono essere originate dall'influenza che gli affetti hanno sui ricordi o sulle percezioni, connotandoli in senso positivo e negativo. In tutti i casi, il ruolo delle passioni dipenderà anche dalla loro intensità.

The extent of collaboration and competition between cognitive and affective systems, and the outcome of the affect when it occurs, depends critically on the intensity of affect¹⁴⁸.

¹⁴⁷ A.R. Damasio, *Descartes' error: emotion, reason, and the human brain*, New York, Putnam, 1994.

¹⁴⁸ C. Camerer, G. Loewenstein and D. Prelec, "Neuroeconomics: how neurosciences can inform economics", cit., p. 30. Dello stesso parere anche Loewenstein, "Out of control: visceral influences on behavior", *Organizational Behavior and human decision processes*, 1996, 65, pp. 272-292.

Se è di bassa intensità, l'affetto avrà un ruolo prevalentemente consultivo, che secondo diverse teorie¹⁴⁹ consiste nel fornire informazioni, o motivazioni, che vengono utilizzate come input per le scelte. A un livello medio di intensità corrisponde invece una maggiore consapevolezza dei potenziali conflitti fra input cognitivi e affettivi, ed è a questi livelli che si osservano i tentativi di esercitare l'autocontrollo che sono stati l'oggetto di molti studi¹⁵⁰. Al livello più alto, perdiamo il controllo.

Finally, at even greater levels of intensity, affect can be so powerful as to virtually preclude decision making. No one “decides” to fall asleep at the wheel, but many people do. Under the influence of intense affective motivation, people often report themselves as being “out of control” or “acting against our own self-interest”¹⁵¹.

Va sottolineato che, al di là della distinzione fra le due diverse tipologie di processi mentali, nell'analisi di Camerer la decisione viene comunque identificata con la razionalità, in quanto è un atto deliberato e consapevole, per cui in assenza di queste due caratteristiche non si può parlare di decisione.

¹⁴⁹ A. R. Damasio, cit; P. Slovic et al., “Risk as analysis and risk as feeling: some thoughts about affect, reason, risk and rationality”, *Risk analysis*, 2004, v. 24, n. 2, pp. 311-322.

¹⁵⁰ Fra i più noti, J. Elster, *Ulysses and the Sirens*, Cambridge, Cambridge University Press, 1977, T.C. Schelling, *Choice and consequence*, Cambridge (Mass.), Harvard University Press, 1984.

¹⁵¹ C. Camerer, G. Loewenstein and D. Prelec, “Neuroeconomics: how neurosciences can inform economics”, cit., p. 31.

L'interpretazione è un'altra forma di interazione fra sistemi neurali, e si è già sottolineato come si tratti di uno degli aspetti più importanti, assieme alla automaticità e alla modularità, del funzionamento del cervello umano. L'idea è che il nostro continuo sforzo di interpretare i nostri comportamenti, di dare una struttura narrativa alla nostra vita mentale, scaturisca da un forte impulso del cervello al conferimento di senso. Si spiega così anche l'illusione di poter sempre trovare il significato di ciò che facciamo, nonostante la maggior parte dell'attività neurale sia in realtà inaccessibile.

Riguardo a questa forma di inaccessibilità cognitiva, gli esperimenti di Libet¹⁵² avrebbero mostrato che il preciso momento in cui diventiamo consapevoli dell'intenzione di compiere un'azione in realtà sarebbe successivo all'iniziale onda di attività cerebrale associata a quell'azione di circa 300 millisecondi. La risposta comportamentale seguirebbe a sua volta la sensazione dell'intenzione di altri 200 millisecondi. Ci sarebbe perciò una lacuna temporale fra l'inizio dell'attività neurale e l'istante in cui sentiamo di aver deciso di compiere l'azione. Nella coscienza viene registrata una regolare combinazione della sensazione di intenzione seguita dal comportamento manifesto, ma dato che l'attività neurale che precede l'intenzione è inaccessibile, sperimentiamo ciò che chiamiamo il libero arbitrio, che non sarebbe altro che la conseguenza dell'impossibilità di identificare qualcosa che sia causa del sentimento dell'intenzione. Poiché il comportamento segue sistematicamente l'intenzione, abbiamo la sensazione che l'intenzione,

¹⁵² B. Libet, "Unconscious cerebral initiative and the role of conscious will in voluntary action", *Behavior and brain sciences*, 1985, 2, pp. 146-159.

espressione del nostro libero arbitrio, sia la causa dell'azione, quando invece sia l'intenzione sia l'azione sarebbero causati da eventi neurali precedenti, inaccessibili alla coscienza.

Riguardo al tema dell'illusorietà del libero arbitrio, l'esperimento di Libet è molto discutibile se non altro perché presuppone l'esistenza di una sorta di centro unificato di comando cui farebbero capo sia il centro della visione sia la facoltà del ragionamento pratico, troppo simile all'omuncolo cartesiano¹⁵³ e che si rivelerebbe palesemente inefficiente. Se intendiamo il processo decisionale come il risultato di diversi eventi neurali distribuiti su più livelli, invece che come un singolo atto istantaneo, la tesi della negazione del libero arbitrio perde consistenza. Questa descrizione è invece più interessante se la leggiamo come una reazione all'inaccessibilità cognitiva, come testimonianza dello sforzo di non rinunciare mai a trovare una spiegazione ai nostri comportamenti che ci confermi nel ruolo di protagonista.

The brain contains a specific cognitive module that binds intentional actions to their effects to construct a coherent conscious experience of our own agency¹⁵⁴.

¹⁵³ D. Dennett, *Freedom evolves*, London, Allen Lane, 2003, pp. 227-242.

¹⁵⁴ P. Haggard, S. Clark e J. Kalogeras, "Voluntary action and conscious awareness", *Nature Neurosciences*, 2002, v. 5, n. 4, pp. 382-385; p. 385.

Capitolo 3 - Discussioni filosofiche sulla razionalità pratica

3.0 Premessa

Le interpretazioni della EUT il più delle volte chiamano in causa – in modo più o meno esplicito – la natura della razionalità della scelta presupposta dalla teoria. Lasciando da parte l’aspetto epistemologico, cioè se la EUT debba essere interpretata come una teoria descrittiva o normativa, si tratterà di esaminare in che modo le discussioni filosofiche sui meccanismi della scelta possano chiarire i presupposti concettuali della EUT, in particolare riguardo al significato e al ruolo della razionalità, alla relazione fra preferenze e scelte, al modo in cui interpretarne gli assiomi, ecc. La principale difficoltà di questa operazione risiede probabilmente nel fatto che la EUT grazie all’identificazione di scelte e preferenze mette fuori gioco il contesto della motivazione, che ha invece un ruolo centrale nelle teorie filosofiche dell’azione. D’altra parte, proprio il fatto di fare della scelta il punto di partenza costituisce sia un punto di forza sia un elemento di debolezza della teoria. Come cercherò di dimostrare, alcuni dei cosiddetti fallimenti della EUT possono dipendere da questa impostazione.

Le discussioni filosofiche intorno alla teoria della scelta (anche non direttamente legate alla EUT) hanno riguardato la rilevanza degli elementi cognitivi nella deliberazione, e cioè le credenze sul contesto e le informazioni disponibili sullo stato del mondo in cui si agisce, ma anche l’imprescindibilità della componente conativa, il desiderio, da molti considerato un fattore

essenziale, in quanto sarebbe l'unico a poter dare l'impulso decisivo per effettuare la scelta. Si parla a questo riguardo anche del ruolo dell'intenzionalità, sia come un terzo fattore, oltre alla credenza e al desiderio, sia in combinazione con la credenza al fine di tradurre in pratica la capacità cognitiva. Queste discussioni portano a contrapporre una lettura cognitivista e una non cognitivista del contesto della scelta: la prima disposta a far coincidere causa e ragione di una scelta e a dare una giustificazione razionale delle motivazioni, la seconda portata a distinguere fra cause e ragioni per sottolineare la possibilità di situazioni in cui la causa di un'azione non è anche la ragione e quindi la motivazione non avrà necessariamente un carattere cognitivo.

Va detto che, al di là del carattere più o meno cognitivo della motivazione, la relazione fra cause e ragioni è al centro della contrapposizione dei due modelli esplicativi dell'azione, quello causalista e quello teleologico. Il confronto fra queste due posizioni – la prima rappresentata da Davidson, la seconda dalla Anscombe – verte sulla natura degli stati mentali e sulla possibilità di considerarli vere e proprie cause delle azioni.

Un altro problema legato alla combinazione e alla distinzione di ruoli fra credenza e desiderio riguarda l'esistenza o meno di una differenza sostanziale fra il piano teoretico e quello pratico, e cioè la possibilità di stabilire un confronto fra la relazione di premesse e conseguenze in un sillogismo teoretico e l'individuazione dell'azione che ci permetterà di raggiungere l'obiettivo in un sillogismo pratico. Se si accetta l'idea di una simmetria fra il piano teoretico e quello pratico, allora la correttezza dell'informazione sul contesto in cui operiamo diventa fondamentale per la scelta dell'azione più

appropriata al raggiungimento dei nostri scopi e sarà tradotta in pratica dall'intenzione. Il presupposto è che il ragionamento che riguarda la presa di decisione è comunque, ovviamente, essenzialmente teoretico. Interrogandosi sulla natura delle *ragioni* di una scelta, Parfit – come vedremo – difende la superiorità delle ragioni normative rispetto alle ragioni motivanti: le prime fanno riferimento ad una razionalità forte della corrispondenza con fatti e valori normativi, mentre le seconde rispecchiano una razionalità cosiddetta procedurale, che stabilisce un legame *solo* psicologico fra le motivazioni dell'agente e le sue scelte.

Un altro punto importante, soprattutto, ma non solo, per le scelte interattive, si trova su una linea di confine fra razionalità ed etica. Nella letteratura sulla scelta si parla di razionalità strumentale o sostantiva; solo nel secondo caso è in discussione anche il valore (interpretato in termini di razionalità) dell'obiettivo o del fine della scelta, nel primo caso si guarda solo alla capacità rappresentativa della scelta, ossia al suo essere lo *strumento* più adeguato per i nostri fini. E' evidente che nel secondo caso la razionalità ha un immediato risvolto etico e, dato che i nostri desideri sono rappresentativi dei nostri obiettivi, il perseguimento di questi ultimi potrà avere un indiscutibile valore normativo di carattere morale. Nel primo caso la razionalità riguarderà invece una sorta di relazione – secondo alcuni paragonabile a quella di coerenza logica e quindi sicuramente normativa – fra la scelta e la soddisfazione procurata dal raggiungimento dell'obiettivo. Non sembra esserci in questo caso nessun giudizio di valore sull'obiettivo in sé, ma solo sulla efficacia, strumentale appunto, della scelta in vista del fine.

L'espressione "razionalità strumentale" (*means-ends reasoning*) nasconde però un possibile equivoco. Nel contesto sopra descritto – che vorrebbe essere rappresentativo dell'uso che più comunemente si fa di questa espressione in una prospettiva utilitaristica – l'attenzione è sull'appagamento del desiderio che assume in un certo senso un valore in sé, ed è perciò indipendente dal contenuto, ossia da ciò che effettivamente lo produce. Di conseguenza, anche in questo caso la razionalità avrà un valore normativo, che consisterà nel soddisfacimento o nell'incremento di benessere (o utilità) che potrà essere conseguito con una data scelta. Si può dire che la ragione è vista ancora una volta come l'*autorità* cui occorre conformarsi.¹⁵⁵ Un'alternativa è quella che Blackburn chiama una concezione della ragione "Humean-friendly", secondo cui le nostre più alte funzioni cognitive possono ovviamente esercitare un controllo su impulsi e desideri ma senza pensare di poterne prescindere, e soprattutto senza vedere nella ragione una fonte indipendente e autosufficiente da cui far scaturire l'azione¹⁵⁶.

3.1 Donald Davidson. Credenza, desiderio, intenzionalità

Nelle spiegazioni delle modalità in cui avviene la scelta sono questi gli elementi cui si fa più spesso ricorso. Cambia però, come si è già accennato, il ruolo che viene loro attribuito.

¹⁵⁵ R. Sugden, "David Hume's Treatise of Human Nature", in *Topoi*, 2008, 27, pp. 153-159

¹⁵⁶ S. Blackburn, *Ruling passions*, p.245; "Précis of Ruling passions", in *Philosophy and Phenomenological Research*, LXXV, 1, 2002.

La teoria dell'azione di Davidson è sicuramente un punto di partenza appropriato, anche perché le sue riflessioni investono spesso direttamente la teoria della scelta. Prima ancora dei suoi più importanti testi filosofici, in un saggio del 1957 scritto con l'economista Suppes, Davidson ha difeso l'applicabilità empirica della teoria normativa della decisione respingendo come fuorviante l'alternativa fra un approccio descrittivo e uno normativo alla teoria e proponendo un modello che si discosta da quelli standard di von Neumann e Morgenstern e di Savage "not in demanding anything different of the rational decision maker, but by demanding less"¹⁵⁷. È sufficiente "indebolire" gli assiomi per migliorarne la capacità predittiva.

Il saggio *Actions, Reasons and Causes*¹⁵⁸ si apre con una domanda : "What is the relation between a reason and an action when the reason explain the action by giving the agent's reason for doing what he did?" Davidson propone di chiamare queste spiegazioni *razionalizations*. La tesi che vuole difendere è "the ancient and common-sense position that razionalization is a species of causal explanation". Come va inteso il termine *razionalizations* è subito chiarito:

¹⁵⁷ D. Davidson, Suppes e Siegel, *Decision-Making: An Experimental Approach*, Stanford, Stanford University Press, 1957, p. 6. E' importante ricordare che nelle misurazioni sperimentali dell'utilità gli autori accettano la proposta di definire l'utilità sui guadagni e sulle perdite piuttosto che sulle posizioni patrimoniali finali. Cfr. D. Kahneman e A. Tversky, *Prospect Theory*, cit., p. 276.

¹⁵⁸ D. Davidson, *Essays on action and events*, p.3-19

A reason rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his action – some feature, consequence, or aspect of the action the agent wanted, desired, prized, held care, thought dutiful, beneficial, obligatory, or agreeable. We cannot explain why someone did what he did simply by saying the particular action appealed to him; we must indicate what it was about the action that appealed ¹⁵⁹.

Secondo Davidson, la ragione primaria (*primary reason*) dell'azione, quella che serve a spiegarla, va ricercata nella coppia formata da desiderio e credenza. Davidson preferisce il termine più generale di *pro-attitude* per il primo elemento della coppia, “some sort of pro attitude toward action of a certain kind”, che comprende

desires, wantings, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values in so far as those can be interpreted as attitudes of an agent directed toward actions of a certain kind ¹⁶⁰.

Il vantaggio del termine *attitude* è di poter essere utilizzato per indicare sia un tratto caratteriale assolutamente costante sia un improvviso capriccio passeggero. Davidson insiste a non intenderlo come se implicasse il forte convincimento che un dato tipo di azioni debba essere portato a termine, valga la pena di essere portato a termine o sia, tutto considerato, desiderabile. E' possibile avere per tutta la vita un desiderio irrefrenabile di fare qualcosa –

¹⁵⁹ *Ivi*, p. 3.

¹⁶⁰ *Ivi*, p. 4

l'esempio di Davidson è "to drink a can of paint" - senza credere che valga la pena farlo. Si tratta di una precisazione che serve al tempo stesso a sottolineare i differenti ruoli di desiderio e credenza e a chiarire il senso della loro combinazione al fine di spiegare un'azione. Così, ad esempio, l'azione di premere il pulsante della radio è spiegata insieme dal desiderio di ascoltare della musica e dalla credenza che, premendo quel pulsante, si accenda la radio.

L'interazione fra credenza e desiderio viene a costituire un sistema di riferimento indispensabile per rendere intellegibile un'azione ed è in questo senso costitutiva della sua razionalità. Allo stesso modo, la possibilità di spiegare la scelta di un agente a partire dalla combinazione di un desiderio (*pro-attitude*) e di una relativa credenza è una condizione indispensabile per poterla considerare razionale. Un punto centrale è che, per Davidson, a qualificare un'azione, marcandone la differenza rispetto ad altri tipi di eventi, è proprio il fatto di poter essere descritta attraverso quegli atteggiamenti intenzionali rappresentati da credenze e desideri. Questo significa che le ragioni di una scelta dipendono dalla descrizione che ne diamo, e che per capire un comportamento facciamo delle ipotesi sugli stati mentali dell'agente, cioè proviamo a interpretarlo¹⁶¹.

When we ask why someone acted as he did, we want to be provided with an interpretation. His behavior seems strange, alien, outré, pointless, out of character, disconnected; or perhaps we cannot even recognize an action in it. When we learn his reason, we have an interpretation, a new description of what

¹⁶¹ M. De Caro, *Dal punto di vista dell'interprete. La filosofia di Davidson*, Roma, Carocci, 1998, p. 81-82.

he did, which fits it into a familiar picture. The picture includes some of the agent's beliefs and attitudes; perhaps also goals, ends, principles, general character traits, virtues and vices¹⁶².

In questo passo Davidson espone con grande chiarezza cosa succede nel momento in cui capiamo le ragioni di un'azione, che inizialmente sembrava incomprensibile, riuscendo a interpretarla. A quel punto vediamo l'azione sotto una nuova luce (*a new description*) e possiamo ricondurla all'interno di un quadro familiare (*a familiar picture*) in cui ritroviamo diversi tipi di atteggiamento intenzionale che ben conosciamo.

Beyond this, the redescription of an action afforded by a reason may place the action in a wider social, economic, linguistic, or evaluative context. To learn, through learning the reason, that the agent conceived his action as a lie, a repayment of a debt, an insult, the fulfillment of an avuncular obligation, or a knight's game is to grasp the point of the action in its setting of rules, practices, conventions, and expectations¹⁶³.

Comprendendone le ragioni afferriamo il senso dell'azione, inserendola nell'appropriato contesto di riferimento che non è dato soltanto dal movente che ci ha spinti o dal risultato che si voleva ottenere, ma anche da quell'insieme di "regole, pratiche, convenzioni, aspettative" che costituiscono lo sfondo di pratiche sociali condivise. Per Davidson però condividere questa impostazione "olistica" propria dell'ultimo Wittgenstein non significa

¹⁶² D. Davidson *op. cit.*, p. 9-10.

¹⁶³ *Ivi*, p. 10.

condividerne anche l'idea che proprio la dipendenza delle ragioni dalle descrizioni delle azioni le differenzi dalle cause che sono invece sicuramente separate dagli effetti. Al contrario, l'esistenza di un legame concettuale fra ragioni e azioni non esclude, ad esempio, la possibilità che queste ultime possano smentire la validità di un'interpretazione, mettendo in discussione il valore di verità delle ragioni che erano state addotte¹⁶⁴. La tesi di Davidson è che le ragioni non acquistano un senso solo all'interno delle descrizioni, ma mantengono una relazione esplicita, causale, con la realtà delle azioni. Solo così possiamo individuare fra le diverse possibili interpretazioni – e quindi fra diverse ragioni possibili – di un'azione, la vera ragione, quella che ha effettivamente causato l'azione¹⁶⁵. Se la ragione è anche la causa dell'azione, la spiegazione che fa riferimento alle ragioni sarà al tempo stesso una spiegazione causale¹⁶⁶. La relazione fra le cosiddette ragioni primarie e le azioni è una relazione fra due eventi che possono essere descritti in termini di ragioni o in termini di cause; nel primo caso si darà una spiegazione teleologica – nell'esempio fatto sopra, il desiderio di ascoltare della musica e la credenza che qualche stazione radio l'avrebbe trasmessa – nel secondo caso lo stesso evento sarà descritto utilizzando un vocabolario physicalistico che descriva l'attività neurofisiologica e motoria implicata nella stessa azione. La

¹⁶⁴ *Ivi*, p. 81: “L'interpretazione che Arthur Chamberlain diede di Hitler durante i colloqui di Monaco – come di una persona di cui ci si potesse fidare – venne drammaticamente invalidata dalle invasioni della Cecoslovacchia e della Polonia”.

¹⁶⁵ “A primary reason for an action is its cause” in D. Davidson, *op. cit.*, p.12

¹⁶⁶ D. Davidson, “Actions, reasons and causes”, 1963 in *Essays on actions and events*,

prima spiegazione fornirà una giustificazione dell'azione, la seconda potrà esemplificare una legge di natura, tuttavia “entrambe [...] daranno conto, spiegandola, di una relazione causale tra due eventi, *che rimane causale indipendentemente da come quegli eventi vengono descritti*”¹⁶⁷. Si tratta però di un parallelismo che non implica la possibilità di ridurre la prima spiegazione alla seconda, dal momento che non ci sono vere e proprie leggi riguardanti relazioni fra eventi mentali e eventi fisici¹⁶⁸. Davidson spiega perché le generalizzazioni che mettono in relazione ragioni e azioni non possono dar luogo a leggi da cui sia possibile ricavare previsioni accurate e affidabili¹⁶⁹.

If we reflect on the way in which reasons determine choice, decision, and behavior, it is easy to see why this is so. What emerges in the *ex post facto* atmosphere of explanation and justification, as *the* reason frequently was, to the agent at the time of action, one consideration among many, *a* reason. Any serious theory for predicting action on the basis of reasons must find a way of evaluating the relative force of various desires and beliefs in the matrix of decision; it cannot take as its starting point the refinement of what is to be expected from a single desire¹⁷⁰.

Due sono i punti rilevanti in questo brano. Il primo riguarda la diversa prospettiva rispetto a una data azione, quella *ex ante* – che riflette il punto di

¹⁶⁷ M. De Caro, *op. cit.*, p. 83.

¹⁶⁸ Id, *Mental Events*

¹⁶⁹ D. Davidson, “Actions, reasons and causes”, in *op. cit.* p. 15-16

¹⁷⁰ *Ivi*, p. 16.

vista dell'agente – in cui sono in gioco le varie possibili ragioni di un'azione e quella *ex post* – che è invece il punto di vista esterno di colui che interpreta un'azione – in cui “emerge” la ragione. E' solo dal punto di vista dell'interprete che è possibile individuare la ragione vera e cioè la causa dell'azione.

Il secondo punto riguarda la presenza di una pluralità di desideri e credenze all'interno della “matrice della decisione”, che impone una valutazione della loro forza relativa. Secondo Davidson, il nesso tra ragioni e azioni può essere rappresentato da un sillogismo pratico che, a partire dalle premesse - costituite da un desiderio che mira al fine dell'azione e da una credenza relativa ai mezzi per conseguire quel fine - dedurrà come conclusione la desiderabilità dell'azione. In risposta alla Anscombe che ne negava il carattere deduttivo, Davidson precisa che questo sillogismo non corrisponde a un “pezzo di ragionamento pratico” e perciò la conclusione non sarà un giudizio sulla desiderabilità di un'azione: “trattandosi, appunto, di un ragionamento *pratico*, non accadrà che l'agente inferisca realmente la conclusione logica dalle premesse; egli piuttosto *agirà* in conformità con tale conclusione”¹⁷¹. Come conseguenza, questo sillogismo avrà il limite di non poter ricostruire nello specifico il ragionamento pratico in cui soppesiamo le diverse ragioni che sono in gioco, e perciò non fornirà un modello teorico “neither for a predictive science of action nor for a normative account of

¹⁷¹ M. De Caro, *op. cit.*, p. 85. La replica di Davidson alla Anscombe in D. Davidson, *op. cit.*, p. 9 nota 4.

evalutative reasoning”¹⁷². Il fatto che la relazione fra ragioni e azioni non sia riconducibile a vere e proprie leggi di natura non basta però a confutare la tesi che si tratti anche di una relazione causale.

Nelle razionalizzazioni, che spiegano le azioni in funzione delle rispettive ragioni, coesistono perciò il carattere teleologico (perché le ragioni ci spiegano il *fine* delle azioni, l’intenzione con cui sono compiute) e quello causale (nel senso della causalità efficiente, propria delle scienze naturali)¹⁷³.

Questa duplicità di carattere – e in particolare l’esistenza di un legame causale fra ragioni e azioni – trova un fondamento ontologico grazie alla teoria dell’identità delle occorrenze che è al centro della filosofia della mente di Davidson, il monismo anomalo, che pur identificando mente e cervello (fisicalismo ontologico) ritiene che i fenomeni mentali non siano completamente riducibili alle leggi dei fenomeni fisici (anomalo = senza leggi). Secondo la teoria dell’identità delle occorrenze, infatti, l’occorrenza di un singolo evento che rientra in un certo tipo di eventi mentali è identica all’occorrenza di un evento che ricade sotto un certo tipo di eventi fisici, ma ciò non implica una corrispondente identità fra i due tipi di eventi, mentali e fisici. Per questo è possibile descrivere una stessa azione – cioè, come si è già detto, un evento intenzionale – in termini mentalistici e in termini fisicalistici, mentre non è possibile ridurre la prima descrizione alla seconda.

¹⁷² *Ivi*, p. 16

¹⁷³ M. De Caro, *op. cit.*, p.83.

3.2 Ancora Davidson: comportamenti irrazionali

E' interessante il fatto che successivamente al saggio del 1963 cui si è fatto finora riferimento, Davidson metta in discussione l'idea di spiegare la relazione fra ragione e azione attraverso un sillogismo pratico, riflettendo su quello che in filosofia è un tipico caso di comportamento "irrazionale", ossia l'omissione di un'azione che sarebbe pienamente giustificata dai desideri e dalle credenze che abbiamo, imputata alla cosiddetta *weakness of will*¹⁷⁴.

In doing *x* an agent acts incontinently if and only if: (a) the agent does *x* intentionally; (b) the agent believes there is an alternative action *y* open to him; and (c) the agent judges that, all things considered, it would be better to do *y* than to do *x*¹⁷⁵.

I casi di debolezza della volontà sembrano contraddire la tesi secondo cui le azioni sono le conclusioni inevitabili che scaturiscono dalle premesse rappresentate da desideri e credenze e perciò richiedono una revisione del ragionamento pratico. Una nuova analisi delle ragioni getta luce sul fatto che i desideri sono sempre condizionati, vuoi dalla valutazione delle circostanze concomitanti, vuoi dalla presenza di una pluralità di ragioni – "*the relative force of various desires and beliefs*"- spesso contrastanti e pertanto la conclusione non potrà non essere a sua volta condizionata. Per risolvere il problema Davidson si serve della nozione di intenzione. Dato che l'azione richiede di necessità una presa di decisione incondizionata l'intenzione farà da

¹⁷⁴ D. Davidson, "How is Weakness of Will Possible", 1969, in *Essays on actions and events*

¹⁷⁵ *Ivip*.22.

intermediario fra ragione primaria e azione¹⁷⁶. Solo nel momento in cui “intendiamo compiere un’azione” siamo incondizionatamente pronti a compierla, “*all things considered*”. Riepilogando, credenze e desideri consentono una prima valutazione condizionata che l’intenzione poi trasforma in una decisione incondizionata a compiere l’azione che a quel punto è desiderabile senza condizioni¹⁷⁷. Quello che accade nel caso di una “*incontinent action*”, come la chiama Davidson, è che l’agente non ha l’intenzione di fare ciò che pure giudica la cosa migliore da fare “*all things considered*”; il fallimento quindi riguarda l’intenzione ad agire in accordo con un giudizio condizionato. Questi comportamenti comunque non falsificano la tesi che le ragioni sono cause di azioni. Il problema di Davidson è però che anche aggiungendo l’intenzione al novero delle cause non si ottiene un’analisi causale completa dell’azione intenzionale, che permetta di considerarle condizioni non solo necessarie ma anche sufficienti. Si tratterebbe cioè di capire quali connessioni fra l’intenzione di compiere una certa azione e l’azione stessa sono necessarie e sufficienti per poterla descrivere come intenzionale¹⁷⁸.

In *Paradoxes of Irrationality*, Davidson identifica due posizioni estreme riguardo a un’azione irrazionale come l’akrasia. A un estremo c’è il *Plato Principle* che considera l’akrasia un vero e proprio paradosso, impossibile da

¹⁷⁶ D. Davidson, *Intending*, in op. cit.

¹⁷⁷ M. De Caro, op. cit. p. 86 sg.

¹⁷⁸ E’ il problema delle catene causali devianti. D. Davidson, *Freedom to act*, in op. cit., p. 78-9.

spiegare, dal momento che “no one willingly acts counter to what he knows to be best”; all’estremo opposto c’è il *Medea Principle* secondo cui “a person can act against his better judgement, but only when an alien force overwhelms his or her will”. Tuttavia, in questo secondo caso l’agente aveva una ragione – “a passion or impulse that overcame his better judgement” - ma la ragione non era la *sua* ragione. La presenza di una causa mentale che non è anche una ragione è una caratteristica comune anche ad altri tipi di irrazionalità.

In standard reason explanations [...] not only do the propositional contents of various beliefs and desires bear appropriate logical relations to one another and to the contents of beliefs, attitude, or intention they help to explain; the actual states of belief and desire cause the explained state or event. In the case of irrationality, the causal relation remains, while the logical is missing or distorted¹⁷⁹.

A questo punto si può dire che se per il *Plato Principle* l’akrasia semplicemente non può esistere, per il *Medea Principle* si può parlare solo di un’azione a-razionale, cioè che cade al di fuori della razionalità dal momento che – essendo originata da una causa che non è anche una ragione – non può essere descritta come un’azione intenzionale.

In un saggio successivo, *Incoherence and Irrationality*, la sua interpretazione dell’irrazionalità è ancora più chiara. Il saggio inizia con un’affermazione secca: “Irrationality, as rationality, is a normative

¹⁷⁹ D. Davidson, “Paradoxes of Irrationality”, 1982, in *Problems of irrationality*, Oxford, Clarendon Press, 2004, p. 179.

concept”¹⁸⁰. Se è così, qual è lo standard di razionalità e, soprattutto, chi lo decide? Esclusa la possibilità di identificare la razionalità con il *mio* insieme di valori e di utilizzarla per stigmatizzare come irrazionali tutti quei comportamenti con cui non concordo, si parlerà allora di “irrazionalità oggettiva” per i casi in cui mi trovo ad agire contro la mia stessa concezione di ragionevolezza e in cui c’è di conseguenza “some sort of inner inconsistency or inchoerence”. Data la natura olistica dei nostri atteggiamenti proposizionali – credenze, desideri, intenzioni – per cui il loro significato dipende dalle relazioni logiche che hanno fra di loro, non potrà però essere giudicata irrazionale una singola credenza. “Irrationality consists ... in inconsistency within a set of beliefs (or beliefs combined with principles)”. Saranno irrazionali quelle credenze che sono incoerenti fra loro, sulla base di principi – come i principi logici fondamentali – ritenuti validi dall’agente. Non c’è bisogno però di accertarsi che l’agente davvero li condivida perché “these are principles shared by all creatures that have propositional attitudes or act intentionally”. Il rispetto degli standard di razionalità è la *conditio sine qua non* per poter avere pensieri, giudizi, intenzioni la cui identità si basa sulle relazioni causali che detengono con il mondo esterno e sulle relazioni che hanno fra di loro. Tali relazioni sono costitutive delle atteggiamenti proposizionali. La credenza che sta per piovere indebolirebbe molto la sua pretesa di essere proprio quella credenza se non avesse la capacità di spingere colui che ha questa credenza e che non vuole bagnarsi a compiere un’azione

¹⁸⁰ D. Davidson, “Incoherence and Irrationality”, in *op. cit.*, 1985, p.189

appropriata come quella di procurarsi un ombrello¹⁸¹. Ma allora non ha senso interrogarsi in generale sulla razionalità di un agente: “rationality is, in its primitive, sense is a condition of having thoughts at all”. Solo a partire dalla condivisione di quei principi basilari di razionalità è possibile sollevare la questione dell’irrazionalità di certe credenze o azioni.

Le tesi di Davidson sull’irrazionalità hanno al centro, in linea con la sua teoria causalista dell’azione, i cosiddetti atteggiamenti proposizionali. Un evento è mentale, e quindi è un’azione, se possiamo darne una descrizione, in certo senso tradurlo, utilizzando proposizioni intenzionali riferite a un desiderio, credenza o intenzione. Questi ultimi sono quindi atteggiamenti intenzionali proprio in quanto hanno contenuto proposizionale. La conseguenza è che non vengono tenuti in considerazione gli eventi mentali che non sono passibili di una descrizione intenzionale di questo tipo¹⁸². Ma se il mentale viene identificato con il solo pensiero proposizionale, il rischio è che si perda qualcosa riguardo alla possibilità di dare una spiegazione dei comportamenti considerati irrazionali.

Prendiamo, ad esempio, un caso di akrasia proposto da Davidson. Vado avanti a discutere di politica con un amico, anche se so che non riuscirò a fargli cambiare idea e non amo il conflitto di opinioni. Le ragioni che mi portano a continuare una discussione che so benissimo non porterà da nessuna parte sono l’irritazione che provo per le sue idee sbagliate e l’impossibilità di

¹⁸¹ *Ivi*, p. 196

¹⁸² M. De Caro, *op. cit.*, pp. 102-3.

resistere alla voglia di farlo ragionare, anche se so benissimo che non ci riuscirò. Ma, quando ci penso a mente fredda, queste ragioni contano sicuramente meno delle ragioni contrarie alla continuazione della discussione. L'irrazionalità di questo comportamento dipende dalla discrepanza fra l'azione o l'intenzione di andare avanti a discutere e le ragioni che io stesso riconosco come rilevanti in quella circostanza (al successo dell'azione)¹⁸³. Confrontando le ragioni per continuare una discussione che io stesso giudico senza speranza e le ragioni - sicuramente più fondate - per indurmi a mettermi fine al più presto evitandomi una sicura frustrazione, si nota che la differenza sta nel fatto che le prime hanno una forte connotazione emotiva, che è strettamente legata alla situazione che sto vivendo in quel momento, mentre le seconde, al contrario, appaiono come più oggettive e soprattutto più facilmente traducibili in pensieri proposizionali. E' proprio la centralità di quest'ultimo fattore nella teoria di Davidson a renderle più "forti" e quindi costitutive del "my own best judgement", che decide della irrazionalità della scelta contraria.

E' evidente che i comportamenti irrazionali di questo tipo sono molto diversi da quelli che violano gli assiomi della EUT. Tuttavia, quando si tratta di stabilire i criteri di razionalità per le scelte, le considerazioni di Davidson sulla teoria della decisione riconfermano l'importanza del concetto di azione intenzionale.

Decision theory deals with intentional choices, and supplies a rationale for such choices. But the general idea that intentional actions can be rationalized is

¹⁸³ D. Davidson, *op. cit.*, p. 193.

no invention of decision theorists; it is part of the everyday concept of intentional action. For an intentional action is an action done with some intention, and this requires that there be a value the agent hopes to realize, and a belief that by acting as he does he has some chance of realizing the value.¹⁸⁴

Il solo fatto di poter descrivere un'azione come intenzionale implica che l'azione soddisfi i criteri minimi di razionalità. Quindi, l'assunzione da parte delle teorie della decisione di un modello coerente di comportamento, riscontrabile nell'insieme complessivo delle scelte di un agente, è una condizione necessaria per poter descrivere – e quindi spiegare – quel comportamento come intenzionale. La ricerca ostinata di una spiegazione è la reazione che abbiamo di fronte a comportamenti che ci appaiono incoerenti.

Thus we hesitate to attribute outright violations of transitivity of preference to an agent, and search instead for an explanation that rationalizes the apparent violation, perhaps by redescribing the choices or the alternatives from the agent's point of view¹⁸⁵.

E' quasi un paradosso, come ribadisce Davidson a proposito delle spiegazioni psicoanalitiche di Freud, che per spiegare l'irrazionalità si debba utilizzare una forma di spiegazione che razionalizza ciò che spiega¹⁸⁶; ma di fatto tendiamo a interpretare i comportamenti di scelta in modo da ricondurre

¹⁸⁴ D. Davidson, "A new basis for decision theory", in *Theory and Decision*, 18, 1985, p. 89.

¹⁸⁵ *Ivi*, p. 90

¹⁸⁶ D. Davidson, "Incoherence and Irrationality", in *op. cit.*, p.190.

anche quelli apparentemente irrazionali all'interno di un modello coerente e razionale.

My point is that if we are intelligibly to attribute attitudes and beliefs, or usefully to describe motion as behavior, then we are committed to finding, in the pattern of behavior, belief, and desire, a large degree of rationality and consistency¹⁸⁷.

Questo vale anche per la comprensione di un linguaggio: l'interprete può comprendere le differenze fra il suo linguaggio e quello del suo interlocutore solo se li accomuna un background molto ampio di credenze e valori e se esistono standard di razionalità e norme di consistenza condivisi¹⁸⁸. Al centro è sempre l'intenzionalità. E' l'intenzionalità, infatti, la proprietà comune alle entità linguistiche e agli stati mentali - come le credenze e i desideri - che ne condividono quelle che sono considerate le tre caratteristiche fondamentali¹⁸⁹. Anzitutto la direzionalità, rappresentativa della sua natura intrinseca, cioè quella di riferirsi a un determinato contenuto. Poi la normatività, per Davidson "the ineluctable normative element in interpretation"¹⁹⁰ costituita dalle norme di coerenza - che impongono il rispetto delle leggi della logica - e di corrispondenza - che riguardano la verità o falsità di un pensiero o di una proposizione. Infine l'olismo proprio degli stati mentali - ma valido anche in

¹⁸⁷ D. Davidson, "Psychology as Philosophy", in *Essays on actions and events*, cit., p. 237

¹⁸⁸ D. Davidson, "A new basis for decision theory", in *Theory and Decision*, 18, 1985, p. 91

¹⁸⁹ M. De Caro, *Azione*, Bologna, Il Mulino, 2008, p. 117-18

¹⁹⁰ D. Davidson, *op. cit.*, p. 92

campo linguistico – che “dipende invece dal fatto che i nessi normativi che li legano gli uni agli altri formano una rete i cui nodi non possono essere isolati: nessuna credenza, per esempio, potrebbe esistere separatamente¹⁹¹”. Queste considerazioni coinvolgono naturalmente anche i comportamenti di scelta. Del resto, Davidson ha più volte sottolineato esplicitamente la stretta affinità fra la teoria dell’interpretazione e la teoria della decisione.

The interpretation of another’s language is a problem very similar to the problem of discovering his subjective probabilities and utilities. Beliefs, desire and meaning are all attitudes towards proposition¹⁹².

In conclusione, si può dire che la teoria di Davidson privilegia gli atteggiamenti intenzionali che hanno contenuto proposizionale perché sono quelli che - costitutivamente - non possono prescindere dal rispetto di standard di razionalità, come coerenza logica e correttezza semantica. Tuttavia, questa impostazione lascia fuori quei comportamenti che, non essendo sostenuti da ragioni che possano essere espresse in forma proposizionale, non possono essere competitivi rispetto agli altri quando si tratta di fare una scelta sulla base del “my own best judgement”.

¹⁹¹ M. De Caro, *op. cit.*, p. 117-18. De Caro sottolinea che anche l’intenzione condivide con gli altri stati mentali la proprietà dell’intenzionalità.

¹⁹² D. Davidson, *op. cit.*, p. 87.

3.3 Derek Parfit e Bernard Williams: ragioni esterne e interne

Il carattere normativo della razionalità pratica è passibile di diverse interpretazioni che coinvolgono sia la relazione fra razionalità teorica e razionalità pratica sia quella fra credenza e intenzione, nel loro ruolo di stati mentali che definiscono le condizioni per l'azione. Partendo da quest'ultimo punto, abbiamo visto che Davidson stabilisce una precisa distanza fra la credenza come valutazione condizionata e l'intenzione che invece segna il passaggio alla decisione incondizionata: la differenza è dovuta al compito che ciascuna svolge. Mentre la credenza ha il compito apparentemente teoretico di confrontare le alternative disponibili e individuare la scelta in grado di appagare il desiderio, l'intenzione certifica la propensione a tradurre in azione il risultato finale della valutazione veicolato dalla credenza, cioè a metterlo in pratica. In definitiva, rappresentano le due fasi del ragionamento o sillogismo pratico, al punto che anche il compito della credenza sembra assumere un chiaro risvolto pratico, funzionale al raggiungimento dell'obiettivo.

E' possibile però accentuare la distanza fra credenza e intenzione. La credenza è portata "to fit the way the world is"¹⁹³, ed è di conseguenza spinta a modificarsi se si scopre che le cose non stanno così. L'intenzione, invece, segue un percorso completamente diverso, e cercherà piuttosto di far sì che sia il mondo ad adeguarsi. Da questa prospettiva, l'intenzione diventa uno stato mentale non cognitivo, dato che non deve preoccuparsi che il suo contenuto proposizionale corrisponda alla realtà. Sottolineare questa caratteristica

¹⁹³ M. Smith, "The Humean Theory of Motivation", in *Mind*, 96, 1987, p. 36-61.

dell'intenzione può anche condurre, come vedremo più avanti, a mettere in discussione la natura normativa della razionalità pratica, ma per il momento può spiegare due diverse accezioni di normatività.

La questione della natura cognitiva o meno dell'intenzione coinvolge quella più generale della relazione fra razionalità teoretica e razionalità pratica. Se si ritiene che la prima sia impegnata a consegnarci una rappresentazione distaccata, impersonale, della realtà, la differenza con la razionalità pratica riguarderà sia l'oggetto di riflessione – rispettivamente *matters of fact* e *matters of value* – sia la prevalenza nel secondo caso del *first-personal point of view*. Se invece si sostiene che anche la razionalità teoretica abbia a che fare con questioni normative – in cui si tratta di valutare la fondatezza delle ragioni per credere – allora si potrà stabilire una sorta di parallelismo o continuità fra le due, al di là delle differenze in termini di oggetto, e cioè, rispettivamente, la verità di una proposizione e la desiderabilità di un'azione. Riguardo poi agli esiti, dato che il ragionamento pratico dà luogo a un'azione intenzionale, intellegibile solo in quanto espressione di stati mentali, sia la ragione teoretica sia quella pratica trattano di atteggiamenti, anche se nel primo caso cambiano le nostre convinzioni e nel secondo caso invece cambiano le nostre intenzioni¹⁹⁴. Questa affinità spiegherebbe anche la possibilità di parlare di irrazionalità sia in ambito teoretico sia in ambito pratico, ogniqualvolta vengono a mancare gli

¹⁹⁴ M. E. Bratman, *Intentions, Plans, and Practical Reason*, Cambridge MA, Harvard University Press, 1987, R. J. Wallace, *Practical reason*, Stanford University of Philosophy, <http://plato.stanford.edu/entries/practical-reason>.

atteggiamenti appropriati, che sarebbero dovuti scaturire dalle credenze o dalle intenzioni corrispondenti.

La difesa della normatività, con le implicazioni di cui si è appena detto, è al centro delle riflessioni di Parfit volte a contestare le tesi esposte da Williams nel suo saggio sulle ragioni esterne e interne¹⁹⁵. In *Reasons and motivation*¹⁹⁶, Parfit si interroga sul significato da attribuire a due proposizioni molto comuni che utilizziamo quando stiamo decidendo di compiere un'azione: "What do we have most reason to want, and do?" e "What is it most rational for us to want, and do?"¹⁹⁷. Per Parfit si tratta di due proposizioni diverse solo nel senso che le ragioni della prima sono basate su fatti – riprendendo il suo esempio, se il nostro albergo va a fuoco avremo una ragione per saltare giù dalla finestra – nella seconda invece la razionalità cui si fa riferimento dipende dalle credenze che abbiamo – se il nostro albergo non è in fiamme, ma noi crediamo che lo sia, la nostra scelta di saltare dalla finestra sarà razionale, ma in realtà non avremmo avuto ragione di farlo. Questo significa che la risposta ai due interrogativi sarà la stessa solo nel caso in cui le nostre credenze siano vere, e che è quindi la credenza a decidere della razionalità di una scelta.

¹⁹⁵ B. Williams, "Internal and external reasons", in *Moral Luck*, Cambridge, Cambridge University Press, 1981

¹⁹⁶ D. Parfit, "Reasons and motivation", in *The Aristotelian Society*, 77 (Supplementary Volume), p. 99-130

¹⁹⁷ D. Parfit, *Reasons and motivation*, p. 99.

Per Parfit il senso ultimo dei termini “ragione” e “razionalità” è il fatto di avere un carattere vincolante – e infatti si tratta di una ragione *normativa* - da cui l’azione scaturisce necessariamente, fatta salva la fondatezza delle nostre credenze. La prima proposizione è quindi quella fondamentale, dato che le nostre credenze mirano ovviamente alla verità, cioè alla corrispondenza con quei fatti su cui ci basiamo – e dobbiamo basarci - se vogliamo avere ragioni per agire. Sulla base di questa analisi delle due proposizioni Parfit distingue quelle che chiama “ragioni normative” dalle “ragioni motivanti”. Se abbiamo le prime e agiamo in base ad esse, a quel punto diventano ragioni motivanti. Vanno comunque considerate indipendenti perché possiamo anche avere solo una delle due, come dimostra l’esempio precedente.

Un secondo esempio mette ben in evidenza la differenza fra la posizione difesa da Parfit - l’esternalismo - e quella con cui Parfit polemizza – l’internalismo – sostenuta da Williams. Si tratta della decisione di prendere una medicina che potrebbe proteggere la nostra salute, prevenendo una qualche malattia futura. Mentre dal punto di vista internalista la persona in questione prenderà la medicina solo se ha a cuore la sua salute futura, dalla prospettiva esternalista la ragione di proteggere la nostra salute non dipende dall’aver a cuore o meno la nostra salute futura. Tornando alle definizioni precedenti, l’esternalista sostiene che la ragione normativa è indipendente dallo stato motivazionale dell’agente, quindi la proposizione “we have a reason to do something” implica

if we knew the relevant facts, and were fully *substantively* rational, we would be motivated to do this thing¹⁹⁸.

A contare davvero è sempre il fatto di “have a reason to do something”, che in questo caso corrisponde a “we all have reasons to protect our health and prevent our own future suffering”. La motivazione è solo la conseguenza di una informazione corretta dei fatti e del nostro essere “fully substantively rational”.

La prospettiva internalista procede invece in direzione opposta. Abbiamo una ragione per agire in un certo modo

if we knew the relevant facts, and deliberated in a way that was *procedurally* rational, we would be motivated to do this thing¹⁹⁹.

In questo caso, quindi, c'è anzitutto una motivazione, a partire dalla quale, seguendo “a sound deliberative route”, veniamo ad avere una ragione per compiere una data azione. Per Williams la motivazione scaturisce dal “the agent’s subjective motivational set”, che comprende non solo i nostri desideri, ma “dispositions of evaluation, patterns of emotional reaction, personal loyalties and various projects, as they may be abstractly be called, embodying commitments of the agent²⁰⁰”. Riprendendo l’esempio precedente, se la cura

¹⁹⁸ *Ivi*, p. 101.

¹⁹⁹ *Ibid*

²⁰⁰ B. Williams, *Internal and external reasons*, *cit*, p

della nostra salute non fa parte del nostro “motivational set” non avremo nessuna ragione per prendere la medicina²⁰¹.

La prima differenza fra le due posizioni sottolineata da Parfit è che mentre il fatto di essere “substantively rational” implica che dobbiamo curarci di certe cose, come la nostra salute, in quanto hanno in valore oggettivo; essere “procedurally rational” significa che dobbiamo deliberare rispettando le regole della logica, ma non richiede di avere un particolare desiderio o aspirazione, come il nostro benessere. In questo secondo caso, la razionalità consiste nell’adozione di un procedimento argomentativo corretto, ma non comporta la conoscenza di “fatti rilevanti”, come il valore intrinseco rappresentato dalla cura della salute.

La distanza fra le due posizioni è ancora più evidente se si guarda alla natura dei due opposti argomenti.

According to Externalists ... what gives us reasons for acting are not facts about our own motivation, but facts about our own or other’s people well being, or about other things that are worth achieving.... Internalists derive conclusion about reasons from psychological claims about the motivation that, under certain conditions, we would in fact have²⁰².

²⁰¹ Non ha caso gli esternalisti preferiscono l’espressione “c’è una ragione per fare x” , mentre gli internalisti utilizzano l’espressione “A ha una ragione per fare x”. La prima espressione sottolinea come la “ragione per” sia – o almeno possa essere – indipendente dal soggetto.

²⁰² D. Parfit, *op. cit.*, p. 102.

Per l'esternalista Parfit è inaccettabile l'idea che le ragioni per agire siano giustificate solo in base alla presenza di fatti psicologici estemporanei come le motivazioni, invece che da "normative claims about what is worth achieving". Nel primo caso "avere una ragione per fare x" equivale ad asserire che "dopo una deliberazione informata e proceduralmente razionale saremo motivati a fare x". Questa proposizione si limita a descrivere uno stato psicologico, perciò può al massimo avere "normative significance" o "normative force", ma non è "relevantly normative" per Parfit che la considera frutto di una sorta di naturalismo normativo. Le ragioni normative devono essere ragioni esterne proprio perché non possono dipendere da motivazioni, credenze o desideri dell'agente, ma devono anzi prescindere e basarsi su verità di fatto o giudizi di valore.

These are reasons to want something, for its own sake, which are provided by facts about this thing. Such reasons we call value-based²⁰³.

Parfit ammette che le verità normative - derivanti da "facts about what is relevantly worth achieving" - a cui fanno appello le teorie esternaliste possono sembrare "metaphysically mysterious, or inconsistent with a scientific world view", ma non ritiene possibile contestare l'irriducibilità dei concetti normativi che non ammettono una spiegazione naturalistica. Dal momento che le ragioni normative hanno lo stesso statuto delle verità normative, è chiaro che non si può pensare di ricondurle alle motivazioni; tuttavia, se siamo razionali, quelle ragioni sicuramente ci motiveranno.

²⁰³ *Ivi*, p. 128.

Nello stesso saggio Parfit descrive la concezione internalista di Williams applicandola non solo alla situazione in cui diventiamo motivati a compiere una data azione ma alla decisione di agire²⁰⁴. Secondo Williams, perché la decisione sia razionale dobbiamo credere di avere “most reason to act in some way” il che vuol dire che “there is a ‘sound deliberative route’ starting from our ‘existing motivations’, by which we could rationally decide to do X”. Per l’alternativa externalista invece, se “X is what we have most reason to do ... we could rationally decide to do X, for this reason, *whatever* our existing motivations”. Anche in questa contrapposizione sembra esserci un conflitto fra la spiegazione di una scelta che verte sulla possibilità di individuare ciò che effettivamente ci ha spinto a compierla e l’imposizione di un carattere di obbligatorietà alla scelta che deriva dall’attribuzione di un giudizio di valore alla ragione per cui l’abbiamo fatta. A differenza di quanto sostenuto dalla teoria causalista di Davidson, nella prospettiva externalista il contesto della spiegazione non ha alcuna relazione con quello della giustificazione.

Nel caso di Parfit, la normatività è una conseguenza della linea di continuità che lega le ragioni per credere alle ragioni per agire. E’ dalla verità della credenza assieme alla capacità di comprendere il valore dell’oggetto – nelle parole di Parfit è la conoscenza dei “relevant facts” - che scaturiscono le ragioni che si possono definire “value-based”. Il ragionamento pratico si configura come capacità di ragionare riguardo a ciò che sarebbe oggettivamente bene fare, e la decisione sarà presa quindi sulla base di questo tipo di riflessione valutativa. Rispetto al binomio formato da desiderio e

²⁰⁴ D. Parfit, *id.*, p. 117

credenza, è quest'ultima ad avere un ruolo decisivo per la razionalità della decisione.

Nella prospettiva internalista di Williams, invece, le ragioni normative coincidono con le ragioni motivanti, e quindi interne, con l'indubbio vantaggio di evitare le forti implicazioni metafisiche sulla natura di ragioni irriducibilmente normative. Facendo leva sulla sua dimensione pratica, la ragione si identifica con la capacità di portare a conclusione le implicazioni di convincimenti di vario tipo – desideri, progetti, ideali, scopi, ma anche preferenze, valutazioni²⁰⁵ - contenuti nell'insieme motivazionale soggettivo dell'agente. La naturale conseguenza di questa impostazione è che - dato che la motivazione è condizione necessaria per avere una ragione - la forza normativa della ragione coinciderà con la sua forza motivante.

Da questo punto di vista l'impostazione di Williams appare, per certi aspetti, vicina a quella di Davidson riguardo alla possibilità di far convergere cause e ragioni dell'agire, con la differenza che mentre per Davidson quello che conta è il punto di vista dell'interprete, per Williams al centro è il punto di vista del soggetto. Data la relazione imprescindibile che hanno con l'insieme motivazionale soggettivo, le ragioni sono normative solo in quanto riconosciute come tali dall'agente stesso.

²⁰⁵ B. Williams, "Internal reasons and the obscurity of blame", in *Making sense of the humanity*, p. 35; "Values, reasons, theory of Persuasion", in *Ethics, rationality and economic behavior*, Oxford, Clarendon Press, 1966, p. 66.

The normative statement ‘He has a reason to do ϕ ’ – introduce the possibility of that reason be an explanation; namely if the agent accepts that claim²⁰⁶.

Riassumendo, se per l’esternalista Parfit è la continuità fra ragione teoretica e ragione pratica e il conseguente ruolo determinante della credenza nella verità e oggettività delle ragioni a certificare la loro normatività, nel caso di Williams è la dimensione pratica dell’insieme motivazionale assieme all’insistenza sulla prospettiva specifica di quel determinato agente a legittimare una diversa versione della normatività delle ragioni per agire.

3.4 Bernard Williams: quale razionalità per le ragioni interne?

Per capire meglio il senso di questa forma di normatività delle ragioni si può partire dalla formulazione ritenuta da Williams stesso più soddisfacente delle sue tesi: “A has a reason to ϕ only if there is a sound deliberative route from A’s subjective motivational set ...to A’s ϕ -ing”²⁰⁷. Al di là della stretta relazione fra ragione e motivazione di cui si è già detto, è importante analizzare natura e caratteristiche sia dell’”S” – abbreviazione usata da Williams per indicare l’insieme motivazionale soggettivo – sia del “percorso deliberativo valido”.

²⁰⁶ B. Williams, “Internal reasons and the obscurity of blame”, in *Making sense of the humanity*, p. 39.

²⁰⁷ B. Williams, “Poscript: Some further notes on internal and external reasons”, in Millgram ed., *Varieties of practical reasoning*, Cambridge MA, MIT Press, 2001, p. 91.

Dalle diverse formulazioni dell'S già riportate è evidente che il desiderio è solo uno degli elementi coinvolti, anche se la sua presenza esplicita la vicinanza al modello humaneo che attribuisce solo al desiderio la capacità di motivare. Ciononostante, la sua caratteristica peculiare rappresentata da una volontà o tensione soggettiva intensa verso qualcosa, investe un po' tutti gli altri elementi dell'S. Nelle parole di Williams: "I have discuss S primarily in terms of desires, and this term can be used, formally, for all elements in S"²⁰⁸. Ed è sempre Williams a sottolineare che "it is important to emphasize the variety of elements that are, on this view, to be included in the agent's S"²⁰⁹. Se è vero che le ragioni per agire sono solo interne – al punto che spesso quelle che sono considerate ragioni esterne non sono in realtà altro che ottimistiche ragioni interne²¹⁰ - è comprensibile l' esigenza di ampliare e diversificare il contenuto delle motivazioni. Perciò, all'interno dell'S, a fianco di elementi volutamente generici come "disposizioni"o "attitudini", è significativa la presenza di "progetti" di vario genere, che richiedono un particolare impegno dell'agente per essere portati a termine. E' evidente che in questo caso si va molto al di là del desiderio del momento che impone un immediato soddisfacimento e si riconosce all'agente la capacità di discriminare fra le azioni o scelte possibili sulla base di un piano prestabilito che impone necessariamente delle priorità. Un altro elemento importante,

²⁰⁸ B. Williams, *Internal and External Reasons*, cit.,

²⁰⁹ B. Williams, "Values, reasons, theory of Persuasion", in *Ethics, rationality and economic behavior*, cit., p. 67.

²¹⁰ B. Williams, *Internal reasons and the obscurity of blame*, cit, p. 40.

comune ai diversi elenchi dei contenuti dell'S proposti da Williams, sono le valutazioni o, meglio, le disposizioni associate al riconoscimento da parte dell'agente di vari tipi di valore²¹¹. In questo modo si sottrae un argomento importante ai sostenitori delle ragioni esterne che rivendicava la necessità del riferimento a valori - *matter of value* - universali e quindi esterni. E' anche vero che, facendo parte dell'insieme motivazionale di un determinato agente, le valutazioni non potranno avere carattere universale. Ma questo non significa però che non possano essere sottoposte a critica, ma soltanto che non si potrà farlo prescindendo del tutto dal peculiare modo di pensare e di vedere le cose di quella determinata persona, cioè dal "mettersi nei suoi panni". Pena il fallimento, ad esempio, di ogni serio tentativo di persuasione. La perdita dell'universalità viene quindi compensata da un incremento della capacità di comprensione di comportamenti che - in nome del valore universale della razionalità - potevano essere semplicemente bollati come irrazionali. In ogni caso, ancora una volta, il punto di vista soggettivo si conferma come un fattore irriducibile.

A fianco dell'ampiezza e della varietà dell'S va sottolineata anche la voluta indeterminatezza associata da Williams ad ogni elemento presente al suo interno. Si è già detto della genericità associate ai progetti, ma le implicazioni della vaghezza e persino dell'ambiguità sono evidenti soprattutto rispetto alle "preferenze", di cui si parla subito dopo aver citato le predisposizioni valutative.

²¹¹ B. Williams, "Values, reasons, theory of Persuasion", *cit.*, p. 71.

It is not the case that everything in the agent's S has already to be formed into preferences; and in so far as it is formed into preferences, those preferences do not necessarily have to satisfy formal condition of completeness. There is no naturalistic reason, based on considerations of psychology or philosophy of mind, to suppose that these indeterminacies are radically reducible, in particular to preference orderings that can be handled by Bayesian techniques²¹².

E' l'esempio più chiaro del carattere costitutivamente non strutturato dell'insieme motivazionale. In base a un'analisi naturalistica o descrittiva non si può sostenere che tutte le nostre propensioni assumono la forma definita che è propria delle preferenze, e le stesse preferenze del resto non raggiungono quasi mai quel livello di stabilità richiesta dall'assioma di completezza. Tanto meno è possibile eliminare del tutto l'indeterminatezza che le caratterizza in vista di un ordinamento univoco, funzionale alla possibilità di darne una trattazione matematica.

Il risultato è non solo quello di offrire una descrizione sicuramente più "realistica" della natura degli elementi che fanno parte dell'S, ma anche di inserirli in una prospettiva dinamica e quindi potenzialmente modificabile: ad esempio, ogni volta che l'agente si rende conto di essersi sbagliato nel pensare di aver una ragione per perseguire una certa linea d'azione.

²¹² *Ivi*, p.67.

The internalist account does leave room for normative correction; it allows, for instance, for the agent to be shown that his deliberation is unsound because it is based upon false information²¹³.

Per capire cosa debba intendersi con “percorso deliberativo valido” occorre anzitutto ammettere l’esistenza nell’S di ogni agente razionale del desiderio di non cadere in errore, e della conseguente generale tendenza ad agire sulla base di informazioni valide. L’ammissione di “a modest amount of prudence” può spiegare la forza vincolante del sentirsi impegnati a procedere in questa direzione, restano tuttavia aperti gli interrogativi sui criteri e sui confini di questa razionalità pratica.

What then are the criteria of deliberative rationality, in which a rational agent presumably does necessarily have an interest? How far do they go beyond the general desire for true information, and minimal constraints of consistency? In particular, how far do they include the kinds of demands that rational decision theory has tended to place on them?²¹⁴

Nel saggio che stiamo esaminando Williams si concentra in realtà su una questione specifica che riguarda il modo in cui la teoria della scelta razionale cerca di razionalizzare le scelte che sembrano violare l’assioma di transitività. La proposta di Broome - di cui si è già detto²¹⁵ - prevedeva la possibilità di

²¹³ *Ibid*, 68.

²¹⁴ *Ivi*, p. 68-9.

²¹⁵ Cfr. capitolo 1. L’esempio portato da Broome di violazione della transitività è il seguente. Di fronte alla scelta fra visitare Roma o andare in montagna Maurice preferisce Roma dato

togliere ogni ambiguità alle preferenze, sforzandosi di riscriverle specificandone i contenuti in modo da ottenere una più accurata rappresentazione delle alternative. Diventa così possibile dare una valutazione diversa di una stessa opzione, a seconda di quale sia l'alternativa con cui viene messa a confronto, dando luogo a una suddivisione in due diverse sotto-opzioni. Per evitare che si tratti di una soluzione del tutto arbitraria che eliminerebbe ogni criterio vincolante di coerenza, le differenziazioni dell'agente dovranno essere giustificate sulla base di considerazioni che chiamano in causa i valori dell'agente. Tuttavia, se è vero che per riuscire a rendere intellegibili quelle scelte incoerenti i valori in questione dovranno essere valori condivisi, è altrettanto vero che saranno pur sempre i valori dell'agente. L'appartenenza dei valori all'S dell'agente conferma ancora una volta la tesi internalista che le ragioni per agire sono un prodotto del suo S.

Il fatto che il percorso deliberativo si presenti in questo caso come un procedimento razionale che coinvolge questioni di valore può essere già un indizio sufficiente per ritenere che si tratti di una tipologia di ragionamento pratico che non è riducibile alla mera razionalità strumentale. La prima definizione che ne dà Williams è infatti in negativo “does not merely involve perceiving means to an end that has already been formulated”²¹⁶. Anche il caso

che ha paura di scalare. Tra restare a casa e andare a Roma preferisce andare a Roma, ma poi preferisce andare in montagna allo stare a casa.

²¹⁶ B. Williams, “Internal reasons and the obscurity of blame”, in *Making sense of the humanity*, p. 38. Esempi di ragionamento deliberativo si trovano anche in *Internal and External reasons*, cit., p.104.

più semplice, infatti, implica una qualche forma di ragionamento pratico, come è quello che ci porta a ritenere che si ha ragione di fare una determinate cosa perché è il modo più “conveniente, economico, piacevole” di soddisfare un qualche elemento presente nel nostro S. Ma non è affatto l’unica forma possibile di deliberazione. Stabilire una scansione temporale per la realizzazione dei diversi desideri oppure, in caso di desideri irreparabilmente conflittuali, considerare quali hanno maggior peso; decidere quale contenuto specifico dare a un progetto adottato in termini generali o con che tipo di contenuto soddisfare un’aspettativa originariamente più generica. Un esempio particolarmente interessante è l’invenzione di alternative.

One of the most important things deliberation does, rather than thinking of means to a fixed end, is to think of another line of conduct altogether, as when someone succeeds in braking out of a dilemma²¹⁷.

Di fronte a questi diversi tipi di ragionamento pratico o deliberativo acquista evidenza l’impossibilità di assegnare dei confini precisi che rendano “fully determinate in general what may count as a sound deliberative route”. Si rimane con un’idea molto vaga di cosa un agente ha ragione di fare. Ma per Williams una concezione così ampia non è affatto un svantaggio che si possa attribuire alla tesi internalista, ma piuttosto un dato di fatto che offre al contrario diversi vantaggi. “It is often vague what one has reason to do”²¹⁸. Lo è ancora di più ripensando all’ultimo esempio, che è sicuramente quanto di più

²¹⁷ *Ivi*, p. 38

²¹⁸ *Ivi*, p. 38.

lontano si possa immaginare da un rigido *means end reasoning*. Mettere fine ad una situazione di stallo mentale che sembra senza vie di uscita è spesso possibile solo inventando una possibilità assolutamente non prevista, e questo dimostra che anche i confini fra immaginazione e razionalità sono vaghi.

If someone is good at thinking about what to do, he or she needs not just knowledge and experience and intelligence, but imagination; and it is impossible that it should be fully determinate what imagination might contribute to a deliberation. This is one reason why it may be indeterminate what exactly an agent has reason to do²¹⁹.

L'altra ragione è quella già presa in esame e riguarda la complementare ampiezza e varietà dell'insieme motivazionale dell'agente di cui si è già detto. Ci sono anche altri vantaggi nell'indeterminatezza costitutiva del procedimento deliberativo. In molti casi si può arrivare a scoprire di avere ragioni che non si sapeva affatto di avere, e così aggiungere nuove ragioni interne al nostro S, ma un supplemento di riflessione può anche farci scoprire di non avere le ragioni che pure credevamo di avere, a causa della falsità di una credenza. L'immaginazione in particolare è di grande aiuto nel farci vedere delle conseguenze o delle implicazioni cui inizialmente non avevamo pensato e che troviamo inaccettabili, ma in cui saremmo incorsi se avessimo portato avanti un progetto che ritenevamo di aver ragione di promuovere.

²¹⁹ Ivi

Naturalmente l'immaginazione può sia far venir meno un desiderio, sia far nascere invece nuove possibilità e nuovi desideri²²⁰.

Varietà e indeterminatezza sono caratteristiche speculari sia dell'insieme degli stati motivazionali sia della tipologia dei procedimenti deliberativi e testimoniano la loro intrinseca soggettività. E' proprio il rifiuto da parte di Williams di attribuire universalità alle ragioni a segnare la distanza con altre forme di internalismo quali quelle rispettivamente della Korsgaard e di McDowell. Pur condividendone la concezione che la normatività delle ragioni coincide con la loro forza motivante, le critiche di questi ultimi alle tesi di Williams riguardano il modo di intendere le ragioni interne attraverso un confronto fra il suo modello di ispirazione humeana e, rispettivamente, quello kantiano della Korsgaard e quello aristotelico di McDowell²²¹.

Nel saggio *Skepticism about Practical Reason* il riferimento della Korsgaard è ai "dubts about the extent to which human action is or could possibly be directed by reason". E questo scetticismo è rivolto sia alla rilevanza delle considerazioni razionali per le decisioni e le scelte (content skepticism) sia alla possibilità per la ragione di motivare (motivational skepticism). Entrambe le forme di scetticismo sono presenti nella filosofia di Hume, nel ritenere che il ragionamento pratico abbia solo il compito di determinare i mezzi appropriati per un dato fine e nell'escludere la possibilità

²²⁰ B. Williams, *Internal and External Reason*, cit., p. 104-105

²²¹ L. Greco, "Bernard Williams e la natura delle ragioni in etica" in *Ethics and Politics*, 2005, 1-15, M. P. Jenkins, *Bernard Williams*, Montreal, McGill-Queen's, 2006, pp. 105-114

che la ragione sia in grado, da sola, di motivare. La tesi della Korsgaard è che quest'ultima forma di scetticismo dipende interamente dalla prima, dal momento che se la ragione ha qualcosa da dire su scelta e azione non può non essere anche motivante²²². Come per Kant, è la convinzione che i principi razionali possono fornire ragioni per agire a renderli fonti di motivazione, indipendentemente dalle passioni. Una tesi quest'ultima che viene preclusa dal confinare, come fa Hume, la ragione pratica al ruolo strumentale che significa non solo “prevent reason from determining ends” ma anche che “reason by itself neither selects nor ranks our ends”²²³. Ed è in questa funzione di guida nel determinare i fini da perseguire che consiste invece la forza normativa della ragione, da cui dipende, a sua volta, la normatività del principio strumentale²²⁴.

L'internalismo di Williams secondo Korsgaard mette sicuramente in discussione la prima forma di scetticismo, respingendo l'identificazione del percorso deliberativo con il solo *means ends reasoning*, ma non la seconda dal momento che le ragioni sono sempre condizionate dall'insieme motivazionale soggettivo. Infatti se, come si è detto, per Williams al di là del ragionamento strumentale i possibili procedimenti deliberativi spaziano dalla riflessione

²²² C.M. Korsgaard, “Skepticism about practical Reason”, in *Varieties of Practical Reasoning*, cit., p. 105-125.

²²³ *Ivi*, p. 104

²²⁴ C. M. Korsgaard, “The Normativity of Instrumental Reason”, in G. Cullity and B. Gaut (eds), *Ethics and Practical Reason*, oxford, Oxford University Press, 1997, p. 215-54. Cfr. la critica di Blackburn in *Ruling passions*, cit., p. 242 sg.

valutativa alla prefigurazione di scenari alternativi ad opera dell'immaginazione fino alla conversione, è anche vero – come sottolinea Korsgaard – che nessuno di essi può prescindere dalla relazione con almeno un elemento dell'S dell'agente.

“...internal reasons are by definition relative to something that he calls the agent's “subjective motivational set”: this follows from the fact that they can motivate. The content of this set is left open, but one kind of thing it will obviously contain is the agent's desires and passions. Internal reasons are reasons reached by deliberation from the subjective motivational set: they can motive us because of their connection to that set”²²⁵.

Del resto, i procedimenti di cui parla Williams riguardano l'ispirazione creativa, l'immaginazione e altre forme di procedimento euristico più che il pensiero razionale in senso proprio. Soprattutto, fra di essi non c'è posto per una razionalità di tipo kantiano identificata con principi formali incondizionati che pretendono di individuare ragioni pratiche universalmente valide. Impersonalità e universalità sono caratteristiche che, nell'ottica internalista di Williams, impedirebbero alle ragioni pratiche di motivare. A rendere incompatibile l'internalismo di Korsgaard con quello di Williams è il fatto che il ruolo rivestito dall'insieme motivazionale soggettivo è invece attribuito dalla Korsgaard a una razionalità pratica intesa come fonte autonoma di normatività e motivazione, costitutivamente indipendente da fattori psicologici e contingenti. Trattandosi poi di una facoltà comune a tutti gli agenti umani, sarebbe in grado di assicurare a tutti sia la prescrittività delle

²²⁵ C. M. Korsgaard, *Skepticism about practical Reason*, cit, p.116.

ragioni sia la loro forza motivante. Se una legge di ragione - le cui pretese di determinare la nostra condotta riconosciamo come valide - non ci motiva, non possiamo trarne un pretesto per esprimere scetticismo nei suoi confronti. E' solo nostro il fallimento nel conformarci alle sue prescrizioni e non fa che mostrare i limiti della nostra razionalità.

The fact that the law [of reason] might not govern conduct, even when someone understood it, is no reason for skepticism: the necessity is in the law, and not in us²²⁶.

Del resto, per Korsgaard, la filosofia può solo dirci come sarebbe essere razionali – sia in campo teoretico sia in quello pratico – ma la misura in cui le persone siano effettivamente spinte da considerazioni razionali non è di sua competenza. Probabilmente è la psicologia a doversene occupare, e le tesi di Williams, non coniugando necessità e universalità delle ragioni, finiscono per attribuire una connotazione solo psicologica alla loro forza motivante.

La critica di McDowell a Williams parte da una prospettiva diversa, ben esemplificata dal titolo del suo saggio: *Might there be external reasons*²²⁷? Da un ottica cognitivista, McDowell intende contestare la convinzione di Williams che possono esserci solo ragioni interne, sottolineandone le conseguenze per le ragioni etiche.

²²⁶ *Op. cit.*, p. 122

²²⁷ J. McDowell, "Might there be External Reasons?" In J. E. J. Altham & R. Harrison (eds) *World, Mind, and Ethics*, Cambridge, Cambridge University Press, 1995, p. 68-85.

The implication of William's skepticism is that ethical reasons are reasons only for those for whom they are internal reasons: only for those who have motivations to which ethical considerations speak, or can be made to speak²²⁸.

Dato che anche per lui la normatività delle ragioni non può non avere anche forza motivazionale, il problema da risolvere per riabilitare le ragioni esterne è ancora quello di capire come possano motivare, cioè come si possa passare dal non essere motivati all'esserlo, prescindendo dagli stati motivazionali già esistenti nel soggetto. Williams riteneva fosse difficile credere che questa transizione potesse essere opera di un ragionamento puro, cioè autosufficiente rispetto all'S dell'agente. McDowell respinge lo scetticismo di Williams, dal momento che non c'è una vera e propria transizione dalla credenza nella verità di una proposizione che riguarda una ragione esterna – A ha una ragione per ϕ – alla motivazione a ϕ . Più semplicemente, se l'agente viene ad avere quella credenza, delibererà correttamente, cioè vedrà le cose nella giusta luce e quindi sarà motivato.

There is no mystery about how the process can be the acquisition, simultaneously, of a way of seeing things and a collection of motivational directions or practical concerns, focused and activated in particular cases by exercises of the way of seeing things²²⁹.

Non c'è niente di misterioso perché la sola condizione è quella di aver ricevuto una educazione etica che, come insegna Aristotele, familiarizzandoci

²²⁸ *Ivi*, p. 68.

²²⁹ *Ivi*, p. 73.

con modi di comportamento adeguati ci consente di assimilarli assieme ai correlati modi corretti di pensiero. In questo modo acquisiamo un habitus che ci dà la capacità di cogliere le cose nella giusta luce, cioè di comprendere la verità delle ragioni esterne ed esserne motivati. Ma se non abbiamo ricevuto una educazione di questo tipo non potremo essere ricettivi nei confronti delle ragioni esterne. Per McDowell, quindi, la condizione per deliberare correttamente non sta, come voleva Williams, nel partire da un elemento dell'S, ma è solo grazie all'educazione che riconosciamo le ragioni, che di fatto sono esterne, e - facendole nostre - le rendiamo interne.

In questa argomentazione, l'agente che ha ricevuto questa adeguata educazione è una sorta di agente ideale, il *phronimos* aristotelico cui fa riferimento Williams nella sua risposta alle tesi di McDowell²³⁰, e il ricorso a un modello ideale è indispensabile per salvaguardare la dimensione critica o normativa del concetto di razionalità. In alternativa a una spiegazione meramente psicologica del processo deliberativo quale è quella di Williams, occorre concepire invece come le cose sarebbero idealmente, mantenendosi sufficientemente indipendenti dalle procedure psicologiche di un agente reale, in modo da avere un modello ideale di riferimento per dare una valutazione critica dei comportamenti effettivi.

Una buona ragione per dubitare della tesi di Williams che la sola razionalità pratica è la razionalità delle ragioni interne, è quindi l'idea che la dimensione critica di questo concetto di razionalità richieda di trascendere i

²³⁰ B. Williams, *Replies* cit., p.185-224.

meri fatti della psicologia individuale²³¹. Dal punto di vista di Williams l'asserzione "A ha a una ragione per ϕ " nella descrizione externalista non mette più in relazione azioni e persone ma tipologie di azione e tipologie di circostanze, come è rivelato dalla formula "in X circostanze c'è una ragione per ϕ ". Ed è proprio l'impersonalità della formula incentrata sull'ideale del *phronimos* a rivelarsi inaccettabile, dal momento che "did not adopt the relativity to that agent's S which is the mark of the internalist view"²³².

Un modello esplicativo che fa riferimento ad agenti ideali, come il *phronimos* di Aristotele, o a principi universali, come l'imperativo kantiano, è del tutto inadeguato se si vogliono spiegare le ragioni di agenti particolari costitutivamente imperfetti. Lo psicologismo non è per Williams un rischio da evitare. "I accept that my account is psychologistic, in the sense that on my view a statement about A's reasons is partly a statement about A's psychology"²³³.

3.5 John Broome: *instrumental reasoning*

La razionalità strumentale richiede agli agenti di scegliere quei mezzi che sono necessari in relazione ai fini che si sono proposti di raggiungere. Questa forma di razionalità è stata a lungo considerata come l'aspetto meno problematico della razionalità pratica, dal momento che non presupponeva

²³¹ J. McDowell, *op. cit.*, p. 77.

²³² B. Williams, *op. cit.*, p. 193.

²³³ *Ivi*, p. 191.

nessun esame razionale dei fini degli agenti. Molti di coloro che condividono lo scetticismo humeano sulla subalternità della ragione sono disposti ad ammettere questa forma di razionalità solo strumentale, resta comunque il problema di spiegare e giustificare il suo carattere normativo. Anche in questo caso, quindi, le differenti interpretazioni riguarderanno il tipo di normatività che compete all'*instrumental reasoning*.

Se il principio strumentale rappresenta una norma vincolante della razionalità pratica, ne consegue che si può esercitare una critica razionale nei confronti delle situazioni in cui è evidente il fallimento di questo tipo di coerenza strumentale richiesto dalla relazione fra mezzi e fini. Una concezione naturalistica di derivazione humeana²³⁴ può sembrare poco compatibile con questo vincolo di oggettiva normatività, mentre una opposta concezione costruttivista di tipo kantiano può avere difficoltà a parlare di necessità nella scelta dei mezzi rispetto al perseguimento di un fine che appare invece svincolato da qualsiasi forma di valutazione. Tuttavia, nel primo caso si può interpretare il principio strumentale come una regola di coerenza interna, subalterna alla scelta del fine da parte dell'agente che, al contrario, è assolutamente insindacabile - e perciò in questo senso soggettiva. Nel secondo caso, la razionalità strumentale può essere ritenuta normativamente vincolante

²³⁴ Una posizione scettica anche verso questa forma di razionalità sembra corrispondere di più alle intenzioni dello Hume storico. Cfr. Dreier 1997, Millgram, 1995. T. Nagel, *The possibility of altruism*, Princeton, Princeton University Press, 1978, c. 1970, p. 64.

solo se fondata su principi normativi che ci guidano nella scelta di determinati fini, stabilendo dei criteri vincolanti per la valutazione di questi ultimi²³⁵.

Le discussioni più recenti hanno riguardato la possibilità che la razionalità strumentale generi “ragioni per” compiere una determinata azione. Il fatto che per il raggiungimento di un determinato fine è necessario un determinato mezzo è una ragione per scegliere quel mezzo? I saggi di John Broome²³⁶ sono di grande aiuto per capire le implicazioni di questo interrogativo e le risposte che si possono proporre.

Broome non condivide la teoria humeana che attribuisce alla ragione un ruolo esclusivamente teoretico. Se in una situazione di pericolo faccio una scelta diversa dalla sola – di cui pure sono a conoscenza - che può salvarmi la vita, la mia scelta è irrazionale²³⁷. Se una ragione pratica esiste occorre però capire quale ruolo svolge e fino a che punto ci guida. La teoria minimalista di derivazione humeana sostiene che il ruolo della ragione si esaurisce “in

²³⁵ C. Korsgaard, *The Normativity of Instrumental Reason*, cit. “The instrumental principle cannot stand alone...Unless something attaches normativity to our ends, there can be no requirement to take the means to them.

²³⁶ I saggi dedicati a questo argomento da J. Broome sono diversi. *Reasons and motivation, The unity of reasoning, Normative requirement, Reason*. I riferimenti sono nelle note successive.

²³⁷ J. Broome, “Reason and Motivation”, *The Aristotelian Society*, 77 (Supplementary Volume), p. 131-147. Su questa stessa base Nagel afferma che per Hume la ragione pratica di fatto non esiste. Cfr. T. Nagel, *The possibility of altruism*, Princeton, Princeton University Press, 1978, pp. 64-65.

forming beliefs about which means are appropriate to our ends”. A quel punto è solo una inclinazione naturale a far sì che quando vogliamo qualcosa e crediamo che compiere una determinata azione sia il solo modo per ottenerla, vogliamo compiere quell’azione. Ma questa ricostruzione del funzionamento del principio strumentale è implausibile, anche perché semplifica troppo.

Normally we have many ends, which conflict to various degrees. Normally we have many acts to choose amongst and normally none of these acts will have perfectly certain results; each may bring about one or more of our ends with various degrees of likelihood. Furthermore, many of our activities interact with the activities of other people, and that adds the complications of strategic planning. How to choose appropriate means to our ends is the subject of decision theory and game theory ²³⁸.

Una situazione così complessa può essere governata solo dalla ragione che - in una prospettiva cognitivista - può portarci a credenze normative riguardo a ciò che dobbiamo volere o fare. Non è facile però trovare dei validi modelli di inferenza in grado di giustificare queste credenze normative. Broome esclude quella che chiama “naïve instrumental reasoning”: “M wants X, and the only way for M to get X is for her to do A. Therefore, M ought to want to do A”²³⁹. Anche se le premesse sono vere può comunque esserci qualcosa che M vuole più di X, oppure la sua avversione verso A - l’azione che le farebbe ottenere X - può essere più forte del desiderio per X, in entrambi i casi la conclusione – M deve voler fare A - non sarebbe vera. Ma se

²³⁸ *Ivi*, p 132-3.

²³⁹ *Ivi*, p. 134.

questi errori si possono evitare sostituendo i desideri “monadici” con le preferenze che possono invece essere confrontabili, resta la possibilità che l’agente in questione abbia preferenze incoerenti, il che porterebbe a conclusioni altrettanto incoerenti. “Instrumentalism is just false: wanting something is not a reason to want it, and consequently not a reason to try and get it”²⁴⁰. D’altra parte, anche se si riuscisse a giustificare la capacità della ragione di portarci a una credenza normativa, il suo ruolo sarebbe comunque teoretico, pur trattandosi di una credenza non sul modo migliore di raggiungere uno scopo, ma su ciò che la ragione richiede di fare. Se il ruolo della ragione deve essere pratico, le nostre credenze normative devono, qualche volta almeno, poter spiegare le nostre azioni: “for reason to be practical, it must be *possible* for people to do what they believe they ought to do because they believe they ought to do it”²⁴¹. Resta da capire come una credenza normativa possa spiegare un’azione. Può essere ancora la ragione a spingerci ad agire in conformità con quella credenza? Abbiamo già detto che la teoria humeana esclude che una credenza da sola possa motivare, e che l’alternativa che ritiene che siano i desideri a fornire ragioni per agire è falsa. Un’altra possibilità è quella che Broome chiama “the ‘natural disposition’ account of rational action”.

²⁴⁰ *Ivi*, p. 137.

²⁴¹ *Ivi*, p. 139. Il corsivo, presente nel testo, serve a distinguere “*the internalism requirement*” – termine ripreso dalla Korsgaard – dal *belief requirement* che implicherebbe la necessità invece della possibilità di fare quello che si ritiene di dover fare.

We may suppose it is simply a feature of most people's psychology that they are disposed to do what they believe they ought to do. Let call this feature 'general practical rationality'²⁴².

Riassumendo, la ragione ha un ruolo che va al di là di quello "minimalista" perché ci fornisce credenze normative, tuttavia è una particolare forma di razionalità – che non è altro che uno stato psicologico - a disporci a fare ciò che crediamo di dover fare. In realtà Broome non trova pienamente soddisfacente quest'ultima ipotesi, perché la ragione – intesa in senso proprio - viene messa da parte troppo presto, prima che "the real decision making begins"²⁴³. Forse bisogna pensare di fare a meno delle credenze normative, anche considerando che ci sono situazioni in cui non servono. Se, ad esempio, siamo di fronte a una scelta fra due azioni da cui otterremo esiti diversi e che hanno valori incommensurabili non riusciremo a decidere razionalmente fra le due azioni formando una credenza normativa su cosa dobbiamo fare. Si può ipotizzare allora che la ragione pratica non comporti credenze normative ma coinvolga direttamente la volontà: "it goes directly from willing an end to willing the means". Si tratterebbe di una concezione per certi aspetti più vicina alla posizioni kantiane della Korsgaard e che però, non potendo far conto su credenze normative, avrebbe maggiori difficoltà a confrontarsi con la complessità del ragionamento strumentale. E' importante perciò chiarire la natura del ragionamento pratico in generale.

²⁴² *Ivi*, p. 141.

²⁴³ *Ivi*, p. 145

3.6 Broome: ragioni e requisiti normativi

Broome è più volte tornato ad affermare che il ragionamento pratico ha un forte elemento cognitivo²⁴⁴ e a difendere la tesi che il ragionamento pratico è strettamente legato a quello teoretico.

Practical reasoning is reasoning whose conclusion-state is an intention. But if you form an intention by explicit reasoning, in which you express the conclusion to yourself, you will also express a belief. Your reasoning must therefore lead to a belief as well as an intention. The result is that theoretical and practical reasoning are entangled together²⁴⁵.

Sia il ragionamento teoretico sia quello pratico riguardano la verità o falsità delle proposizioni che ne costituiscono il contenuto. Ad essere diversi sono gli atteggiamenti verso la verità, “the attitude of taking as true” nel primo caso, e “the attitude of being set to make true” nel secondo²⁴⁶. Questo stretto legame fra ragionamento teoretico e pratico è anche un antidoto contro lo scetticismo riguardo al ragionamento pratico, perché può garantire, al ragionamento strumentale per esempio, la stessa forza normativa del *modus ponens*. E’ importante però chiarire in cosa consista la normatività della razionalità pratica. Ritenendo inadeguata una concezione “forte” di razionalità pratica intesa come dovere incondizionato (“ought”), Broome distingue due

²⁴⁴ J. Broome, “Unity of reasoning?” In *Disputatio*, 11, 2008,

²⁴⁵ *Ivi*, p. 12

²⁴⁶ “Normative practical reasoning”, *Proceeding of the Aristotelian Society*, 2001, p. 175-93, p. 176

possibili interpretazioni “deboli” identificate rispettivamente con le espressioni “have a reason to” e “follow a normative requirement”.

It is easy to confuse a reason and a normative requirement because both involve a weakening of ought. A reason is a weakening of ought; it is an ought *pro tanto* – an ought as far as it goes. If you have a reason to do something, and no reason not to do it, then you ought to do it. A normative requirement is also a weakening of ought; it is weakened by being made relative. It is easy to muddle the two types of weakening²⁴⁷.

Gli aggettivi usati da Broome per descrivere la differenza fra ragione e requisito sono che mentre la ragione è debole (slack) ma assoluta, il requisito normativo è rigido (strict) ma relativo. Bisognerà capire quale delle due riesce a interpretare correttamente il tipo di relazione che lega premesse e conclusioni in un sillogismo pratico. Il modello più semplice di sillogismo pratico è quello che va dall'intenzione verso un fine a quella verso un mezzo che si crede necessario per il raggiungimento di quel fine. Nell'esempio portato da Broome²⁴⁸: dall'intenzione di aprire una bottiglia di vino (prima premessa) e dalla credenza che per farlo occorra procurarsi un cavatappi (seconda premessa) segue come conclusione l'intenzione di procurarsi un cavatappi. Se intendiamo la normatività nei termini di “una ragione per” dovremmo dire che il sillogismo ci consente di ricavare dall'intenzione di

²⁴⁷ “Are intentions reasons?” In *Practical Rationality and Preferences: Essays for David Gauthier*, C Morris, A. Ripstein ed., Cambridge, Cambridge University Press, 2001, p. 98-120.

²⁴⁸ “Normative requirement”, in *Ratio*, 1999, pp. 398-419.

aprire il vino e dalla credenza che per farlo serva un cavatappi la conclusione normativa dell'intenzione di procurarselo ("detached conclusion") che a questo punto acquista una forza normativa indipendente dalle due promesse che pure ne costituiscono "la ragione per". Tuttavia, se intesa in questo modo come "reasons relation" la relazione fra premesse e conseguenze non è vincolante ("strict") al punto che non procurarsi il cavatappi sia un comportamento sicuramente fuori luogo. Questo perché la "ragione per" è comunque una ragione *pro tanto*, dal momento che può coesistere con una ragione contraria, e a causa di questa sua "debolezza" (*slack*) non può supportare una relazione "strict" da cui possa scaturire una conclusione normativa. Se è intesa invece come "normative requirement", la normatività non sarà nel conseguente, ma nella implicazione logica fra premesse e conclusione, e non avrà a che fare con le ragioni dell'agente, ma con una caratteristica strutturale del ragionamento. E questa dimensione della normatività, secondo Broome, è quella più adeguata al ragionamento strumentale.

Instrumental reasoning brings you to take appropriate means to your ends, and it is not paralysed if your ends happen to be ones you should not have... It would not be possible if reasoning had to generate a reason for its conclusion. But in fact reasoning simply determines a normative requirement: that the premises require the conclusion²⁴⁹.

La normatività della relazione strumentale, quindi, si limita a individuare i mezzi più appropriati per soddisfare un fine, senza che ne scaturisca una

²⁴⁹ *Ivi*, p. 419.

conclusione normativa riguardo alla effettiva adozione di quei mezzi, dal momento che i fini possono non essere quelli che si dovrebbero avere e in tal caso non costituiscono una ragione per l'adozione dei mezzi. "The rightness of correct instrumental reasoning cannot depend on the rightness of the end"²⁵⁰. La tesi che la razionalità vada interpretata in termini di requisiti e non di ragioni vale sia per il ragionamento teoretico sia per quello pratico. Non è vero infatti che essi abbiano a che fare rispettivamente con ragioni per credere e ragioni per intendere: in entrambi i casi il ragionamento riguarda il soddisfacimento di requisiti normativi fra credenze e intenzioni²⁵¹. Riguardo al ragionamento pratico, la tesi che neppure il fatto di adottare un fine basti a renderlo una ragione per perseguirlo ha diversi sostenitori. Secondo Joseph Raz, questa tesi non fa che alimentare una sorta di "mito della razionalità strumentale" mentre, al contrario, né intenzioni né fini in quanto tali possono costituire o fornire ragioni²⁵². Bratman utilizza il termine "bootstrapping" per riferirsi a una ragione che sembra venir fuori dal nulla, data la difficoltà di sostenere che il ragionamento strumentale debba generare una ragione come conclusione²⁵³. La sola cosa che si può dire secondo Broome è che "reasoning

²⁵⁰ J. Broome, "Practical reasoning", in *Reason and Nature: Essays in the Theory of Rationality*, J. Bermudez and A. Miller, ed., Oxford, Oxford University Press, 2002, p. 96.

²⁵¹ J. Broome, "Reasons", in *Reason and Value: Essays on Moral Philosophy of Joseph Raz*, Oxford University Press.

²⁵² J. Raz, "The Myth of Instrumental Rationality", *Journal of Ethics & Social Philosophy*, n. 1, 2005.

²⁵³ J. Broome, "Have we reason to do as rationality requires? A Comment on Raz", *Journal of Ethics & Social Philosophy*, Symposium 1, 2005.

simply determines a normative requirement: that the premises require the conclusion”²⁵⁴.

Una volta identificata con i requisiti, possiamo dire che la razionalità è normativa? La risposta di Broome è che in un senso si può considerarla automaticamente normativa, in quanto la razionalità è un sistema di requisiti o regole e pertanto genera una nozione di correttezza: “following the rules is correct according to the rules”²⁵⁵. Quindi, se “normativo” ha a che fare con norme, regole o correttezza la risposta è affermativa: i requisiti sono normativi appunto perché stabiliscono degli standard di correttezza. Se invece ci chiediamo se, data una regola o un requisito, dobbiamo seguirlo, la risposta di Broome è problematica: non ci sono buoni argomenti per sostenerlo. Escludendo comunque la possibilità di essere completamente scettici riguardo alla normatività e quindi riconoscendo che ci sono cose che dobbiamo fare, credere, sperare, si può provare a difendere la normatività almeno su basi strumentali. I requisiti di razionalità potrebbero essere normativi perché rispettarli è un modo per conseguire alcune delle cose che dobbiamo conseguire. Anche in questo caso però si incontrano delle valide obiezioni, come la possibilità di avere una falsa credenza riguardo a ciò che dobbiamo conseguire o la difficoltà di dimostrare che essere razionali sia sempre

²⁵⁴ J. Broome, “Normative requirements”, *Ratio*, 12, 1999, p. 419.

²⁵⁵ J. Broome, “Is Rationality normative?”, *Disputatio*, 2007, pp. 161-78.

strumentalmente efficace. Difficoltà che portano Broome a dichiararsi agnostico su questo argomento²⁵⁶.

Naturalmente il ragionamento strumentale ha un ruolo centrale nella teoria della scelta, ma Broome trova insoddisfacente anche la descrizione di ragionamento strumentale offerta dalla teoria della decisione. Una breve ricostruzione della sua argomentazione mostrerà quali sono a suo avviso le ragioni del fallimento. La premessa di Broome è che, in questo caso, la teoria della decisione va intesa non come teoria sulla struttura delle preferenze - che non implica un ragionamento - ma come teoria sulla struttura del bene, che al rispetto degli assiomi aggiunge la prospettiva teleologica propria dell'utilitarismo.

I cannot envisage any process of reasoning that could take you from a preference to an intention. A preference can cause an intention, but I do not see how this causal process could be one of reasoning....The theory can easily be divorced from preferences, to serve as a theory about the structure of goodness instead²⁵⁷.

L'utilità che consegue dal risultato di una decisione ne rappresenta la bontà, il che significa che un risultato è migliore di un altro se e solo se ha maggiore utilità. Dato che ogni azione ha una sua utilità attesa, fra tutte le

²⁵⁶ "It still seems intuitively plausible to me that rationality is normative. Moreover, it seems plausible to me that this is so for a broadly instrumental reason. But I am sorry to say I simply cannot find an argument". (*Ivi*, p. 177).

²⁵⁷ "Practical Reasoning" in *cit.* p. 102.

azioni disponibili la migliore sarà quella che ha la più alta utilità attesa. La teoria della decisione descrive il ragionamento strumentale come volto a determinare cosa è meglio fare subordinatamente a una determinata intenzione. Ad esempio, “subordinatamente all’intenzione di Paolo di comprarsi una casa, la cosa migliore per Paolo è chiedere un prestito”. Ma perché questa credenza - che la cosa migliore è chiedere un prestito - giustifichi l’intenzione di farlo è necessario un presupposto teleologico che imponga la scelta dell’alternativa migliore. In questo modo l’intenzione di comprare una casa normativamente richiede di chiedere un prestito. Nella teoria della decisione però questo tipo di proposizioni soggette a condizioni derivano da probabilità condizionate che non implicano il condizionale materiale: “se Paolo compra una casa, deve chiedere un prestito”. La descrizione del ragionamento strumentale è insoddisfacente perché per stare in piedi ha bisogno del sostegno della teleologia, devo presumere che Paolo agisca per amore del bene e non in particolare per comprare una casa. Il perseguimento del bene è invece “merely *constrained*” dalla condizione di star comprando una casa. Da solo, infatti, l’obiettivo di comprare una casa non è in sé sufficiente a convalidare il ragionamento, che deve perciò fare appello al bene in generale per avere un obiettivo esterno a cui mirare. Ma ciò significa che la teoria della decisione non ci dà una descrizione soddisfacente del ragionamento strumentale: “the best means to an end is not the same as what is best, given the end”²⁵⁸. Il ragionamento strumentale non può dipendere dalla teleologia, non può fare appello a un obiettivo esterno di bene. Il suo obiettivo

²⁵⁸ *Ivi*, p. 108.

è interno: scegliere il mezzo migliore è normativamente richiesto dal fine stesso e non dal perseguimento del bene. Rimane vero però, come si è già detto, che nessuna conclusione normativa riguardo a ciò che dobbiamo fare o che abbiamo ragione di fare può derivare dalla relazione di requisito razionale.

In definitiva, solo nei casi più semplici di ragionamento strumentale - quelli in cui crediamo che un particolare mezzo sia necessario per raggiungere un determinato fine - la correttezza del ragionamento è assicurata dalla logica deduttiva. Quando invece, come accade molto più spesso, crediamo che un particolare mezzo sia quello migliore per raggiungere quel fine, il ragionamento strumentale non può essere interpretato come ragionamento normativo.

3.7 *Ruling passions*: la plasticità della ragione

Deliberation is an active engagement with the *world*, not a process of introspecting our own consciousness of it²⁵⁹.

Come vedremo, questa affermazione racchiude un punto centrale delle argomentazioni con cui Blackburn delinea il suo punto di vista su due temi: la natura esterna o interna delle ragioni e la questione della normatività, in particolare quella, generalmente considerata ineluttabile, del ragionamento strumentale.

²⁵⁹ S. Blackburn, *Ruling Passions: a theory of practical reasoning*, Oxford, Clarendon Press, 1998, p. 254. Corsivo nel testo.

La prima cosa sottolineata da Blackburn riguardo alle ragioni è la loro natura relazionale: sono sempre ragioni *per* qualcosa. La relazione può coinvolgere proposizioni, fatti, intenzioni, azioni che possono essere ragioni per altre proposizioni, desideri, intenzioni o azioni. Così, ad esempio, dalla percezione di un fatto all'interno di un contesto segue la formazione di un motivo o intenzione. A ognuna di queste relazioni corrisponde nella mente un movimento "guidato" da uno stato mentale (che esprime il primo elemento della relazione) a un secondo. Si parla di ragione proprio perché è un movimento "guidato", non si basa su una sorta di libera associazione, e quindi implica una forma di ragionamento. Dire che p è una ragione per q significa esprimere approvazione per il tipo di guida della mente indicato dalla relazione.

Una buona ragione per una determinata scelta può però non essere la "ragione" di un agente, a causa di una sua, non importa se e quanto colpevole, cecità. Se il comportamento di un agente non solo non va nella direzione giusta, ma neppure è in grado di seguire l'indicazione di chi lo spinge a farlo, rimane pur sempre passibile di critica, e per Blackburn è solo in questo senso che si può dire che le ragioni sono esterne. Infatti, ciò che diciamo nell'ambito della valutazione non può essere determinato dalle preoccupazioni contingenti di un agente.

Pur condividendo la prospettiva internalista di Williams, Blackburn non condivide tutte le sue conclusioni. Discutendo il caso di un uomo che picchia la moglie e mostra di non avere risorse interne sufficienti per comportarsi in modo diverso, Williams riteneva difficile sostenere che ci fossero comunque buone ragioni per indurlo a smettere di farlo. Secondo Blackburn, invece,

anche in quel caso ci sono ottime ragioni per smettere ed è solo quell'agente, forse a causa di una mente troppo corrotta, a non vederle. In questo senso si può dire che le ragioni sono esterne, dal momento che lo sono i punti di partenza dei percorsi guidati della mente, ma gli internalisti sono comunque nel giusto, perché è altrettanto vero che noi non possiamo prescindere dalle condizioni contingenti in cui ci troviamo per spiegare le nostre azioni ed eventualmente per approvarle. L'internalismo di Williams in quel caso gli fa correre l'inutile rischio di incorrere nel relativismo, che di per sé non è affatto una conseguenza della concezione internalista.

La concezione kantiana ha invece come premessa l'idea di una ragione autosufficiente, che pretende piena autorità sugli stati appetitivi e che per questo può, da sola, individuare i comportamenti da perseguire. Dalla prospettiva naturalistica di Blackburn questa concezione “enchant our natures²⁶⁰” poiché non dà una descrizione veritiera del nostro comportamento. Il compito della ragione, come descritto da Hume, è quello di fornirci informazioni sugli aspetti salienti della situazione in cui ci troviamo ad agire, permettendoci di procedere ad ulteriori deduzioni e inferenze al riguardo. In questo modo otteniamo ulteriori informazioni su quali saranno le conseguenze di certe azioni, quali probabilità hanno di verificarsi e quali implicazioni ne potranno conseguire. Oppure, all'inverso, possiamo individuare le azioni che con maggiore probabilità potranno produrre i risultati a cui miriamo.

²⁶⁰ S. Blackburn, “Précis of Ruling passions”, in *Philosophy and Phenomenological Research*, LXV, 1, 2002, p. 122.

Reason's office is to represent the world to us as it is. But then, how we react to that situation, and that includes how we react to it ethically, is another matter. It is a matter of a dynamic response – the formation of passions, attitudes, policies, or intentions – and the most clear sighted apprehension of the situation is no guarantee that this side is functioning well. The nature of our dynamic response shows our passionate nature, or sensibility. We typically express this structure by voicing ethical remarks, saying what is to be done, or felt, or avoided²⁶¹.

È la descrizione che Broome trovava insoddisfacente perché riteneva che in questo modo la ragione si ferma troppo presto, non partecipa alla decisione, non è investita di alcuna responsabilità riguardo alla azione vera e propria. Secondo Blackburn, invece, una volta raccolte le informazioni ed effettuate le previsioni, la fase dinamica dell'azione – “how we react”- vede coinvolte le passioni, le attitudini, le intenzioni ma non più la ragione. Di conseguenza, anche se, grazie alla ragione, conosciamo gli aspetti essenziali del contesto in cui ci troviamo ad agire, non c'è nessuna garanzia che la dinamica dell'azione sarà altrettanto efficace; la responsabilità in questo caso andrà attribuita alla passione o alla volontà piuttosto che alla ragione. In realtà, anche quando raccogliamo informazioni siamo guidati dalle nostre inclinazioni, da ciò a cui teniamo – *passions* and *concerns* sono i termini più usati da Blackburn - e sono loro a spingerci a dare rilievo a determinati aspetti del contesto, quelli che così diventano salienti e che per questo ci preoccupiamo di esaminare più a fondo.

²⁶¹ S. Blackburn, *Ruling Passions: a theory of practical reasoning*, cit., p. 240.

L'opinione di Blackburn è che la concezione sentimentalista di Hume non implichi affatto la messa al bando della ragione. Quando utilizziamo il termine 'ragionevole' ci riferiamo a un comportamento per quanto possibile affrancato dai nostri limiti cognitivi e cioè da una serie di fattori come l'ignoranza, l'incapacità di comprendere una situazione, l'assenza di lungimiranza, la mancanza di attenzione per il punto di vista degli altri. Ma è pur sempre alla luce di un desiderio di fare la scelta migliore o di una passione per la verità che quei fattori ideali sono ritenuti espressione di ragionevolezza. La loro reputazione non proviene da una fonte indipendente di autorità. E infatti, se è vero che generalmente preferiamo prendere decisioni sulla base di un'ordinata raccolta di informazione e di un'attenta riflessione sugli esiti possibili, in alcuni casi siamo pronti a riconoscere più importanza alla spontaneità o ad accorgerci che un eccesso di informazioni in certe situazioni distorce la nostra capacità di giudizio.

Questo "Hume-friendly use of reason" è agli antipodi dell'accezione kantiana che intende la ragione come completamente indipendente da ogni inclinazione e desiderio e per questo in grado di essere fonte di prescrizioni normative. Apparentemente quest'ultima descrizione sembra più vicina alla fenomenologia della decisione, a quello che sentiamo nel momento della scelta, mentre nel modello humeano l'agente sembra presentarsi con un profilo troppo passivo, in balia di passioni e preoccupazioni che è incapace di governare. In realtà il sé di Hume, concepito come determinato dalla sua natura empirica, sembra passivo solo nel confronto con la visione kantiana di un sé noumenico decisamente separato dal sé empirico.

You, when you deliberate, are whatever you are: a person of tangled desires, conflicting attitudes to your parents, inchoate ambitions, preferences, and ideals, with an inherited ragbag of attitudes to different actions, situations, and characters. You do not manage, ever, to stand apart from all that²⁶².

Nella lettura naturalistica di Blackburn, appare irrealistico presentare, come fa Kant, un soggetto completamente libero dalle sue inclinazioni naturali e acquisite, anche se è sicuramente una fantasia attraente quella di pensare di poter “trascendere” questi dati di fatto che ci riguardano. Korsgaard ha riproposto questa tesi kantiana, individuando proprio nella struttura riflessiva della coscienza umana la capacità di prendere le distanze dai nostri impulsi e desideri. La replica di Blackburn è che possiamo tenerci a distanza da un particolare impulso solo sulla base di altri desideri o di un’altra cosa che per noi è più importante, ma non possiamo farlo da una prospettiva “trascendentale”, completamente indipendente da qualsiasi desiderio o passione. La dipendenza dal carattere contingente delle passioni e dei valori comporta la rinuncia all’autonomia della ragione, e al suo ruolo di legislatore universale di norme di condotta.

Secondo Blackburn l’errore che si commette nell’opporci alla tradizione humanea è quello di identificare la prospettiva deliberativa con quella che prende in esame i desideri e li valuta di fronte al tribunale indipendente della

²⁶² *Ivi*, p.252.

ragione. La conseguenza di questa “oggettificazione dei desideri²⁶³” è di dare un’interpretazione errata del processo deliberativo.

Typically, in deliberation what I do pay attention to are relevant *features* of the external world: the cost of the alternatives, the quality of food, the durability of the cloth, the fact that I made a promise. I don’t *also* pay attention to my own desires. My own concerns and dispositions determine which features I notice and how I react to them ²⁶⁴.

In questa descrizione fuorviante della situazione deliberativa i nostri atteggiamenti vengono passivamente messi a confronto e sottoposti al giudizio della ragione. In realtà, ad essere oggetto d’esame nel processo deliberativo non sono desideri e passioni, ma il contesto e gli aspetti salienti della scelta. Come recita la citazione riportata all’inizio del paragrafo, la deliberazione riguarda un nostro impegno attivo con il mondo, non un qualche procedimento introspettivo della nostra coscienza. Gli esempi portati da Blackburn per sottolineare la centralità della relazione con il mondo sono convincenti: l’ultima cosa che vogliamo fare quando ci domandiamo come sfuggire al traffico, o se muovere la torre in una partita a scacchi, è distogliere la mente dal traffico o dalla scacchiera. Le questioni decisionali su qual è la cosa è più importante o coerente da fare vertono sempre su aspetti del mondo e non sui nostri desideri.

²⁶³ Carla Bagnoli, in “Blackburn on the normative question”, *Iride*, 30, 2000, p. 7.

²⁶⁴ S. Blackburn, cit, p. 253-4, i corsivi sono nel testo.

La metafora della nave utilizzata da Blackburn è sicuramente molto efficace per illustrare le differenze fra la concezione humeana e quella kantiana. Per Hume, la nave è in mano all'equipaggio che rappresenta le varie e molteplici inclinazioni, passioni, sentimenti, e la rotta è determinata dalla risoluzione delle pressioni contrastanti dell'equipaggio. Quando una sola voce ha prevalso e la decisione è stata presa ci sono diversi scenari possibili: le voci discordanti possono essere spazzate via o starsene in disparte pronte a farsi sentire alla prossima occasione, oppure cercare comunque di esercitare una qualche influenza sulla direzione che è stata presa. Alcuni membri dell'equipaggio possono essere più costanti e affidabili di altri; quelli etichettati come "buon senso" o "solerzia" possono avere anche il compito di incoraggiare quanti sono in difficoltà o di spronare i più indolenti. La presenza di elementi di questo tipo renderà più o meno felice la combinazione dell'intero equipaggio, ed avrà ovviamente un'influenza determinante sulla navigazione. Per Kant, l'equipaggio resta lo stesso, ma sul ponte di comando c'è un'altra voce – "a voice with ultimate authority and ultimate power"²⁶⁵ – quella del Capitano che impersona la ragion pura pratica, separata – e quindi libera – da ogni inclinazione. È lui ad avere la responsabilità della rotta e a

²⁶⁵ *Ivi*, p. 246.

comandare l'equipaggio²⁶⁶, ed è ancora soltanto lui a garantire la stabilità della rotta. Il suo comando è espressione di una volontà assolutamente autonoma, non soggetta alla legge di causalità, e dunque libera dalle influenze del contesto in cui opera. Infatti, questa che Kant chiama "causalità della libertà" appartiene al livello noumenico e perciò non può essere rilevata empiricamente, anche se in questo modo possiamo correre il rischio di confondere l'assenza di un'esperienza di causalità con l'esperienza dell'assenza di causalità²⁶⁷. Dal momento che non conosciamo e non possiamo conoscere tutte le forze che informano il nostro agire, pensiamo di sapere che queste forze sono assenti.

Del resto, anche evitando il passaggio al livello noumenico e interpretando l'immagine della ragione capitano come espressione del "first-person deliberative stance²⁶⁸", cioè della prospettiva propria dell'agente, resterebbe comunque inalterata la rappresentazione dell'equipaggio, sempre costitutivamente subalterno al cospetto del capitano. Ed è proprio la tesi della passività delle passioni a far nascere l'esigenza di un punto di vista superiore e

²⁶⁶ Carla Bagnoli, in "Blackburn on the normative question", *Iride*, 30, 2000, pp. 8-14, rileva come Blackburn non sottolinei che quella della nave è una metafora dell'autorità, non del controllo, in quanto il compito del capitano è di comandare in modo autorevole piuttosto che imporre semplicemente il proprio controllo sull'equipaggio. A mio parere però in entrambi i casi resta il ruolo di subordinazione dell'equipaggio rispetto al capitano, ed è tale ruolo che qui interessa sottolineare.

²⁶⁷ Blackburn riporta qui una osservazione di Schopenhauer, *op. cit.*, p. 260.

²⁶⁸ H. Allison, *Kant's Theory of Freedom*, Cambridge, Cambridge University Press, 1990, p. 3.

indipendente che stia a fondamento della deliberazione, autore di quell'atto di autolegislazione universale che per Korsgaard è la fonte della normatività.

How can an act of self-legislation be the source of normativity? Legislation, one would have supposed, like promising, is an act possible only to those who *already* think in normative terms²⁶⁹.

L'obiettivo di Blackburn nell'interrogarsi sul perché ci debba essere un simile fondamento della normatività è di arrivare a mettere in discussione la stessa normatività. L'idea che ci sia qualcosa come “*the normative question*” è un'illusione che deriva dal considerare inclinazioni, desideri, passioni “free of normativity, in themselves alien to reason” e perciò bisognosi di legittimazione da una fonte completamente distinta e indipendente che trasformi i desideri in doveri senza l'aiuto di altri desideri. Ma se le inclinazioni non fossero trattate in questo modo non ci sarebbe più una singola questione normativa: ve ne sarebbero tante quante sono le prospettive deliberative che di volta in volta possiamo adottare, a seconda di quali siano i nostri desideri o le nostre preoccupazioni. “How can attitudes and desires possibly claim equality with, or sovereignty over, the force of reason?”²⁷⁰. A parere di Blackburn, questo interrogativo deriva da un modo sbagliato di rappresentare le cose. Ovviamente, esercitiamo un controllo sui nostri desideri, in qualche misura almeno, ma in modo diverso dall'imperativo categorico: lo facciamo, piuttosto, richiamando alla mente qualcosa a cui teniamo di più.

²⁶⁹ S. Blackburn, *op. cit.*, p. 257. Corsivo nel testo.

²⁷⁰ S. Blackburn, “Précis of Ruling passions”, *cit.*, p. 123.

Nella prospettiva “Hume-friendly”, la ragione non sarà impegnata a reprimere desideri e inclinazioni, ma piuttosto a mobilitarne altri che possano prendere il loro posto.

When I think of the morning, I suppress or destroy the desire to drink more in the evening. But that is only because I care sufficiently about having a clear head or avoiding the hangover in the morning. If that care or concern goes in holiday, thinking has no effect on the progress of the party²⁷¹.

Non c'è una particolare scelta che può essere considerata come espressione della ragione, la ragione determina la comprensione che abbiamo della situazione in cui ci troviamo, cioè l'input – nell'esempio riportato le conseguenze di bere troppo alla festa – ma non il modo in cui l'input diventa output, in questo caso la decisione di non bere troppo. Lo stesso Hume chiama ragione proprio questa capacità di essere lungimiranti nelle scelte, quando la voce della prudenza o dell'esperienza ci spinge a dare più peso alle conseguenze a lungo termine, correggendo l'indubbio vantaggio che ha sull'immaginazione l'idea di un piacere imminente. In ogni caso, la preoccupazione per il mio futuro a lungo termine non può essere considerata come l'unica ragionevole: lo è soltanto rispetto a un dato modello di vita, rappresentativo di un ideale sicuramente non vincolante per tutti, e che andrà sempre messo a confronto con altri ideali e quindi altre preoccupazioni. Allo stesso modo, possiamo privilegiare la posizione dell'agente perfettamente informato, dell'agente in grado di calcolare in modo intelligente costi e

²⁷¹ *Ivi*, p. 123.

benefici, o che ha una piena consapevolezza di tutti i dettagli della sua situazione. Nel farlo, stiamo semplicemente esprimendo quella che è la nostra concezione di vita, che in questo caso privilegia la prudenza invece della spontaneità.

In questa prospettiva humanea, anche la normatività espressa dalle ragioni esterne può essere, come si è già anticipato, riconsiderata. Basta fare attenzione a non farsi confondere dall'ambiguità dell'espressione "X ha una ragione per φ " che può essere intesa sia come una considerazione descrittiva sulla psicologia di X, sia come una considerazione normativa riguardo a ciò che dovrebbe comunque motivare X, indipendentemente dalle sue effettive motivazioni. Entrambe le interpretazioni, una volta distinte, possono essere condivise. Possiamo quindi sostenere con convinzione l'idea che qualsiasi agente dovrebbe sentirsi motivato ad agire in maniera corretta – ritornando sull'esempio di Williams – anche se non manifesta alcun desiderio in tal senso. Ovviamente, nel sostenerlo mettiamo in campo i nostri valori e in base ad essi giudichiamo brutale, sconsiderata, spregevole la persona che non riesce a comportarsi bene. Ma giudicarle tale persona irrazionale significherebbe invece rivestire di una falsa oggettività la nostra convinzione sulla motivazione che un agente dovrebbe avere. In questo caso, infatti, dire che "X ha una ragione per comportarsi diversamente" porterebbe a confondere i due piani – descrittivo e valutativo – che vanno invece mantenuti distinti e a pensare le ragioni come qualcosa che è lì da qualche parte e che noi dobbiamo essere in grado di recepire. Williams ammette un concetto di "ragione per φ " che può anche non essere presente nell'effettivo complesso motivazionale di X, ma che ne farebbe parte se X deliberasse in modo perfetto, cioè

rimuovendo l'ignoranza, di cui tutti hanno interesse a fare a meno. Per Blackburn esiste quindi un perfetto "Hume-friendly use" delle ragioni esterne.

For saying that is merely insisting on the property of the verdict that sometimes a person should be otherwise, whatever his or her motivational profile. It is not investing that verdict with a spurious necessity²⁷².

L'imperativo che impone di adattare i mezzi ai fini – cioè la razionalità strumentale – è visto da Blackburn come una sorta di cavallo di Troia. L'idea è che il fallimento nella scelta dei mezzi appropriati per raggiungere un dato fine sia una manifestazione di irrazionalità. Acconsentire alla natura normativa della *means ends rationality* può essere il punto di partenza per aprire la strada a una concezione della razionalità pratica di tipo kantiano. Se è vero che c'è qualcosa di sbagliato se un agente dice di avere un determinato fine ma poi esita ad adottare i mezzi che servono a raggiungerlo, non si tratta necessariamente di un difetto di razionalità. Infatti, quello che facciamo di fronte a un caso del genere è di cercare una diversa interpretazione, per esempio ricercando nascosti desideri o credenze che possano razionalizzare l'esitazione dell'agente. I risultati di questo processo ermeneutico possono essere sconcertanti oppure non risolutivi dal momento che non siamo così trasparenti a noi stessi come ci piacerebbe essere.

Neither the person who fails to adapt means to ends, nor the akratic, let alone the person who treats the other people as mere means to their ends, is usefully called irrational. They demand more subtle diagnoses, and those diagnoses

²⁷² *Ruling passions*, cit., p. 266.

concern particular flaws either in their doxastic states, or, more often, in their appetitive states²⁷³.

Blackburn sottolinea inoltre che è sicuramente sbagliata l'opinione molto diffusa che Hume ammetta solo il ragionamento strumentale. Non soltanto perché Hume attribuisce questi fallimenti alla volontà o alla passione piuttosto che alla razionalità, ma perché anche occuparsi dei fini merita di essere considerato ragionamento. Infatti, posso criticare i fini di qualcun altro sulla base dei miei valori oppure anche i miei stessi fini sulla base di altri valori che hanno per me maggior peso. Per Hume è impossibile criticare uno dei miei fini solo nel caso in cui si tratti del mio unico oggetto di valore o di desiderio, e ogni altra cosa sia desiderata solo come mezzo per quel fine. In questo caso ci mancherebbe il punto di vista da cui esprimere dubbi sugli obiettivi, e questo è il solo modo in cui possiamo dare valutazioni. Si tratta comunque di una situazione limite perché non siamo così *single-minded*.

Blackburn fa anche notare che nel ragionamento strumentale c'è un fraintendimento che contribuisce a fare confusione. Il condizionale “se intendiamo un fine, dobbiamo intendere i mezzi” può dare l'impressione che con l'antecedente si individui uno stato mentale, e quindi la questione sembra essere se il fatto di avere quello stato mentale dia una qualche ragione per supporre che dobbiamo intendere i mezzi. In realtà, nel contesto della deliberazione l'apparente riferimento a uno stato mentale è incidentale, non c'è un'inferenza da uno stato mentale a un piano per realizzarlo, ma solo

²⁷³ *Ivi*, p. 133.

un'inferenza dall'ipotesi di un fine da raggiungere alla proposta di un piano per riuscirci.

Infine, non va sottovalutato il fatto che l'accusa di irrazionalità possa anche avere dei risvolti non graditi. Se, invece di prestare attenzione alla particolare natura degli errori delle persone, preferiamo parlare di irrazionalità ogni volta che il loro modo di pensare ci sembra difettoso, o comunque poco efficiente rispetto al nostro, non rischiamo solo di perdere di vista degli aspetti e delle distinzioni importanti. Giudicando qualcuno "irrazionale" tendiamo a considerare tale irrazionalità un difetto irrimediabile, e mostriamo di essere poco interessati alla possibilità che ha di correggersi e di migliorarsi. Laddove invece, secondo Blackburn, una descrizione più attenta e una comprensione più approfondita dell'errore può indicarci le strategie più efficaci per porvi rimedio. Il che dimostra che la ragione è malleabile almeno quanto i sentimenti.

Capitolo 4 - Robert Sugden: dalla non expected utility theory alla behavioral economy

4.0 Premessa

In questo capitolo verrà esaminata una concezione della teoria della scelta razionale formulata da un economista – Robert Sugden – che ha preso le distanze dalla teoria classica della scelta (EUT) contestandone la definizione di razionalità e attribuendo il suo fallimento empirico a una insoddisfacente descrizione del contesto di scelta. In prima istanza, dalla sua riflessione scaturisce una revisione della teoria classica, la regret theory, che si propone una più accurata analisi del contesto motivazionale. Successivamente, una lettura fortemente naturalistica della teoria dell'azione di Hume lo porterà a reinterpretare il significato della dimensione consequenzialista della scelta e a contestare il carattere normativo e strumentale della razionalità pratica, fino ad escludere un qualsiasi ruolo significativo della razionalità nei processi decisionali: una prospettiva che lo porterà ad aderire alla behavioral economics.

La reazione degli economisti alle violazioni degli assiomi della EUT messe in evidenza dai problemi e paradossi degli anni '50 e successivamente dagli esperimenti “mentali” e di laboratorio sono state diverse. Le principali tesi difensive messe in campo per salvaguardare la teoria sono state tre.

Una linea di difesa è stata quella di negare la rilevanza dei risultati sperimentali perché non si basavano su scelte reali, e questo bastava a renderli

poco credibili. Gli argomenti utilizzati da questa linea difensiva riguardavano le vincite monetarie proposte negli esperimenti, che erano troppo basse per motivare effettivamente le persone, o la non ripetibilità delle scelte che impediva agli agenti di imparare, dopo un certo numero di decisioni simili, a comportarsi in maniera razionale, cioè a conformarsi agli assiomi della EUT. Una diversa linea di difesa ha sostenuto che la teoria della scelta razionale costituisce la base per le teorie che prevedono i comportamenti nell'aggregato²⁷⁴, pertanto non viene messa in discussione da esperimenti che riguardano scelte individuali. Una difesa plausibile però solo se le violazioni osservate negli esperimenti fossero trascurabili oppure occasionali, mentre i dati raccolti parlano di deviazioni consistenti e sistematiche.

La tesi argomentativa ritenuta più forte contro le sfide portate dalle ricerche sperimentali è quella che rivendica il carattere normativo della EUT, che la rende inattaccabile anche di fronte alle osservazioni di scelte reali²⁷⁵. Una teoria normativa si occupa di indicare come le persone dovrebbero scegliere se vogliono comportarsi razionalmente, e non ha bisogno di presumere che la gente comune agisca sempre in modo razionale²⁷⁶. Quindi, la

²⁷⁴ G. Stigler e G. S. Becker, "De gusti bus non est disputandum", *American Economic Review*, 1977, 67, 2, pp. 76-90, p. 76.

²⁷⁵ Savage sostenne questa posizione dopo le dimostrazioni di Allais. M. Allais and O. Hagen, editors, *Expected Utility Hypotheses and the Allais Paradox*, Dordrecht, Reidel, 1979.

²⁷⁶ R. Sugden, How people choose, in S. H. Heap et al., *The Theory of Choice: a critical guide*, Oxford, Blackwell, 1992, p. 48-49.

razionalità limitata basterà a spiegare i comportamenti devianti. In realtà spesso - anche dopo essere stati convinti di aver fatto una scelta incompatibile con gli assiomi della teoria - non intendiamo tornare indietro e sentiamo di poter sostenere il nostro comportamento con delle buone ragioni. In tal caso, l'argomento che porta a bollare come irrazionali le scelte che violano un assioma potrebbe trascurare un qualche fattore che invece per l'agente razionale è determinante. Neppure l'interpretazione normativa della EUT sembra riuscire a preservarla dalle sfide portate dai risultati sperimentali: “we should always be suspicious of attempts to isolate theories from evidence that might refute them²⁷⁷”.

I risultati sperimentali non provocarono solo reazioni difensive. Nello stesso periodo in cui gli psicologi cognitivi provavano a ridefinire i fondamenti della decisione attraverso ricerche sperimentali sui meccanismi dell'intelligenza umana, molti economisti provarono a modificare qualche assioma della teoria in modo da ottenere previsioni più vicine ai comportamenti reali. La diversità dei percorsi corrisponde alle diverse caratteristiche delle rispettive metodologie e nel caso degli economisti le proposte di riforma di qualche assioma della teoria potevano consentire di salvaguardarne il nucleo centrale, grazie anche all'aggiunta di qualche ipotesi che consentisse di dar conto dei comportamenti devianti. Nel primo capitolo si è detto, ad esempio, dei tentativi in tal senso compiuti – fra la fine degli anni '70 e i primi anni '80 - da MacCrimmon e Chew e da Machina, i primi conferendo una forma più debole all'assioma di indipendenza (Weighted

²⁷⁷ *Ivi*, p. 49.

Expected Utility), Machina invece rielaborando la teoria in modo da farne almeno del tutto (Non-Linear Expected Utility). La proposta di Sugden, la regret theory, è in un certo senso più radicale, perché elimina sia l'assioma di equivalenza sia l'assioma di transitività, ma è soprattutto diverso il modo in cui Sugden si confronta con la EUT e le riflessioni sulla definizione di razionalità che ne scaturiscono.

4.1 Regret Theory

Nell'analisi di Sugden le violazioni sistematiche degli assiomi mostrate dagli esperimenti mettono profondamente in discussione la teoria della scelta razionale, dato che è convinzione comune che gli assiomi rappresentino l'essenza del comportamento razionale in condizioni di incertezza²⁷⁸. La sua ipotesi è che le violazioni possano essere dovute all'influenza esercitata sulle scelte da altri fattori importanti che però sono stati trascurati dalla EUT. La teoria di Sugden si propone di costruire un modello alternativo che tenga conto di questi fattori e sia così in grado di spiegare le violazioni.

La prima tesi della regret Theory è l'esistenza di una *choiceless utility*, e cioè l'utilità che un agente ricaverebbe dalla conseguenza x se ne facesse esperienza senza averla scelta. Si tratta di un concetto di utilità definito indipendentemente dalla scelta, e conforme all'approccio utilitaristico classico.

²⁷⁸ G. Loomes e R. Sugden, "Regret Theory: an alternative theory of rational choice under uncertainty", *The Economic Journal*, 92, 1982, pp. 805-824, p. 805.

What we understand by “choiceless utility” is essentially what Bernouilli and Marshall understood by “utility” – the psychological experience of pleasure that is associated with the satisfaction of desire. We believe that is possible to introspect about utility, so defined, and that it is therefore meaningful to talk about utility being experienced in choiceless situations²⁷⁹.

La nostra capacità introspettiva ci suggerisce anche che l’esperienza psicologica di piacere che proviamo per il verificarsi di una particolare conseguenza – come risultato di una nostra scelta in condizioni di incertezza – dipenderà non solo dalla natura di quella conseguenza, ma anche dal confronto con una diversa conseguenza, se la nostra scelta fosse stata diversa. Se in funzione dello stato del mondo che si è verificato, quest’ultima conseguenza si rivelasse più desiderabile di quella prodotta dalla nostra scelta, rimpiangeremo di non aver scelto diversamente. In caso contrario, ci compiaceremo della scelta fatta. In una giornata nuvolosa, tornare indietro a prendere l’ombrello ci avrà fatto perdere un po’ di tempo, ma se verrà a piovere avremo motivo di rallegrarci della nostra prudenza, proprio pensando a quello che sarebbe successo se non l’avessimo fatto.

La teoria di Sugden include i concetti di rimpianto e compiacimento per mezzo di una funzione di utilità modificata che consente di interpretare i due opposti sentimenti come decremento o incremento di utilità, intesa come esperienza psicologica.

²⁷⁹ *Ivi*, p. 807.

To formulate regret and rejoicing in this way is to assume that the degree to which a person experiences these sensations depends only on the choiceless utility associated with the two consequences in question – “what is” and “what might have been” – and is independent of any other characteristics of these consequences²⁸⁰.

Si presume che l'individuo che sperimenti rimpianto o compiacimento provi a prevedere questi sentimenti e ne tenga conto quando prende una decisione in condizioni di incertezza. Pertanto, le sue scelte saranno volte a massimizzare l'utilità *modificata* attesa, senza bisogno di ritenerlo l'unico obiettivo coerente con l'essere un individuo razionale.

Posta a confronto con la prospect theory, che per spiegare le violazioni della EUT deve assumere molte più cose – l'arrotondamento delle probabilità per eccesso o per difetto e l'esclusione di quelle più basse, una sopravvalutazione delle basse probabilità e una corrispettiva sottovalutazione di quelle alte – la regret theory ha l'indubbio vantaggio della semplicità: riesce infatti a spiegare molti dei risultati sperimentali sulla base di pochissimi assunti. Le spiegazioni che le due teorie propongono vanno in direzioni diverse. La prospect theory attribuisce le violazioni a un trattamento distorto delle informazioni che abbiamo al momento di prendere una decisione, la regret theory invece alla prefigurazione delle diverse reazioni psicologiche che potrò sperimentare, in conseguenza delle mie scelte. Se la *choiceless utility* e il *regret* non possono essere gli unici fattori che influenzano le scelte, sono di certo due fattori molto significativi.

²⁸⁰ *Ivi*, p. 808.

Un'altra differenza importante rispetto alla prospect theory è l'indisponibilità a identificare la razionalità con gli assiomi della teoria della scelta razionale e ad attribuire i fallimenti a errori percettivi o di calcolo. La regret theory non solo può spiegare i comportamenti di scelta che violano gli assiomi, ma è anche in grado di mostrare che non sono affatto comportamenti irrazionali.

We do not doubt that in reality misperceptions and miscalculations occur, and sometimes in systematic rather than random ways. Nonetheless, our inclinations as economists is to explain as much human behavior as we can in terms of assumptions about rational and undeceived individuals. Thus we believe that regret theory does more than predict certain systematic violations of conventional expected utility theory: it indicates that such behavior is not, in any meaningful sense of the word, irrational²⁸¹.

La tesi di Sugden è che si può spiegare il comportamento umano, anche quello che viola gli assiomi di indipendenza o di transitività, senza mettere in discussione la razionalità e la lucidità degli individui.

La sua è una posizione di rottura verso quella che lui stesso definisce come una sorta di tregua accettata da molti teorici di entrambe le parti, con l'eccezione di Allais. I termini di questa tregua vedono da un lato i difensori della EUT come Savage riaffermare il forte carattere normativo degli assiomi come fondamenti della scelta razionale, pur ammettendo i limiti predittivi della teoria. Infatti, lo stesso Savage di fronte al paradosso di Allais, dopo aver

²⁸¹ *Ivi*, p. 819

riconosciuto come un errore la prima scelta e averla sostituita con una diversa, dichiarava di continuare a sentirsi istintivamente attratto dalla prima scelta sbagliata. Dall'altro lato, i sostenitori di teorie alternative come Kahneman e Machina, che continuano comunque a ritenere che le violazioni degli assiomi portino a conseguenze inaccettabili, che gli stessi agenti, una volta consapevoli dell'errore, vorranno correggere. Quindi, sulla forza normativa degli assiomi e sulla conseguente irrazionalità dei comportamenti "devianti" i pareri concordano.

We shall challenge the idea that the conventional axioms constitute the only acceptable basis for rational choice under uncertainty. We shall argue that is no less rational to act in accordance with regret theory, and that conventional expected utility theory therefore represents an unnecessarily restrictive notion of rationality²⁸².

In difesa della razionalità degli assiomi di equivalenza e di transitività della EUT c'è l'idea che il valore assegnato a una determinata azione A dipende solo dall'interazione fra il grado di probabilità che assegniamo al verificarsi delle conseguenze dell'azione A da un lato, e l'ordine di preferenze dell'agente – compresa la sua attitudine al rischio – dall'altro. L'utilità attesa assegnata a un'azione sarà quindi indipendente dal numero e dalla natura delle azioni alternative disponibili e a decidere sarà l'ordine di preferenze completo e transitivo di tutte le azioni. Ne segue che il tipo di conseguenze che ne derivano non è tenuto in nessuna considerazione: se ogni azione è valutata indipendentemente, non importa se le conseguenze di una data azione – venuta

²⁸² G. Loomes e R. Sugden, "Regret Theory", cit., p. 820.

meno l'incertezza e realizzatosi un determinato stato del mondo - possano o meno competere con le conseguenze che sarebbero derivate da una qualsiasi altra azione che avrei potuto scegliere.

Se le conseguenze contano solo in termini di distribuzione delle probabilità, tutte le azioni che condividono la stessa distribuzione di probabilità avranno la stessa utilità attesa e dovranno essere considerate equivalenti ai fini della scelta (assioma di equivalenza). Allo stesso modo la forza dell'assioma di transitività deriva dal supporre che il valore di scegliere una determinata azione sia indipendente dalla natura e dalla combinazione delle azioni che vengono contemporaneamente scartate. Dire che preferisco A a B non significa soltanto che la scelta A dall'insieme $\{A, B\}$ è preferita rispetto alla scelta B dall'insieme $\{A, B\}$, ma implicherebbe anche che scegliere A dall'insieme $\{A, B, C\}$ è preferito a scegliere B dall'insieme $\{A, B, C\}$.

Secondo Sugden, invece, gli assiomi di equivalenza e di transitività non sono affatto condizioni necessarie della scelta razionale in condizioni di incertezza. L'esperienza comune del rimpianto e del compiacimento dimostra l'illegittimità di questi argomenti, perché significa che il valore di ogni scelta dipende dal confronto delle sue conseguenze con il numero e la natura delle conseguenze delle altre scelte disponibili. Il fatto che molte persone provino questi sentimenti e che quando prendono una decisione in condizioni di incertezza provino a prevederli e a tenerne conto sono, come si è detto, i due assunti fondamentali della Regret theory.

In relation to the first assumptions, it seems to us that psychological experiences of regret and rejoicing cannot properly be described in terms of the concept of rationality: a choice may be rational or irrational, but an experience is just an experience. As far as the second assumption is concerned, if an individual does experience such feelings, we cannot see how he can be deemed irrational for consistently taking those feelings into account²⁸³.

Un fattore importante che la EUT non prende in considerazione e che influenza le scelte, è quindi la capacità di prevedere i sentimenti di rimpianto e compiacimento, e il modello della Regret theory invece ne dà conto, mostrando che i comportamenti che violano gli assiomi a causa di quei sentimenti non possono essere ritenuti irrazionali. Non solo non può essere messa in dubbio la razionalità del comportamento di un agente che nel fare la scelta tiene conto di quei sentimenti, ma, al contrario, sarebbe irrazionale non tenerne conto. Ciò non significa che comportarsi in accordo con gli assiomi tradizionali sia irrazionale, ma che quegli assiomi legittimano una definizione inutilmente restrittiva e dunque troppo restrittiva – *unnecessarily* è l'aggettivo usato da Sugden – di comportamento razionale.

4.2 Disappointment theory

Il rimpianto non è l'unica esperienza psicologica ad influire sul comportamento di scelta, la disappointment theory²⁸⁴ individua nel

²⁸³ *Ivi*, p. 820.

²⁸⁴ La disappointment theory viene proposta alla metà degli anni '80 da David E. Bell e poi sviluppata da Loomes e Sugden con una diversa formulazione del disappunto e con diverse

disappunto²⁸⁵ e nell'opposto sentimento di esultanza due esperienze emotive che rivestono un ruolo importante.

It is suggested that, when considering any uncertain prospect, an individual forms some *prior expectation* about that prospect. After the uncertainty is resolved, the individual experiences one particular consequence of the prospect: if that consequence falls short of the prior expectation, then in addition to the utility derived from the consequence itself, the individual also experiences some degree of disappointment; whereas if the consequence is better than the prior expectation, the individual also feels some measure of elation²⁸⁶.

Questa volta l'attenzione è focalizzata sugli esiti alternativi che sarebbero potuti essere migliori o peggiori di quelli effettivamente sperimentati e, a seconda dei casi, si proverà un sentimento di esultanza o di delusione. Si cerca di massimizzare l'"*expected satisfaction*", e si è portati a includervi la previsione dei sentimenti che si proveranno *ex post* rispetto alla scelta, in questo caso la previsione di un possibile sentimento di delusione o di esultanza. La soddisfazione o la delusione che deriva da una conseguenza dipenderanno in parte dal confronto fra quella conseguenza e altre

implicazioni rispetto a quelle di Bell, mentre la regret theory viene elaborata indipendentemente sia da Bell sia da Loomes e Sugden nel 1982. La sistemazione completa della disappointment theory anche dal punto di vista formale è di Faruk Gul in "A theory of disappointment aversion", *Econometrica*, 1991, v. 59, n. 3, pp. 667-686.

²⁸⁵ Ho tradotto "disappointment" con "disappunto" o "delusione" a seconda dei contesti.

²⁸⁶ G. Loomes e R. Sugden, "Disappointment and dynamic consistency in choice under uncertainty", *Review of Economic Studies*, 1986, LIII, pp. 271-282, p. 271. Corsivo nel testo.

conseguenze della stessa azione in differenti stati del mondo. Nella regret theory invece si trattava del confronto fra la conseguenza che si è effettivamente concretizzata e le conseguenze di altre azioni nello stesso stato del mondo. Dopo aver vinto il secondo premio di una lotteria, posso sentirmi delusa pensando che, se si fosse realizzato un diverso stato del mondo, avrei potuto vincere il primo premio. Delusione e rimpianto sono due diversi tipi di malessere emotivo che si possono sperimentare quando ci si trova a riflettere su “what might have been”²⁸⁷.

La disappointment theory può spiegare le violazioni del principio della cosa sicura, cioè dell’assioma di indipendenza. La tendenza a evitare delusioni è alla base del comportamento di scelta cui fa riferimento il paradosso di Allais, e cioè la preferenza per una vincita sicura rispetto a una vincita più alta ma a rischio, e al contempo la preferenza per una vincita più alta con probabilità più bassa, se anche l’altra opzione è a rischio. Può spiegare anche un comportamento evidenziato da un esperimento condotto da Slovic e Tversky²⁸⁸, e cioè come mai molte persone che hanno violato l’assioma d’indipendenza, come nel paradosso di Allais, continuino a preferire la scelta che hanno fatto, anche dopo che ne sia stata dimostrata l’incoerenza.

Secondo Sugden, i due modelli teorici del rimpianto e del disappunto possono essere elementi complementari nell’elaborazione di una teoria

²⁸⁷ *Ivi*, p. 281.

²⁸⁸ P. Slovic e A. Tversky, “Who accepts Savage’s axiom?”, *Behavioral Science*, 19, pp. 368-373.

alternativa della scelta razionale. Si può quindi continuare a parlare di scelta *razionale* anche incorporando nella teoria sentimenti come *regret* e *disappointment*²⁸⁹.

Notice also that regret theory assumes not merely that an individual will regret a decision *after* it has turned out badly, but that he will foresee the possibility of such regret *before* he chooses. In other words, at the same time as I am choosing to bet on the spin of the roulette wheel, I am convinced that I shall regret this choice if I happen to lose²⁹⁰.

Questa precisazione vuole rispondere alla perplessità che può suscitare una reazione emotiva verso un esito che si è verificato indipendentemente dalla mia volontà. Una volta fatta la scelta, quale sia l'esito che si verificherà è completamente fuori dal mio controllo, non è allora forse irrazionale preoccuparmene? Un'analisi dello stato mentale del rimpianto dovrebbe togliere ogni dubbio alla possibilità di considerarlo irrazionale e convalidare il giudizio negativo sul criterio di razionalità dell'EUT.

²⁸⁹ In certe circostanze *regret* e *disappointment* possono essere complementari: "they may be complementary rather than mutually exclusive" in G. Loomes e R. Sugden, "Testing for regret and disappointment in choice under certainty", *The Economic Journal*, 1987, 87, pp. 118-129, p. 120.

²⁹⁰ R. Sugden, "Regret, recrimination and rationality", *Theory and Decision*, 1985, 19, pp. 77-90, p. 78.

4.3 Un diverso concetto di razionalità: *non-expected utility theory*

Secondo Sugden, nel sentimento del rimpianto si possono individuare due elementi distinti. Il primo è il semplice desiderio di aver fatto una scelta diversa: ora che so quale stato del mondo si è realizzato, col senno di poi so che avrei fatto meglio a non fare la scelta che ho fatto*. Questo desiderio non implica nessun giudizio sull'erroneità della scelta fatta a suo tempo. Se scelgo di fare un viaggio in macchina invece che in treno e resto coinvolto in un incidente di cui non ho colpa, posso rimpiangere la scelta fatta anche se non avevo modo di sapere quello che sarebbe successo. L'altro elemento del rimpianto è il rimorso (*self-recrimination*): se si ritiene di aver preso una decisione sbagliata a causa di un errore di giudizio ce ne rammaricheremo, anche se poi è andato tutto bene. Scegliendo di tornare a casa in macchina dopo una cena in cui ho bevuto troppo invece di prendere un taxi, il giorno dopo posso rammaricarmi per aver corso un rischio inutile, nonostante non si sia fatto male nessuno e abbia anche risparmiato i soldi del taxi.

La difesa della razionalità del comportamento spiegato in base al rimpianto faceva riferimento alla prima delle due nozioni di rimpianto, cioè semplicemente al desiderio di aver fatto una scelta diversa. L'argomento era che il criterio di razionalità non si può applicare ai desideri o ai sentimenti in generale, mentre sarebbe irrazionale – se possiamo prevedere quali saranno le nostre reazioni alle diverse conseguenze - non tenerne conto; perciò secondo Sugden “a theory of rational choice ought to take such psychological

propensities as given²⁹¹”. Ma è possibile includere anche il sentimento del rimorso nella teoria senza rimettere in dubbio la razionalità della scelta?

Un’analisi più approfondita serve a mettere in luce come l’intensità del rimpianto dipende da qualcosa di più del semplice confronto fra “what is” e “what might have been”, perché dipende anche da quanto posso biasimarmi per la mia decisione. Ripensando alle diverse situazioni in cui si prova rimpianto per decisioni prese si può riscontrare una differenza fra una situazione in cui avevo solo una generica opportunità di scegliere diversamente e quella in cui invece so che avrei potuto effettivamente fare una scelta diversa, e infatti possono facilmente immaginare di averla fatta. Per questo, secondo Sugden, il rimpianto richiede immaginazione, e l’intensità di questo sentimento è legata alla vividezza con cui riesco a immaginare un diverso corso di eventi.

Il rimorso ha un legame significativo anche con i diversi concetti di probabilità. Proviamo rimorso per una decisione basata su un giudizio di probabilità soggettiva perché possiamo interrogarci sulla validità di quel giudizio. Al contrario, una probabilità oggettiva non può essere causa di rimorso se poi le cose vanno male: non avrò rimorsi per aver scommesso “testa” se poi esce “croce”, anche se posso chiedermi se sia stato saggio accettare di scommettere e provare rimorso per averlo fatto.

In realtà, sono molte le dimensioni del rimpianto che ne possono determinare l’intensità e tutte sembrano dipendere dalla possibilità di

²⁹¹ *Ivi*, p. 79.

difendere di fronte a noi stessi la decisione presa, giudicandola comunque ragionevole. Bisogna anche tener conto che le scelte che si affrontano nella vita di tutti i giorni sono molto più complicate di quelle utilizzate negli esperimenti.

In everyday life will probably have involved ill-defined feasible sets, multi-dimensional consequences and unknown probabilities²⁹².

In molte circostanze ci troviamo a dover decidere cosa preferire senza poter consultare un ordine preesistente di preferenze o applicare una massima generale della cui validità siamo convinti o conformarci a una consuetudine. La scelta in condizioni di incertezza è in questi casi ancora più difficile perché già al momento della scelta è difficile essere convinti che sia quella la scelta giusta, e di conseguenza il rimorso diventa molto più probabile. Se le cose sono andate male, desidererò di aver fatto la scelta che invece ho scartato, perché ora so che sarebbe stata migliore, e non mi consolerà molto il pensiero di averla fatta a ragion veduta. Di fatto non era così, e quindi questo è un caso in cui la presenza dell'elemento "*self-recrimination*" accanto al desiderio che le cose fossero andate diversamente ha un ruolo significativo, e può essere prevista.

So the regret theory should be understood as a theory in which individuals foresee self-recrimination. If it is accepted that people's preferences may be incomplete, foreseeing self-recrimination is not a contradiction in terms, for the

²⁹² *Ivi*, p. 93.

reason I have just outlined. But can a theory of this kind be called a theory of *rational choice*²⁹³?

I motivi per una risposta negativa da parte dei teorici della EUT sono due. Anzitutto, il presupposto che le preferenze possano essere incomplete: ci si può trovare ad affrontare problemi di scelta senza avere delle preferenze chiaramente definite a guidarci. Per l'EUT l'incompletezza delle preferenze equivale a una forma di irrazionalità. Inoltre, potrebbe esserci un classico paradosso: se la regret theory è una teoria della scelta razionale, chi agisce in accordo con essa può giustificare la sua scelta proprio sulla base di questo accordo, ma a quel punto non potremmo più considerarla una scelta difficile.

La risposta per Sugden sta nel fatto che la regret theory non è una teoria della scelta razionale nello stesso senso in cui lo è l'EUT. Una persona che agisce in base a quest'ultima ha un ordine di preferenze perfetto e completo, cioè un piano d'azione per fronteggiare ogni possibile contingenza all'interno di una classe molto ampia di problemi di scelta. Gli assiomi, grazie al loro contenuto normativo, possono guidare gli agenti nella predisposizione del piano d'azione permettendo di individuare il modo – che non è ovviamente l'unico – per rendere le scelte coerenti fra loro. Gli assiomi hanno invece poco da dire nel caso di scelte singole (*one-off choice*).

Al contrario, la regret theory è al suo meglio proprio in caso di assenza di un piano d'azione di questo tipo, il che significa che un comportamento conforme alla teoria non è la conseguenza di una consapevole pianificazione.

²⁹³ *Ivi*, p. 95. Corsivo nel testo.

La teoria descrive come tendiamo a comportarci quando non sappiamo quale sia il comportamento giusto.

Regret theory describes some of the regularities in human behaviour that occur because people sometimes don't know how they should choose. If this is so, it is clear that the people whose choice are being explained cannot themselves see the theory as a means of completely justifying their choices; the whole problem is they cannot find a course of action that they can completely justify to themselves²⁹⁴.

E' interessante che Sugden torni a sottolineare non solo che il campo d'elezione della regret theory è quello delle scelte difficili, ma soprattutto che la difficoltà non sia da ricercarsi nei limiti della razionalità, quanto piuttosto nell'impossibilità di giustificare la propria scelta: non ci sono sufficienti ragioni per la scelta. Neppure la teoria può dare una mano per giustificarle, infatti è essenzialmente una teoria descrittiva e non prescrittiva; ciononostante i modelli di scelta che descrive non sono irrazionali, meglio "are not *contrary* to reason"²⁹⁵.

È qui la differenza con la EUT. Per quest'ultima i comportamenti che violano gli assiomi sono irrazionali. Morgenstern si spinge a dire che nessuno li violerebbe consapevolmente, perciò quando questo succede va considerato una sorta di errore. Per considerarlo un comportamento irrazionale basta che chi viola gli assiomi possa essere persuaso di aver commesso un errore.

²⁹⁴ *Ivi*, p. 96.

²⁹⁵ *Ibid* Corsivo nel testo.

Per la regret theory si può non essere capaci di giustificare completamente le proprie scelte senza che questo significhi che si è commesso un errore. Al massimo si può ritenere che una scelta che contravviene agli assiomi della EUT – come nel caso del paradosso di Allais²⁹⁶ – sia erronea, se la persona crede di avere un piano d’azione in base al quale compie le sue scelte. Se però non ha un piano che le indichi come dovrebbe comportarsi, troverà difficile fare una scelta, e prevederà di potersene pentire. Non si può sostenere che si sbaglia a considerarla una scelta difficile o a prevedere di potersene pentire; se poi nella scelta tiene conto *coerentemente* del prevedibile sentimento di rimorso (*self-recrimination*) che potrà provare, perché si dovrebbe persuaderla a correggere il suo comportamento?

L’irrazionalità sarebbe allora la conseguenza dell’assenza di un piano di azione che potrebbe facilitarci le cose prescrivendo con chiarezza quale deve essere il nostro comportamento. L’idea di tacciare di irrazionalità chi non sa come comportarsi, dato che non conosce le sue preferenze riguardo al particolare problema che ha di fronte, sembra difficile da accettare.

In other words, it is irrational to find choices difficult; it is irrational not to know what one should do. But since we all know that human beings often do

²⁹⁶ Il “common ratio effect”, uno dei due tipi di violazione della EUT descritti da Allais, è coerente sia con la regret theory sia con la disappointment theory, ma le spiegazioni fornite dalle tue teorie sono abbastanza differenti. Cfr. G. Loomes e R. Sugden, “Testing for regret and disappointment in choice under certainty”, cit., p. 121.

find choices difficult, this amounts to saying that homo sapiens is an irrational species. Such a definition of rationality would seem remarkably unhelpful²⁹⁷.

La distanza fra la prospect theory e la regret theory sta nel fatto che la seconda proponendosi come teoria alternativa alla EUT ne mette contestualmente in discussione il concetto di scelta razionale. Infatti, mentre Kahneman sembra piuttosto voler ridimensionare le aspettative dell'omo oeconomicus mostrando le illusioni percettive e cognitive cui è soggetto, nella prospettiva della razionalità limitata di Simon, Sugden mette invece in luce il ruolo di fattori trascurati dalla EUT, e cioè delle emozioni. L'influenza delle emozioni, responsabili dei "fallimenti" della razionalità, porta con sé la necessità di un riesame del concetto di razionalità utilizzato dalla EUT, un riesame che può avvalersi di contributi sia di economisti sia di filosofi.

4.4 I fondamenti filosofici della scelta razionale

Perché sia così importante capire cosa si intenda esattamente quando si parla di "scelta razionale" è detto a chiare lettere da Sugden.

The theory of rational choice has a central place in modern economics. In mainstream economics, explanations are regarded as 'economic' to the extent that they explain the relevant phenomena in terms of the rational choices of individual economic agents²⁹⁸.

²⁹⁷ *Ivi*, p. 97-8.

²⁹⁸ R. Sugden, "Rational choice: a survey of contributions from economics and philosophy", *The Economic Journal*, 1991, 101, pp. 751-785, p. 751.

Vale la pena di tornare a esaminare più da vicino la differenza fra le due interpretazioni della teoria economica - e di conseguenza della teoria della scelta - su cui si dividono gli economisti. Alcuni economisti sostengono la natura puramente descrittiva della teoria economica che avrebbe lo scopo di prevedere il comportamento umano; la validità della teoria si fonda sulla corrispondenza con i comportamenti reali che osserviamo: di conseguenza, se le previsioni non concordano con i dati osservativi, si dovrà cercare una teoria migliore. In questa prospettiva la teoria della scelta razionale si limita a descrivere certe regolarità del comportamento umano, e il concetto di razionalità equivale a quello di coerenza, un agente è razionale nella misura in cui le sue scelte sono coerenti fra loro. Questa concezione è però poco interessante agli occhi di Sugden, dal momento che la coerenza riguarda le scelte “when viewed from the perspective of some theory²⁹⁹”. A questo punto l’espressione “teoria della scelta razionale” risulterebbe tautologica: un comportamento coerente con il modello postulato da una teoria della scelta, e costruito sulla base di regolarità riscontrate nell’osservazione dei comportamenti comuni, sarà considerato razionale.

La teoria dell’utilità attesa è invece una teoria della scelta razionale che ha un contenuto normativo: ci dice come dobbiamo scegliere in quanto agenti razionali. Se ha anche forza predittiva è perché gli uomini tendono ad agire razionalmente. Per esaminare il concetto di razionalità della scelta, occorre partire dai fondamenti filosofici di questa teoria, ma il comune riferimento a Hume può suscitare qualche perplessità.

²⁹⁹ *Ivi*

It is often said that the theory of rational choice used within economics embodies an instrumental conception of rationality – a conception whose classic statement can be found in the work of Hume. [...] But it is an open question whether the theory of rational choice, as we know it today, can be given a consistent foundation in terms of a Humean conception of rationality³⁰⁰.

Il presupposto di una teoria della scelta razionale è che le scelte possano essere influenzate dalla ragione, e una risposta alla domanda su come questo avvenga viene dalla concezione strumentale della razionalità, ossia la *means-end rationality* comunemente associata alla teoria della motivazione di Hume. Il dubbio di Sugden è se si possa davvero ritenere valida questa associazione, se quella di Hume sia una concezione strumentale della razionalità nel senso della EUT.

La teoria della motivazione di Hume è stata oggetto di diverse interpretazioni. Sugden fa riferimento in particolare all'interpretazione di Michael Smith³⁰¹, che è incentrata sulla distinzione categoriale fra le due componenti dell'agire, credenza e desiderio, in base alla diversa "direction of fit" con il mondo. La credenza è uno stato mentale con una direzione di adattamento che va dalla mente al mondo, dato che è vera quando il suo contenuto corrisponde a come stanno le cose nel mondo, mentre il desiderio è uno stato mentale con un'opposta direzione di adattamento, poiché si realizza facendo in modo che il mondo gli corrisponda. Questa distinzione categoriale,

³⁰⁰ *Ivi*, p. 752.

³⁰¹ Oltre al saggio "The humean theory of motivation", in *Mind*, 1987, v. 96, pp. 36-61; *The moral problem*, Oxford, Blackwell, 1994, pp. 92-129.

secondo Sugden, spiegherebbe perché i desideri non possono derivare da credenze.

When we ask what motivates a person's action, we are necessarily working in the realm of desire: the motive for an action must involve some kind of goal, and 'having a goal' is a mental state with which the world must fit³⁰².

In realtà l'interpretazione di Smith mira soprattutto a riformare l'approccio sentimentalista di Hume, di cui critica in particolare la tesi che i desideri non possono essere oggetto di critica razionale. La prospettiva della *direction of fit* si propone in un certo senso di "razionalizzare" il desiderio, che viene ridefinito come uno stato mentale con un dato contenuto proposizionale, una concezione che non corrisponde alla descrizione che ne dà Hume.

A passion is an original existence, or, if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification. When I am angry, I am actually possest with the passion, and in that emotion have no more a reference to any other object, than when I am thirsty, or sick, or more than five foot high³⁰³.

Attribuire al desiderio un contenuto proposizionale significa renderne possibile la traduzione in giudizi che potrebbero determinare la correttezza o meno del desiderio, cioè stabilire se si tratti di desideri razionali o

³⁰² R. Sugden, "Rational choice", cit. p. 753.

³⁰³ D. Hume, *A treatise of human nature*, cit., p. 415.

irrazionali³⁰⁴. Infatti, la lettura humeana di Smith³⁰⁵ sostiene una concezione disposizionale del desiderio in base alla quale il desiderio comprende al suo interno degli elementi di carattere valutativo proprio della credenza che possono certificarne la razionalità, con la conseguenza di attenuare molto la differenza fra le due componenti, basata invece sul carattere cognitivo della sola credenza.

We might say that the desire to \emptyset is also, inter alia, a disposition to believe, under certain conditions, that \emptyset -ing is prima facie desirable (= desired by me); or that there is a reason to \emptyset ³⁰⁶.

Smith in effetti si propone esplicitamente di “correggere” la teoria di Hume, dal momento che trova inaccettabile l’idea che non sia possibile dare una valutazione razionale dei desideri, ma in questo modo fraintende la teoria della motivazione di Hume.

Per Hume la ragione, da sola, non può mai essere motivo di un’azione della volontà e non può mai contrapporsi alla passione nella guida della volontà³⁰⁷; il motivo quindi sarà sempre una passione che la ragione potrà

³⁰⁴ L. Greco, *L’io morale. David Hume e l’etica contemporanea*, Napoli, Liguori Editore, 2008, p. 153.

³⁰⁵ Una lettura condivisa fra gli altri anche da J. Bricke, *Mind and morality: an examination of Hume’s moral psychology*, Oxford, Clarendon Press, 1996. CFR. T. Nagel, *The possibility of altruism*, Princeton, Princeton University Press, 1978.

³⁰⁶ M. Smith, “The humean theory of motivation”, cit, p. 53-53.

³⁰⁷ D. Hume, *A treatise of human nature*, 1740, II ed. By P. H. Nidditch, Oxford, Clarendon Press, 1988, p. 413.

guidare individuando le connessioni fra gli oggetti che hanno suscitato l'impulso emotivo e estendendo così il campo di indagine agli oggetti che hanno caratteristiche simili. Nella situazione di scelta, l'impulso ad agire, ossia il motivo, è sempre una passione che nel momento in cui viene tradotta in atto può essere affiancata da un giudizio che l'aiuti a scegliere i mezzi adeguati al fine da lei designato. Hume distingue nettamente i ruoli che svolgono passione e ragione e ne definisce con chiarezza la diversa natura: la prima ha il compito pratico di dare inizio all'azione – la passione che si traduce in atto – mentre la ragione svolge invece un compito di carattere teoretico, assicurandosi sia della fondatezza della passione – cioè dell'effettiva esistenza dell'oggetto che l'ha suscitata – sia della adeguatezza dei mezzi per raggiungere il fine, senza che però ciò implichi che ci sia un unico modo per farlo.

Di conseguenza, per Hume l'irrazionalità è imputabile in ultima analisi piuttosto alla credenza che all'azione. Un'azione è irrazionale quando il desiderio è accompagnato da una falsa credenza, come quando una passione si fonda sulla supposizione dell'esistenza di un oggetto che in realtà non esiste; oppure quando nel tradurre in atto una passione scegliamo mezzi inadeguati allo scopo. Su quest'ultimo punto il commento di Sugden è il seguente.

Thus reason is to be seen as an instrument for achieving ends that are not themselves given by reason. We may say that an act is irrational if it is not *the*

best means of achieving the ends that the actor himself had a view when choosing the act³⁰⁸.

Quest'interpretazione fa dipendere l'irrazionalità da un ragionamento che non sia in grado di individuare il mezzo migliore per raggiungere i fini che l'agente si propone con una determinata scelta. Hume però parla solo di *mezzi insufficienti* allo scopo.

When in exerting any passion in action, we chuse means insufficient for the design'd end, and deceive ourselves in our judgment of causes and effects. Where a passion is neither founded on false suppositions, nor chuses means insufficient for the end, the understanding can neither justify nor condemn it. [...] Tis as little contrary to reason to prefer even my own acknowledg'd lesser good to my greater, and have a more ardent affection for the former than the latter. A trivial good may, from certain circumstances, produce a desire superior to what arises from the greatest and most valuable enjoyment³⁰⁹.

La parte omessa nella citazione è quella in cui Hume utilizza i noti due esempi volutamente paradossali che servono a sottolineare l'impossibilità di valutare la razionalità di una passione: non è contrario alla ragione preferire la distruzione del mondo intero al graffio del mio dito, o scegliere la mia completa rovina pur di evitare il più lieve dolore a un indiano o a una persona che mi è del tutto sconosciuta. Di conseguenza, il ruolo della ragione sembra essere solo strumentale e non valutativo, e perciò non indica *il mezzo migliore*

³⁰⁸ R. Sugden, "Rational choice", cit. p. 753. Corsivo mio.

³⁰⁹ D. Hume, *A treatise of human nature*, p. 416.

per ottenere il fine, non può giudicare la scelta di un bene che si sa essere inferiore, e neppure valutare la differenza fra diversi tipi di desideri. Ciononostante, non si può fare a meno di provare un forte senso di irragionevolezza di fronte a quegli esempi di scelta e la loro stessa paradossalità è indubbiamente volta a solleccitarlo.

Sugden ritiene che ci siano almeno due aspetti della teoria della scelta razionale che non sono compatibili con la concezione di Hume. Il primo aspetto è dato dalla tesi che, una volta definite le preferenze di una persona, la ragione sia sempre in grado di indicare cosa fare, per cui le scelte si possono considerare determinate. La teoria infatti prescrive un unico corso d'azione per ogni contingenza, con la sola eccezione dell'indifferenza. Tuttavia, proprio perché, come si è detto, la ragione non indica il mezzo migliore, questa determinazione (*determinacy*) non si trova nella teoria della motivazione di Hume. Riguardo al secondo aspetto, per la EUT le preferenze dovrebbero soddisfare certe condizioni di coerenza, come la transitività, ma la concezione di Hume – come è evidente anche nella citazione sopra riportata – non implica affatto il rispetto di queste condizioni.

Per ovviare a queste difficoltà, secondo Sugden, i presupposti della teoria della scelta razionale non dovrebbero essere considerati assiomi della razionalità ma premesse di carattere psicologico sulla natura del desiderio, come nella teoria economica utilitarista di fine XIX, ad esempio di Jevons, che fondava la scelta su una teoria psicologica incentrata sulla sua quantificazione monodimensionale dei piaceri e delle pene, in base a criteri di durata e

intensità³¹⁰. Questa concezione, insieme alla teoria della motivazione di Hume, permetterebbe di formulare una teoria della scelta strumentalmente razionale, internamente coerente e determinata.

Nella ricostruzione di Sugden la presa di distanze da questa concezione avviene a partire dagli anni '30 del XX secolo, quando ci si libera della psicologia utilitarista per abbracciare la più ambiziosa strategia di fondare la teoria sugli assiomi di razionalità.

An alternative foundation for the theory of choice was sought in a purer concept of rationality, free of any psychological assumptions. The culmination of this programme was Savage's *Foundations of Statistic* (1954)³¹¹.

Si tratta di un concetto di razionalità più puro perché condivide la normatività dei principi della logica. Per valutare la teoria che Savage costruisce sulla base di questa concezione bisogna brevemente ritornare sul significato che viene attribuito ai termini cruciali di probabilità, preferenza e scelta. La probabilità viene definita in termini di preferenza, dato che per la sua concezione soggettiva della probabilità quest'ultima misura la fiducia nella verità di una proposizione, la preferenza per x indica quindi anche la maggiore fiducia per il verificarsi del corrispondente evento A. Rifiutando l'idea di utilizzare l'introspezione per conoscere le preferenze, queste ultime andranno a loro volta interpretate in termini di scelta. Anche se gli assiomi

³¹⁰ W.S. Jevons, *The theory of political economy*, 1871, rist. Harmondsworth, Penguin Books, 1970.

³¹¹ R. Sugden, "Rational choice", cit. p. 757.

sono formulati in termini di preferenza, è la scelta il concetto fondamentale; diventa perciò cruciale che le preferenze siano complete per essere autorizzati ad assumere che le scelte rivelino le preferenze e così tradurre le regole di coerenza delle preferenze in quelle di coerenza delle scelte. La completezza delle preferenze è un requisito della razionalità, dal momento che solo a questa condizione le scelte possono essere determinate, e la determinatezza sembra essere una caratteristica della razionalità da cui non si può prescindere.

Sugden ha buon gioco a mostrare che in realtà possono esserci situazioni di scelta in condizioni di incertezza in cui le opzioni non sono commensurabili, e dunque non siamo in grado di stabilire un ordine di preferenza. In tal caso la scelta potrebbe essere casuale, o influenzata dall'effetto framing del problema di scelta. La via d'uscita possibile per la teoria di Savage sarebbe allora quella di descrivere questo comportamento come indifferenza fra le due scelte; ma se si introduce una terza possibilità, si vede che l'indifferenza a quel punto potrebbe valere solo con una delle due opzioni precedenti, e l'assioma della transitività verrebbe violato. Il punto cruciale è che in una situazione del genere l'agente non saprebbe fornire una ragione adeguata per la sua scelta; potrebbe anche ammettere che è stato un aspetto del framing a farla propendere per una delle opzioni, e che in una rappresentazione diversa del problema la sua scelta sarebbe potuta essere diversa. In questo modo, senza pretendere di offrire una ragione per la sua scelta, si limiterebbe a ripercorrere il percorso mentale che l'ha portata a fare una data scelta in una situazione in cui non vi erano ragioni sufficienti a sostenerla. Per la teoria humeana della motivazione, questo comportamento non potrebbe essere ritenuto irrazionale. poiché nasce da un desiderio che,

proprio in quanto dato psicologico, non può essere oggetto di un giudizio di razionalità.

The truth is that her desire do not provide her with enough data from which to work out what she would do. [...] We must admit the possibility that reason is like one of those algorithms that operations researchers are trained to use: a useful instrument, but only for decision problems that happen to have the right structure³¹².

Per una concezione strumentale della razionalità la completezza non è un requisito necessario delle preferenze razionali; gli assiomi della teoria di Savage invece sono troppo forti per poter essere giustificati dalla concezione humane della razionalità.

Savage interpreta la razionalità come coerenza fra le scelte (with one another) e non come coerenza delle scelte con un dato sistema di credenze e desideri. Il punto di partenza non sono le valutazioni dell'utilità e delle probabilità, per arrivare poi a calcolare cosa sarebbe razionale scegliere; al contrario, il punto di partenza sono le scelte.

He (Savage) starts with a consistent pattern of choices and then derives measures of utility and probability from these choices. Certain patterns of choices are deemed to be irrational, by virtue of their internal inconsistency, and quite independently of any reference to the chooser's mental states. This does not fit easily with the idea of reason being the slave of the passions³¹³.

³¹² *Ivi*, p. 760

³¹³ *Ivi*, p. 760.

Secondo Sugden, non si può parlare in questo caso di razionalità strumentale. La differenza fondamentale con la razionalità strumentale è l'assenza di una relazione con gli stati mentali dell'agente. Savage non prende in considerazione desideri e credenze, l'unico requisito è che le decisioni di una persona soddisfino certe condizioni di coerenza reciproca. La coerenza quindi riguarda solo le relazioni fra le scelte e con le corrispettive conseguenze, che a loro volta sono prese in considerazione isolatamente, escludendo ogni possibilità di confrontarle. Le violazioni della transitività mettono però in evidenza molte esperienze di scelta che non trovano spiegazione, perché nella ristretta definizione di conseguenza di Savage non trovano posto considerazioni o sentimenti dell'agente su come le cose sarebbero potute andare, se la scelta fosse stata diversa. In realtà, se scommetto su un cavallo che poi non vince, la conseguenza non è solo la riduzione della mia ricchezza, ma è anche l'esperienza di un doloroso sentimento di rimpianto che nasce dal confronto fra il mio attuale stato di ricchezza e quello che avrei potuto avere se non avessi scommesso.

On a Humean view of rationality, regret is just another kind of passion, to which reason must be a slave: there is no sense in which the feeling of regret can be called reasonable or unreasonable³¹⁴.

La regret theory si basa sull'idea che l'agente non dia importanza solo a ciò che ottiene ma anche a ciò a cui nel contempo rinuncia. In questo caso la descrizione di una conseguenza dovrà includere tutto ciò che è rilevante per

³¹⁴ *Ivi*, p. 763.

determinare le preferenze di un agente. Inoltre, l'assioma di transitività può essere difeso come una proprietà necessaria della razionalità solo se un aspetto rilevante come può essere il rimpianto non viene tralasciato nella descrizione delle conseguenze, perché solo in questo caso abbiamo buone ragioni per aspettarci che le preferenze di un agente razionale rispettino la transitività.

Alla razionalità assiomatica di Savage, Sugden contrappone una razionalità strumentale che permette di interpretare la situazione di scelta a partire dalla relazione con gli stati mentali dell'agente, restituendo importanza al percorso motivazionale, e quindi alla rappresentazione di scenari alternativi e alla costruzione delle preferenze. Una posizione, almeno in apparenza, vicina a quella di Bernard Williams, e alla sua difesa del carattere costitutivamente non strutturato dell'insieme motivazionale³¹⁵.

La razionalità strumentale è, secondo Sugden, la concezione della razionalità di Hume, ma in realtà si tratta di una interpretazione di Hume abbastanza controversa. Thomas Nagel, ad esempio, si esprime in termini assolutamente negativi riguardo all'esistenza stessa di una concezione della razionalità pratica nella teoria di Hume. La sua tesi è che la caratteristica determinante della razionalità pratica sia quella di essere "a reason to do or want something. [...] To hold, as Hume did, that the only proper rational criticism of action is a criticism of the beliefs associated with it, is to hold that practical reason does not exist"³¹⁶. Il ruolo della ragione rimarrebbe per Hume

³¹⁵ Cfr. *supra*, cap. 3 par. 3.

³¹⁶ T. Nagel, *The possibility of altruism*, cit., p. 64.

un ruolo esclusivamente teoretico, del tutto inadeguato al compito normativo di indicarci cosa fare o cosa volere e perciò del tutto insoddisfacente per Nagel³¹⁷.

Riassumendo, la dimensione naturalistica della filosofia di Hume è incompatibile con la natura normativa della razionalità della scelta propria dell'EUT. La sua teoria della motivazione è invece compatibile con una razionalità pratica meno “restrittiva” rispetto a quella dell'EUT, e cioè la razionalità strumentale, che per Sugden è al servizio delle passioni ed è l'unica a svolgere il ruolo di causa o motivo dell'azione, individuando i percorsi più efficienti per raggiungere lo scopo, l'oggetto della passione. Questa razionalità strumentale non è determinata, cioè non indica un unico determinato percorso per raggiungere l'obiettivo, non è vincolata al rispetto di criteri di coerenza, proprio perché dipende dagli stati mentali dell'agente, cioè da quegli aspetti psicologici della motivazione da cui la EUT avevo preso le distanze per concentrarsi invece sulle scelte che potevano essere considerate costitutivamente determinate.

La confusione intorno al concetto di razionalità strumentale si spiega col fatto che quella che per Sugden è l'autentica razionalità strumentale di Hume è stata recepita dal pensiero economico sotto la particolare angolazione – di fatto deformante - prodotta dal “principio di utilità” che, secondo Bentham, dà espressione ai requisiti di razionalità. Il principio di utilità approva o disapprova una qualsiasi azione a seconda della sua capacità di incrementare il

³¹⁷ J. Broome è dello stesso parere. Cfr. cap. 3.

piacere, il benessere o felicità dell'agente, in questo modo però gli stati mentali che contano non sono quelli che causano l'azione (i motivi della scelta) – come nella teoria di Hume – ma quelli che conseguono dall'azione. L'utilità è vista come una qualità che tutte le azioni possiedono e che può essere misurata in modo unitario, così ogni problema di scelta razionale diventa un esercizio di massimizzazione. La razionalità strumentale nel senso di Bentham si identifica quindi con la massimizzazione dell'utilità.

Bentham's account is in terms of mental states which *result* from actions, and not in terms of the desires which, in Hume's theory, provide the motive power for action. This gives reasons a more active role. Bentham makes it rational to do what maximizes happiness over the rest of one's life, even if one's current desires point in a different direction. To this extent, reason can overrule passion, although only for the sake of achieving a greater overall balance of "benefit, advantage, pleasure, good, or happiness. [...] This crucial feature of Bentham's utilitarianism is not an implication of Hume's theory of practical reason"³¹⁸.

Nella concezione utilitarista il ruolo della ragione cambia profondamente rispetto alla teoria di Hume e diventa "più attivo" perché, fintantoché la ragione è schiava delle passioni e queste ultime hanno un'origine indipendente, non potrebbe essere "a requirement of reason"³¹⁹, che le passioni siano passibili di rappresentazione matematica, e dunque misurabili in termini

³¹⁸ M. Hollis e R. Sugden, "Rationality in action", *Mind*, 1993, v.102, 405, pp. 1-35, p. 5-6. Corsivo nel testo.

³¹⁹ *Ivi*, p. 6.

di utilità. Dal punto di vista di Hume, se i desideri di un agente non hanno le proprietà di commensurabilità e di consistenza logica richieste dalla ragione per identificare l'azione migliore, tanto peggio per la ragione.

Questa ricostruzione richiede alcune precisazioni. Se è vero che l'utilitarismo è estraneo alla teoria della motivazione di Hume, l'attenzione sugli stati d'animo che conseguono dall'azione è più difficile da biasimare – anche la regret theory tiene in conto gli stati d'animo che possono derivare da una certa scelta – a meno di non sottolineare l'esclusivo interesse per l'utilità. La strumentalità della razionalità in chiave utilitarista comprende quindi sia consequenzialismo sia il welfarismo³²⁰. La ragione ha poi sicuramente un ruolo più attivo nella teoria di Bentham – nel senso che può legittimamente contrapporsi alle passioni imponendone una trattazione quantitativa – ma, come si dirà più avanti, anche nella teoria di Hume c'è posto per una ragione niente affatto inerte.

Nella ricostruzione di Hollis e Sugden il percorso compiuto dalla teoria economica per arrivare allo sviluppo della teoria dell'utilità attesa di Von Neumann e Morgenstern (1947) ha quindi un primo punto di svolta nell'approccio utilitaristico di Bentham e nella successiva matematizzazione della teoria dell'utilità marginale. Le basi psicologiche dell'utilitarismo creano comunque dei problemi sia riguardo alla effettiva possibilità di uniformare la molteplicità in base al criterio dell'utilità, sia per il carattere

³²⁰ E' la relazione fra consequenzialismo e welfarismo messa in luce da Sen: cfr. A.K. Sen, *On Ethics and Economics*, Oxford, Blackwell 1987.

monodimensionale del concetto di utilità inteso come piacere. Il tentativo di evitare queste complicazioni senza rinunciare a quegli assunti psicologici porta gli economisti marginalisti a intenderli non come verità universali ma come semplificazioni sufficientemente realistiche per gli scopi della teoria economica³²¹.

Il secondo punto di svolta è nell'approccio completamente diverso a cui ha dato inizio Pareto, mostrando che la teoria della scelta poteva fare a meno delle tesi psicologiche sull'utilità, pur mantenendo la struttura matematica della teoria utilitarista.

We may describe the person's behaviour as utility-maximizing, but this is merely a convenient language for expressing the idea that the person chooses in accordance with a given set of preferences. Summing up this approach, Pareto claims: "The theory of economic science thus acquire the rigour of rational mechanics; it deduces its results from experience, without bringing in any metaphysical entity" (1972, p. 113)³²².

A portare a compimento la teoria mainstream della scelta razionale, sono la teoria delle preferenze rivelate di Samuelson, che definisce gli assiomi riguardo ai comportamenti di scelta – in particolare l'assioma di consistenza logica fra le scelte – e l'estensione dell'approccio di Pareto alle scelte in condizione di rischio ad opera di Von Neumann e Morgenstern con la teoria dell'utilità attesa. A questo punto la sintesi operata da Savage rafforza il

³²¹ Hollis e Sugden portano gli esempi di Jevons e di Edgeworth, *op. cit.* p. 5.

³²² *Ivi*, p. 6.

carattere normativo della teoria equiparando gli assiomi ai principi della logica e ribadendo la riduzione delle preferenze alle scelte introdotta da Samuelson.

Notice that Savage has made no explicit assumptions about desire or beliefs. He has required only that a person's decisions satisfy certain conditions of mutual consistency. This, then, is not a theory of instrumental rationality. [...] Savage retains a utilitarian mathematics while abandoning not only the utilitarian psychology but also the broadly Humean theory of practical reason on which this mathematics previously relied. The notion of utility has thus been bleached of all psychological content and we are left with an abstract scheme of quasi-logical relations³²³.

Una volta venuto meno ogni riferimento psicologico dal concetto di utilità, la razionalità è interpretata come semplice coerenza fra le decisioni, in un astratto schema di relazioni semilogiche in cui non c'è più posto per una teoria della razionalità strumentale.

4.5 Oltre la razionalità strumentale: Hume e la *behavioral economics*

Nella teorizzazione più recente di Sugden, la rivisitazione della teoria di Hume continua a svolgere un ruolo cruciale per spiegare i comportamenti di scelta e costruire un'alternativa alla EUT, che ora per Sugden è rappresentata dalla behavioral economics, con la conseguenza che anche la razionalità strumentale va messa in discussione. I titoli dei due articoli che hanno entrambi al centro la teoria di Hume sono rivelatori: "Why rationality is *not* a

³²³ *Ivi*, p. 7.

consequence of Hume's theory of choice³²⁴” uscito nel 2005 e l'anno dopo “Hume's non-instrumental and non-propositional decision theory³²⁵”. L'intento di Sugden è quello di contestare l'interpretazione più diffusa della filosofia di Hume, che gli attribuisce una teoria della decisione in chiave strumentalista – al punto che “philosophers often use the term ‘Humean’ as a synonym for ‘instrumental’³²⁶ –, per sostenere che in realtà Hume nega che si possa attribuire razionalità alle azioni, e che la sua teoria dell'azione è compatibile con la behavioural economics.

La risposta all'interrogativo posto dal titolo del primo articolo è nel fatto che la teoria della scelta ricavabile dai testi di Hume non implica che gli agenti siano razionali, almeno nel senso della teoria della scelta razionale. Al contrario, se le loro scelte violano gli assiomi e la spiegazione causale di queste scelte non riesce ad essere anche una giustificazione – non possono addurre ragioni che spieghino perché le abbiano fatte – i loro comportamenti sono non-razionali.

Hume non parla quasi mai di “preferenze” ma di desideri che, come tutti gli stati mentali, sono intesi “not as long-standing dispositions but as short-lived events in the mind”³²⁷, una caratterizzazione che mette in evidenza la

³²⁴ In *The European Journal of the History of Economic Thought*, 2005, 12:1, pp. 113-118.

³²⁵ In *Economics and philosophy*, 2006, 22, pp. 365-391.

³²⁶ R. Sugden, “Hume's non-instrumental and non-propositional decision theory”, 2006, cit., p. 366.

³²⁷ R. Sugden, “Why rationality is *not* a consequence of Hume's theory of choice”, 2005, cit., p. 115.

differenza con le preferenze di cui parla la EUT – stabili e determinate – e può spiegare l’eventuale mutevolezza e instabilità delle scelte che ne conseguono.

Although Hume says that desire is ‘founded on’ and ‘arises from’ pain and pleasure, he does not claim that the desire to perform an action is always induced by a settled belief that that action will bring about pleasurable consequences, or that the strength of desire is correlated with the corresponding degree of expected pleasure, rather, his hypothesis is that the idea of pleasure tends to induce the feeling of desire³²⁸.

In questa ricostruzione del meccanismo di scelta non c’è posto per la solida credenza (*belief*) che quell’azione produrrà piacevoli conseguenze, il desiderio è indipendente da ogni valutazione sulle conseguenze, e la sua intensità non ha niente a che vedere con il livello di piacere che ci si aspetta di provare, ma varia a seconda della vivacità degli stati mentali presenti in quel momento, e può essere influenzata da associazioni di idee del tutto indipendenti dal piacere che ne seguirà. Sottolineando che l’ipotesi di Hume è che sia piuttosto l’idea di piacere a indurre il desiderio, Sugden vuole sostenere che manca una relazione strumentale o consequenziale che giustifichi - quindi “razionalizzi” - la scelta in base al piacere che ne deriverà. Quindi la relazione fra il desiderio e la scelta non è necessariamente una relazione razionale: il primo è la causa, ma non necessariamente anche la ragione in grado di giustificare la scelta. Le relazioni causali fra stati mentali, basate sull’associazione di idee, possono indurre varie forme di

³²⁸ *Ibid*

comportamenti non-razionali, come sono quelli che la behavioral economics si propone di spiegare.

Il secondo articolo è quello in cui Sugden esamina più a fondo la teoria di Hume, col proposito di mostrare che in Hume si trovano sia una spiegazione naturalistica dei processi mentali che sono alla base delle decisioni umane, sia una spiegazione teoretica della formazione di giudizi normativi riguardo alle decisioni; quest'ultima spiegazione è inoltre, per importanti aspetti, indipendente dalla prima. Quanto la questione della razionalità sia al centro di questa rilettura di Hume è evidente da un'interessante osservazione di Sugden a proposito di quegli interpreti – come John Broome e Christine Korsgaard³²⁹ – che, pur contestando che la teoria di Hume sia una teoria della razionalità strumentale, sono portati proprio per questo a considerarla poco credibile. Soprattutto se, come è il caso di Broome, presuppongono la validità di una qualche variante della teoria della scelta razionale.

They may entertain the possibility that there is *more* to rationality than the internal consistency of preferences, but they find it hard to make sense of the idea that there could be *less*³³⁰.

³²⁹ I riferimenti riguardano in particolare il saggio di J. Broome “Can a Humean be a moderate?”, in *Ethics out of economics*, Cambridge, Cambridge University Press, 1999, pp. 68-87 e quello di C. Korsgaard, “The normativity of instrumental reason”, in *Ethics and practical reason*, cit. Entrambi ritengono che nella teoria di Hume non c'è posto per una razionalità pratica.

³³⁰ R. Sugden, “Hume’s non-instrumental and non-propositional decision theory”, 2006, cit., p. 369. Corsivo nel testo.

L'idea che la razionalità non si riduca alla coerenza interna fra la preferenze “rivelate” nelle scelte, implica la possibilità di intendere la razionalità come fonte di giudizi di valore sulla bontà delle scelte, che avrebbero così una giustificazione. Alla base può esserci l'idea che un agente razionale desidera qualcosa in quanto crede che sia buono; si tratta quindi della possibilità di identificare i valori con i desideri³³¹. Avendo adottato la prospettiva naturalistica di Hume, questa possibilità non può essere presa in considerazione; occorre invece dimostrare che anche la razionalità minima richiesta, la semplice coerenza logica fra le preferenze, non è sostenibile. Nell'interpretazione soggettivistica della teoria della decisione, le proposizioni che esprimono le preferenze devono rispettare i principi della logica, come la transitività, anche se non riguardano la realtà oggettiva ma solo stati mentali soggettivi.

I shall say that a theory of mental states is propositional if it treats mental states as representing, expressing or involving propositions, and if the logical or conceptual consistency of such propositions can be used to test the rationality of the corresponding mental states³³².

Solo se è possibile “tradurre” gli stati mentali in proposizioni che ne esprimono il contenuto, la coerenza logica di queste proposizioni dimostrerà la razionalità degli stati mentali. Questo approccio è condiviso, secondo Sugden,

³³¹E' la tesi *desire as belief*, una tesi antihumeana secondo cui alcune credenze possono anche motivare. Contro questa tesi si pronuncia D. Lewis, “Desire as belief”, *Mind*, 1988, v. 97, n. 387, pp. 323-332 e “Desire as belief II”, *Mind*, 1996, v. 105, n. 418, pp. 303-313.

³³²R. Sugden, 2006, cit., p. 371. Corsivo nel testo.

dalla teoria della decisione razionale e dalla filosofia analitica, mentre, come intende dimostrare, Hume offre una teoria completamente opposta.

Il primo argomento è di carattere metodologico: la teoria della mente è presentata come una psicologia sperimentale, che ha l'obiettivo di descrivere i processi mentali, come i ragionamenti. Le operazioni di dimostrazione e di inferenza induttiva descritte in termini naturalistici equivalgono nel caso della dimostrazione alla capacità di formarsi un'immagine mentale – cosa è o non è concepibile –, nel caso dell'inferenza induttiva alla scoperta che la relazione causale non è che la percezione di una congiunzione costante, ma comunque contingente. In questa prospettiva la ragione umana non potrà mai essere considerata fonte di autorità.

Il secondo argomento riguarda il carattere dinamico attribuito alla mente nella teoria di Hume - “Hume is concerned not only with what it is for us to think something, but also with *when* we think it³³³” -, sia la teoria della scelta sia la filosofia analitica si occupano invece solo del contenuto della mente, concepito come uno stock di idee e percezioni. Per la teoria della scelta in particolare, le preferenze dell'agente sono considerate come un insieme di proposizioni sulla desiderabilità relativa di tutte le potenziali opzioni. Di fronte a un particolare problema, l'agente rintraccerà da questo insieme di preferenze quelle rilevanti per il caso in questione, e la sua razionalità sarà dimostrata dalla coerenza dell'insieme delle sue preferenze. Nella teoria di Hume, invece, i pensieri sono in uno stato di flusso costante, le operazioni

³³³ *Ivi*, p. 373. Corsivo nel testo.

mentali non consistono nella consultazione di un preesistente insieme di contenuti mentali, ma nella transizione da uno stato mentale all'altro sulla base del principio di associazione, i cui criteri sono quelli di somiglianza, contiguità, causalità. Proprio il carattere transitorio degli stati mentali sembra ostacolare la possibilità di convertire in proposizioni queste associazioni fra impressioni o idee.

L'ultimo argomento riguarda appunto il carattere pre-linguistico, e quindi non proposizionale, di emozioni e volizioni. Dato che gli stati mentali sono in continuo flusso e che la percezione di un oggetto dipende dalla posizione in cui ci troviamo rispetto ad esso, c'è bisogno del linguaggio per dare stabilità alle nostre percezioni; ma ciò non toglie che, nella prospettiva humeana, i sentimenti precedono sempre il linguaggio. Per ribadire con forza questa priorità, Sugden introduce l'esempio della passione indiretta dell'orgoglio descritta nel *Treatise* insieme all'umiltà come “simple and uniform impressions³³⁴”, mostrando l'inadeguatezza dell'interpretazione che ne dà Davidson³³⁵, quando sostiene che la passione dell'orgoglio consiste in realtà in un giudizio esprimibile nell'enunciato “A è orgoglioso che p”, dove p è una proposizione. Nella ricostruzione di Davidson l'enunciato combina il resoconto di un sentimento – l'orgoglio – con l'asserzione di una credenza che riguarda un oggetto o una qualità valutata positivamente, e che per qualche circostanza è in relazione con il soggetto che prova orgoglio. Secondo

³³⁴ D. Hume, *A treatise of human nature*, cit., pp. 277-80.

³³⁵ D. Davidson, “Hume's cognitive theory of pride”, in *Essays on actions and events*, cit., pp. 365-382.

Davidson, quella credenza espressa nella proposizione è sia la causa del sentimento di orgoglio sia la ragione per cui si ha quel sentimento. Sugden ha buon gioco a mostrare come in realtà per Hume l'orgoglio sia un'impressione semplice in cui non c'è nulla da analizzare, perciò solo un'esperienza diretta potrà farci capire di che si tratta. Nell'interpretazione di Sugden il sentimento è legato all'idea associata ad esso, “but the feeling does not imply a belief in the truth of any proposition about the world. It's just a feeling. As a matter of empirical psychology, the feeling is caused by the thought³³⁶”. Si tratta di un'associazione fra un pensiero e un sentimento che in un dato momento coesistono nella mia mente, un pensiero che non è una credenza stabile ma un'idea che mi è passata per la mente, che è causa del sentimento – e quindi dell'azione - ma non ne dà una giustificazione razionale.

La critica alla prospettiva cognitivista di Davidson offre a Sugden l'occasione per delineare con chiarezza la sua posizione. Se i sentimenti sono indipendenti dalle ragioni lo sono anche dai contenuti – per questo possiamo riconoscere la distinta qualità affettiva dell'orgoglio o della paura indipendentemente da ciò di cui siamo orgogliosi o di cui abbiamo paura.

If we think propositionally, it seems obvious to say that object are desired for properties that they possess. Is it credible to treat desire as a generic feeling that can be associated with many different ideas and which can be recognized as

³³⁶ R. Sugden, 2006, cit., p. 377.

a felt experience independently of what the desire is for? Perhaps surprisingly, some psychological and neurological evidence seems to suggest exactly this³³⁷.

In questo modo il divorzio fra cause e ragioni auspicato da Sugden sembra approdare più che al sentimentalismo di Hume a una concezione fortemente riduzionista, che finisce per mettere da parte non solo una spiegazione in termini di razionalità strumentale o proposizionale, ma una spiegazione *tout court*.

Una volta dimostrato che la teoria di Hume è con ogni evidenza estranea sia all'analisi proposizionale delle preferenze sia a una concezione strumentale della razionalità, la tesi humeana secondo cui, come si è detto, l'idea di piacere tende a indurre il desiderio che a sua volta mette in moto la volontà, può servire a gettare luce su quei comportamenti giudicati irrazionali dalla EUT.

Sugden prende in esame alcune situazioni di scelta in cui vengono proposti dei comportamenti che non possono essere rappresentati né come massimizzazioni del benessere, né come preferenze rivelate che soddisfano gli assiomi di coerenza della teoria della scelta razionale, ma risultano invece coerenti con l'approccio di Hume. La prima riguarda l'incoerenza temporale, ossia l'influenza della contiguità temporale sul desiderio, che rende le preferenze dipendenti dal contesto in modi che sarebbero considerati irrazionali da una "proposition-based theory"³³⁸. È il caso di un

³³⁷ *Ivi*, 381.

³³⁸ R. Sugden, "David Hume's *Treatise of Human Nature*", *Topoi*, 2008, 27, pp. 153-59, p. 155.

capovolgimento delle preferenze che dipende da un fattore temporale apparentemente irrilevante, e cioè dal momento in cui viene fatta la scelta. Nella teoria della scelta la differenza in termini di date non è una ragione accettabile per determinare una differenza in termini di preferenze. Se, nonostante la decisione di sottopormi a un'operazione presa molti mesi prima, a pochi giorni dalla data – in assenza di nuove informazioni – sono fortemente propensa a posporre la data, dal punto di vista della teoria della scelta sto rivelando un'incoerenza temporale nelle mie preferenze. Nella concezione humeana, invece, i desideri sono solo sentimenti, non traducibili in proposizioni, pertanto la forza con cui desidero qualcosa è causata dalla sua vicinanza, “but closeness is not the reason for the strenght of the desire, it is only the cause³³⁹”. Non si può quindi parlare né di errore, né di irrazionalità dei desideri, come sono invece inclini a fare i teorici della scelta razionale.

Un caso diverso è quello del conflitto fra opposti desideri che, in base all'assioma di completezza della EUT, può sempre essere risolto in modo da formare una “all-things-considered” preferenza. Hume analizza diverse tipologie di conflitti irrisolti fra passioni interpretandoli in base alla teoria dell'associazione di idee.

Upon the whole, contrary passions succeed each other alternately, when they arise from different objects: They mutually destroy each other, when they proceed from different parts of the same: And they subsist both of them, and

³³⁹ R. Sugden, 2006, cit., p. 384.

mingle together, when they are deriv'd from the contrary and incompatible chances or possibilities, on which any one object depends³⁴⁰.

Nella concezione dinamica della mente la temporanea esistenza di uno stato mentale può causare la temporanea esistenza di un altro, e la possibilità che si verifichi una situazione di equilibrio o di disequilibrio fra i due stati è solo una questione di carattere empirico. Questo significa che uno stato di indecisione fra due opposte passioni o desideri può rimanere irrisolto; ci si può trovare di fronte a un problema di scelta senza alcuna preferenza già definita o desiderio “all-things considered” rispetto alle opzioni fra cui bisogna scegliere. Anche in questo caso, il modello della scelta razionale sconta il fatto di non essere basato su ipotesi empiriche sul funzionamento della mente, ma su una concezione aprioristica della mente razionale intesa come un deposito di proposizioni consistenti, con la conseguenza di andare sistematicamente incontro al fallimento se utilizzato per prevedere i comportamenti reali.

Secondo Sugden, nel *Treatise*, accanto alla teoria che descrive i meccanismi psicologici che stanno dietro ai comportamenti di scelta e indipendentemente da essa, si trova anche una spiegazione teoretica della formazione di giudizi normativi riguardo alle scelte, e cioè una teoria che, non dando per scontata la validità dei concetti normativi, spiega i meccanismi che stanno dietro ai nostri giudizi su vizi e virtù. Ed è questa teoria della virtù a includere elementi di razionalità strumentale, spiegando, ad esempio, come un tratto di carattere venga considerato razionale se mostra di essere utile alla

³⁴⁰ D. Hume, *Treatise*, cit., p. 443.

stessa persona che lo ha o alla società in cui vive. E ancora, il fatto che gli oggetti possano essere più o meno preferibili in se stessi, indipendentemente dalle associazioni di idee che evocano in determinate circostanze, fa sì che a quanti si dimostrino capaci di perseguire ciò che è in se stesso preferibile, sia riconosciuta una virtù chiamata razionalità.

When we consider any object at a distance, all their minute distinctions vanish, and we always give the preference to whatever is in itself preferable, without considering its situation and circumstances. This give rise to what in an improper sense we call *reason*, which is a principle, that is often contradictory to those propensities that display themselves upon the approach of the object³⁴¹.

Il commento di Sugden – “this is not his theory of what people in fact desire, nor of how people in fact make decisions³⁴²” – serve a ribadire che, nel sistema teoretico di Hume, i giudizi sulla razionalità delle azioni sono distinti e indipendenti dai principi psicologici che di fatto governano le decisioni ed è proprio questa la prospettiva che può rivelarsi più promettente per la teoria economica.

For the last 50 years, it must be said, most explanatory work in economics has assumed that agents act instrumentally with respect to preferences that satisfy standard conditions of consistency. [...] Still, we may yet find that the best explanations of decision-making are, as Hume’s theory is, based on psychological assumptions which make no reference to rationality. Current

³⁴¹ D. Hume, *Treatise*, cit., p. 536.

³⁴² R. Sugden, 2006, cit., p. 389.

work in behavioural economics is premised on the credibility of the latter possibility³⁴³.

Su queste basi Sugden contesta più in generale la validità dell'approccio volto a giustificare dei principi normativi facendo appello all'autorità della ragione. Una critica che investe anche chi come Sen respinge la tesi che il benessere possa essere misurato in base al soddisfacimento delle preferenze e propone in alternativa l'idea di uno "spazio valutativo" delle opportunità (*capability approach*). In questo caso il criterio di giudizio è l'assunto che gli individui hanno *ragione* di attribuire valore a queste capacità, anche se di fatto non le desiderano. Giudicando infondata la pretesa della ragione di porsi come autorità morale, Sugden ritiene di perseguire la prospettiva humeana affermando che le regole morali sono convenzioni del linguaggio che usiamo per descrivere i nostri sentimenti di approvazione e di biasimo, cosicché possono esserci azioni che sono approvate in una società ma biasimate in un'altra. In tutte le società la morale tenderebbe ad avere un carattere tendenzialmente utilitaristico, il che spiega il riproporsi, ma solo nel contesto della normatività etica, della razionalità strumentale.

4.6 Hume. Il ruolo delle passioni calme

La rigidità della contrapposizione che Sugden individua nella teoria di Hume fra il piano della psicologia empirica, impegnata a dare spiegazioni sul funzionamento dei meccanismi di scelta, e il piano dell'etica normativa,

³⁴³ *Ivi*, p. 390.

incentrata sul convenzionalismo di una razionalità ora sì proposizionale e strumentale, mi sembra poco convincente.

Distinguere nettamente tra l'indagine sperimentale della natura umana e l'etica non ha molto senso all'interno di una prospettiva humeana, poiché gli oggetti esaminati – gli esseri umani – sono caratterizzati, fin dall'inizio, in termini di ragionevolezza, socialità e attività³⁴⁴.

I due piani non sono affatto contrapposti nella teoria di Hume e soprattutto l'alternativa non è, come pensa Sugden, tra una razionalità strumentale³⁴⁵ o nessuna razionalità. L'ipotesi interpretativa di Blackburn di una forma di ragionevolezza che non si ponga come fonte di autorità, ma come conquista di “a steady and general point of view³⁴⁶” – la “Hume-friendly reason” – permette di recuperare una parte importante delle riflessioni di Hume, modificando in modo sostanziale l'interpretazione che ne ha dato Sugden.

Anche l'ipotesi sulla natura convenzionale della ragione in Hume può essere ridimensionata, se quella che, erroneamente secondo Hume³⁴⁷, viene

³⁴⁴ L. Greco, *L'io morale. David Hume e l'etica contemporanea*, cit., p. 131.

³⁴⁵ Della presenza in Hume di una ragione non strumentale è convinto Magri: “Reason motivates us to impartial morality as a practical principle that is distinct from instrumental rationality in that it can assess the value of final object of choices and is capable of opposing passions and desires”. In T. Magri, “Natural obligation and normative motivation in Hume’s Treatise”, *Hume Studies*, 1996, v. XXII, n.2, pp. 231-253, p.241.

³⁴⁶ D. Hume, *Treatise*, cit., p. 581-2.

³⁴⁷ Ivi, p. 417.

confusa con la ragione è in realtà una “passione calma”, che non si contrappone alle altre passioni ma si limita invece a correggerle. Caratteristica delle passioni calme - fra cui vanno inclusi i sentimenti morali con cui esprimiamo approvazione per ciò che è moralmente buono e disapprovazione per ciò che è moralmente cattivo - è di saper esercitare un influsso persistente e duraturo sulla volontà³⁴⁸.

Men often act knowingly against their interest: For which reason the view of the greatest possible good does not influence them. Men often counter-act a violent passion in prosecution of their interests and designs: Tis not therefore the present uneasiness alone, which determine them. In general we may observe, that both these principles operate on the will; and where they are contrary, the either of them prevails, according to the *general* character or *present* disposition of the person. What we call strength of mind, implies the prevalence of the calm passions above the violent; tho' we may easily observe, there is no man so constantly possess'd of this virtue, as never on any occasion to yield to the solicitations of passion and desire³⁴⁹.

Per Hume, la ragione – “what in an improper sense we call *reason*³⁵⁰” – si identifica con la capacità di scegliere dopo ponderata riflessione, in modo da

³⁴⁸ L. Greco, *L'io morale. David Hume e l'etica contemporanea*, cit., p. 137.

³⁴⁹ Ivi, p. 418. Corsivi nel testo.

³⁵⁰ La resistenza di Hume a utilizzare il termine ragione si può forse spiegare con la sua ostilità nei confronti di quanti vedono nella ragione la facoltà distintiva degli esseri umani e aspirano a una morale oggettiva basata su verità assolute che la ragione ci permetterebbe di conoscere.

affrancarsi dai limiti cognitivi come la scarsa lungimiranza, o riuscire a superare la difficoltà di orientarsi nelle situazioni difficili. Prestando ascolto alla voce della prudenza, possiamo sforzarci di correggere la forza che esercita sull'immaginazione l'idea di un piacere imminente.

This natural infirmity I may very much regret, and I may endeavour, by all possible means, to free myself from it. I may have recourse to study and reflexion within myself; to the advice of friends; to frequent meditation, and repeated resolution: And having experience'd how ineffectual all these are, I may embrace with pleasure any other expedient, by which I may impose a restraint upon myself, and guard against this weakness³⁵¹.

Quindi, il punto cruciale della teoria di Hume è che la ragione non va considerata come una facoltà autonoma, dal momento che si presenta sempre sotto un aspetto passionale: “a general calm determination of the passions, founded on some distant view or reflection³⁵²”.

³⁵¹ Ivi, p. 536-7.

³⁵² Ivi, p. 583.

Conclusioni in forma di sommario

Il focus del lavoro è stato quello di esaminare il modo in cui discipline diverse che utilizzano metodologie diverse – teoria economica, psicologia e filosofia – affrontano il tema della scelta, e la possibilità di una teoria della scelta razionale. Le implicazioni di carattere psicologico e filosofico della concezione dell'agente razionale su cui si basa la EUT sono risultate evidenti fin dalla presentazione della teoria dell'utilità attesa. In particolare, dall'esame delle sue caratteristiche teoriche, e dai concetti chiave su cui è basata: utilità, preferenza, probabilità. Per capire il significato che rivestivano per la teoria gli assiomi di razionalità, si è rivelato poi determinante il confronto fra i diversi modi di interpretare sia questi concetti sia la sua metodologia e la teoria stessa, di volta in volta considerata prevalentemente descrittiva, predittiva o normativa. Come ho sottolineato, gli intenti che la teoria dell'utilità attesa si proponeva erano diversi e difficilmente compatibili. Puntare sulla forza predittiva di una teoria assiomatica comporta necessariamente che si tralascino quei comportamenti che meno facilmente potevano essere rimodellati sulla base delle sue esigenze. La violazione degli assiomi ha permesso di esaminare più a fondo i suoi fondamenti e di vagliare la solidità delle diverse interpretazioni della teoria sulla base della capacità di trovare risposte attraverso proposte di riforma della EUT, ma anche elaborando interpretazioni alternative per spiegare i fallimenti mostrati dai paradossi e dagli esperimenti. L'impossibile compromesso fra interpretazioni normative e descrittive ha portato alla divaricazione fra i fautori della teoria assiomatica – che pure ammettono la sua limitata applicabilità – e i fautori di un'alternativa

rappresentata dalla behavioural economics e dalla neuroeconomics, di cui sono stati però sottolineati alcuni limiti metodologici.

Dalla prospettiva della psicologia cognitiva, la scarsa capacità di descrivere i comportamenti reali mostrata dalla EUT rendeva necessaria una sua rielaborazione, che comporta fra l'altro la reintroduzione delle componenti psicologiche della scelta rimosse dal modello standard. L'attenzione ai meccanismi cognitivi sottesi alla scelta e all'influenza esercitata dai fattori emotivi, oltre a permettere di dare conto dei comportamenti devianti, è stata la base per un approccio diverso al problema della scelta: approccio che utilizza una metodologia descrittiva e un apparato concettuale in grado di dar ragione delle distorsioni percettive e valutative in cui incorrono le stime sia dell'utilità sia della probabilità. Il che però, come sia è visto, non comporta necessariamente il rifiuto del modello EUT, che conserva un ruolo di modello normativo di riferimento e consente di misurare il grado di scostamento di molti dei nostri comportamenti, indotto anche dalle nostre limitate capacità computazionali, che ci spingono a utilizzare ogni sorta di scorciatoie mentali per ridurre i costi computazionali ma anche emozionali delle scelte più impegnative.

La tesi nell'ambito della psicologia cognitiva che ci è sembrata più convincente a questo riguardo è quella che non esclude in linea di principio la possibilità di governare più efficacemente le situazioni, interrompendo l'attivazione automatica di procedure mentali, grazie a una riflessione più attenta e ragionata. In quest'ottica, è di sicuro interesse l'ipotesi di una razionalità procedurale a due velocità nonché l'individuazione delle diverse forme di interazione fra meccanismi cognitivi e affettivi messe in luce fra gli

altri da Colin Camerer. Sollevano invece molte perplessità l'eccessivo entusiasmo per l'uso tecniche di neuroimmagine all'interno delle neuroscienze. A tutt'oggi il contributo che riescono a dare sembra soffrire di alcuni limiti metodologici, indotti anche dalla sopravvalutazione dei risultati reali o potenziali offerti da queste "mappe statistiche" del funzionamento del cervello. In primo luogo, si intravede il rischio di un'impostazione rigidamente individualista che, tenendo poco conto delle componenti sociali della scelta, rischia di lasciarsi tentare da interpretazioni riduzioniste che promettono di fare a meno di ogni sorta di razionalità pratica comunque declinata. In secondo luogo, si avverte la mancanza di una riflessione epistemologica sul corretto uso di queste nuove tecniche, che sottolinei come una buona caratterizzazione dei fenomeni mentali sia sempre una premessa indispensabile per una loro corretta categorizzazione.

In questo lavoro, la disamina dalla prospettiva filosofica ha avuto anzitutto il compito di chiarire i presupposti concettuali irrinunciabili di una teoria della scelta. Così, nell'analisi di Davidson, viene riconosciuto il ruolo imprescindibile che spetta alla motivazione, rappresentata dall'interazione fra *belief* e *desire* e che permette di rendere intellegibile, e dunque razionale, un'azione. La distinzione fra una descrizione dell'azione in termini di cause o di ragioni serve in primo luogo a chiarire che solo nel secondo caso si dà una spiegazione che è anche una giustificazione dell'azione e, in secondo luogo, a chiudere la strada al riduzionismo fisicalista. Confrontandosi nel suo caso esplicitamente con il fallimento della EUT, Davidson comunque difende una razionalità normativa che si identifica con un certo grado di coerenza logica indispensabile per descrivere un'azione, e che rappresenta un requisito

indispensabile per comprendere un'azione. In definitiva, il criterio per giudicare la razionalità delle motivazioni è dato dalla sua traducibilità in pensieri proposizionali, anche se è un criterio che finisce per penalizzare gli atteggiamenti emotivi che risultano difficilmente traducibili in proposizioni coerenti e ben strutturate.

La divergenza di vedute fra Derek Parfit e Bernard Williams su “ragioni esterne” e ragioni interne”, ha messo in luce molti aspetti importanti legati alla contrapposizione fra un approccio normativo e uno naturalistico sulla razionalità. A mio avviso, gli argomenti utilizzati da Parfit in difesa delle ragioni esterne rivelano la preoccupazione, solo in parte condivisibile, di garantire autorevolezza e universalità alle ragioni. Mentre non soddisfa affatto la soluzione di rendere le ragioni irriducibili e indipendenti dalle instabili motivazioni psicologiche, concependole come “value based”. Sulla stessa linea è anche l'economista John Broome, che mette a confronto due modi di declinare la razionalità, esprimendo la sua preferenza per i più vincolanti (*strict*) “requisiti normativi” rispetto alle vaghezza (*slack*) delle “ragioni per”, e rivelando la volontà di rendere i meccanismi decisionali conformi a quelli della logica. A garanzia della normatività, l'alternativa alla logica è anche per lui il ricorso a valori assoluti. Infatti, il punto più interessante delle argomentazioni con cui cerca di dimostrare la natura intrinsecamente normativa del ragionamento strumentale – *means-ends reasoning* – è quello in cui ammette che ciò richiederebbe che la teoria della decisione razionale avesse il sostegno della teleologia, in quanto si dovrebbe presumere che lo scopo dell'agente sia il bene, inteso come valore assoluto, invece che il conseguimento di un obiettivo particolare.

Mentre la normatività poggia sulle regole vincolanti della logica o sul riferimento a valori che, per ammissione dello stesso Parfit, non riescono a nascondere una misteriosa origine metafisica, una prospettiva opposta è invece quella su cui si muovono Bernard Williams e Simon Blackburn. Nel confronto con Parfit, la concezione di Williams, sostenendo che le ragioni consistono in un valido percorso deliberativo (*sound deliberate route*) a partire dall'“agent's subjective motivational set” restituisce importanza al punto di vista dell'agente. La perdita di universalità implicita in questa ottica mi sembra sicuramente compensata dall'incremento di comprensione verso comportamenti etichettati come “irrazionali”. Una risposta dal versante filosofico alla concezione delle preferenze come coerenti e stabili propria della EUT è quella rappresentata dalla tesi di Williams sul carattere costitutivamente non strutturato dell'insieme motivazionale, che vede nell'indeterminatezza delle maggior parte delle nostre propensioni un ostacolo alla possibilità di rappresentarle sotto forma di un ordinamento di preferenze.

Il richiamo a Hume che lega Williams e Blackburn è poi evidente nella comune ostilità verso l'idea che la razionalità strumentale sia l'unica forma di razionalità. Nel caso di Williams, si tratta di sottolineare come il percorso deliberativo sia un procedimento razionale comprensivo, ad esempio, della capacità di immaginare alternative e di riflettere in maniera ponderata (“a modest amount of prudence”) e che è anche in grado di tener conto delle questioni di valore. Tutte caratteristiche che lo allontanano decisamente dalla razionalità strumentale. Di derivazione humeana sono i termini chiave del discorso di Blackburn - *passions* e *concern* - che ci segnalano gli aspetti salienti del contesto in cui ci troviamo a compiere scelte o a prendere decisioni

e ci spingono così a servirci della ragione – di ogni tipo di ragionamento - per esaminarli più a fondo.

La descrizione che dà Blackburn della presa di decisione è del resto quanto mai lontana dalla stabilità e coerenza richiesti all'agente razionale:

You, when you deliberate, are whatever you are: a person of tangled desires, conflicting attitudes to your parents, inchoate ambitions, preferences, and ideals, with an inherited ragbag of attitudes to different actions, situations, and characters. You do not manage, ever, to stand apart from all that ³⁵³.

La figura dell'agente perfettamente informato, in grado di calcolare in modo intelligente costi e benefici, o che ha sotto controllo tutti i dettagli della sua situazione, rappresenta un ideale che sicuramente non può essere considerato vincolante. È un ideale che non andrà assolutizzato, per cui al massimo sarà espressione di una concezione di vita che dà valore più alla prudenza che alla spontaneità.

Come, a mio avviso, ha ben argomentato Blackburn, non si tratta di rinunciare alla razionalità *tout court*, ma solo di mettere da parte l'idea che la razionalità possa essere autonoma dalle passioni e avere un ruolo di legislatore universale di norme di condotta. Saper distinguere fra il piano valutativo-normativo e quello descrittivo, senza negare la validità di entrambi, significa ad esempio biasimare un comportamento sconsiderato, evitando di definirlo irrazionale. Con l'accusa di irrazionalità non facciamo in realtà che ammantare

³⁵³ *Ivi*, p.252.

di una falsa oggettività la nostra convinzione sulla motivazione che un agente dovrebbe avere.

Un importante punto a favore della prospettiva naturalista di Blackburn – ancora una volta condivisa dallo psicologismo di Williams – è di riuscire ad evitare il rischio del riduzionismo individualista in cui rischiano di incorrere le neuroscienze, rammentandoci che la deliberazione riguarda un nostro impegno attivo con il mondo, non un qualche procedimento introspettivo della nostra coscienza. Del resto, comprendere le ragioni di una scelta significa anche saperla inserire in un contesto più ampio - che non è dato solo dai motivi che ci hanno spinti o dalle conseguenze cui aspiravamo - ma anche dall'insieme di aspettative, regole di comportamento e norme sociali che costituiscono lo sfondo delle nostre pratiche condivise.

In definitiva, la prospettiva della teoria economica ha il sicuro merito di rimarcare che non si può fare a meno di avere dei criteri di razionalità che garantiscano la coerenza dei nostri comportamenti perché altrimenti - come anche la filosofia di Davidson mette bene in luce – tali comportamenti rischierebbero di diventare incomprensibili. D'altra parte, lo sforzo interpretativo dei filosofi che condividono un'ottica naturalista restituisce alla razionalità una preziosa elasticità mentale - confermata dagli studi sul funzionamento del cervello - e dal momento che il *sense making* è, secondo gli psicologi, una delle forme di interazione fra i processi affettivi e cognitivi, trova conferma anche l'idea di Hume di una ragione che non è altro che una "passione calma".

Bibliografia

Allais, M. and O. Hagen, editors, *Expected Utility Hypotheses and the Allais Paradox*, Dordrecht, Reidel, 1979.

Allingham, M., *Choice Theory: a very short introduction*, Oxford, Oxford University Press, 2002.

Allison, H., *Kant's Theory of Freedom*, Cambridge, Cambridge University Press, 1990.

Anand, P., *Foundations of rational choice under risk*, Oxford, Clarendon Press, 1993.

Bagnoli, C., in "Blackburn on the normative question", *Iride*, 30, 2000, pp. 8-14.

Benini, A., "Intervista", *Humana.Mente*, v. 9, 2009, pp. 159-64.

Benini, A., *Che cosa sono io. Il cervello alla ricerca di sé stesso*, Milano, Garzanti, 2009.

Bermudez, J. and A. Miller, ed, *Reason and Nature: Essays in the Theory of Rationality*, Oxford, Oxford University Press, 2002.

Bianchi, C., *La pragmatica del linguaggio*, Roma-Bari, Laterza, 2003.

Binmore, K., *Natural Justice*, Oxford, Oxford University Press, 2005.

Binmore, K., *Game theory: a very short introduction*, Oxford, Oxford University Press, 2007.

Blackburn, S., "Précis of Ruling passions", in *Philosophy and Phenomenological Research*, LXV, 1, 2002.

Blackburn, S., *Ruling Passions: a theory of practical reasoning*, Oxford, Clarendon Press, 1998.

Blackburn, S., *Hume*, London, Granta, 2008.

Bonner, J., *Introduction to the theory of social choice*, Baltimore, The John Hopkins University Press, 1986.

Bratman, M. E., *Intentions, Plans, and Practical Reason*, Cambridge MA, Harvard University Press, 1987.

- Bricke, J., *Mind and morality: an examination of Hume's moral psychology*, Oxford, Clarendon Press, 1996.
- Broome, J., "Can a Humean be a moderate?", in *Ethics out of economics*, Cambridge, Cambridge University Press, 1999, pp. 68-87.
- Broome, J., "Are intentions reasons?" In *Practical Rationality and Preferences: Essays for David Gauthier*, C Morris, A. Ripstein ed., Cambridge, Cambridge University Press, 2001, p. 98-120.
- Broome, J., "Have we reason to do as rationality requires? A Comment on Raz, *Journal of Ethics & Social Philosophy*, Symposium 1, 2005.
- Broome, J., "Is Rationality normative?", *Disputatio*, 2007, pp. 161-78.
- Broome, J., "Normative practical reasoning", *Proceeding of the Aristotelian Society*, 2001, p. 175-93.
- Broome, J., "Practical reasoning", in *Reason and Nature: Essays in the Theory of Rationality*, J. Bermudez and A. Miller, ed., Oxford, Oxford University Press, 2002.
- Broome, J., "Reason and Motivation", *The Aristotelian Society*, 77 (Supplementary Volume), p. 131-147.
- Broome, J., "Reasons", in *Reason and Value: Essays on Moral Philosophy of Joseph Raz*, Oxford, Oxford University Press 2004, pp. 28-55.
- Broome, J., "Unity of reasoning?" In *Disputatio*, 11, 2008.
- Broome, J., *Weighing Goods*, Oxford, Blackwell, 1991.
- Broome, J., "Normative requirement", in *Ratio*, 1999, pp. 398-419.
- Bruni, L. R. Sugden, "The road not taken: how psychology was removed from economics and how it might be brought back", *The Economic Journal*, 2007, 117, pp. 146-173.
- Camerer, C., "Behavioral economics: reunifying psychology and economics", *Proceeding of the National Academy of Sciences of the United States of America*, 1999, v.96, n.19, pp. 10575-10577, p. 10575.
- Camerer, C., G. Loewenstein and D. Prelec, "Neuroeconomics: how neurosciences can inform economics", *Journal of Economic Literature*, 2005, 9-64.
- Camerer, C., G. Loewenstein, D. Prelec, "Neuroeconomics: why economics needs brains", *Scandinavian Journal of Economics*, 2004, 106 (3), pp. 555-79.

- Cherniak, C., *Minimal rationality*, Cambridge, Mass, The MIT Press, 1986.
- Chew, S. and K.R. MacCrimmon, "Alpha Utility Theory, lottery composition, and the Allais paradox", *Working paper*, University of British Columbia, Vancouver, 1979.
- Cullity, G. and B. Gaut (eds), *Ethics and Practical Reason*, Oxford, Oxford University Press, 1997.
- Damasio, A.R., *Descartes' error: emotion, reason, and the human brain*, New York, Putnam, 1994.
- Davidson, D., *Actions, reasons and causes*, 1963, in *Essays on actions and event*, Oxford, Clarendon Press, 1980, pp. 3-19.
- Davidson, D., *How is Weakness of Will Possible*, 1969, in *Essays on actions and events*, cit., pp. 21-42.
- Davidson, D., *Freedom to act*, 1973, in *Essays on action and events*, cit., pp. 63-82.
- Davidson, D., *Psychology as Philosophy*, 1974, in *Essays on actions and events*, cit., pp. 229-244.
- Davidson, D., *Hume's cognitive theory of pride*, 1976, in *Essays on actions and events*, cit., pp. 277-290.
- Davidson, D., *Intending*, 1978, in *Essays on actions and events*, cit., pp. 83-102.
- Davidson, D., *Essays on action and events*, New York, Oxford University Press, 1980.
- Davidson, D., *Paradoxes of Irrationality*, 1982, in *Problems of irrationality*, cit., pp. 169-187.
- Davidson, D., "A new basis for decision theory", in *Theory and Decision*, 18, 1985.
- Davidson, D., *Incoherence and Irrationality*, 1985, in *Problems of irrationality*, Oxford, Clarendon Press, 2004, pp. 189-198.
- Davidson, D., Suppes, P. e Siegel, S., *Decision-making: An experimental approach*, Stanford, Stanford University Press, 1957.
- De Caro, M., *Dal punto di vista dell'interprete. La filosofia di Davidson*, Roma, Carocci, 1998.
- De Caro, M., *Azione*, Bologna, Il Mulino, 2008.

- Dennett, D., *Freedom evolves*, London, Allen Lane, 2003.
- Donatelli, P., *La filosofia morale*, Bari-Roma, Laterza, 2001.
- Egidi, M., “The dual process account of reasoning: historical roots, problems, perspectives”, *CEEL Working paper*, n. 6, 2007.
- Egidi, M., L. Marengo, “Cognition, institutions, near decomposability: rethinking Herbert Simon's contribution”, *CEEL, Working Paper*, 2002, n.6, pp. 1-13.
- Ellsberg, D., “Risk, ambiguity and the Savage's axioms”, *The Quarterly Journal of Economics*, 1961, v.75, n.4, pp. 643-669.
- Elster, J., *Ulysses and the Sirens*, Cambridge Mass., Cambridge University Press, 1977.
- Elster, J., *Sour Grapes: Studies in the Subversion of Rationality*, Cambridge Mass., Cambridge University Press, 1983.
- Elster, J., ed., *The multiple self*, Cambridge Mass., Cambridge University Press, 1986.
- Elster, J., “Emotions and Economic Theory”, *Journal of Economic Literature*, 1998.
- Farina, F., Hahn, F., Vannucci, S., *Ethics, rationality and economic behavior*, Oxford, Clarendon Press, 1966.
- Frank, R. H., *Passions within Reasons*, New York, Norton, 1988.
- Frank, R. H., *Microeconomia*, Milano, McGraw-Hill Italia, 1998.
- Frank, R. H., *What Price the Moral High Ground?*, Princeton, Princeton University Press, 2004.
- Fodor, J., *The modularity of mind*, Cambridge (MA), MIT Press, 1983.
- Friedman, M. e L. Savage, “The utility analysis of choices involving risk”, *The Journal of political economy*, 1948, 56, pp. 279-304.
- Frixione, M., *Come ragionamo*, Roma-Bari, Laterza, 2007.
- Gilovich, T., *How we know what isn't so: the fallibility of human reason in everyday life*, New York, Free Press, 1991.
- Gilovich, T., D. Griffin, D. Kahneman (eds.), *Heuristics and Biases: the psychology of intuitive judgement*, New York, Cambridge University Press, 2002.
- Giroto, M., *Il ragionamento*, Bologna, Il Mulino, 1994.

Glimcher, P. W. et al., *Neuroeconomics. Decision making and the brain*, London, Elsevier, 2009.

Greco, L., “Bernard Williams e la natura delle ragioni in etica” in *Ethics and Politics*, 2005, 1-15.

Greco, L., *L'io morale. David Hume e l'etica contemporanea*, Napoli, Liguori Editore, 2008.

Grether, D. e C. Plott, “Economic theory of choice and the preference reversal phenomenon”, *The American Economic Review*, 1979, pp. 623-638.

Gul, F., “A theory of disappointment aversion”, *Econometrica*, 1991, v. 59, n. 3, pp. 667-686.

Haggard, P., S. Clark e J. Kalogeras, “Voluntary action and conscious awareness”, *Nature Neurosciences*, 2002, v. 5, n. 4, pp. 382-385.

Hargreaves Heap, S. et al., *The Theory of choice. A critical guide*, Oxford, Blackwell, 1992.

Hausman, D. M., McPherson M. S., *Economic Analysis, Moral Philosophy, and Public Policy*, Cambridge, Cambridge University Press, 1996.

Hogarth, M., e Reder M. W., ed., *Rational Choice. The Contrast between Economics and Psychology*, Chicago, The University of Chicago Press, 1986.

Hollis, M. e R. Sugden, “Rationality in action”, *Mind*, 1993, v. 102, n. 405, pp. 1-35.

Holt, C. A., “Preference reversals and the independence axiom”, *American Economic Review*, 1986, 76, pp. 508-15.

Hume, D., *A treatise of human nature*, 1740, II ed. By P. H. Nidditch, Oxford, Clarendon Press, 1988.

Jenkins, M. P., *Bernard Williams*, Montreal, McGill-Queen's, 2006.

Jevons, W.S., *The theory of political economy*, 1871, rist. Harmondsworth, Penguin Books, 1970.

Kahneman, D., “New challenges to the rationality assumption” in *Journal of Institutional and Theoretical Economics*, 2000, pp. 18-36. Poi in *Choices, Values and Frames*, cit. pp. 758-774.

Kahneman, D., *Maps of bounded rationality: a perspective on intuitive judgement and choice*. Prize lecture, December 8, 2002.

(http://nobelprize.org/nobel_prizes/economics/laureates/2002/kahnemann-lecture.pdf).

Kahneman, D., A. Tversky. “Judgement under Uncertainty: heuristics and biases”, in *Science*, 185, n. 415 (1974), p. 1124-1131.

Kahneman, D., A. Tversky, “Prospect Theory. An analysis of decision under risk”, *Econometrica*, 1979 p. 263.

Kahneman, D., A. Tversky (eds.), *Choices, Values, and Frames*, New York, Cambridge University, 2000.

Karni, E., Z. Safra, ‘Preference reversals’ and the observability of preferences by experimental methods, *Econometrica*, 1987, 55, 675–85.

Kahneman, D., P. Slovic e A. Tversky (eds.), *Judgement under Uncertainty: heuristics and biases*, Cambridge (Ma.), Cambridge University Press, 1982.

Kahneman D., P. Wakker, R. Sarin, “Back to Bentham. Exploration of experienced utility”, *The Quarterly Journal of Economics*, v. 112, n. 2, 1997, pp. 375-405.

Keynes, J. N., *The scope and method of political economy*, London, Macmillan, [1890], 1917.

Korsgaard, C., *The Sources of Normativity*, Cambridge Mass., Cambridge University Press, 1996.

Korsgaard, C. M., “The Normativity of Instrumental Reason”, in G. Cullity and B. Gaut (eds), *Ethics and Practical Reason*, oxford, Oxford University Press, 1997, p. 215-54.

Korsgaard, C.M., “Skepticism about practical Reason”, in *Varieties of Practical Reasoning*,cit.

Legrenzi, P., *Prima lezione di scienze cognitive*, Roma-Bari, Laterza, 2005.

Lewin, S.B., “Economics and psychology: lessons for our own day from the early twentieth century”, *Journal of Economic Literature*, 1996, v. XXXIV, pp. 1293-1323.

Lewis, D., “Desire as belief”, *Mind*, 1988, v. 97, n. 387, pp. 323-332.

Lewis, D., “Desire as belief II”, *Mind*, 1996, v. 105, n. 418, pp. 303-313.

Libet, B., “Unconscious cerebral initiative and the role of conscious will in voluntary action”, *Behavior and brain sciences*, 1985, 2, pp. 146-159.

- Lichtenstein, S. and P. Slovic, "Reversals of preferences between bids and choices in gambling decisions", *Journal of Experimental Psychology*, 1971, vol. 89, pp. 46-55.
- Lichtenstein, S. and P. Slovic, *The construction of preferences*, Cambridge, Cambridge University Press, 2006.
- Lindman, H., "Inconsistent preferences among gambles", *Journal of Experimental Psychology*, 1971, vol. 89, pp. 390-7.
- Loewenstein, G., "Out of control: visceral influences on behavior", *Organizational Behavior and human decision processes*, 1996, 65, pp. 272-292.
- Loewenstein, G., "Emotions in economic theory and economic behavior", *The American Economic Review*, 2000, v. 90, n. 2, pp. 426-432.
- Loewenstein, G., ed., *Exotic preferences. Behavioral Economics and Human Motivation*, Oxford, Oxford University Press, 2007.
- Loomes, G. e R. Sugden, "Regret Theory: an alternative theory of rational choice under uncertainty", *The Economic Journal*, 92, 1982, pp. 805-824.
- Loomes, G. e Sugden R., "A rationale for preference reversal", *American Economic Review*, 1983, v. 73, n. 3, pp.428-432.
- Loomes, G. e R. Sugden, "Disappointment and dynamic consistency in choice under uncertainty", *Review of Economic Studies*, 1986, LIII, pp. 271-282.
- Loomes, G. e R. Sugden, "Testing for regret and disappointment in choice under certainty", *The Economic Journal*, 1987, 87, pp. 118-129.
- Loomes, G., Starmer, R. Sugden, "Preference reversal: information processing effect or rational non-transitive choice", *The Economic Journal*, 1989, 99, pp. 140-51.
- Luchins, A.S. and E.H. Luchins, "New experimental attempts in preventing mechanization in problem-solving", *The Journal of General Psychology*, 1950, 42, pp. 279-291.
- Machina, M., "Choice under uncertainty: problems solved and unsolved", *Journal of Economic Perspectives*, 1987.
- Machina, M., "Expected utility"analysis without the independence axiom", *Econometrica*, 1982, 50, pp. 277-323.
- Magri, T., "Natural obligation and normative motivation in Hume's Treatise", *Hume Studies*, 1996, v. XXII, n.2, pp. 231-253.

- McDowell, J., "Might there be External Reasons?" In J. E. J. Altham & R. Harrison (eds) *World, Mind, and Ethics*, Cambridge, Cambridge University Press, 1995, p. 68-85.
- McFadden, D. , Smith V. L., Kahneman D., *Critica della ragione economica*, a cura di M. Motterlini e M. Piattelli Palmarini, Milano, Il Saggiatore, 2005.
- Mele, A. R. , Rawling P. eds. *The Oxford Handbook of Rationality*, New York, Oxford University Press, 2004.
- Millgram, E. H. ed., *Varieties of practical reasoning*, Cambridge Mass., MIT Press, 2001.
- Motterlini, M., "Un dilemma per il consenso informato", *Kos*, luglio 2006.
- Nagel, T., *The possibility of altruism*, Princeton, Princeton University Press, 1978.
- Nagel, T., *The View from Nowhere*, New York, Oxford University Press, 1986.
- Nozick, R., *Philosophical Explanations*, Cambridge Mass., Harvard University Press, 1981.
- Nozick, R., *The Nature of Rationality*, Princeton, Princeton University Press, 1993.
- Novarese, M., Rizzello S., *Economia sperimentale*, Milano, Bruno Mondatori, 2004.
- Pareto, V., *Manuale di economia politica*, Milano, Società editrice libraria, 1906, rist. Roma, Bizzarri, 1965.
- Parfit, D., "Reasons and motivation", in *The Aristotelian Society*, vol. 77 (Supplementary Volume), 1997, p. 99-130.
- Parfit, D., *Reasons and Persons*, Oxford, Clarendon Press, 1984.
- Piattelli Palmarini, M., *Psicologia ed economia della scelta*, Torino, Codice edizioni, 2005.
- Putnam, H., *The Collapse of the Fact/Value Dichotomy and Other Essays*, Cambridge Mass., Harvard University Press, 2004.
- Quiggin, J., "A theory of anticipated utility", *Journal of Economic Behavior and Organization*, 1982, 3(4), pp. 323-43.
- Raz, J., "The Myth of Instrumental Rationality", *Journal of Ethics & Social Philosophy*, n. 1, 2005.

- Resnik, M. D., *Choices: an introduction to Decision Theory*, Minneapolis, University of Minnesota Press, 1986.
- Rizzello, S., *L'economia della mente*, Bari-Roma, Laterza, 1997.
- Roncaglia, A., *La ricchezza delle idee. Storia del pensiero economico*, Bari-Roma, Laterza, 2001.
- Samuelson, P., "A note of pure theory of consumer's behaviour", in *Economica*, 1938.
- Savage, L., *The Foundations of Statistics*, New York, Wiley, 1954.
- Schneider, W. e R. Shiffrin, "Controlled and automatic human information processing: I. Detection, search, and attention", *Psychological Review*, 1977, 84, pp. 127-190.
- Searle, J., *Rationality in Action*, Cambridge, Mass, The MIT Press, 2001.
- Segal, U., "Does the preference reversal phenomenon necessarily contradict the independence axiom?", *American Economic Review*, 1988, 78, 233–6.
- Seidl, C., "Preference reversal", *Journal of Economic Surveys*, 2002, pp.621-655.
- Sen, A. K., "Behaviour and the concept of preference", *Economica*, 1973, pp. 241-259.
- Sen, A. K., "Rational fools", 1977, in *Philosophy and Public Affairs*, 6, pp.317-44.
- Sen, A. K., *Choice, Welfare and Measurement*, Oxford, Basil Blackwell 1982.
- Sen, A. K., *On Ethics and Economics*, Oxford, Basil Blackwell, 1987.
- Sen, A. K., *Rationality and Freedom*, Cambridge, Mass., Harvard University Press, 2002.
- Sen, A. K., *The Idea of Justice*, London, Allen Lane, 2009.
- Simon, H.A., "How to decide What to Do", *Bell Journal of Economics*, 1978, IX, n.2, pp.494-507.
- Simon, H. A., "From substantive to procedural rationality", in *Philosophy and economic theory* edited by F. Hahn and M. Hollis, New York, Oxford University Press, 1979.
- Simon, H. A., *Reason in Human Affairs*, Stanford, Stanford University Press, 1984.

Simon, H.A., "Bounded Rationality in Social Science: Today and Tomorrow", , in *Mind & Society*, 2000, I, n.1, pp.25-40.

Simon, H.A., *Models of Bounded Rationality*, Cambridge Mass, Cambridge University Press, 2000.

Slovic, P. e A. Tversky, "Who accepts Savage's axiom?", *Behavioral Science*, 19, pp. 368-373.

Slovic, P. et al., "Risk as analysis and risk as feeling: some thoughts about affect, reason, risk and rationality", *Risk analysis*, 2004, v. 24, n. 2, pp. 311-322.

Slovic, P., Choice between equally valued alternatives, *Journal of experimental psychology: human perception and performance*, 1975, p. 281.

Slovic, P., D. Griffin, A. Tversky, Compatibility effects in judgment and choice, 1990, in R. M. Hogarth (ed.) *Insights in Decision Making: theory and applications*, Chicago, The University of Chicago Press, 1990, pp. 5-27.

Smith, M., "The humane theory of motivation", in *Mind*, 1987, v. 96, pp. 36-61.

Smith, M., *The moral problem*, Oxford, Blackwell, 1994.

Stein, E., *Without good reason. The rationality debate in philosophy and cognitive science*, Oxford, Clarendon Press, 1996.

Stigler, G., e G. S. Becker, "De gusti bus non est disputandum", *American Economic Review*, 1977, 67, 2, pp. 76-90.

Stroud, B., *Hume*, London, Routledge, 1977.

Sugden, R., "Regret, recrimination and rationality", *Theory and decision*, 1985, 19, pp. 77-90.

Sugden, R., "Rational choice: a survey of contributions from economics and philosophy", *The Economic Journal*, 1991, 101, pp. 751-785.

Sugden, R., "Why rationality is *not* a consequence of Hume's theory of choice", *The European Journal of the History of Economic Thought*, 2005, 12:1, pp. 113-118.

Sugden, R., "Hume's non-instrumental and non-propositional decision theory", *Economics and philosophy*, 2006, 22, pp. 365-391.

Sugden R., "David Hume's *Treatise of Human Nature*", *Topoi*, 2008, 27, pp. 153-59.

Thaler, R., "Toward a positive theory of consumer choice", *Journal of Economic Behavior & Organization*, 1980, 1, pp. 39-60.

- Tversky, A., "Elimination by aspect: a theory of choice", *Psychological Review*, 1979, pp. 281-99.
- Tversky, A., D. Kahneman, *The framing of decisions and the psychology of choice*, Science, 1981, v. 211, pp.453-8, p. 453.
- Tversky, A., S. Sattah, P. Slovic, "Contingent weighting in judgement and choice", *Psychological review*, 1988, pp. 371-384.
- Tversky, A. e R. Thaler, "Anomalies: preference reversal", *Journal of Economic Perspectives*, 1990, pp. 201-211.
- Wallace, R. J., *Practical reason*, Stanford University of Philosophy, <http://plato.stanford.edu/entries/practical-reason>.
- Walsh, V. C., *Rationality, Allocation and Reproduction*, Oxford, Clarendon Press, 1996.
- Williams, B., *Problems of the Self*, Cambridge Mass., Cambridge University Press, 1973.
- Williams, B., "Values, reasons, theory of Persuasion", *Ethics, rationality and economic behavior*, Oxford, Clarendon Press, 1966.
- Williams, B., "Internal and external reasons", in *Moral Luck*, Cambridge, Cambridge University Press, 1981.
- Williams, B., "Internal reasons and the obscurity of blame", in Id., *Making sense of the humanity*, Cambridge, Cambridge University Press 1995, pp. 35-45.
- Williams, B., "Postscript: Some further notes on internal and external reasons", in Millgram ed., *Varieties of practical reasoning*, Cambridge MA, MIT Press, 2001.
- Yaary, M., "The dual theory of choice under risk", *Econometrica*, 1987, 55, pp.95–115.
- Zajonc, R.B., "Feeling and thinking: preferences need no inferences", *American Psychologist*, 1980, 35, pp. 151-75.