

I MODI DELLA RAZIONALITÀ

a cura di
Massimo Dell'Utri e Antonio Rainone

Volume pubblicato con i contributi del Dipartimento di Scienze Umane e Sociali dell'Università di Napoli "L'Orientale", del Dipartimento di Scienze Umanistiche e Sociali dell'Università di Sassari e di Sardegna Ricerche (L.R. 7 agosto 2007, n. 7).

MIMESIS EDIZIONI (Milano – Udine)
www.mimesisedizioni.it
mimesis@mimesisedizioni.it

Collana: *Filosofie* n. 468
Isbn: 9788857538310

© 2016 – MIM EDIZIONI SRL
Via Monfalcone, 17/19 – 20099
Sesto San Giovanni (MI)
Phone: +39 02 24861657 / 24416383
Fax: +39 02 89403935

INDICE

PREFAZIONE	7
1. REGOLE LOGICHE E RAZIONALITÀ UMANA <i>di Mario Piazza</i>	11
2. L'UNIVERSO INCERTO DELLA RAGIONE UMANA <i>di Dario Antiseri</i>	29
3. RAZIONALITÀ SCIENTIFICA E PLAUSIBILITÀ <i>di Carlo Cellucci</i>	47
4. TEORIE NATURALISTICHE DELLA RAZIONALITÀ <i>di Federico Laudisa</i>	65
5. RAZIONALITÀ, RAGIONAMENTO E COGNIZIONE <i>di Vincenzo Crupi</i>	81
6. COGNIZIONE E RAZIONALITÀ NEGLI ANIMALI NON UMANI <i>di Luca Forgiione</i>	99
7. RAZIONALITÀ PRATICA: DELIBERAZIONE E SCELTA DEGLI SCOPI <i>di Antonio Rainone</i>	115
8. RAZIONALITÀ E SCIENZE SOCIALI: ORIGINI, SVILUPPI E STATO DELL'ARTE DI UN ARTICOLATO DIBATTITO METODOLOGICO <i>di Albertina Oliverio</i>	133
9. RAZIONALITÀ ED EMOZIONI IN ETICA <i>di Mario De Caro e Massimo Marraffa</i>	151
10. RAZIONALITÀ E VERITÀ <i>di Massimo Dell'Utri</i>	169

11. TRADUZIONE, INTERPRETAZIONE E RAZIONALITÀ <i>di Francesca Ervas</i>	187
12. LINGUAGGIO, COMUNICAZIONE, RAZIONALITÀ <i>di Francesco Ferretti e Ines Adornetti</i>	203
AUTORI	221
INDICE DEI NOMI	225

RAZIONALITÀ, RAGIONAMENTO E COGNIZIONE¹

1. Introduzione: una mappa teorica

Secondo Cartesio, nelle *Regulae ad directionem ingenii*, «ogni errore che può capitare agli uomini [...] non avviene mai per una cattiva inferenza» (Descartes, 1628; tr. it. p. 151). Se talvolta traiamo conclusioni scorrette – si spiega nel seguito – ciò dipende dal fatto che non abbiamo vagliato con sufficiente attenzione le premesse di cui ci siamo serviti. Ma per Cartesio i principi generali che di fatto guidano il ragionamento umano ordinario sono pressoché infallibili. A un paio di secoli di distanza, l'empirista John Stuart Mill esprimeva però un punto di vista molto diverso: «le inferenze sbagliate [...] sono assolutamente inevitabili se non a prezzo di un'intensa coltivazione della facoltà del pensiero [e] sono deplorabilmente frequenti anche dopo che gli uomini abbiano raggiunto il loro grado di cultura più alto» (Mill, 1843; tr. it. pp. 975-76). Questa radicale divergenza di vedute – che riprendo dall'utile discussione di Diego Marconi in un bel libro di alcuni anni fa (Marconi, 2001, pp. 104-105) – illustra i due estremi di un continuo di posizioni teoriche riguardo alla razionalità umana. Semplificando un po', possiamo dire che per il “conservatore” Cartesio la ragione umana va bene così com'è: in quanto strumento inferenziale, funziona esattamente come dovrebbe. Per il “riformista” Mill è vero il contrario: l'errore è un dato sistematico del ragionamento umano spontaneo e può essere evitato soltanto con opportune misure correttive.

Per valutare quale di queste due opposte visioni della ragione umana sia la più adeguata, è naturale fare un salto in avanti nel tempo. Dopotutto, lo studio descrittivo del ragionamento umano si è pienamente sviluppato soltanto nel secolo scorso, con la psicologia sperimentale e la nascita delle scienze cognitive contemporanee. L'accumulazione dei dati empirici e il

1 L'elaborazione di questo contributo rientra nelle attività di ricerca comprese nel progetto FIRB *Strutture e dinamiche della conoscenza e della cognizione* (MIUR, D11J12000470001).

progresso di queste discipline dovrebbero in effetti aver ridotto drasticamente il novero delle posizioni plausibilmente difendibili a proposito della ragione umana e dei suoi limiti. O almeno così ci si potrebbe aspettare. Ebbene, per quanto naturale, questa aspettativa sembra invece disattesa. Gli attuali orientamenti sono certamente più sofisticati ed empiricamente informati, ma la loro varietà resta paragonabile a quella che corre fra le citazioni di Cartesio e di Mill più sopra. Vediamo brevemente come si collocano le posizioni più rilevanti.

All'estremo cartesiano, c'è chi sostiene che l'irrazionalità umana *non può* essere sperimentalmente dimostrata. La formulazione classica di questo punto di vista si trova in un celebre articolo del logico e filosofo L. Jonathan Cohen, *Can Human Irrationality Be Experimentally Demonstrated?* (Cohen, 1981). Secondo Cohen il ragionamento umano può rivelarsi effettivamente scorretto soltanto a causa di episodici fallimenti di esecuzione (*performance*), paragonabili per esempio agli errori di calcolo in aritmetica. Ma errori di questo tipo hanno una rilevanza scientifica molto limitata e soprattutto non contraddicono l'idea di fondo di una *competenza* inferenziale pienamente adeguata della mente umana. Questo punto di vista implica che la razionalità umana possa essere difesa con argomenti *a priori* e lascia ben poco spazio alla considerazione dei risultati dello studio sperimentale del ragionamento. Esso ha suscitato alcune forti reazioni critiche, ma ha conservato una certa influenza, specialmente in filosofia.

Una posizione meno estrema riconosce che la natura della razionalità umana è una questione contingente ed empirica, ma preserva l'aspetto più centrale della tesi di Cartesio. Secondo questo approccio, quindi, lo studio empirico del ragionamento umano *potrebbe* almeno in linea di principio documentare chiare e significative deviazioni da principi validi di razionalità, ma di fatto avviene il contrario, e ciò dà appunto sostegno empirico a un ottimismo di stampo cartesiano. A partire da questo fondamentale presupposto, si diramano due grandi varianti a seconda che si accetti o meno come norma valida ciò che Edward Stein (1996, pp. 214 e ss.) ha definito, con una fortunata etichetta, la «visione standard della razionalità». La visione standard della razionalità comprende i principi della logica classica, della teoria della probabilità e della teoria dell'utilità attesa. Possiamo dunque distinguere, per comodità, un orientamento razionalista standard e uno non-standard. L'approccio bayesiano alla cognizione è un importante esempio di razionalismo del primo tipo: secondo i suoi esponenti, è possibile rendere conto dei dati empirici disponibili rappresentando il ragionamento umano in condizioni di incertezza in base a opportuni modelli

probabilistici razionalmente fondati (Chater *et al.*, 2010; Griffiths, Tenenbaum, 2011).

Ma esiste, abbiamo detto, anche un razionalismo non-standard, secondo il quale la razionalità del comportamento osservato può essere adeguatamente difesa, sì, ma soltanto andando oltre i modelli normativi classici di logica, probabilità e utilità. Gli esponenti di questa posizione ritengono quindi che certi modelli non-standard del ragionamento rappresentino in molti casi *criteri normativi migliori* di quelli tradizionali per valutare il ragionamento. Anche questa forma di razionalismo è esemplificata da una varietà di approcci, caratterizzati da ciò che ciascuno di essi propone in alternativa alla visione standard. Un esempio molto influente è il programma di ricerca di Gerd Gigerenzer e collaboratori sulle euristiche frugali ed ecologicamente efficienti (Gigerenzer, Hertwig, Pachur, 2011).

Anche nel campo anti-cartesiano, più vicino alla posizione esemplificata da Mill, si trovano varianti più o meno moderate. Un'idea centrale è che di errori di ragionamento veri e propri ne esistano eccome, sistematici, ricorrenti e utili – anzi fondamentali – per la comprensione scientifica della cognizione umana. Questo presupposto è condiviso, per esempio, dalla teoria dei modelli mentali di Philip Johnson-Laird e collaboratori (Johnson-Laird, 2006; Legrenzi, Girotto, 1996). Ma l'esempio da manuale del punto di vista anticartesiano è probabilmente il programma di ricerca di Amos Tversky e Daniel Kahneman (Gilovich, Griffin, Kahneman, 2002), che ha anche determinato la nascita dell'odierna “economia cognitiva” (Motterlini, Guala, 2011). Questi studiosi descrivono le procedure euristiche del ragionamento umano come scorciatoie cognitive, generalmente efficienti ma tali da produrre, in condizioni specifiche, sostanziali violazioni dei principi di razionalità. È bene osservare che la visione standard della razionalità (logica, probabilità e utilità classiche) resta il riferimento normativo prevalente per questo insieme di posizioni.

Questa panoramica è ampia, ma resta inevitabilmente sommaria. Una gran quantità di precisazioni e approfondimenti sarebbe opportuna. Per fare solo un esempio, è noto che tanto Kahneman quanto Gigerenzer hanno presentato i loro rispettivi lavori come sviluppi dell'idea fondamentale di Herbert Simon della “razionalità limitata” (*bounded rationality*). Eppure, nello schema più sopra (così come nelle loro esplicite prese di posizione), i due risultano schierati su fronti teorici opposti. Come si spiega questa discrepanza? Si tratta in effetti di due interpretazioni significativamente diverse del carattere “limitato” della razionalità umana. Per Kahneman la razionalità umana è essenzialmente imperfetta: vincolata com'è dalle risorse

se e dalla struttura del cervello umano, è governata da principi qualitativamente distinti (le euristiche) da quelli dei modelli formali di ragionamento corretto e per questo è soggetta a errori sistematici. Per Gigerenzer, d'altra parte, la mente umana è "limitata" soltanto in contrapposizione all'illimitata accuratezza (nella stima dei parametri considerati) e capacità (nell'esecuzione dei calcoli) presupposte da molti modelli normativi tradizionali. Ma ciò di per sé non penalizza essenzialmente la qualità del ragionamento umano – anzi la favorisce, almeno in alcuni ambienti incerti e statisticamente rumorosi (Gigerenzer, Brighton, 2009).

Non sono mancati, specialmente in filosofia, i tentativi di mettere ordine in questo panorama teorico variegato e conflittuale (Samuels, Stich, Bishop, 2002). Nonostante questi lodevoli sforzi, però, si può ben dire che le "guerre della razionalità" sono tutt'altro che finite (Labinaz, 2013). Nel seguito di questo capitolo, propongo di affrontare il tema riprendendo la domanda posta dall'ultra-razionalista Cohen nel titolo del suo discusso articolo di ormai 35 anni fa. Diversamente dal caso di Cohen, però, non si tratterà per noi di una domanda retorica.

2. Come è possibile dimostrare sperimentalmente l'irrazionalità umana?

Come si può stabilire se si è verificato un errore di ragionamento? Cominciamo considerando un esempio illustrativo. Due amici, Anna (*A*) e Bruno (*B*), decidono di passare una serata al casinò. Entrati da poco, sono ancora fra i tavoli da gioco a chiacchierare. A un certo punto, però, *B* sembra distratto: si è accorto che al tavolo della roulette più vicino è appena uscito l'ultimo di una sequenza ininterrotta di cinque rossi (*RRRRR*). "Ehi" – dice allora *B* – "che coincidenza: era improbabile!" *A* non è d'accordo: i colori dei numeri estratti alla roulette – pensa *A* – rappresentano eventi indipendenti, e un rosso (proprio come un nero) ha una probabilità di $18/37$ di uscire a ogni giro (si tratta di una roulette francese). Ma se è così, per il calcolo della probabilità, nessuna sequenza alternativa (per esempio, *RBBRB*) ha una probabilità di occorrenza più alta di cinque rossi di fila, contrariamente a quel che sta dicendo *B*. La probabilità è infatti sempre di $(18/37)^5$ (e ancora più bassa se si include fra gli esiti lo zero). Ora, se foste nei panni del personaggio *A*, concludereste che il vostro amico ha fatto un *errore di ragionamento*? Supponiamo che *A* tragga questa conclusione. Ebbene, ci sono esattamente *tre motivi* per cui l'idea che *B* abbia *ragionato male*

potrebbe in realtà essere scorretta. (Uno dei punti seguenti non esclude gli altri, ma la lista è esaustiva).

(i) *Problema del criterio*. Per prima cosa, *A* ha applicato le regole di una precisa teoria formale (il calcolo della probabilità). Si tratta di un criterio di riferimento adeguato per valutare la correttezza di giudizi soggettivi relativi a eventi singoli? In termini generali, mettere in discussione la norma di ragionamento adottata rappresenta un'obiezione quanto meno legittima, e la scelta di *A* deve quindi presupporre un'opportuna giustificazione.

(ii) *Problema del disaccordo*. Forse *B* non sottoscrive qualcuna delle premesse iniziali di *A*; per esempio, l'indipendenza fra i giri successivi della roulette. Mettiamo: *B* pensa che, nel lancio subito successivo a un rosso, sia davvero più probabile ottenere un nero che un altro rosso. Sebbene si tratti di una convinzione per quanto ne sappiamo falsa, questa idea non determinerebbe affatto, di per sé, un errore di ragionamento, ma piuttosto un disaccordo fattuale: le premesse di *A* e *B* sarebbero in contrasto, ed entrambi potrebbero quindi aver ragionato correttamente (in base a principi normativi condivisi, in questo caso proprio il calcolo della probabilità), ma a partire da presupposti differenti.

(iii) *Problema del fraintendimento*. Infine, siamo sicuri di sapere che cosa intendeva dire *B*? Di nuovo, consideriamo un'ipotesi specifica: forse *B* voleva dire che una sequenza uniforme di cinque rossi era meno probabile di una combinazione mista di rosso e nero della stessa lunghezza, che si può ottenere in molti modi diversi (*RRRRB*, *RRRBR*, *RRRBB*, ...). Ma questo è del tutto corretto anche in base alle premesse e ai calcoli di *A*! Di nuovo, quindi, *B* non avrebbe fatto alcun errore di ragionamento. *A* avrebbe piuttosto, e semplicemente, frainteso l'osservazione di *B*, senza che vi fosse in questo caso alcuna reale divergenza di opinioni.

Negli studi di psicologia del ragionamento, lo sperimentatore si trova spesso in una posizione simile a quella del personaggio *A*. Egli assume che i partecipanti (come *B*, nella nostra storiella illustrativa) si servano di certe premesse – fornite esplicitamente nella stessa descrizione del problema sperimentale, oppure antecedentemente note – e chiede loro di indicare una risposta fra diverse possibili (per esempio, un giudizio quantitativo, oppure una scelta fra opzioni alternative). Di regola, tanto le premesse quanto le risposte sperimentali ammesse *non sono* espresse nel linguaggio formale di una specifica teoria normativa del ragionamento (per esempio, la teoria della probabilità), ma piuttosto nel linguaggio naturale. Per mettere in relazione le premesse e le risposte dei partecipanti con la teoria normativa che considera rilevante, lo sperimentatore dovrà servirsi di opportune ipotesi ausiliarie, spesso tacite. Proprio come *A* con *B*, egli dovrà quindi attribuire

al partecipante certe premesse formalmente ben definite e, analogamente, tradurre la sua risposta in un linguaggio teorico (come quello della teoria matematica della probabilità). In breve, l'argomentazione che conduce a una diagnosi di (ir)razionalità ha una struttura logica articolata. La cosa non è affatto sorprendente: che l'interpretazione dei dati scientifici richieda tipicamente l'impiego di una molteplicità di ipotesi è un'acquisizione della filosofia della scienza che risale all'inizio del Novecento (Duhem, 1906). Qui ne abbiamo un'esemplificazione specifica nello studio del ragionamento. Ma per la nostra discussione ne deriva una conseguenza di importanza cruciale. Se, e solo se, è possibile escludere i tre potenziali problemi elencati più sopra – quelli che abbiamo chiamato *del criterio*, *del disaccordo* e *del fraintendimento* –, allora può essere fondata la conclusione che una certa risposta indichi una violazione della razionalità. Solo a queste condizioni, infatti, si può dire che il partecipante abbia effettivamente contraddetto un principio di ragionamento valido e rilevante, e si può porre il problema teorico di spiegare l'occorrenza di questo errore (si veda la *Figura 1* per un'illustrazione).

Non è facile, però, che queste condizioni siano effettivamente soddisfatte, e per questo le presunte dimostrazioni dell'irrazionalità umana sono talvolta accolte con scetticismo, o almeno con comprensibile cautela. Nel seguito, per ovvie ragioni di spazio, potremo concentrarci con attenzione su un solo esempio, che useremo come caso di studio. Si tratta di un fenomeno molto noto e discusso, ma non sempre adeguatamente compreso: la *fallacia della congiunzione*.

3. Un caso di studio: la fallacia della congiunzione

In uno dei loro esperimenti più noti e discussi, Tversky e Kahneman (1983, p. 297) escogitarono la descrizione di un personaggio fittizio di nome Linda: single, schietta e molto intelligente, laureata in filosofia e intensamente impegnata negli anni degli studi in attività e manifestazioni pacifiste e per l'uguaglianza dei diritti. I partecipanti coinvolti dovevano valutare se fosse «più probabile» l'ipotesi che Linda fosse diventata «un'impiegata di banca e un'attivista del movimento femminista» o «un'impiegata di banca», e un'ampia maggioranza indicò la prima opzione. A prima vista, è come dire che estrarre da un mazzo di carte l'asso di cuori è più probabile che estrarre un asso (uno qualunque). È illogico, perché ci sono quattro assi e quello di cuori è solo uno di essi. Analogamente, chiunque sia una bancaria femminista è anche una bancaria. Più in generale, per la teoria della probabilità, la congiunzione di due condizioni (asso e cuori, bancaria e femminista) non può

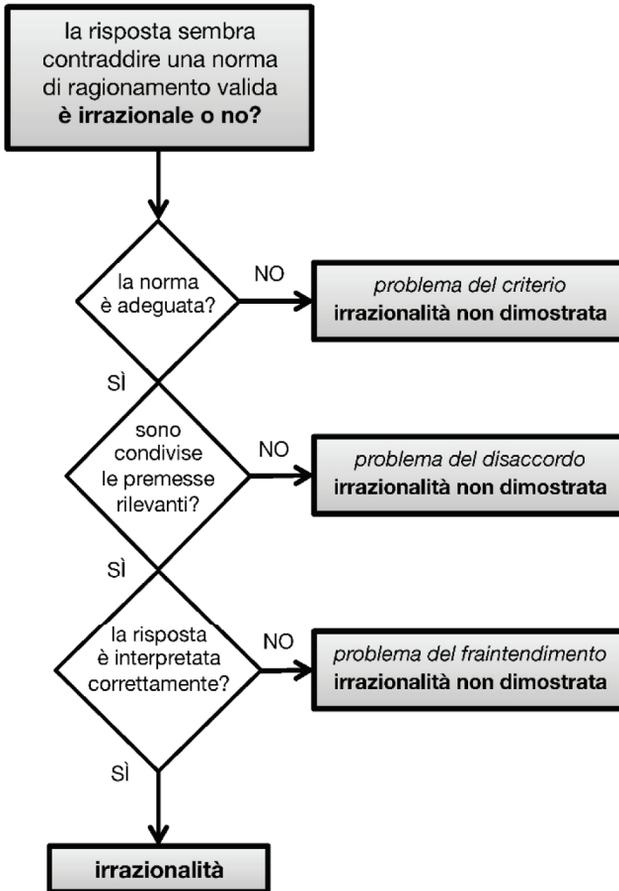


Figura 1. Dimostrare l'irrazionalità

mai essere più probabile di una sola di esse. In simboli, $P(x\&y) \leq P(x)$ è un teorema del calcolo della probabilità, anche detto *regola di congiunzione*. La validità della regola riguarda qualsiasi x e y e si estende anche a un qualunque insieme di condizioni antecedentemente date o presupposte, z , cioè $P(x\&y|z) \leq P(x|z)$. Ma allora questi esperimenti di Tversky e Kahneman hanno dimostrato un caso di irrazionalità, cioè la violazione di un principio valido del ragionamento corretto?

La questione è stata dibattuta in dozzine di articoli (a oggi sono più di 100), ma intanto l'esempio di Linda si è diffuso ben al di là della letteratu-

ra psicologica come illustrazione dei limiti della ragione umana (Gould, 1992). La diagnosi di irrazionalità è stata molto contestata fra logici e filosofi (per es., Levi, 1985), ma non solo (Hertwig, Gigerenzer, 1999). Le obiezioni sollevate sono numerose, ma prima di analizzarle dovremo chiarire meglio la base dei dati sperimentali in discussione. Negli studi sulla fallacia della congiunzione occorre infatti tenere presente tre aspetti distinti di fondamentale importanza metodologica.

(i) *Il design sperimentale: tra soggetti (between-subject) vs. entro i soggetti (within-subject)*. Corrisponde alla distinzione di Tversky e Kahneman (1983, p. 297) fra test “indiretti” e “diretti”. Nei test indiretti (o *tra* soggetti), ci sono *due* gruppi di partecipanti di una stessa popolazione (per esempio, studenti universitari) e a ciascun gruppo viene presentato solo *uno* dei due enunciati rilevanti: la congiunzione (in particolare, “Linda è una bancaria e un’attivista femminista”) oppure il congiunto isolato (“Linda è una bancaria”). La regola di congiunzione prescrive che la media dei giudizi di probabilità della congiunzione nel primo gruppo *non* debba essere maggiore di quella del congiunto isolato nel secondo gruppo. Nei test diretti (o *entro* i soggetti), invece, a *uno* stesso partecipante viene chiesto di giudicare la probabilità di *entrambi* gli enunciati. I test indiretti mostrano se i partecipanti hanno «una disposizione a rispondere alle due domande in modo incoerente», nel qual caso essi non deriverebbero «le loro risposte da una struttura [probabilisticamente] coerente» di stime o di credenze quantitative (Kahneman, Tversky, 1996, p. 587). Nei test diretti, d’altra parte, si può vedere se i partecipanti esprimono *di fatto* una risposta probabilisticamente incoerente.

(ii) *Struttura velata vs. trasparente*. Anche questa è una distinzione già presente in Tversky e Kahneman (1983, pp. 298-99), che riguarda però esclusivamente i test diretti (entro i soggetti). Quando la struttura è *velata* (*subtle*), la congiunzione e il congiunto isolato sono inclusi in una lista più estesa di enunciati da valutare (in certe versioni classiche, per esempio, sono otto in totale), alcuni dei quali hanno un contenuto del tutto differente (per esempio “Linda lavora in una libreria e segue corsi di yoga”). In un test *trasparente*, invece, non ci sono elementi “riempitivi” (*fillers*).

(iii) *Il compito sperimentale: stima vs. ordinamento vs. scelta*. Negli studi *tra* soggetti è inevitabile che ai partecipanti venga chiesta una *stima quantitativa* della probabilità dell’enunciato che viene loro presentato (per esempio: “quanto è probabile a tuo giudizio che Linda sia un’impiegata di banca?”). Negli studi *entro* i soggetti, invece, sono possibili tre procedure ben diverse. In alternativa a una stima quantitativa, si può anche chiedere ai partecipanti di ordinare (dal più al meno probabile) gli enunciati presentati, oppure dare loro un semplice compito di scelta (“indica il più probabi-

le fra i seguenti”). Il compito di ordinamento è il più naturale per rilevare l’occorrenza dell’effetto in studi a struttura velata perché, anche in presenza di diversi enunciati, si ottiene comunque una valutazione comparativa fra la congiunzione e il singolo congiunto di interesse. In studi trasparenti che comprendono soltanto due enunciati (la congiunzione e un congiunto isolato) è di fatto automatico che il compito sia di scelta. (Si veda la *Tabella 1* per una presentazione riassuntiva delle combinazioni possibili).

		test indiretti (tra soggetti)	test diretti (entro i soggetti)	
			velati (con <i>fillers</i>)	trasparenti (senza <i>fillers</i>)
tipo di compito	scelta		—	T&K (1983), pp. 299-300
	ordinamento		T&K (1983), pp. 297-8, 301	T&K (1983), pp. 302, 306
	stima	T&K (1983), pp. 295-8, 307-8	—	T&K (1983), pp. 308-9

Tabella 1. Varianti di procedure sperimentali negli studi sulla fallacia della congiunzione. Come si vede, molte delle combinazioni possibili – ma non tutte – si trovano già nel lavoro originale di Tversky e Kahneman (1983).

4. Il problema del criterio

Gli errori cognitivi sono stati spesso paragonati alle illusioni percettive, e l’analogia è in effetti istruttiva sotto molti punti di vista. In particolare, esistono due modi diversi per far “scompare” un’illusione percettiva. Il primo consiste nel rimuoverne i fattori determinanti; il secondo nell’introdurre elementi che contrastano l’illusione indirizzando il giudizio verso la soluzione corretta. La *Figura 2* illustra i due casi in riferimento alla classica illusione di Ebbinghaus. Con la fallacia della congiunzione si osserva qualcosa di simile. Nello scenario di Linda, l’effetto è chiaramente determinato da una relazione psicologicamente saliente fra la descrizione iniziale del personaggio e il secondo congiunto, cioè l’idea che Linda sia un’attivista femminista. Tversky e Kahneman, come noto, indicano tale relazione con il termine “rappresentatività” (ma si veda anche Tentori, Crupi, Russo, 2013). Così, se dalla descrizione del personaggio sono eliminate tutte le caratteristiche rilevanti (dicendo solo, per esempio, “Linda è una donna di 31 anni”), le determinanti dell’effetto vengono meno e la fallacia “scompare” (Tversky, Kahneman, 1983, p. 305).

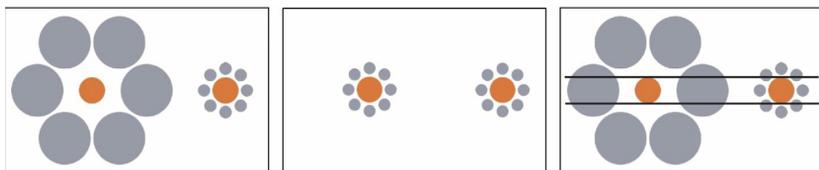


Figura 2. L'illusione di Ebbinghaus (a sinistra) e due modi diversi di farla "scompare": (i) rimuovere le sue determinanti (immagine centrale), e (ii) indirizzare il giudizio verso la valutazione corretta (a destra).

Ma qual è l'analogo del secondo metodo, quello esemplificato a destra in *Figura 2*, in cui vengono messi in particolare evidenza elementi utili a evitare l'errore? Ecco un esempio:

200 donne selezionate per partecipare a un sondaggio hanno in comune le caratteristiche seguenti. Hanno in media 30 anni, sono single e molto intelligenti. Si sono laureate in filosofia. Da studentesse si sono interessate dei temi della discriminazione e della giustizia sociale, partecipando anche a manifestazioni pacifiste.

Ti chiediamo di stimare la frequenza di questi eventi:

- Quante di queste 200 donne sono impiegate di banca? _____ su 200
- Quante di queste 200 donne sono impiegate di banca e attiviste femministe? _____ su 200

In questa versione abbiamo un test diretto e trasparente con un compito di stima. Inoltre, gli elementi statistici rilevanti sono rappresentati in un formato *frequentista* (come rapporti fra numeri interi in una classe di riferimento), anziché come probabilità o percentuali. Dei 43 partecipanti cui fu sottoposto il problema solo il 7% commise la fallacia della congiunzione, stimando il secondo numero maggiore del primo (Hertwig, Gigerenzer, 1999, p. 291). Si tratta di un risultato interessante: la fallacia della congiunzione, proprio come molte illusioni percettive, può essere fatta "scompare" anche introducendo condizioni che favoriscono l'elaborazione di un giudizio corretto. Ma questo è sufficiente a mostrare che la fallacia della congiunzione non è dopotutto una vera fallacia? Gigerenzer ha sostenuto di sì, ma per farlo ha dovuto combinare questo dato con un'importante argomentazione teorica riconducibile, nella nostra terminologia, al *problema del criterio*. Vediamo di che cosa si tratta.

Per sostenere che la fallacia della congiunzione sia effettivamente un caso di irrazionalità occorre senz'altro presupporre che la regola di congiunzione sia una norma valida per i compiti sperimentali considerati. Nel contestare questa premessa, Gigerenzer ha notato, correttamente, che gli esperimenti tradizionali (come la versione originaria del problema di Linda) riguardavano solitamente la probabilità di eventi singoli, senza una classe di riferimento omogenea e ben definita. Gigerenzer ha inoltre difeso una *interpretazione frequentista* della probabilità, secondo la quale l'assegnazione di valori di probabilità a eventi di questo tipo è semplicemente insensata. Un frequentista, quindi, *non* riconoscerebbe affatto la regola di congiunzione come un principio valido di razionalità per i compiti sperimentali classici studiati da Tversky e Kahneman (1983) e con ciò respingerebbe anche l'idea che in tali compiti si manifesti un errore di ragionamento: «la regola di congiunzione applicata a un evento singolo non è una norma generale di ragionamento corretto» (Gigerenzer, 1996, p. 593).

Naturalmente, anche il frequentista deve riconoscere la regola di congiunzione come valida nell'ambito di applicazione appropriato della probabilità statistica, cioè appunto quando si tratti di frequenze opportunamente riferite a classi di casi. E tuttavia, prosegue l'argomento, una formulazione frequentista «fa in gran parte scomparire» l'illusione cognitiva (Gigerenzer, 2005, p. 204). Ricapitolando, l'obiezione pertinente *non* è semplicemente che in certe condizioni la fallacia della congiunzione scompare. È normale che ciò possa avvenire con le illusioni cognitive come con quelle percettive, e non dimostra certo che non sono illusioni. Il problema che stiamo discutendo riguarda piuttosto la validità del criterio di razionalità impiegato, cioè la regola di congiunzione della teoria della probabilità. Secondo Gigerenzer, non è un principio vincolante nei casi solitamente considerati (cioè nei compiti probabilistici tradizionali), mentre è ampiamente rispettato proprio quando dovrebbe esserlo (cioè se si adotta un formato di presentazione frequentista). Ma è proprio così? Consideriamo questa versione del problema di Linda:

Supponi che ci siano 1000 donne che corrispondono a questa descrizione [quella di Linda]. Ti chiediamo di ordinare gli elementi di questa lista dal più al meno frequente nelle persone comprese in questo gruppo di 1000 donne:

- è un insegnante
- è un'impiegata di banca
- ha l'hobby della lettura
- ama le feste
- è un medico con due figli
- è una psicologa

- è una fumatrice di sigarette
- è una contabile che ha l'hobby della scrittura
- è un avvocato
- è un'impiegata di banca attiva nel movimento femminista
- ha l'hobby del pattinaggio.

Nei termini definiti più sopra, si tratta di un test diretto e velato con un compito di ordinamento. Nonostante l'impiego di una formulazione frequentistica, Steven A. Sloman *et al.* (2003, p. 306) hanno osservato che la fallacia della congiunzione era prevalente: il 70% dei partecipanti giudicava l'enunciato “è un'impiegata di banca attiva nel movimento femminista” più probabile di “è un'impiegata di banca”. In un problema identico, ma con una formulazione probabilistica standard, la percentuale degli errori *non era* significativamente differente. Inoltre, Tentori, Bonini e Osherson (2004, p. 472) hanno ottenuto lo stesso tipo di risultato in un test diretto *e trasparente* con un compito di scelta. È chiaro quindi che la fallacia della congiunzione resta ben presente anche con un formato frequentista, cioè in un caso in cui la validità normativa del principio formale di riferimento non è controversa. Sloman *et al.* (2003) hanno anche mostrato che la variabile principale che determina una riduzione dell'effetto (fino a meno del 35%) è piuttosto il passaggio da un compito di ordinamento a un compito di stima (lo stesso tipo di compito, è da notare, assegnato anche nella versione di Hertwig e Gigerenzer più sopra).

5. Il problema del disaccordo

È possibile ricostruire come razionali le risposte dei partecipanti in base a considerazioni legittime ma differenti da quelle che vengono solitamente attribuite loro? Un argomento di questo tipo si trova in una sofisticata analisi di epistemologia bayesiana presentata da Bovens e Hartmann (2003). Per cominciare, si suppone che per i partecipanti la descrizione di Linda sia un dato di fatto noto; chiamiamolo *e* (per *evidence*). Nel modello di Bovens e Hartmann, inoltre, “Linda è un'impiegata di banca” (*b*) e “Linda è un'impiegata di banca e un'attivista femminista” (*b&f*) vengono considerati *resconti* di due distinte fonti di informazione (per esempio, testimoni) che possono essere più o meno affidabili. Ora, secondo Bovens e Hartmann, i dati in *e* implicano valutazioni di affidabilità molto differenti per queste due ipotetiche fonti. In particolare, se una fonte riporta solamente *b* (chiamiamo *r* la descrizione di questo fatto), allora la descrizione di Linda, *e*, suggerisce

che la fonte stessa è probabilmente *poco affidabile*. Per contro, se una fonte riferisce $b \& f$ (chiamiamo r^* questo fatto), la probabilità che la fonte sia affidabile è maggiore, proprio in base a e . Ma a questo punto, dimostrano Bovens e Hartmann, è formalmente possibile che la probabilità dell'enunciato b riferito da una fonte probabilmente inaffidabile sia minore della probabilità di $b \& f$ riferito da una fonte probabilmente affidabile, cioè $P(b|e \& r) < P(b \& f|e \& r^*)$. Se questa fosse la corretta rappresentazione formale del giudizio dei partecipanti, la regola di congiunzione *non* verrebbe violata, perché i dati alla base delle due stime di probabilità sono differenti.

L'analisi di Bovens e Hartmann (2003) è teoricamente affascinante, ma si imbatte in almeno una formidabile difficoltà. Nelle procedure sperimentali *indirette* si possono in effetti attribuire ai due distinti gruppi di partecipanti che devono valutare, rispettivamente, b vs. $b \& f$ le diverse premesse $e \& r$ vs. $e \& r^*$. Ma in un test *diretto*, anche volendo preservare l'idea delle due fonti di incerta affidabilità, ciascun partecipante avrebbe ovviamente a disposizione tutti quei dati insieme (cioè la congiunzione di e , r e r^*) e il suo compito sperimentale sarebbe quello di confrontare $P(b|e \& r \& r^*)$ con $P(b \& f|e \& r \& r^*)$. Resta quindi il fatto che, giudicando la seconda quantità più grande della prima, la maggioranza dei partecipanti di questi studi contraddice la regola di congiunzione, anche nella ricostruzione di Bovens e Hartmann (2003).

6. Il problema del fraintendimento

Il problema del possibile fraintendimento delle risposte dei partecipanti è stato di gran lunga il più discusso in riferimento alla fallacia della congiunzione. Le due critiche più importanti riguardano l'interpretazione del termine 'probabile' e l'interpretazione del congiunto isolato. Le considereremo brevemente entrambe.

Il termine 'probabile', nel linguaggio naturale, è indubbiamente ambiguo in molti contesti. Di conseguenza, nell'affermare che "Linda è un'impiegata di banca e un'attivista femminista" è "più probabile" di "Linda è un'impiegata di banca", non è detto che le persone stiano esprimendo la relazione matematicamente falsa $P(b \& f) > P(b)$. Forse stanno valutando, in relazione a questi due enunciati, una proprietà V differente dalla probabilità matematica standard e tale che per V non vale un principio analogo alla regola di congiunzione, cosicché asserire che $V(b \& f) > V(b)$ non è affatto irrazionale. Un esempio importante, già considerato da Tversky e Kahneman (1983, p. 312), è quello in cui V rappresenta il *valore informativo*:

L'enunciato "l'inflazione sarà fra il 6% e il 9% alla fine dell'anno" potrebbe essere una previsione di maggior valore rispetto a "l'inflazione sarà fra il 3% e il 12%", anche se la probabilità che si verifichi il secondo è più alta. Una buona previsione è un compromesso fra una stima puntuale, destinata a essere certamente smentita, e in intervallo con una credibilità del 99,9%, che è spesso troppo ampio.

Si noti che questa possibile critica è logicamente distinta dal *problema del criterio*. Chi la sostiene, infatti, non ha alcun bisogno di contestare o limitare la validità normativa della regola di congiunzione. Al contrario, egli suggerisce che le persone in realtà si conformino a tale principio: per rendersi conto che le loro risposte *non* lo contraddicono occorrerebbe però darne un'interpretazione più raffinata, che tenga conto appunto dell'ambiguità nel termine 'probabile'.

Tuttavia, la fallacia della congiunzione rimane prevalente anche quando il termine 'probabile' è ommesso del tutto dalla presentazione del compito. Lo abbiamo già visto più sopra nella versione impiegata da Sloman e collaboratori, in cui si chiedeva di valutare *quanto fossero frequenti* certe caratteristiche (o combinazioni di caratteristiche) in una popolazione di riferimento. Per di più, anche in test diretti e trasparenti, i partecipanti sperimentali si mostrano spesso disposti a *scommettere* sull'occorrenza di una congiunzione di eventi più che sull'occorrenza di uno solo di essi a parità di condizioni economiche. Tversky e Kahneman (1983) hanno già documentato questo risultato con il problema classico di Linda (56% di scommesse su *b&f* anziché su *b*, p. 300). Più di recente, Bonini, Tentori e Osherson (2004) l'hanno replicato con scommesse reali su eventi futuri come "fra giugno e novembre dell'anno prossimo, per ridurre gli incidenti automobilistici, il governo italiano lancerà una campagna di informazione [*x*] e punirà più duramente le infrazioni [*y*]". I partecipanti potevano suddividere una cifra "virtuale" fra diverse possibilità. Alla fine del periodo di tempo stabilito, avrebbero ricevuto vincite *reali* esattamente corrispondenti alle loro ripartizioni per gli eventi che si erano effettivamente verificati. Nel problema del traffico, per esempio, puntarono in media 2,46 euro su *x* e ben 3,29 euro su *x&y*. Si noti che, in un caso come questo, puntare *qualsunque* cifra sulla congiunzione anziché su un congiunto è una scelta impossibile da razionalizzare per un agente intenzionato a vincere.

Un'ultima fondamentale obiezione riguarda poi l'interpretazione del congiunto isolato. Nel caso di Linda, per esempio, chiedere una comparazione di probabilità fra *b&f* e *b* può apparire innaturale dal punto di vista dei principi pragmatici della comunicazione, proprio in virtù dell'ovvia re-

lazione di implicazione logica (o di inclusione insiemistica) fra le due proposizioni. Per ristabilire il carattere cooperativo dell'interazione i partecipanti sarebbero quindi spinti a reinterpretare il congiunto isolato "Linda è un'impiegata di banca", rendendolo informativo in misura comparabile alla congiunzione "Linda è un'impiegata di banca e un'attivista femminista" e tale da rappresentare uno stato di cose effettivamente alternativo a quest'ultima. L'ipotesi più plausibile (Adler, 1984) sarebbe quindi che in questo contesto "Linda è un'impiegata di banca" venga interpretato come "Linda è un'impiegata di banca e non è un'attivista femminista" ($b \& \text{non-}f$). Le risposte dei partecipanti esprimerebbero così l'idea che $P(b \& f) > P(b \& \text{non-}f)$, senza alcuna violazione della regola di congiunzione.

Anche questa è un'osservazione legittima e importante. È da notare però innanzitutto che non si applica ai test indiretti, cioè agli esperimenti tra gruppi distinti di partecipanti, perché in questi studi uno dei due gruppi considera il congiunto isolato (per esempio, "Linda è un'impiegata di banca") senza che la congiunzione venga neppure menzionata. Non si dà quindi alcun elemento pragmatico che debba suggerire una reinterpretazione. Nei test diretti, d'altra parte, un metodo sperimentale molto efficace per neutralizzare la potenziale interferenza dei fattori pragmatici in questione consiste nell'includere *esplicitamente* $b \& \text{non-}f$ fra gli enunciati da valutare. È un accorgimento che si trova, per esempio, in Wedell e Moro (2008) (ma già anche in Tentori *et al.*, 2004). In un test diretto e trasparente con l'usuale descrizione di Linda, i partecipanti dovevano indicare l'enunciato più probabile fra:

- Linda è un'impiegata di banca
- Linda è un'impiegata di banca ed è un'attivista femminista
- Linda è un'impiegata di banca e non è un'attivista femminista.

Nella scelta fra queste opzioni, reinterpretare il congiunto isolato "Linda è un'impiegata di banca" come $b \& \text{non-}f$ non solo non è richiesto da considerazioni pragmatiche, ma sarebbe al contrario del tutto anomalo, implicando tra l'altro una plateale violazione della massima di non-ridondanza. Nel lavoro di Wedell e Moro (2008, p. 117) la percentuale di fallacie con questa versione fu del 67%, statisticamente indistinguibile da una condizione di controllo con le due sole opzioni usuali inclusa nello stesso studio.

7. Conclusioni

Documentare sperimentalmente l'irrazionalità umana non solo è possibile, ma è un fatto sostanzialmente acquisito in almeno un caso molto noto, che abbiamo analizzato con attenzione. Da ciò non consegue, d'altra parte, che tutti o anche solo molti degli errori cognitivi che sono stati identificati nello studio del ragionamento siano effettivamente tali. In ciascuno di questi casi, spesso non meno interessanti di quello qui discusso, è necessaria un'analisi altrettanto approfondita per trarre conclusioni ben fondate. Lo schema che abbiamo seguito qui parte da alcune considerazioni metodologiche generali e può essere opportunamente applicato ad altri esempi controversi. Il fatto che gli esiti possano variare da caso a caso non deve preoccupare: non tutte le diagnosi di irrazionalità risultano pienamente convincenti alla luce dei potenziali problemi che abbiamo caratterizzato (criterio normativo, disaccordo fattuale, o fraintendimento) (si veda, per esempio, Crupi, Tentori, Lombardi, 2009). Ciò indica semmai che questo schema può effettivamente contribuire alla comprensione dei dati disponibili, orientare la riflessione teorica e anche suggerire ulteriori sviluppi di ricerca. In effetti, il confronto fra visioni differenti ha svolto un importante ruolo euristico, alimentando la crescita della conoscenza del ragionamento umano e dei suoi limiti sia dal punto di vista teorico sia sul piano sperimentale. Così, mentre da un lato non sembra che le "guerre della razionalità" siano destinate a concludersi in tempi brevi, il dibattito sulla razionalità ai confini fra filosofia e psicologia si è rivelato scientificamente fruttuoso, ed è probabile che continui a esserlo.

Riferimenti bibliografici

- Adler, J. (1984), *Abstraction is Uncooperative*, in «Journal for the Theory of Social Behaviour», 14, pp. 165-81.
- Bonini N., Tentori K., Osherson D. (2004), *A Different Conjunction Fallacy*, in «Mind & Language», 19, pp. 199-210.
- Bovens L., Hartmann S. (2003), *Bayesian Epistemology*, Oxford University Press, Oxford.
- Chater N., Oaksford M., Hahn U., Heit E. (2010), *Bayesian Models of Cognition*, in «Wiley Interdisciplinary Reviews: Cognitive Science», 1, pp. 811-23.
- Cohen L.J. (1981), *Can Human Irrationality Be Experimentally Demonstrated?*, in «Behavioral and Brain Sciences», 4, pp. 317-31.
- Crupi V., Tentori K., Lombardi L. (2009), *Pseudodiagnosticity Revisited*, in «Psychological Review», 116, pp. 971-85.

- Descartes R. (1628), *Regulae ad directionem ingenii* (tr. it. *Regole per la guida dell'intelligenza*, a cura di L. Urbani Ulivi, Fabbri, Milano 2000).
- Duhem P. (1906), *La théorie physique. Son objet et sa structure* (tr. it. dalla seconda ed. 1914 *La teoria fisica: il suo oggetto e la sua struttura*, Bologna, Il Mulino 1978).
- Elio R., ed. (2002), *Common Sense, Reasoning, and Rationality*, Oxford University Press, Oxford.
- Gigerenzer G. (1996), *On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky*, in «Psychological Review», 103, pp. 592-96.
- (2005), *I think, Therefore I Err*, in «Social Research», 72, pp. 195-218.
- Gigerenzer G., Brighton H. (2009), *Homo heuristicus: Why Biased Minds Make Better Inferences*, in «Topics in Cognitive Science», 1, pp. 107-43.
- Gigerenzer G., Hertwig R., Pachur T., eds. (2011), *Heuristics: The Foundations of Adaptive Behavior*, Oxford University Press, New York.
- Gilovich T., Griffin D., Kahneman D., eds. (2002), *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press, New York.
- Gould, S.J. (1992), *The Streak of Streaks*, in *Bully for Brontosaurus: Further Reflections in Natural History*, Penguin Books, London, pp. 463-69.
- Griffiths T.L., Tenenbaum J.B. (2011), *Predicting the Future as Bayesian Inference: People Combine Prior Knowledge with Observations when Estimating Duration and Extent*, in «Journal of Experimental Psychology: General», 140, pp. 725-43.
- Hertwig R., Gigerenzer G. (1999), *The «Conjunction Fallacy» Revisited: How Intelligent Inferences Look like Reasoning Errors*, in «Journal of Behavioral Decision Making», 12, pp. 275-305.
- Johnson-Laird P.N. (2006), *How We Reason*, Oxford University Press, New York.
- Kahneman D., Tversky A. (1996), *On the Reality of Cognitive Illusions*, in «Psychological Review», 103, pp. 582-91.
- Labinaz, P. (2013), *La razionalità*, Carocci, Roma.
- Legrenzi P., Girotto V. (1996), *Mental Models in Reasoning and Decision Making Processes*, in Oakhill, Garnham (1996), pp. 95-118.
- Levi I. (1985), *Illusions about Uncertainty*, in «British Journal for Philosophy of Science», 36, pp. 331-40.
- Marconi D. (2001), *Filosofia e scienza cognitiva*, Laterza, Roma-Bari.
- Mill J.S. (1843), *A System of Logic Ratiocinative and Inductive* (tr. it. *Sistema di logica deduttiva e induttiva*, UTET, Torino 1988).
- Motterlini M., Guala F., a cura di (2011), *Mente, mercati, decisioni: introduzione all'economia cognitiva e sperimentale*, UBE, Milano.
- Oakhill J., Garnham A., eds. (1996), *Mental Models in Cognitive Science*, Taylor and Francis, Milton Park.
- Samuels R., Stich S., Bishop M. (2002), *Ending the Rationality Wars: How to Make Disputes about Human Rationality Disappear*, in Elio (2002), pp. 236-68.
- Sloman S.A., Over D., Slovak L., Stibel J.M. (2003), *Frequency Illusions and Other Fallacies*, in «Organizational Behavior and Human Decision Processes», 91, pp. 296-309.

- Stein E. (1996), *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*, Clarendon Press, Oxford.
- Tentori K., Bonini N., Osherson D. (2004), *The Conjunction Fallacy: A Misunderstanding about Conjunction?*, in «Cognitive Science», 28, pp. 467-77.
- Tentori K., Crupi V., Russo S. (2013), *On the Determinants of the Conjunction Fallacy: Probability vs. Inductive Confirmation*, in «Journal of Experimental Psychology: General», 142, pp. 235-55.
- Tversky A., Kahneman D. (1983), *Extensional vs. Intuitive Judgment: The Conjunction Fallacy in Probability Judgment*, in «Psychological Review», 90, pp. 293-315.
- Wedell D.H., Moro R. (2008), *Testing Boundary Conditions for the Conjunction Fallacy: Effects of Response Mode, Conceptual Focus and Problem Type*, in «Cognition», 107, pp. 105-36.