

TENDENZA ALLA CONFERMA, USO DELL'EVIDENZA E SCOPERTA INDUTTIVA NEL COMPITO 2-4-6

Alberto Greco

Laboratorio di Psicologia, Dipartimento di Filosofia, Università di Genova

1. INTRODUZIONE

Il compito 2-4-6, originariamente proposto da Wason (1960), continua ad essere utilizzato nelle ricerche riguardanti le strategie dei soggetti che tentano di scoprire induttivamente una regola scegliendo essi stessi gli esempi da testare (Tweney e al, 1980; Gorman e Gorman, 1984; Tukey, 1986). Come si sa, essendo la regola-target molto generale (tre numeri qualsiasi in ordine crescente), di solito succede che i soggetti ipotizzino regole più particolari di quella, testando esempi che risultano sempre seguirla e finendo con il procurarsi molte conferme per le loro ipotesi. L'interesse di questo compito sta dunque nel fatto che riproduce una situazione in cui si hanno spiegazioni dei dati empirici facilmente confermate ma sbagliate, una situazione frequente nell'indagine scientifica.

Wason fin dall'inizio ha ritenuto che le difficoltà nella performance sarebbero dovute ad una strategia inadeguata, in particolare viziata dal cosiddetto *confirmation bias*, cioè dalla tendenza alla conferma delle proprie ipotesi (o, per converso, al mancato uso della disconferma laddove sarebbe preferibile). Questa interpretazione è suggerita dal fatto che i soggetti che risolvono il compito alla prima enunciazione di ipotesi risultano essere quelli che esibiscono il maggiore uso di una strategia di disconferma. Sia Wason che gli altri autori hanno evidenziato come questo risultato appaia in linea con le note posizioni epistemologiche di Karl Popper (1959), secondo il quale la ricerca scientifica

progredisce essenzialmente quando si sia in grado di falsificare le ipotesi piuttosto che confermarle.

Il punto centrale delle indagini che adottano il compito 2-4-6 è dunque di mostrare che *l'adozione di una strategia inadeguata* (viziata dal confirmation bias o, reciprocamente, dalla mancata adozione della disconferma) *ostacola la scoperta della regola*. Nella presente indagine si intende discutere questo assunto e mettere in rilievo alcuni problemi e alcune incoerenze del paradigma sperimentale che fa uso del compito 2-4-6 di Wason.

Il fatto che le difficoltà dei soggetti nel compito siano da attribuire alla dimensione conferma-disconferma, intanto, non è chiaramente sostenuto da tutte le risultanze empiriche. Ad esempio Tweney et al. (1980) hanno tentato di modificare le strategie dei soggetti utilizzando istruzioni che facevano loro presente la possibilità di usare una strategia di disconferma, ma, pur riuscendo a favorire tale strategia, non sono riusciti ad ottenere un miglioramento nella performance. Invece Gorman e Gorman (1984), seguendo una metodologia simile, con esplicite istruzioni alla disconferma, ma evitando di fornire un feedback all'enunciazione di ipotesi provvisorie da parte del soggetto (come già nell'esperimento di Wason e Katzman citato in Wason e Johnson-Laird, 1968) hanno ottenuto risultati che sosterebbero il ruolo decisivo della disconferma nella performance. Secondo Gorman e Gorman, se i soggetti sono obbligati a decidere da sé se un'ipotesi è corretta, le istruzioni che inducono alla disconferma sono efficaci anche per il miglioramento della performance nella scoperta della regola-target. Al contrario, da un nostro precedente esperimento (Greco, 1989) non risulta che vi sia un nesso fra l'uso della disconferma e la scoperta della regola (in tale esperimento, però, veniva normalmente dato il feedback all'enunciazione delle varie ipotesi).

Questi risultati contraddittori possono dipendere anche dal fatto che spesso nelle ricerche che hanno utilizzato il compito 2-4-6 sono state apportate leggere modifiche a questo o quell'aspetto di esso ma non se ne è riesaminato il contesto complessivo. Inoltre in ognuna di queste repliche sono stati utilizzati modi diversi di presentare il compito, non sempre equivalenti dal punto di vista della rappresentazione del problema da parte dei soggetti. Anche se il compito è chiaramente inadeguato quale emulazione del contesto di ricerca scientifica (Klayman e Ha, 1987; Moneta, 1987) e dello stesso ragionamento induttivo quotidiano (Wetherick, 1962), ci sembra che valga ancora la pena di riflettere su alcune grossolane semplificazioni che questo paradigma conteneva. Esporremo ora due gruppi di questioni che a nostro parere costituiscono importanti fonti di difficoltà per questo paradigma empirico. Il primo gruppo di problemi deriva dalla definizione operativa dei concetti di conferma-disconferma adottata di solito; il secondo gruppo riguarda l'artificiosità del compito.

2. DEFINIZIONE OPERATIVA DI "CONFERMA" E "DISCONFERMA"

Confermare per Wason significa proporre triplette di numeri *compatibili con l'ipotesi che un soggetto in un dato momento ha in mente* e, viceversa, *disconfermare* proporre triplette *incompatibili* con tale ipotesi. Dunque la conferma e la disconferma sono definite, e di conseguenza misurate, facendo riferimento alla compatibilità o meno degli esempi testati rispetto alla *corrente ipotesi* del soggetto. Questa definizione comporta almeno i seguenti due problemi.

a) Vi sono casi in cui non si può chiaramente individuare quale sia l'ipotesi che il soggetto ha formulato o addirittura se ne abbia formulata una; in questi casi risulta difficile valutare se un esempio sia disconfermante oppure no e mettere in relazione la scoperta o non scoperta della regola con una strategia precisa.

b) Identificare la disconferma o falsificazione con il contraddire la propria ipotesi è comunque limitativo, perché disconfermare, più in generale, vuol dire semplicemente tentare di dimostrare di avere sbagliato. In altri contesti, un soggetto può scegliere esempi in contrasto con la propria ipotesi allo scopo di confermarla e viceversa scegliere esempi congruenti con essa allo scopo di falsificarla. Là cosa importante è l'aspetto soggettivo, cioè che cosa il soggetto *si aspetti* nel proporre un esempio.

Esamineremo più in dettaglio questi due problemi nelle seguenti sezioni.

2.1. Formulazione delle ipotesi

Il primo problema deriva dal fatto che la definizione della conferma - disconferma sopra riportata presuppone l'esistenza e la chiara identificabilità di una ipotesi dietro ogni esempio proposto dai soggetti per la verifica. Il paradigma di Wason (e successori) dà semplicemente per scontato che i soggetti usino sempre quella strategia consistente nel formulare un'ipotesi, testare tale ipotesi attraverso gli esempi scelti e nel caso in cui si rivelasse sbagliata sostituirla con un'altra (cfr. Tukey, 1986, che ha suggerito, per interpretare le strategie dei soggetti in questo compito, anche altri modi che non si rifanno soltanto all'epistemologia popperiana).

Questa strategia di esame successivo di ipotesi è quella che Bruner et al. (1956) hanno definito di *scanning* e che era oggetto di particolare interesse per Wason quando ha inventato il compito 2-4-6. Wason ha esplicitamente detto che il compito era stato inventato con l'intento di riprodurre la situazione di "induzione per enumerazione" che si presenta quando si adotta la strategia di *esame successivo* (cfr. Wason e Johnson-Laird, 1972, pp.245 sgg. tr.it.). Non si è dunque presa neppure in considerazione l'eventualità che i soggetti usino quell'altra strategia (detta di *focusing* da Bruner) consistente nel focalizzare l'attenzione su un attributo (come ad es. "pari" o "limite superiore") o un operatore (come ad es. "sommare 2") per poi testare esempi in cui lo stesso viene modificato o sostituito.

Di fatto l'indagine empirica mostra che invece questa strategia è preferita molto spesso dai soggetti (cfr. Greco, 1989).

Da queste osservazioni emerge allora che se, come si è soliti fare, il *confirmation bias* viene valutato e quantificato considerando il numero di esempi congruenti con l'*attuale ipotesi*, nei casi in cui in realtà non ci sia *nessuna ipotesi* ma solo una sistematica variazione di attributi, o una fluttuazione fra *scanning* e *focusing* o addirittura una esplorazione casuale, è difficile ottenere una simile misura senza operare una forzatura sui protocolli. Infatti, quando un soggetto continua a cambiare l'attributo su cui si basa nella scelta degli esempi da testare, si è costretti a considerare ciascuna tripletta enunciata come motivata da una *nuova ipotesi*, mentre in realtà il processo di ragionamento compiuto dal soggetto è totalmente diverso. Ad esempio, mentre nel caso dello *scanning* un'ipotesi viene normalmente abbandonata solo quando si è scoperto che è sbagliata, se si segue una strategia di *focusing* il test di un attributo può essere abbandonato momentaneamente, anche se non è stato eliminato, per essere eventualmente ripreso in seguito. Inoltre con una strategia di *focusing* l'attributo indicato nella scelta dell'esempio da testare può non essere compatibile con l'esempio iniziale (2-4-6), in quanto non scaturisce in realtà da un'ipotesi su quella tripletta ma può rivelare una indagine orientativa sulla rilevanza di qualche attributo o operatore.

In alcune precedenti indagini sperimentali, come si è visto, si è cercato di modificare le istruzioni per fare in modo che i soggetti adottassero una strategia di disconferma, ma non si è mai tentato di facilitare l'adozione della strategia dell'esame di ipotesi successive (*scanning*) e di favorire una codificazione chiara, prima di tutto per gli stessi soggetti, di tali ipotesi ed un controllo della loro corrispondenza o meno con gli esempi testati. Da una precedente indagine sperimentale (Greco, 1989), oltre al fatto che i soggetti impegnati in questo compito mostrano spesso di non essere in grado di formulare delle ipotesi, è emerso anche che, nei casi in cui comunque lo fanno, non riescono a farlo con chiarezza, né riescono a seguire una strategia coerente. Dall'evidenza empirica risulta che frequentemente i soggetti formulano esempi da testare senza essere tuttavia in grado di riferire con esattezza quale ipotesi li motivi, o riferendo ipotesi vaghe, confuse, contraddittorie con l'evidenza già ottenuta o che addirittura non motivano affatto gli esempi scelti. Queste osservazioni, che ricordano l'ipotesi dell'esistenza di una "conoscenza inespressa" (Polanyi, 1966), inducono a ritenere che le relazioni colte nei dati possano essere codificate a diversi livelli di esplicitzza, essendo possibile andare dalla vaga impressione alla formulazione in termini chiari, precisi e completi. D'altra parte, il contesto sperimentale classico non richiedeva una tale razionalità ed anzi non si è mai preoccupato di controllare esplicitamente se le istruzioni inducessero a formulare ipotesi, in che modo queste fossero codificate e come i soggetti valutassero la compatibilità o meno degli esempi testati rispetto alle ipotesi stesse. E' possibile, ad esempio, che i soggetti, se non costretti ad esprimere verbalmente le proprie ipotesi prima della scelta delle triplette da testare in realtà poi compiano questa scelta senza formulare nessuna ipotesi e basandosi su criteri alternativi. E' anche possibile che la strategia di *scanning* non risulti essere quella preferibilmente adottata dai soggetti, ma a

questa conclusione non si può giungere se il contesto empirico non favorisce almeno una codificazione chiara delle ipotesi ed un controllo della corrispondenza fra queste ultime e gli esempi testati. D'altra parte, se questa condizione non è rispettata, come si è visto, non è possibile neppure valutare correttamente se un soggetto intende confermare o disconfermare, perché tale intendimento presuppone il confronto fra un'aspettativa e un dato.

2.2. La falsificazione dal punto di vista soggettivo

Un secondo problema che scaturisce dalla definizione di *conferma-disconferma* data da Wason induce a mettere in discussione il presupposto, in essa implicito, della necessità che l'esempio scelto dal soggetto per il test sia in contrasto con la sua attuale ipotesi.

Klayman e Ha (1987) hanno chiarito un aspetto del compito 2-4-6 che lo rende particolare ed al tempo stesso è alla base di una confusione a proposito della "disconferma". Questi autori hanno infatti mostrato che è importante prima di tutto esaminare le relazioni fra la regola ipotizzata e la regola da scoprire (target), che non necessariamente sono sempre quelle del compito 2-4-6 in tutti i compiti induttivi. Nel compito di Wason le regole ipotizzate sono quasi sempre casi particolari della regola-target, ma in altri compiti è possibile ipotizzare regole più generali oppure regole parzialmente in intersezione con la regola-target. In questi casi in cui le relazioni fra la regola ipotizzata e la regola-target sono diverse, la strategia di disconferma può risultare improduttiva ed il proporre esempi in contrasto con la propria ipotesi può non essere una strategia altrettanto valida. Addirittura, se la regola ipotizzata fosse *più generale* di quella target (ad esempio: regola da scoprire = "numeri pari consecutivi"; regola ipotizzata = "numeri crescenti di 2"), l'unico modo per falsificare l'ipotesi sarebbe formulare esempi che seguono la regola ipotizzata avendo come risposta che non seguono la regola-target (ad es. numeri crescenti di 2 che non siano anche pari consecutivi, come 3-5-7). In altri termini, per un ragionamento eliminativo sarebbe richiesta in questo caso una strategia di *conferma* perché quella di disconferma non potrebbe mai falsificare l'ipotesi (esempi in contrasto con l'ipotesi sarebbero numeri *non* crescenti di 2, sempre in contrasto con la regola "pari consecutivi" e dunque si avrebbe invariabilmente una risposta "no" equivalente ad una verifica dell'ipotesi).

Le osservazioni di Klayman e Ha sono importanti perché dimostrano che l'adozione di una strategia di disconferma nel senso di Wason non è sempre e comunque la scelta più razionale. Infatti dal momento che per *disconferma* si possono intendere due cose diverse (testare casi che contraddicono le proprie ipotesi oppure che non seguono la regola-target), bisogna capire che l'importante ai fini della falsificazione non è di per sé il contraddire la propria ipotesi (o l'attributo attualmente considerato), bensì dimostrare che è sbagliata e quindi eliminarla. Nel caso del compito 2-4-6 le due cose tendono a coincidere, perché in esso la regola ipotizzata risulta con ogni probabilità un caso particolare di quella

target. Solo in queste condizioni effettivamente ha ragione Wason: l'unico modo per falsificare un' ipotesi è formulare esempi in contrasto con essa e sentirsi rispondere che seguono la regola-target.

Un altro caso in cui l'ipotesi potrebbe essere falsificata sarebbe il formulare esempi compatibili con essa nella speranza di ottenere la risposta "no, non segue la regola-target". Nel compito 2-4-6 tale risposta è oggettivamente impossibile, ma forse, se si è interessati alla strategia del soggetto, è importante ed opportuno considerare anche quali esempi siano falsificanti dal punto di vista *soggettivo*. Come aveva osservato Wetherick già nel 1962, in realtà non si può sapere se un soggetto elimini o no consapevolmente un'ipotesi senza chiedergli per ogni tripletta da lui testata se si aspetti che sia positiva o negativa (cioè se si aspetti o no che segua la regola-target). Anche Gorman e Gorman (1984) hanno sottolineato che è importante distinguere fra esempi *predetti* come non conformi alla regola-target ed esempi *effettivamente* risultati non conformi. La "disconferma" può dunque essere rivelata in un soggetto non soltanto considerando se propone esempi in contrasto con le sue ipotesi (che era il criterio di Wason) ma anche considerando che cosa *si aspetti* nel proporre un esempio. Ad es. proporre "8-10-12" per testare l'ipotesi "numeri crescenti di 2" sarebbe una strategia di disconferma per un soggetto che si aspettasse che venga risposto "no, non segue la regola-target", ed in effetti una tale risposta eliminerebbe l'ipotesi; viceversa, proporre "1-10-20" sempre per testare "numeri crescenti di 2" e aspettandosi la risposta "no" sarebbe in realtà una strategia di conferma della propria ipotesi anche se l'esempio scelto è in contrasto con essa.

In base a queste considerazioni, sarebbe dunque opportuno che l'indice di disconferma fosse calcolato soltanto considerando i casi in cui davvero un soggetto intende vedere se l'ipotesi che ha formulato è *falsa* o, quando non ne avesse formulata una (abbiamo già accennato all'importanza di questo caso), se gli attributi che ha individuato in "2-4-6" sono *irrilevanti*. Questi casi si riducono ai seguenti due: tripletta congruente con l'ipotesi (o con l'attributo individuato) ma con risposta attesa "no"; tripletta non congruente con l'ipotesi (il caso considerato da Wason) purchè con risposta attesa "si". Come si è visto in precedenza, nel primo di questi due casi la risposta attesa è impossibile ad ottenersi nel compito 2-4-6, ma esso può considerarsi falsificante dal punto di vista soggettivo, dell' intenzione, anche se oggettivamente non può condurre ad alcuna falsificazione.

3. ARTIFICIOSITA' DEL COMPITO

In questa sezione esporremo degli aspetti che rendono il compito 2-4-6 troppo artificioso non soltanto per consentire un'adeguata emulazione delle strategie adottate nelle situazioni che richiedono induzione, ma anche solo per giustificare la conclusione che le difficoltà nella scoperta della regola siano dovute all'adozione di una strategia sbagliata.

a) Come notato nel già citato lavoro di Gorman e Gorman (1984), consentire l'enunciazione di diverse ipotesi ed il fornire a ciascuna di esse un feedback può trattenere i soggetti dal ragionare in prima persona sulle conseguenze di un'ipotesi e in particolare sulla possibilità di eliminarla solo sulla base dell'evidenza piuttosto che per averlo avuto detto. Comunque questo feedback non è naturale: ad esempio allo scienziato non viene mai detto se le sue ipotesi sono giuste o sbagliate ma egli deve trovare questa risposta interrogando la natura.

b) Le istruzioni standard non forniscono al soggetto criteri chiari e uniformi per la decisione di ritenere accettabile - e quindi enunciabile - la regola; dire la regola "quando si è sicuri" per alcuni può voler dire quando è "plausibile", per altri quando è "convincente", per altri ancora (molto pochi) quando è "dimostrata". Le istruzioni fanno riferimento alla "sicurezza", non alla necessità logica, ma poi si interpretano i risultati per dimostrare l'esistenza di fallacie logiche nel ragionamento dei soggetti. Più in generale, il soggetto non viene incoraggiato a costruirsi un "modello mentale" del compito (Johnson-Laird, 1983; Green, 1989): ad esempio, quali tipi di regole sono possibili e quali relazioni ci sono fra le varie regole ipotizzate.

c) Non si tiene conto delle attese dei soggetti riguardo al modo di concettualizzare il compito. Queste attese concettuali interferiscono in vari modi sull'esecuzione del compito: ad esempio possono influire sull'analisi dell'esempio dato inizialmente quando il soggetto è alla ricerca di relazioni fra i tre numeri. Si può considerare queste attese come risultanti dall'attivazione di schemi o *frames* nel senso di Minsky (1975), con l'importante caratteristica di contenere informazioni di *default* per quegli aspetti che non siano stati specificati esplicitamente.

Un esempio di attesa di questo tipo è quella riguardante il livello di generalità della regola. Le ricerche della Rosch (1975) hanno mostrato che, in assenza di informazioni contrarie, per default viene assunto che gli elementi da categorizzare appartengano ad una categoria "basilare" piuttosto che super-ordinata o subordinata. Nel compito di Wason ciò si traduce in una maggiore plausibilità di regole come "numeri pari" piuttosto che "numeri qualsiasi in ordine crescente" o "numeri pari crescenti minori di 10 con differenza di 2 fra le cifre". Un'altra attesa, di cui non ci siamo occupati in questa indagine, può riguardare la rilevanza o meno dell'ordine degli elementi quale relazione fra gli stessi. In questo caso l'ordine per default non appare rilevante in quanto la tripletta è presa di solito come un insieme di numeri legati da qualche relazione ma enunciabili in un ordine qualsiasi.

In termini più generali, le attese dei soggetti interagiscono con il modo in cui il compito è presentato. Ad esempio, come è stato osservato da Legrenzi e Rumiati (1985), il fatto di presentare tre numeri *pari crescenti di due* "dice più del necessario" per una regola generale e "meno del necessario" per una regola molto specifica. Se si assume che in questo contesto possano valere le stesse regole che valgono negli scambi comunicativi, questa informazione può indurre i soggetti in inganno perché viola il cosiddetto "imperativo di Grice": il destinatario di un

messaggio assume che il mittente non dica nulla di più del necessario. Inoltre, i soggetti possono essere fuorviati dal fatto di avere a che fare con un esperimento di psicologia e che questo genere di serie numeriche caratterizza spesso certi test di intelligenza, ove ci si aspetta che il soggetto continui la serie avendo colto non una regola arbitrariamente scelta ma quella più "naturale".

Queste osservazioni indicano dunque che nell'utilizzare il compito di Wason non solo non si è data la giusta importanza ad un'analisi degli schemi usati dai soggetti per comprendere il compito ma anche che si è trascurata l'importanza della formulazione delle istruzioni di presentazione e soprattutto degli "esempi preliminari" che di solito vengono proposti per spiegarlo. Su queste attese concettuali, però, si può anche influire modificando opportunamente tali istruzioni.

4. INDAGINE EMPIRICA

4.1. Metodologia dell'indagine

Nella parte empirica dell'indagine, pur mantenendo inalterato l'originario compito 2-4-6 per avere la possibilità di operare dei confronti con le ricerche precedenti, sono state introdotte delle modifiche nelle istruzioni e delle modifiche nell'analisi dei risultati. Di queste ultime parleremo quando esporremo i risultati stessi, mentre ora esamineremo le modifiche alle istruzioni, che sono state volte a superare le difficoltà metodologiche sopra esposte.

a. Allo scopo di favorire, se possibile, una codificazione in termini espliciti e completi delle ipotesi, è stato richiesto in via preliminare ai Ss. di scrivere tutte le regole che venissero loro in mente (come già Tukey, 1986) e di farlo nel modo più chiaro possibile.

b. Per superare il problema della mancanza di corrispondenza fra ipotesi dichiarate ed esempi scelti, si è cercato di assicurarsi che i soggetti, nello scegliere gli esempi da testare, fossero consapevoli della coerenza o non coerenza degli esempi stessi con l'ipotesi o con l'attributo individuato. Ai Ss. è stato dunque richiesto di motivare per iscritto la scelta di ogni esempio, oltre che di indicare quale feedback si attendessero da ciascuno.

c. Per fare in modo che la decisione di enunciare la regola facesse riferimento esclusivamente alla necessità logica, ai Ss. è stata consentita dalle istruzioni un'unica enunciazione della regola stessa, solo quando essi ritenessero che fosse dimostrata (come in Gorman e Gorman, 1984); questo accorgimento aveva anche lo scopo di eliminare l'innaturale feedback su eventuali precedenti tentativi di enunciazione di ipotesi.

d. Per valutare l'influenza dell'esempio preliminare sul livello di generalità della regola, in tale esempio sono state usate regole più o meno generali a seconda del sottogruppo di appartenenza dei soggetti.

C'è da aggiungere che il compito è stato presentato come scoperta della legge che governa la risposta fornita da un computer quando nelle tre celle presenti sul monitor si trovino i numeri opportuni. In questo modo la regola viene più facilmente considerata una "legge di natura", meno legata all'arbitrio dello sperimentatore.

4.2. Procedura

I Ss. sono 50 studenti che partecipano volontariamente alla ricerca. 25 Ss., scelti a caso, costituiscono il gruppo di controllo e per loro il compito si svolge secondo la procedura originaria di Wason; la descrizione seguente riguarda dunque gli altri 25 Ss., che costituiscono il gruppo sperimentale.

Il compito viene presentato come la scoperta della regola incorporata nel programma del computer che, a seconda dei numeri presenti nelle tre celle visualizzate sul monitor, alla pressione di un tasto da parte del S. risponde scrivendo "SI" o "NO". Il monitor utilizzato dal S. si trova in una sala attigua a quella in cui è collocata l'unità centrale e il S. stesso non ha a disposizione una tastiera ma solo un grosso pulsante; i numeri da inserire nelle celle vengono comunicati allo sperimentatore che provvede alla modifica. Sul monitor del S. rimangono visualizzate man mano tutte le triplette fornite ed i rispettivi feedback; non vi compare alcun altro messaggio. Ai Ss. viene fornita la possibilità di scrivere qualsivoglia annotazione. Le istruzioni informano della natura della prova con un esempio preliminare che utilizza nomi di animali (cane, lupo, iena). Per 15 Ss., scelti casualmente, la regola è *generale* (tre animali qualunque); per gli altri 10 Ss. la regola è più particolare; per 5 è *basilare* (tre animali quadrupedi) e per 5 è *specifica* (tre animali quadrupedi carnivori). Per il resto le istruzioni sono uguali per tutti i Ss. del gruppo, richiedendo di pensare dapprima tutte le regole ipotizzabili e di scrivere ognuna di esse nel modo più chiaro possibile (viene specificato che dovranno essere comprensibili perché serviranno ad altri soggetti). Ugualmente per ogni tripletta formulata dovrà essere scritto il motivo della scelta e quale esito sia atteso dal test. I Ss. vengono avvertiti che l'enunciazione della regola dovrà essere fatta sulla base di solide ragioni che *dimostrino* la sua validità e che non sarà possibile poi enunciare altre regole.

4.3. Risultati e discussione

In primo luogo si è esaminato se le modifiche introdotte abbiano avuto qualche effetto sulla strategia dei soggetti. Dal momento che le istruzioni miravano ad ottenere dai soggetti del gruppo sperimentale la formulazione di ipotesi più articolate e strategie più coerenti, ci si attendeva che essi avrebbero adottato decisamente la strategia globale dell'esame successivo di ipotesi (*scanning*). Per valutare questa possibilità, è stato calcolato un indice in grado di rivelare la quantità di esempi che ripropongono per il test lo stesso attributo o operatore.

L'idea alla base della misura operativa che si è scelta per questa valutazione parte dall'osservazione che una strategia globale di scanning, essendo centrata su un'ipotesi, tende a far presentare un certo numero di esempi motivati dallo stesso attributo, finché l'ipotesi non cade, mentre una strategia di focusing conduce a variare continuamente gli attributi che motivano la scelta delle triplete. Si è perciò elaborata una misura, detta *indice di perseveranza*, della quantità di esempi che vengono presentati per testare uno stesso attributo, escludendo la prima volta in cui l'attributo stesso è proposto. Tale indice è stato ottenuto utilizzando la seguente formula:

$$\frac{\sum_{i=1}^n a_i - n}{\sum_{i=1}^n a_i} \cdot \sum_{i=1}^n a_i^2$$

(dove a sono le frequenze dei singoli attributi e n il numero degli attributi) cioè la differenza fra il numero di attributi totali e il numero di attributi diversi rapportata al numero di attributi totali. La formula è completata dal prodotto di questo rapporto per la sommatoria dei quadrati delle singole frequenze per ogni attributo, che serve come correttivo per dare un peso maggiore al numero di volte in cui uno *stesso attributo* è riproposto (altrimenti non si potrebbe distinguere, a parità di numero di attributi, fra il caso in cui gli attributi stessi abbiano pari frequenza e il caso in cui qualcuno abbia una frequenza molto elevata e qualche altro una frequenza molto bassa).

Si è stabilito di considerare i valori dell'indice di perseveranza, così calcolati, che risultassero al di sotto di quelli medi, come indicativi di una struttura globale tendente al focusing e quelli superiori alla media come indicativi di una tendenza verso lo scanning. Ciò premesso, se esaminiamo ora i risultati possiamo osservare che sono stati contrari alle aspettative: nonostante le modifiche alle istruzioni, non si sono riscontrate differenze significative fra gli indici di perseveranza dei due gruppi (v. tab. 1a). In entrambi, tali indici si sono mantenuti piuttosto bassi, segnalando la preferenza per una strategia centrata sull'attributo, o di focusing. Da questo risultato sembrano prendere consistenza i dubbi sul fatto che i soggetti abbiano sempre qualche ipotesi da verificare. Se le cose stanno così, appare preferibile non vedere la strategia globale di un soggetto in termini rigidi come centrata sull'ipotesi ma considerare tale variabile quasi come un continuum, cioè con possibilità di oscillazioni anche all'interno dello stesso protocollo. Nelle analisi successive abbiamo perciò evitato di considerare, come di solito, le *ipotesi* quali variabili dipendenti più importanti e abbiamo invece analizzato i dati in termini di *attributi* o *operatori* (vedi sez. 2.1).

Questa modalità di analisi è stata adottata, ad esempio, quando si è cercato di vedere se il diverso trattamento abbia avuto qualche effetto riguardo all'esplicitzza ed articolazione delle eventuali ipotesi. Non è semplice valutare empiricamente tale variabile, ma un indice sufficientemente indicativo può essere il numero di *attributi congiunti* menzionati dai soggetti nel proporre ogni tripletta da testare. Infatti se ad esempio un soggetto esprime la sua ipotesi, nel testare "8-10-12", menzionando gli attributi "pari" e "crescenti di 2", questa formulazione si può considerare più esplicita di una che menzioni solo "pari", in quanto in quest'ultimo caso il soggetto, pur testando di fatto entrambi gli attributi, mostrerebbe di essere consapevole solo di uno dei due. L'indurre i soggetti a formulare preliminarmente un certo numero di ipotesi in termini verbali, per iscritto, quindi in maniera articolata, ha avuto l'effetto di far aumentare in misura altamente significativa rispetto al gruppo di controllo il numero di attributi diversi menzionati nel proporre le triplette da testare (tab. 1b). Se si mette in relazione questo effetto con i risultati relativi al successo, emerge che il numero di attributi diversi menzionati (e quindi il grado di articolazione delle ipotesi) per i soggetti che scoprono la regola è significativamente superiore rispetto a quello di coloro che non vi arrivano (v. tab. 2a).

Si ricorderà che lo scopo principale della parte empirica della presente indagine era volto a considerare la possibilità che la tendenza alla conferma sia di ostacolo per il successo nel compito. Come accennato in precedenza, il gruppo sperimentale e quello di controllo sono stati differenziati non soltanto per la procedura ma anche per l'analisi dei protocolli. Per quanto riguarda i soggetti appartenenti al gruppo di controllo, sono state calcolate misure della conferma o disconferma che seguono i criteri standard di Wason. E' stato cioè calcolato un *indice di disconferma*, consistente nel rapporto - per ogni soggetto - fra il numero degli esempi incompatibili con l'ipotesi corrente o con l'attributo attualmente considerato e il totale di esempi prodotti (cfr. Gorman e Gorman, 1984, p.646), indipendentemente dalla risposta che il soggetto si attende. Analogamente, considerando la proporzione di esempi compatibili con l'ipotesi, è stato calcolato un *indice di conferma* (che ovviamente risulta essere il complemento all'unità del precedente). L'indice di disconferma è stato calcolato anche separatamente per i soggetti che hanno risolto il compito e per quelli che non lo hanno risolto. Per questo gruppo, con questo trattamento e questa analisi dei dati, risulta dunque che i soggetti che risolvono disconfermano significativamente più degli altri (v. tab. 2b).

Un risultato analogo sembrerebbe emergere per il gruppo sperimentale calcolando l'indice di disconferma allo stesso modo (tab. 2c); la disconferma risulta superiore rispetto all'altro gruppo soprattutto fra i soggetti che non risolvono. Questo è in linea, in sostanza, con quanto emerge dal confronto della tab. 1c: la manipolazione sembra aver prodotto più disconferma ma non più successo. Anche fermandosi qui nell'analisi, dunque, già si potrebbe ritenere che siano fondate le conclusioni a cui erano giunti Tweney e coll. (1980) e quanto era emerso da un nostro precedente esperimento (Greco, 1989), cioè che l'uso di strategie di conferma o disconferma è irrilevante ai fini della scoperta della regola.

Vi sono comunque altri indizi che conducono alla stessa conclusione. Poiché, come si è detto, nel presente esperimento era stato richiesto ai soggetti del gruppo sperimentale di indicare fra l'altro, per ogni tripletta formulata per il test, quale risposta si attendessero, per questi soggetti è stato possibile ottenere una misura della conferma o disconferma diversa da quella finora considerata. Questo indice *riveduto* non è basato sul numero di esempi compatibili o incompatibili con l'attuale ipotesi ma considera "disconfermanti" gli esempi che rispondano ai seguenti requisiti (vedi sez. 2.2.): a) sono in contrasto con l'ipotesi o meglio con l'attributo esaminato in un dato momento e la risposta attesa dal soggetto è "SI"; b) non sono in contrasto con tale ipotesi o attributo e la risposta attesa dal soggetto è "NO". Ovviamente risultano "confermanti" gli altri esempi. Se il confronto fra il punteggio di conferma dei soggetti che risolvono e di quelli che non risolvono il compito viene fatto utilizzando questo indice *riveduto*, i soggetti che risolvono sembrano disconfermare leggermente più degli altri, ma non esiste più una differenza significativa (v. tab. 2d).

Questo fatto può essere spiegato se si esaminano in maniera analitica i punteggi che contribuiscono agli indici di disconferma calcolati nei due modi. Come si vede dalla tab. 3, le triplette più frequenti, dopo quelle *compatibili con attesa SI*, sono quelle *incompatibili con attesa NO*: utilizzando la procedura di Wason questo genere di esempi non poteva essere rilevato e dunque non poteva essere correttamente considerato un caso di ricerca di conferma; in quanto *incompatibili*, invece, queste triplette andavano anzi ad incrementare in maniera non realistica il punteggio di disconferma.

Sulla base di questi rilievi a nostro parere si può dubitare di quei dati empirici che misurano le strategie di conferma o disconferma senza tenere conto delle attese dei soggetti e rivedere le conseguenze che ne traggono riguardo al rapporto fra tali strategie e il successo. In particolare, questi risultati avvalorano ancora una volta le conclusioni delle ricerche sopra citate (Tweney e coll., 1980; Greco, 1989) che mettono in dubbio la connessione fra conferma o disconferma e scoperta della regola.

Le variabili più rilevanti, se ve ne sono, vanno dunque ricercate altrove. A titolo esplorativo, si può prendere in considerazione la possibilità che il successo sia legato all'uso di strategie più globali della conferma o disconferma. In effetti, il fatto che l'indice di perseveranza dei soggetti che hanno risolto il compito risulta significativamente inferiore a quello dei soggetti che non lo hanno risolto (v. tab. 2e) può indicare che i soggetti che individuano la regola preferiscono una strategia centrata sull'attributo, o di focusing. D'altra parte questo non sarebbe sorprendente, dal momento che già Bruner et al. (1956) avevano osservato che questa strategia richiede minori risorse cognitive, in termini di capacità mnemonica e inferenziale, rispetto all'altra. Ciò che sorprende è come mai le ricerche che hanno adottato il paradigma di Wason abbiano dato per scontato l'uso della strategia di scanning.

L'ultima valutazione relativa ai risultati di questa indagine empirica ha riguardato l'incidenza della manipolazione del livello di generalità dell'esempio preliminare. Questo cambiamento non ha prodotto un miglioramento significativo

nella performance (v. tab. 4a), ma i soggetti dei sottogruppi il cui esempio preliminare era *basilare* e *specifico* hanno esaminato un numero di attributi significativamente superiore (tab. 4b e 4c) e hanno proposto più frequentemente, quale regola finale, una regola che comprendeva quella corretta ma più specifica (ad es. numeri crescenti e pari, ecc.; v. tab. 4e). Questi ultimi risultati probabilmente sono da considerarsi provvisori e non generalizzabili dato l'esiguo numero di soggetti che, per contingenti motivi di organizzazione dell'esperimento, hanno potuto essere inclusi in questi sottogruppi. Ci sembra però che, anche se non si può dire che i soggetti il cui esempio era generale siano stati facilitati, si può affermare che gli altri sono stati di certo ostacolati. Un dato che colpisce particolarmente (tab. 4d) è l'altissima proporzione di esempi negativi (cioè che non seguono la regola da scoprire) incontrati dai soggetti del sottogruppo il cui esempio era *specifico*, che pure non è servita a facilitarli, proprio perché essi erano fuorviati dall'attesa di una regola molto specifica.

5. CONCLUSIONE

Nella presente indagine abbiamo considerato alcuni problemi che inducono a dubitare della tesi secondo la quale gli esperimenti basati sul compito 2-4-6 dimostrano che la tendenza alla conferma (*confirmation bias*) ostacola il processo di scoperta induttiva. Le ragioni di tale dubbio risiedono essenzialmente in alcune difficoltà dell'usuale impostazione metodologica di tale compito. In particolare, non appare sostenibile parlare di conferma o disconferma solo facendo riferimento a delle *ipotesi* sulla regola quando appare evidente che i soggetti tendono semplicemente a non ragionare esaminando ipotesi successive (cioè ad adottare una strategia di *scanning*), neppure quando - come nella presente indagine empirica - venivano incoraggiati e facilitati a farlo. Sembra, piuttosto, preferibile valutare la conferma o disconferma riferendosi agli attributi o operatori esaminati man mano negli esempi testati (senza considerarli necessariamente "ipotesi" sulla regola) e tenendo conto di quali risposte i soggetti si attendono dai loro test.

I risultati indicano che comunque non appare esserci relazione fra l'uso della conferma o disconferma ed il successo o l'insuccesso nel compito. Si osservi che ciò non equivale a dire che non è dimostrata l'esistenza del *confirmation bias*; anzi, misurando tale variabile nel modo sopra proposto (cioè considerando confermantici anche le triplette incompatibili con gli attributi esaminati quando la risposta attesa sia "no") appare evidente che la tendenza alla conferma è superiore a quanto emergeva con l'analisi tradizionale. Tuttavia continua a sfuggire che cosa in realtà faciliti o ostacoli nella scoperta della regola, ammesso che vi sia qualcosa oltre la pura "fortuna", come erano giunti ad affermare Tweney e coll. (1980) concludendo il loro esperimento II.

Una conclusione di questo genere però, a nostro parere, è inevitabile solo se si continua a trascurare di compiere un'analisi complessiva del processo cognitivo che è messo in atto nei compiti che richiedono un'indagine induttiva. Si tratta di

un processo (che potremmo definire di "uso dell'evidenza" o, se si preferisce, "stile indagativo") che, si può supporre, ha un'elaborazione indipendente dal ragionamento induttivo vero e proprio, poiché la sua adozione è legata anche a presupposti impliciti, stili individuali, ecc. che non sono direttamente connessi con quanto si sta specificamente ricercando. Nonostante la sostanziale mancanza di analisi di questo processo, non si esitava ad affermare che i risultati degli esperimenti con il compito 2-4-6 dimostravano che la tendenza alla conferma ostacolasse la scoperta della regola. Ci sembra invece che un'indagine sulle reali cause di difficoltà o di facilitazione, in questo come in simili contesti, dovrebbe tener conto delle procedure che i soggetti mettono in atto nell'affrontarlo e delle aspettative che possono fuorviarli.

In termini più generali, il problema è in sostanza di studiare come interagiscono attese, modelli mentali, modalità di gestione delle risorse cognitive, procedure, per costituire quei sistemi standardizzati di uso dell'evidenza che si chiamano "strategie". Ciò che è soprattutto importante è tenere presente che le strategie generali (di scanning/focusing o di conferma/falsificazione) non sono opzioni esterne al compito che vengono scelte dai soggetti per guidarne lo svolgimento ma fanno parte integrante della rappresentazione del compito stesso. In questo contesto più ampio si comprendono meglio alcune questioni che sono state discusse nel presente lavoro, come la difficoltà di valutare l'intenzione di confermare o disconfermare in riferimento a "ipotesi" piuttosto che ad "attese": in questo caso, ad esempio, le strategie potrebbero essere meglio concettualizzate come quei particolari tipi di schemi detti script (Schank e Abelson, 1977) e la difficoltà potrebbe dipendere dal fatto che a seconda del tipo di script utilizzato sono attese solo certe evidenze e d'altra parte ogni script prevede solo un certo uso delle evidenze stesse. Ci sembra che una futura esplorazione, condotta in questo senso, dei compiti di scoperta induttiva possa condurre a risultati interessanti.

TABELLA 1. Confronto fra il gruppo sperimentale e quello di controllo.

	S	C	t	g.l.	p
a. Indice di perseveranza	4.79	6.75	-0.54	48	0.59 (bilat.)
b. Numero di attributi diversi	5.4	3.7	2.47	48	0.001
c. Successo (percentuale)	44.0	40.0	0.28	48	0.78 (bilat.)

S = gruppo sperimentale; C = gruppo di controllo

TABELLA 2. Confronto fra Ss. che risolvono e Ss. che non risolvono il compito.

	m	R	NR	t	g.l.	p
a. Numero attributi diversi (gruppo sperimentale)	5.4	6.7	4.5	1.87	23	0.037
b. Indice di disconferma (gruppo controllo)	19.0	31.4	5.6	2.60	23	0.002
c. Indice di disconferma (gruppo sperimentale)	24.0	32.0	18.7	1.39	23	0.089
d. Indice di disconf. riveduto (gruppo sperimentale)	18.4	21.6	16.3	0.77	23	0.22
e. Indice di perseveranza	4.8	2.2	6.4	-1.37	23	0.091

R = Ss. che risolvono; NR = Ss. che non risolvono; m = media

TABELLA 3. *Composizione percentuale delle triplete testate nel gruppo di controllo e relative attese.*

Strategia		R	NR	m	Analisi	
test	attesa				trad.	rived.
C	SI	54.8	71.2	64.6	con.	con.
C	NO	13.1	10.2	11.3	con.	fal.
I	SI	8.5	6.1	7.1	fal.	fal.
I	NO	23.6	12.5	17.0	fal.	con.
Totale		100	100	100		

C = compatibili; I = incompatibili

R = Ss. che risolvono; NR = Ss. che non risolvono; m = media generale

con. = confermanti; fal. = falsificanti

TABELLA 4. *Confronto fra gruppi con istruzioni di diverso livello di generalità.*

	Regola				F	Test stat.	
	g	b	s			g.l.	p
a. Successo (proporzione)	0.33	0.60	0.40	0.52	2		non sig.
b. Numero medio attributi menzionati complessivamente	5.87	8.20	9.20	3.78	2		< 0.03
c. Numero medio attributi diversi menzionati	4.13	6.80	7.80	4.55	2		< 0.02
d. Proporzione di esempi negativi incontrati	14.5	19.0	40.2	4.68	2		< 0.02
e. Insuccesso finale con ipotesi più specifica di quella corretta	g	bs	t	g.l.	p		
	0.07	0.30	-1.57	23	0.06		

g = regola generale; b = regola basilare; s = regola specifica;

bs = regola basilare o specifica

BIBLIOGRAFIA

- BRUNER J.S., GOODNOW J.J., AUSTIN G.A., BROWN R.W. (1956).
A study of thinking. New York, Wiley & Sons. Trad. ital. *Il pensiero*, Roma, Armando, 1969.
- GORMAN M.E., GORMAN M.E. (1984).
 A comparison of disconfirmatory, confirmatory and control strategies on Wason's 2-4-6 task. *Quarterly Journal of Experimental Psychology* 36A 629-648.
- GRECO, A. (1989).
 Confirmation bias e feedback nel compito 2-4-6. *Ricerche di Psicologia* 115-130.
- GREEN D.W. (1989).
Confirmation bias, problem-solving and cognitive models. (In corso di pubbl.)
- JOHNSON-LAIRD P.N. (1983).
Mental models. Cambridge, Cambridge University Press. Trad. ital. *Modelli mentali*, Bologna, Il Mulino, 1988.
- KLAYMAN J., HA Y.W. (1987).
 Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review* 94 211-228.
- LEGRENZI P., RUMIATI R. (1985).
 Regole in Kripke: sviluppi sperimentali in psicologia. *Nuova civiltà delle macchine*, 3-4, 55-59.
- MINSKY M. (1975).
 A framework for representing knowledge. In Winston, P.H. (ed.), *The psychology of computer vision*, New York, Mc Graw-Hill.
- MONETA G.B. (1987).
 Are psychological experiments on hypothesis-testing strategies a good simulation of scientific problem-solving? *Epistemologia* X 29-54.
- POLANYI M. (1966).
The tacit dimension. Tr.it.: *La conoscenza inespresa*, Roma, Armando, 1979.
- POPPER K.R. (1959).
The logic of scientific discovery. New York, Basic Books. Trad. ital. *La logica della scoperta scientifica*, Torino, Einaudi, 1970.
- ROSCH E. (1975).
 Cognitive representation of semantic categories. *Journal of Experimental Psychology: General* 104 129-233.
- SCHANK R.C., ABELSON R.P. (1977).
Scripts plans goals and understanding. Hillsdale, N.J., Erlbaum Ass.
- TUKEY D.D. (1986).
 A philosophical and empirical analysis of subjects' modes of inquiry in Wason's 2-4-6 task. *Quarterly Journal of Experimental Psychology* 38A 5-33.
- TWENEY R.D., DOHERTY M.E., WORNER W.J., PLISKE D.B., MYNATT C.R. (1980).
 Strategies of rule discovery in an inference task. *Quarterly Journal of Experimental Psychology* 32 109-123.
- WASON P.C. (1960).
 On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12 129-140.
- WASON P.C., JOHNSON-LAIRD P.N. (1968).
 On the failure to eliminate hypotheses... A second look. In Wason P.C., Johnson-Laird P.N. (Eds.), *Thinking and reasoning*, Middlesex, Penguin.
- WASON P.C., JOHNSON-LAIRD P.N. (1972).
Psychology of reasoning. London, Batsford. Trad. ital. *Psicologia del ragionamento*, Firenze, Martello Giunti, 1977.

WETHERICK N.E. (1962).

Eliminative and enumerative behavior in a conceptual task. *Quarterly Journal of Experimental Psychology* 14 246-249.

SUMMARY. *Hindrances due to confirmation bias in inductive discovery of a rule is questioned. It is claimed that, as confirmation bias is usually assessed by considering subject's examples consistent with its current hypothesis, confirmation cannot be recognized when a subject has not such a hypothesis (as in focusing strategy) or has not clearly stated it. Instead, subjects' expectations should be more considered to better assess strategy, since it is possible to propose examples consistent with one's hypothesis in order to disconfirm it, or, conversely, to propose inconsistent ones to confirm. More, task setting do not allow subjects to build a clear task representation, and disregards subjects' previous expectations. An experiment is presented in which instructions were modified to partly overcome these problems. Results show that using confirmatory or disconfirmatory strategy is not so crucial in succeeding as using focusing strategy is; further, the importance of preliminary instructions in rule discovering is showed. The need for an overall analysis of cognitive processes of using of evidence is eventually maintained.*

RIASSUNTO. *Le indagini che adottano il compito 2-4-6 intendono mostrare che la scoperta induttiva di una regola è ostacolata dalla tendenza dei soggetti a scegliere esempi che confermano le proprie ipotesi. Nel presente lavoro vengono esposti alcuni problemi che inducono a mettere in dubbio tale assunto. Due primi aspetti sono connessi al fatto che la tendenza alla conferma di solito è valutata considerando "confermanti" gli esempi compatibili con l'ipotesi che un S. stava seguendo quando li ha proposti: in primo luogo, non è possibile dire se un esempio sia addotto a conferma o a disconferma nei casi in cui tale ipotesi non può essere individuata con chiarezza; in secondo luogo, ciò che comunque rende un esempio confermate o falsificante è quello che il S. si aspetta nel proporlo e difatti è possibile scegliere esempi in contrasto con la propria ipotesi per confermarla o viceversa esempi congruenti con essa ma per falsificarla. Un secondo gruppo di problemi è connesso con l'artificiosità del compito. In particolare si osserva che le istruzioni non favoriscono la costruzione da parte dei Ss. di un chiaro modello mentale e che non si tiene conto delle attese concettuali che essi hanno già riguardo a compiti di questo genere. Viene poi presentato un esperimento nel quale si è tentato di superare questi limiti metodologici. I risultati indicano che l'uso di strategie di conferma o falsificazione non influisce sul successo; appaiono invece rilevanti l'uso di una strategia globale di focusing e la manipolazione delle attese dei soggetti tramite le istruzioni preliminari. In conclusione si sostiene che non si può esaminare l'incidenza delle strategie di conferma o disconferma sulla scoperta induttiva senza un'analisi complessiva del processo di uso dell'evidenza, che tenga conto sia delle modalità di rappresentazione del compito (di cui le strategie sono solo una parte) che del modo in cui il compito stesso viene formulato.*