

ESTRATTO

RICERCHE DI PSICOLOGIA

Trimestrale diretto da Marcello Cesa-Bianchi

FAE Riviste s.r.l.

V.le Monza 106 - 20127 Milano - Sped.abb.post.gr. 4°/70



CONFIRMATION BIAS E FEEDBACK NEL COMPITO 2-4-6

CONFIRMATION BIAS AND FEEDBACK IN 2-4-6 TASK

di Alberto Greco

Laboratorio di Psicologia
Dipartimento di Filosofia
Università degli Studi di Genova

Introduzione

La presente indagine riguarda alcuni aspetti del *ragionamento induttivo consapevole*, cioè quel ragionamento che parte dall'esame di dati per scoprire una regola o legge che ne spieghi le realizzazioni. (E' escluso, quindi, il processo di generalizzazione che permette di scoprire intuitivamente o inconsapevolmente una regola, come avviene ad esempio nell'apprendimento di una lingua). Uno dei motivi che rende particolarmente interessante questo tema e che ha contribuito allo svilupparsi di indagini su di esso è indubbiamente il fatto che il ragionamento induttivo è alla base dell'attività cognitiva utilizzata nel fare scienza.

Il contesto sperimentale classico entro cui si svolgono le ricerche su questo tema prevede che ai soggetti venga presentato un esempio, consentendo o chiedendo loro esplicitamente di formulare ipotesi sulla regola; la correttezza di tali ipotesi potrà poi essere provata esaminando ulteriori esempi che seguono o non seguono la regola. Secondo alcune procedure questi ulteriori esempi da sottoporre a verifica sono scelti dallo sperimentatore, mentre talvolta è richiesto allo stesso soggetto di formularli. In ogni caso lo sperimentatore ad ogni formulazione di un esempio fornisce un *feedback* consistente nel comunicare se esso segue o no la regola-*target*. Questo feedback costituisce la fonte di "esperienza" per i soggetti impegnati nella prova, un'esperienza paragonabile a ciò che per uno scienziato è il risultato di un esperimento.

Fra le ricerche che adottano la procedura di richiedere allo stesso soggetto la formulazione degli esempi da verificare, è ormai classica la serie a cui ha dato il via Wason (1960), che utilizzava il compito "2-4-6". L'interesse di tale compito sta nel fatto che, essendo la regola molto generale (tre numeri in ordine crescente), è difficile per il soggetto trovare triplette che smentiscano ipotesi su regole più specifiche e molto più ovvie e immediate. Il compito, dunque, consente di analizzare le strategie utilizzate quando si avrebbero facilmente spiegazioni che reggono al

confronto con l'esperienza ma che sono sbagliate: un caso che può verificarsi con una certa facilità nel corso dell'indagine scientifica. Lo stesso Wason e quasi tutti gli autori successivi hanno evidenziato l'affinità fra questo risultato e le note posizioni epistemologiche di Karl Popper (1959), secondo le quali nel progresso della ricerca scientifica ha un ruolo cruciale la falsificazione delle ipotesi piuttosto che la loro conferma.

Nelle intenzioni di Wason, dunque, il punto di interesse principale del compito non era tanto nel procedimento induttivo in quanto tale, ma nella strategia che i soggetti avrebbero adottato. Wason trovò che la maggior parte dei soggetti in questo compito esibiva una "tendenza alla conferma" (*confirmation bias*), piuttosto che alla disconferma, nello scegliere gli esempi da sottoporre a test. Si osservi che *confermare* o *disconfermare* per Wason significa esclusivamente proporre triplete di numeri rispettivamente compatibili o incompatibili con l'ipotesi che un soggetto in un dato momento ha in mente.

Presupposti della ricerca

Se, come si è visto, Wason era interessato originariamente soprattutto alle strategie utilizzate dai soggetti più che alla scoperta della regola, anche le successive ricerche che hanno ripreso il paradigma di Wason sono state essenzialmente mirate a cogliere i motivi per cui la strategia di falsificazione si innesca oppure no. Non è detto, però, che la strategia (di conferma o falsificazione) sia l'unico fattore che influisce sulla scoperta della regola. Si possono formulare anche altre ipotesi di lavoro su tali fattori, includendovi ad esempio la possibilità di seguire due ipotesi contemporaneamente, anziché una sola, come suggeriva Platt (1964) a proposito della ricerca scientifica. Questa possibilità è stata presa in esame da una ricerca di Tweney et al. (1980) che si inseriva nell'ambito di una serie (Mynatt et al., 1977, 1978) volta ad approfondire il fenomeno della tendenza alla conferma (*confirmation bias*).

Nei primi tre esperimenti riportati, questi autori hanno tentato di manipolare tramite istru-

zione le strategie dei soggetti, inducendo una consapevolezza della possibilità di disconferma (nel senso di Wason, come test di esempi in contrasto con la propria ipotesi), oppure istruendo i soggetti a lavorare su più ipotesi contemporaneamente. Con queste manipolazioni, tuttavia, pur riuscendo a modificare le strategie dei soggetti, essi non sono riusciti a migliorare le prestazioni nella scoperta della regola. Ne emergeva che è possibile indurre attraverso le istruzioni l'uso di una strategia di falsificazione, ma che questo di per sé non garantisce il successo nella scoperta della regola. Per di più, se si istruiscono esplicitamente i soggetti a seguire due ipotesi contemporaneamente, anziché una sola, seguendo il suggerimento di Platt, ciò riesce particolarmente difficile ai soggetti ed anzi peggiora la loro performance.

C'era tuttavia una quarta condizione sperimentale che consentiva ai soggetti di Tweney e coll. di scoprire la regola molto più agevolmente: quando il compito veniva presentato (seguendo un'idea che già era stata di Wetherick, 1962) come la ricerca delle due regole seguite da esempi classificabili in due categorie attraverso due etichette verbali (DAX e MED, essendo DAX la categoria delle triplete di numeri in ordine crescente - la classica regola-target - e MED tutte le altre) si otteneva un miglioramento notevole nella scoperta della regola al primo tentativo, sebbene il *confirmation bias* non fosse eliminato. In queste condizioni, secondo gli autori della ricerca, i soggetti potevano seguire con naturalezza due ipotesi alternative e l'uso di due etichette avrebbe permesso di valutare correttamente le informazioni negative: nel caso in cui gli esempi scelti non appartenessero alla categoria DAX (la stessa di 2-4-6) questa informazione diventava rilevante per la categoria MED anziché essere un dato da ignorare in quanto "sbagliato". Sulla base dei risultati dell'intera serie di esperimenti Tweney e coll. concludevano che ai fini del successo nel compito non importa tanto la strategia seguita (conferma o disconferma) quanto la modalità di concettualizzazione del problema: nell'ultimo esperimento la performance migliorava in quanto i soggetti potevano riferire la conferma o la non conferma otte-

nuta all'una o all'altra di due ipotesi alternative.

Se si esamina la procedura di Tweney e coll. da un punto di vista metodologico, l'effetto risulta ascrivibile in sostanza alla manipolazione del *feedback*, cioè del modo in cui viene risposto dallo sperimentatore nel momento in cui il soggetto comunica una tripletta-test. Tuttavia gli stessi autori hanno riconosciuto che il motivo di questo fenomeno non era del tutto chiaro ed hanno rinviato la discussione ad indagini future. Si può congetturare che tale manipolazione probabilmente desse modo ai soggetti di sfruttare le potenzialità cognitive offerte dal linguaggio o dalla categorizzazione, ma in realtà quale effetto avesse esattamente non era chiaro.

Uno dei motivi di questa mancanza di chiarezza, a nostro parere, era una confusione presente nella procedura di Tweney e coll.. Infatti, dato il modo in cui il *feedback* veniva fornito in quell'esperimento, l'effetto poteva essere dovuto ad una qualunque di tre variabili che non era possibile separare: il fatto che il *feedback* stesso implicasse una "categorizzazione", oppure che facesse uso di "etichette verbali", o ancora che non avesse il significato di "sbagliato" evitando l'uso di una negazione. Gli autori hanno considerato ora l'una ora l'altra di queste variabili senza evidenziare l'incidenza di ciascuna. Accettando provvisoriamente il presupposto che il successo nel compito sia legato, in qualche modo da chiarire, al tipo di *feedback* ottenuto dai soggetti per le triplette da loro testate, abbiamo allestito un esperimento in cui queste variabili fossero esplicitamente separate. Lo scopo dell'esperimento, tuttavia, andava al di là del semplice proposito di evidenziare quale di queste variabili incidesse sul risultato, in quanto si proponeva di analizzare anche in termini più generali le relazioni fra istruzioni, strategie e successo.

Procedura

48 soggetti (dopo una prova preliminare volta all'accertamento della perfetta comprensione delle istruzioni e che seguiva la procedura di Tweney e

coll. usando triplette di nomi di animali) sono stati assegnati casualmente a 4 condizioni sperimentali e sottoposti al compito 2-4-6 ricevendo, a seconda del gruppo di appartenenza, un diverso feedback quando enunciavano le triplette di numeri da loro pensate per controllare la propria ipotesi sulla regola. Le istruzioni iniziali erano ovviamente congruenti con le diverse modalità di svolgimento dell'esperimento. Ai Ss. assegnati al gruppo 1 veniva dato un feedback che sottolineava la correttezza o non correttezza delle triplette scelte, usando le esplicite etichette verbali GIUSTO/SBAGLIATO. Nella condizione 2 veniva replicato l'esperimento IV di Tweney e coll. presentando esplicitamente il compito come un problema di scelta fra due categorie, contrassegnate anche qui da due etichette verbali (DAX/MED). Nella condizione 3 veniva, come per il primo gruppo, sottolineato l'aspetto di correttezza o non correttezza degli esempi scelti ma il feedback veniva fornito in termini non verbali, usando due segnali luminosi di diverso colore e facendo attenzione a non definirli mai verbalmente nel corso dell'esperimento. Ai Ss. appartenenti al gruppo 4, infine, il feedback, come nella condizione 2, classificava gli esempi in due categorie, alle quali stavolta si faceva riferimento in termini non verbali usando i due segnali luminosi.

Risultati e discussione

Le variabili dipendenti prese in considerazione (v. tab. 1) sono:

a) la proporzione di soggetti che scoprono la regola (*successo*);

b) i valori medi dell' *indice di disconferma* calcolato per ciascun protocollo; tale indice consiste nel rapporto fra il numero di triplette incompatibili con l'ipotesi considerata dal soggetto quando le proponeva e il numero totale di triplette del protocollo (questo rapporto è preferibile al numero assoluto di triplette disconfirmatorie per evitare di ritenere che disconfermino di più Ss. che semplicemente propongono molte triplette: per questa misura cfr. Gorman e Gorman, 1984, p.646);

Tab. 1 - Risultati generali

feedback	gruppo			
	1	2	3	4
verbale	*	*		
non-verbale			*	*
correttezza	*		*	
categoriale		*		*

	1	2	3	4
a. Successo	12.5	14.6	10.4	12.5
b. Indice di disconferma	17.0	22.6	11.0	8.9
c. Proporz. di esempi negativi	8.7	12.5	14.2	19.8

Tab. 2 - Dati raggruppati

	condizione		media gen.	condizione	
	V	NV		COR	CAT
a. Successo	54.2	45.8	50.0	45.8	54.2
b. Ind. disconf.	19.8	9.9	14.9	14.0	15.8
c. Prop. es. neg.	10.6	17.0	13.8	11.4	16.2

V = verbale ; NV = non verbale
COR = correttezza ; CAT = categoriale

Le differenze fra le medie riportate in corsivo
sono significative (v. tab. 4)

c) i valori medi della proporzione di esempi negativi calcolata per ciascun protocollo: si tratta del rapporto fra il numero di triplette che non seguono la regola-target e il numero totale di triplette (la ragione dell'uso del rapporto è la stessa indicata al punto b).

I risultati non consentono di chiarire in maniera decisiva se per la scoperta della regola sia più importante l'uso della disconferma o la natura del feedback ricevuto. Infatti, da una parte, la manipolazione del feedback sembra produrre un effetto di rilievo sulla strategia dei soggetti, nel senso che coloro i quali hanno ricevuto il feedback verbale hanno ottenuto nell'indice di disconferma valori significativamente più alti rispetto agli altri (tabb. 2b, 4b). Inoltre, i soggetti che hanno ricevuto il feedback verbale e che risolvono hanno un indice di disconferma significativamente superiore rispetto a quelli che non risolvono (tab. 4d) mentre ciò non accade nel caso di feedback non verbale (tab. 4e). D'altra parte, non risulta esservi una rilevante differenza nella proporzione di soggetti che hanno ricevuto il feedback verbale e che risolvono rispetto ai soggetti delle altre condizioni (tabb. 3a, 4a). In realtà, dal punto di vista del successo non si sono riscontrate differenze significative fra nessuna delle quattro condizioni (tab. 1a), anche se la tendenza è quella prevista (cioè una leggera prevalenza del successo nella condizione 2, che replicava il IV esperimento di Tweney e coll.).

Questi dati inducono a ritenere dunque che l'uso di etichette verbali per il feedback favorisca lo sviluppo di strategie di falsificazione, senza tuttavia che ciò faciliti la scoperta della regola, almeno non nei termini vistosi riscontrati da Tweney e coll.. Ancora una volta, piuttosto, sembra che non vi sia un nesso diretto fra la strategia e la soluzione come nei primi tre esperimenti di quegli autori; una conferma di questo fatto emerge anche se si valutano i risultati dell'esperimento nel complesso, indipendentemente dalle manipolazioni compiute nelle quattro condizioni: i soggetti che scoprono la regola sono esattamente la metà ma il loro indice di disconferma è nettamente e significativamente superiore a quello degli altri (v. tab. 4c).

Tab. 3 - Successo e disconferma nelle condizioni verbale/non verbale

a. Proporz. successo condiz.verbale			condiz.non verb.			Totale gener.
Totale	SR	SNR	Totale	SR	SNR	
50.0	27.1	22.9	50.0	22.9	27.1	100.0

b. Indice di disconferma condiz.verbale			condiz.non verb.			Media gener.
Media	SR	SNR	Media	SR	SNR	
19.8	31.3	6.1	9.9	12.0	8.2	14.9

SR = soggetti che risolvono;
SNR = soggetti che non risolvono

Le differenze fra le medie riportate in corsivo sono significative (v.tab.4)

Tab. 4 - Significatività delle differenze

	Media	V	NV	t	gl	sig.
a.Successo	50.0	54.2	45.8	0.57	46	0.29
b.Indice di disconf.	14.9	19.8	9.9	1.52	46	0.067 *

	Media	SR	SNR	t	gl	sig.
c.Indice di disconf.	14.9	22.5	7.3	2.46	46	0.009 *
d.Indice di disconf. nella condiz. verb.	19.8	31.3	6.1	2.44	22	0.011 *
e.Indice di disconf. condiz.non verb.	9.9	12.0	8.2	0.61	22	0.27
f.Proporz. es. neg.	13.8	22.8	4.8	4.93	46	<0.0001 *

V = Condiz. verbale; NV = Condiz. non verbale
SR = soggetti che risolvono;
SNR = soggetti che non risolvono

* Valori significativi al livello prescelto (p < 0.10)

Una variabile che invece appare legata in maniera inequivocabile al successo in questo compito è la *proporzione di esempi negativi* (cioè di quelle triplette che non seguono la regola-target), nettamente superiore fra i soggetti che risolvono (tab. 4f). Dall'esame dei protocolli, del resto, emerge chiaramente che fra tutti i soggetti impegnati nell'esperimento coloro che hanno avuto l'opportunità di formulare almeno un esempio negativo sono stati esattamente la metà (24 su 48) e di questi ben 20 sono giunti alla regola, mentre solo altri 4 soggetti sono riusciti a farlo senza aver mai incontrato un esempio negativo. I dati del nostro esperimento non consentono dunque di valutare l'effetto della manipolazione di Tweney e coll., per il semplice fatto che è sì possibile sostenere, come questi autori fanno, che la categorizzazione o l'uso di etichette verbali permetteva ai soggetti di non perdere informazioni preziose classificandole come "sbagliate", ma solo nel caso in cui gli esempi scelti fossero negativi. Molti soggetti non giungono a formulare nessuno di questi esempi e quindi non possono usufruire di tali informazioni.

Dai dati sopra esposti sembra dunque lecito assumere che l'opportunità di trovare almeno un esempio negativo sia un fattore essenziale per la riuscita nel compito. Ma è anche interessante analizzare in che modo i soggetti che hanno risolto il compito hanno avuto l'opportunità di trovare la prima tripletta negativa: si scopre allora che ben 13 su 20 (65 %) l'hanno formulata mentre seguivano un'altra ipotesi, ipotesi che nel 69 % dei casi non implicava in nessun modo quella corretta. Insomma l'informazione decisiva sembra essere stata ottenuta praticamente per caso. Tweney e coll. erano giunti ad una conclusione simile quando, concludendo il loro esperimento II, affermavano che il miglior predittore del successo sembrava essere la "fortuna".

In realtà, se è difficile individuare in maniera conclusiva le condizioni che favoriscono la scoperta della regola in questo compito, da una analisi qualitativa dei protocolli emergono certe particolarità del compito stesso che meritano di essere approfondite in quanto costituiscono sicuri ostacoli per il successo. In primo luogo si osserva molto spesso che i soggetti testano esempi

senza aver formulato chiaramente alcuna ipotesi. In molti casi questi esempi nascono come manipolazione, più o meno sistematica, delle triplette precedenti, con lo scopo di esplorare semplicemente l'effetto della variazione di qualche attributo (ad es. vedere se c'è un limite superiore, se i numeri dispari hanno influenza, ecc.). In altri termini, i soggetti talvolta preferiscono una strategia centrata sull'attributo (quella che Bruner, 1956, definiva *focusing*) piuttosto che sull'ipotesi. Wason aveva inventato il compito 2-4-6 per riprodurre e studiare meglio dei problemi tipici della strategia di *scanning* (cfr. Wason e Johnson-Laird, 1972, pp.245 sgg. tr.it.) ma tutti coloro che hanno utilizzato tale compito non hanno mai preso seriamente in considerazione il fatto che un certo numero di soggetti, a volte cospicuo, preferisce invece usare una strategia di *focusing* (eccetto Tukey, 1986, che ha chiaramente messo in evidenza questo problema ed ha distinto le strategie dei suoi soggetti centrate sull'ipotesi da quelle centrate sull'attributo o sull'evidenza).

L'analisi qualitativa dei protocolli rivela anche altri esempi di mancanza di corrispondenza fra le ipotesi dichiarate e gli esempi scelti per la loro verifica. Infatti, quando si controlli quale ipotesi possa essere effettivamente confermata (o falsificata) da ogni tripletta proposta dai soggetti, emerge con netta evidenza la tendenza dei soggetti a formulare esempi che contengono caratteristiche non previste dall'ipotesi in corso di verifica. Succede tipicamente che un soggetto enunci l'ipotesi "numeri pari" e poi scelga la tripletta 8-10-12, che sono non soltanto numeri pari ma in cui è conservata la *differenza di 2* fra le cifre e l'*ordine crescente*; una tripletta come 8-2-12, ad esempio, invece avrebbe testato l'ipotesi "numeri pari" in maniera più mirata ed informativa.

Nel compito "2-4-6" sembra dunque in primo luogo che i soggetti non sempre desiderino o siano in grado di formulare ipotesi e comunque, quando lo fanno, che siano incapaci di attenersi strettamente all'ipotesi formulata non riuscendo a liberarsi dell'interferenza di altre ipotesi implicite, che sembrano aver ricavato automaticamente dai dati di partenza senza esserne consapevoli. Queste osservazioni alimentano il sospetto che la

manca di razionalità nella strategia seguita dai soggetti possa essere ascritta al fatto che il contesto sperimentale non richiede tale razionalità in maniera chiara ed esplicita, piuttosto che ad un *bias* del ragionamento (e questa è anche una notevole differenza fra il compito 2-4-6 e l'induzione tipica dell'indagine scientifica: cfr. Moneta, 1987).

I problemi principali che devono essere superati da chi vuole far uso dei risultati ottenuti in questo compito per dimostrare la presenza del *confirmation bias* nel ragionamento induttivo sono proprio legati da una parte all'estrema particolarità e innaturalità della situazione, dall'altra al fatto che un'analisi dei risultati compiuta solo in termini di strategie (confirmatorie o disconfirmatorie) si rivela parziale in quanto non comprende un'analisi complessiva dei processi cognitivi implicati in questo tipo di attività.

Ad esempio, si può supporre che l'analisi dell'esempio iniziale sia influenzata da alcune attese relative al modo di concettualizzare il compito, delle specie di *frames* (nel senso di Minsky, 1975) che contengono informazioni "per difetto" su quanto non precisato esplicitamente. Tali attese potrebbero riguardare sia il livello di generalità della regola (in assenza di informazioni contrarie, si assume che gli elementi da categorizzare appartengano ad una categoria basilare, piuttosto che super-ordinata o sub-ordinata - com'è noto fin dalle ricerche della Rosch, 1975), sia la rilevanza o meno dell'ordine degli elementi quale relazione fra gli stessi (una tripletta di numeri, per difetto, è considerata un insieme di elementi comunque considerabili). Non si può trascurare neppure il fatto (messo in evidenza da Legrenzi e Rumati, 1985) che l'aver scelto, quale primo esempio, una successione di numeri pari crescenti di due "dice più del necessario" per una regola molto generale e meno del necessario per una regola molto specifica, violando così l'"imperativo" che secondo Grice vale negli scambi comunicativi umani e che si può supporre valga anche nella situazione di laboratorio. D'altra parte, in tale situazione il soggetto incontra (o crede di incontrare) uno "psicologo" e si aspetta di dover esibire una prestazione simile a quella tipicamente adottata per i test di intel-

ligenza, contesto nel quale esiste un solo modo "naturale" di continuare una serie come 2-4-6.

In conclusione, appare evidente che per accertare la consistenza di tali dubbi sarebbe opportuno che la situazione sperimentale venisse modificata in modo da prendere in considerazione in maniera più esplicita variabili finora trascurate. In particolare, bisognerebbe controllare il modo in cui le istruzioni inducono a formulare le ipotesi, le modalità in cui i soggetti codificano tali ipotesi e valutano la coerenza o difformità degli esempi scelti per il test rispetto alle stesse ipotesi. Bisognerebbe anche analizzare i processi cognitivi più generali messi in atto dai soggetti in questo come in altri compiti simili, evitando di limitarsi alla rilevazione dell'adozione di strategie di conferma o falsificazione, che costituiscono solo un aspetto di tali processi. Stiamo prendendo in considerazione questi aspetti in altre indagini in corso di svolgimento.

Riassunto

Secondo Tweney et al. (1980) la scoperta della regola nel compito 2-4-6 sarebbe più legata al modo in cui il problema viene concettualizzato piuttosto che all'uso di una strategia di disconferma. Questi autori hanno ottenuto un notevole miglioramento nella performance presentando il compito come la ricerca delle due regole seguite da esempi classificabili in due categorie, identificate tramite due etichette verbali. In questa procedura tuttavia non era chiaro se la facilitazione fosse dovuta all'introduzione di una categorizzazione, all'uso di etichette verbali o, ancora, al fatto che veniva consentito di considerare rilevanti le informazioni negative anziché ritenerle semplicemente "sbagliate".

I risultati di un esperimento in cui queste variabili sono state esplicitamente separate non consentono di affermare che la manipolazione del feedback abbia effetti sostanziali sul successo nel compito, anche se risulta che l'uso di etichette verbali favorisce la disconferma. La scoperta della regola appare invece più legata all'opportunità di incontrare esempi negativi, cioè che non la seguono, ma che cosa favorisca, a sua volta, tale opportunità oltre il mero caso non appare chiaro.

Un'analisi qualitativa dei protocolli rivela che le indagini che hanno utilizzato finora il compito di Wason si basano su presupposti (quali l'adozione della strategia di *scanning*, la corrispondenza fra ipotesi ed esempi testati, ecc.) tutt'altro che scontati. In particolare vengono evidenziati alcuni aspetti che rendono innaturale il contesto e viene sottolineata la necessità di un'analisi che non si limiti alla rilevazione dell'adozione di strategie di conferma o disconferma ma venga allargata anche ad altri processi cognitivi.

Abstract

According to Tweney et al. (1980), the discovery of the rule in Wason's 2-4-6 task would be more connected to the conceptualization of the problem than to the use of a disconfirmatory strategy. These authors have obtained a substantial improvement in the performance by presenting the task as the search for the two rules followed by examples which could be classified in two categories, identified by means of two verbal labels. In this procedure, however, it was not clear what produced the facilitation: whether the introduction of a categorization, or the use of verbal labels, or, again, the fact that the subjects were allowed to consider negative information as relevant instead of simply dismissing it as "wrong".

The results of an experiment in which these variables have been explicitly separated do not allow us to claim that the feedback manipulation has substantial effects on being successful in the task, although using verbal labels helps the disconfirmation. The discovery of the rule seems, instead, more linked to the possibility of finding negative

examples (i.e., which do not follow it), but it is not clear what, besides simple chance, contributes to such a possibility.

A qualitative analysis of protocols reveals that investigations adopting Wason's task are based on some presuppositions (such as the adoption of the scanning strategy, the correspondence between hypotheses and tested triples, etc.) that are not obvious at all. In particular, some aspects which make the context unnatural are stressed, and the need of an analysis which is not limited to pointing out the adoption of confirmatory or disconfirmatory strategies is underlined. It is also claimed that such an analysis should be expanded to other cognitive processes.

Bibliografia

Bruner, J.S., Goodnow, J.J., Austin, G.A., Brown, R.W., 1956. *A study of thinking*. Wiley & Sons, New York. Trad. ital. *Il pensiero*, Armando, Roma, 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.

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*, Mc Graw-Hill, New York.

Moneta, G.B., 1987. Are psychological experiments on hypothesis-testing strategies a good simulation of scientific problem-solving? *Epistemologia*, X, 29-54.

Mynatt, C.R., Doherty, M.E., Tweney, R.D., 1977. Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology*, 29, 85-95.

Mynatt, C.R., Doherty, M.E., Tweney, R.D., 1978. Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology*, 30, 395-406.

Platt, J.R., 1964. Strong inference. *Science*, 146, 347-353.

Popper, K.R., 1959. *The logic of scientific discovery*. New York, Basic Books. Trad. ital. *La logica della scoperta scientifica*, Einaudi, Torino, 1970.

Rosch, E., 1975. Cognitive representation of semantic categories. *Journal of Experimental Psychology: General*, 104, 129-233.

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., 1972. *Psychology of reasoning*. London, Batsford. Trad. ital. *Psicologia del ragionamento*, Martello Giunti, Firenze, 1977.

Wetherick, N.E., 1962. Eliminative and enumerative behavior in a conceptual task. *Quarterly Journal of Experimental Psychology*, 14, 246-249.