

# IL TAVVU- test per misurare l'adeguatezza della valutazione della valutazione universitaria

Giovanni Battista Flebus

Università di Milano-Bicocca

Questa ricerca parte dall'ipotesi che, l'esame universitario è oggetto di percezione affettiva e cognitiva, legata a valutazioni di comportamenti, propri e altrui, interconnessa con degli scopi, delle procedure, degli stati emotivi e motivazionali che aiutano o interferiscono nella vita degli studenti universitari. Una serie di domande a scelta multipla raccoglie queste ipotesi e propone delle risposte fra cui scegliere quella che meglio descrive il rispondente.

L'ipotesi metodologica soggiacente è che il processo di valutazione degli esami universitari è misurabile con il metodo dei punteggi ottimali (Guttman, 1954; Flebus, 1999) e si può costruire una scala, psicometricamente adeguata, utilizzando le risposte a scelta multipla in cui l'adeguatezza della valutazione universitaria è stabilita a posteriori, coi metodi statistici

## Metodo

Quindici domande sono state formulate sulla base di interviste a studenti, che investigavano valutazioni cognitive, affettive, motivazionali ecc. Il formato è quello della risposta a scelta multipla, che nella letteratura corrente è utilizzabile solo con test a prestazione *massima* (dove si conosce la risposta esatta, contrapposti a test di prestazione *tipica*). L'applicazione del metodo dei punteggi ottimali invece consente la scalatura di ogni singolo item, e prescinde da un'attribuzione a priori della quantificazione, come succede nelle scale Likert o altre scale di valutazione. Il calcolo del punteggio ottimale è identico alla prima dimensione dell'analisi delle corrispondenze. Le differenze sono esposte in un articolo (Flebus, 1999). SPSS fornisce anche il coefficiente alfa di Cronbach per ogni dimensione estratta dalle risposte e il coefficiente di discriminazione, che permette di valutare il contributo di ogni singolo item alla dimensione ricercata. Oltre alle domande della ricerca, sono state chieste valutazioni con domande semplici, nel formato Likert o come scale

di valutazione, per lo scopo di validare i risultati con item dal contenuto chiaro ed evidente.

## Partecipanti

Sono 860 studenti, per la maggior parte di Milano-Bicocca, reclutati in modo informale da studenti di Psicologia che hanno ricevuto credito per la loro partecipazione alla ricerca.

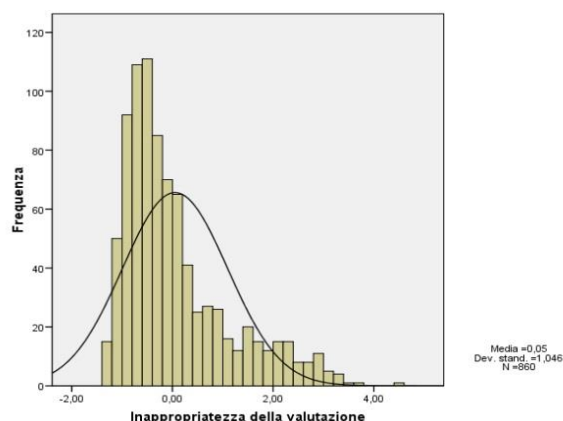
## Analisi dei dati

L'applicazione del programma CATPCA di SPSS alle risposte di 860 studenti su 15 domande ha generato una dimensione ipotetica, con alfa di Cronbach pari a 0,80. Ogni singola risposta ottiene la quantificazione riportata nei grafici, e trasforma una misurazione a livello nominale in una scala a intervalli. Un piccolo numero di partecipanti sono stati esclusi per il carattere estremo delle loro risposte. Le quantificazioni di ogni domanda sono riportate nei grafici, assieme all'intervallo di fiducia che permette di valutare direttamente il numero di risposte (un intervallo di fiducia ampio indica che ci sono poche risposte, viceversa un intervallo ristretto segnala un numero molto grande di studenti che hanno scelto quella risposta). Nei grafici le risposte alle 14 domande buone sono riordinate in ordine **crescente** di inadeguatezza, in modo da facilitarne la lettura.

Per verificare l'unifattorialità della dimensione, è stata condotta un'analisi fattoriale, che dimostra che una sola dimensione è presente nei dati.

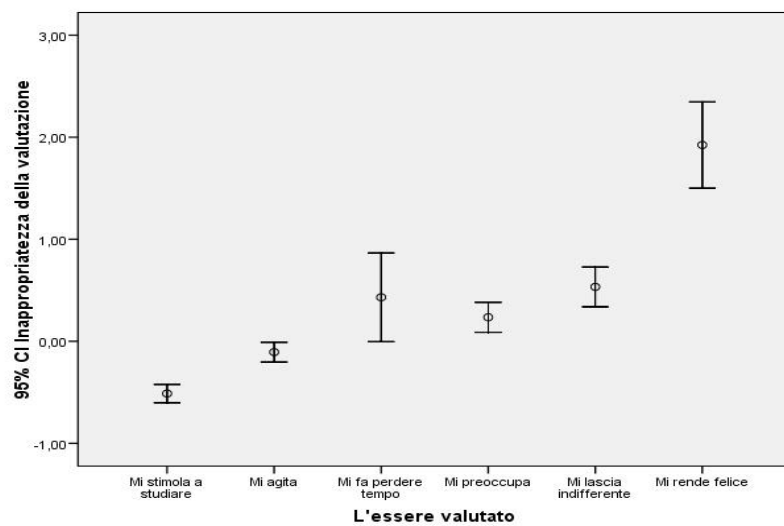
Analisi di fedeltà. Per verificare con metodi standard la fedeltà della scala, si riportano le correlazioni item totale e il coefficiente alfa se l'item è eliminato. Tutti i 14 item sono utili alla produzione della dimensione.

Distribuzione del punteggio. Il grafico 15 riporta la distribuzione del campione di 860 studenti. La distribuzione non è perfettamente normale, e forse una correzione per eventuali applicazioni statistiche sarebbe da esplorare.



## Validazione con criteri esterni.

Le domande usate come criteri esterni al test vi erano quelle riguardanti le reazioni a un buon voto, a un cattivo voto, e altre descrizioni. I grafici riportano le medie di chi sceglie le varie alternative, usando sempre gli intervalli di fiducia.



Infine, lo studio ha esaminato alcune differenze fra maschi e femmine, fra studenti in corso e fuori corso e le differenze fra vari corsi di studio.

## Interpretazione dei risultati

Il metodo dei punteggi ottimali consente di usare semplici risposte non quantificate a priori e produce per ciascuna di esse una quantificazione

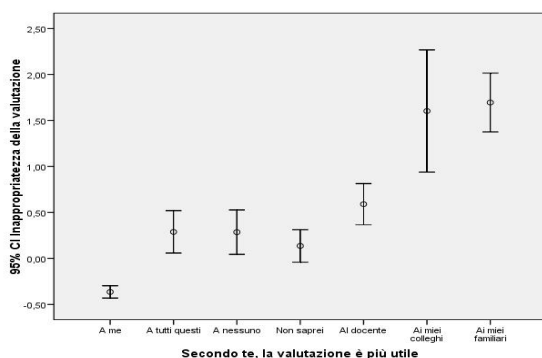
che deve essere interpretata alla luce dei risultati globali. I punteggi negativi indicano nel presente caso la posizione di appropriatezza e quelli positivi indicano l'inappropriatezza o inadeguatezza della valutazione della valutazione universitaria. La tabella qui accanto rappresenta la prima domanda, con le risposte, e le quantificazioni ottenute. Per esempio, la

Q2 Secondo te, la valutazione è più utile

Punti: Coordinate

Categoria	Frequenza	Coordinate del centroide
		Dimensione 1
1 A me	446	-,401
2 Al docente	98	,542
3 Ai miei familiari	47	1,662
4 Ai miei colleghi	6	1,037
5 A tutti questi	80	,218
6 A nessuno	56	,242
7 Non saprei	90	,136
Mancante	1	

Normalizzazione principale per variabile.



domanda sulle persone a cui è più utile la valutazione riceve una grandissima quota di risposte: *a me*. La quantificazione è pari a -0,40 e indica la posizione più adeguata. All'opposto, la risposta *ai miei colleghi* o *ai miei familiari* ricevono quantificazioni di elevata inadeguatezza, mentre la risposta *Non saprei* si attesta sullo zero. I grafici

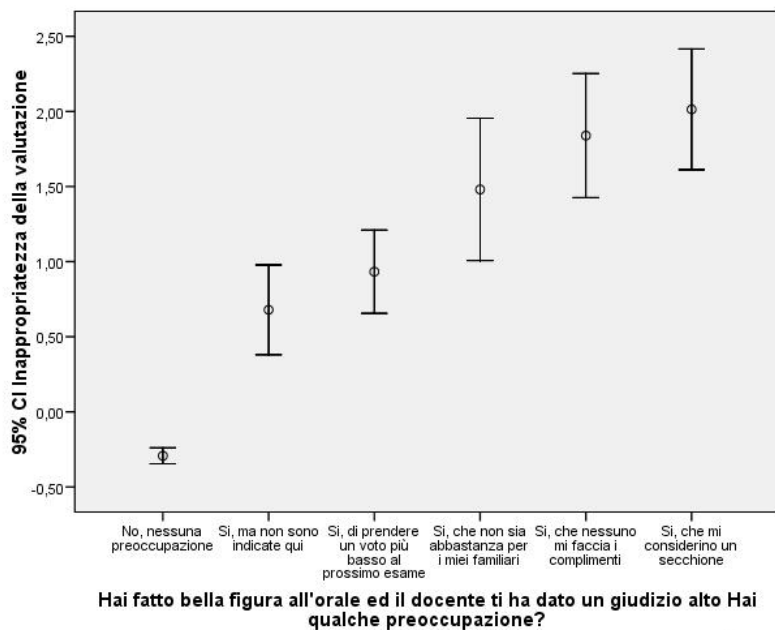
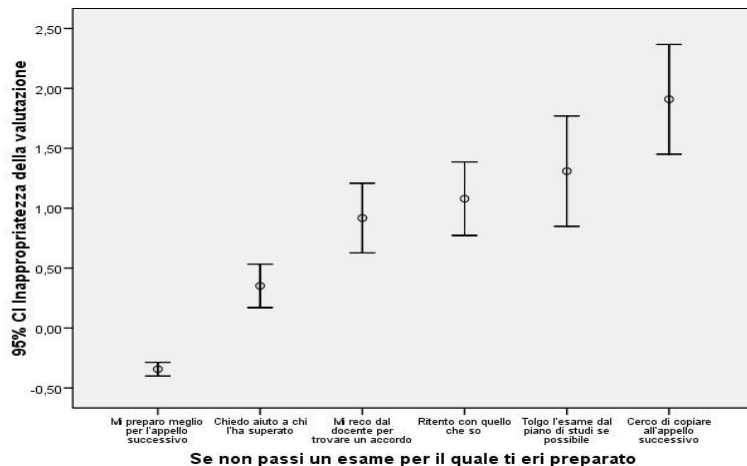
visualizzano le quantificazioni per le altre domande.

Mentre per la maggior parte degli studenti la valutazione è uno stimolo per studiare, per alcuni la reazione è di apparente felicità o indifferenza, di preoccupazione.

La bocciatura ad un esame produce reazioni di desistenza, di tentativi di imbroglio (cerco di copiare) o di indifferenza (ritento con quello che so).

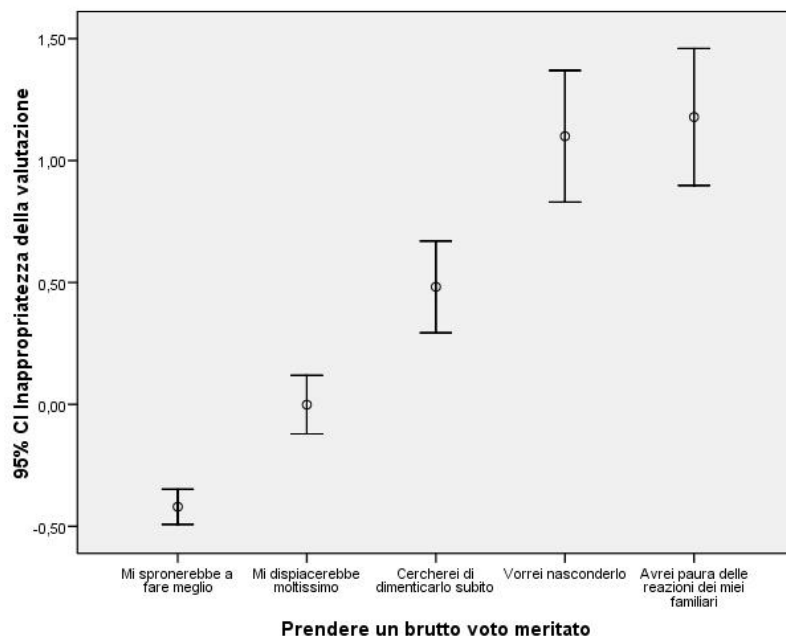
La valutazione adeguata invece di

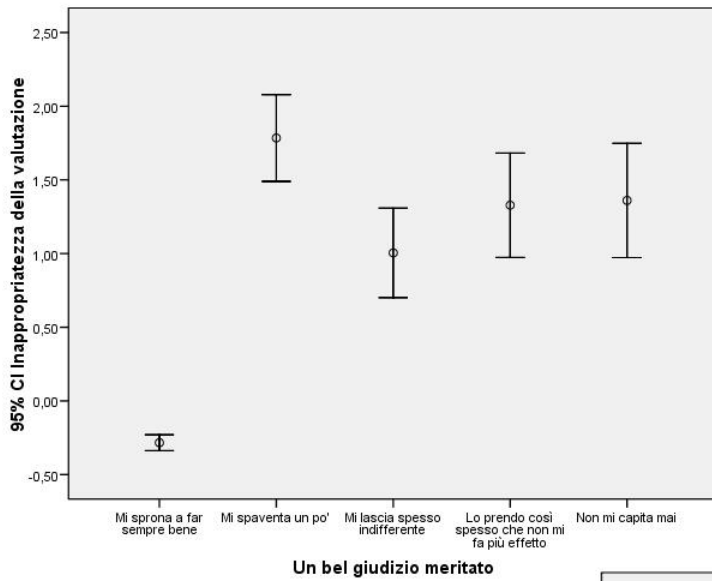
quello che è successo dopo una bocciatura, come vuole il buon senso, è quello di studiare di più per la prossima volta. Anche chiedere aiuto a chi lo ha superato riceve un punteggio di adeguatezza, anche se inferiore.



L'inadeguatezza della valutazione si esprime come preoccupazione di rifiuto degli altri in seguito al giudizio alto (mi considereranno un secchione), o una mancanza di aspettativa di gratificazione sociale (nessuno mi faccia i complimenti) oppure che non sia un risultato abbastanza buono per i familiari.

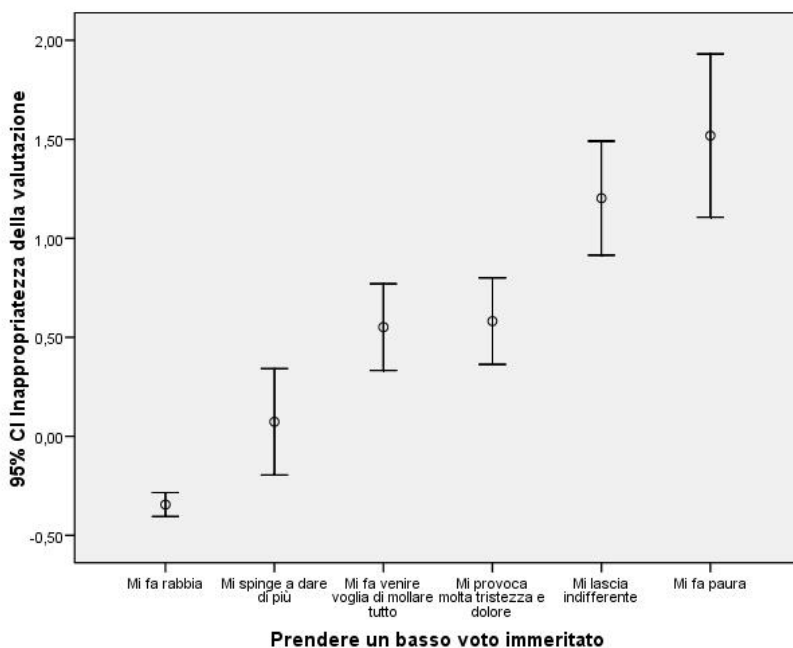
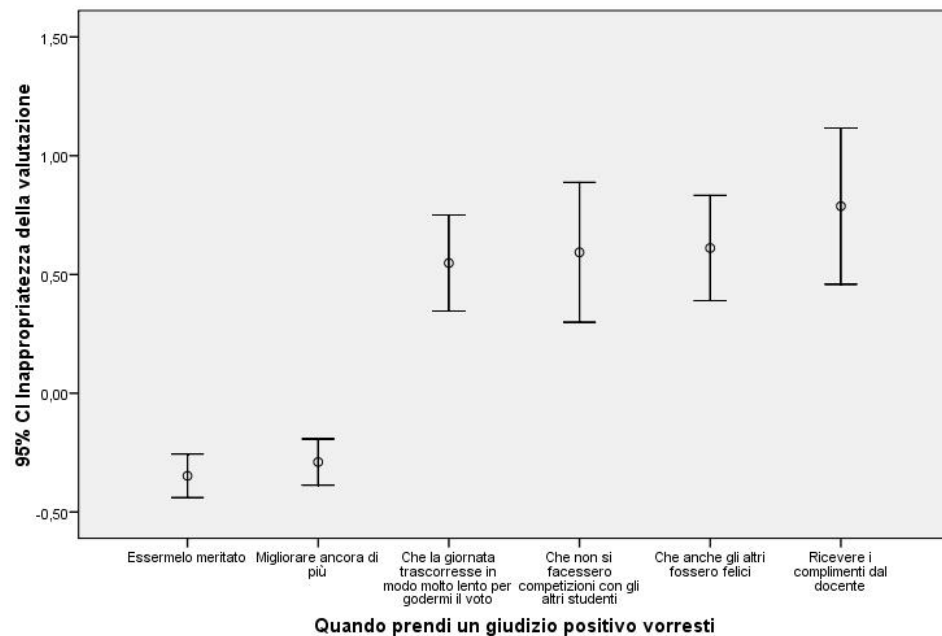
La reazione ad un brutto voto può essere di paura e di evitamento per alcuni, mentre per quelli che hanno la corretta percezione, il brutto voto è uno stimolo a fare meglio.

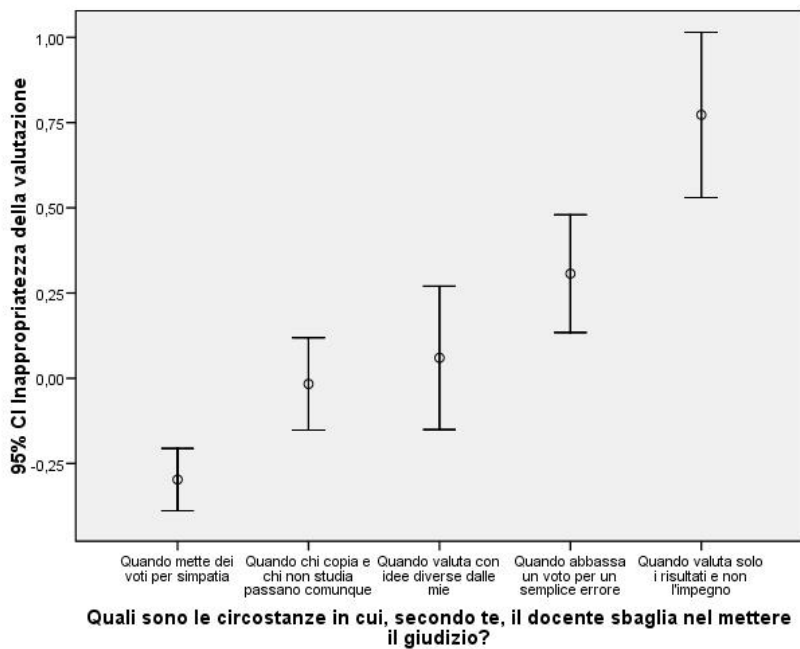




Un giudizio positivo produce delle aspettative che riguardano gli altri, anziché restare confinati alla propria persona e responsabilità. Lo studente inadeguato si aspetta complimenti dal docente, che anche gli altri siano felici o che non ci sia competizione.

Interessante la reazione al voto basso immeritato: l'adeguatezza va con la rabbia e l'inadeguatezza va con la paura, l'indifferenza, la tristezza o la demotivazione, quindi reazioni improprie davanti ad un voto basso.



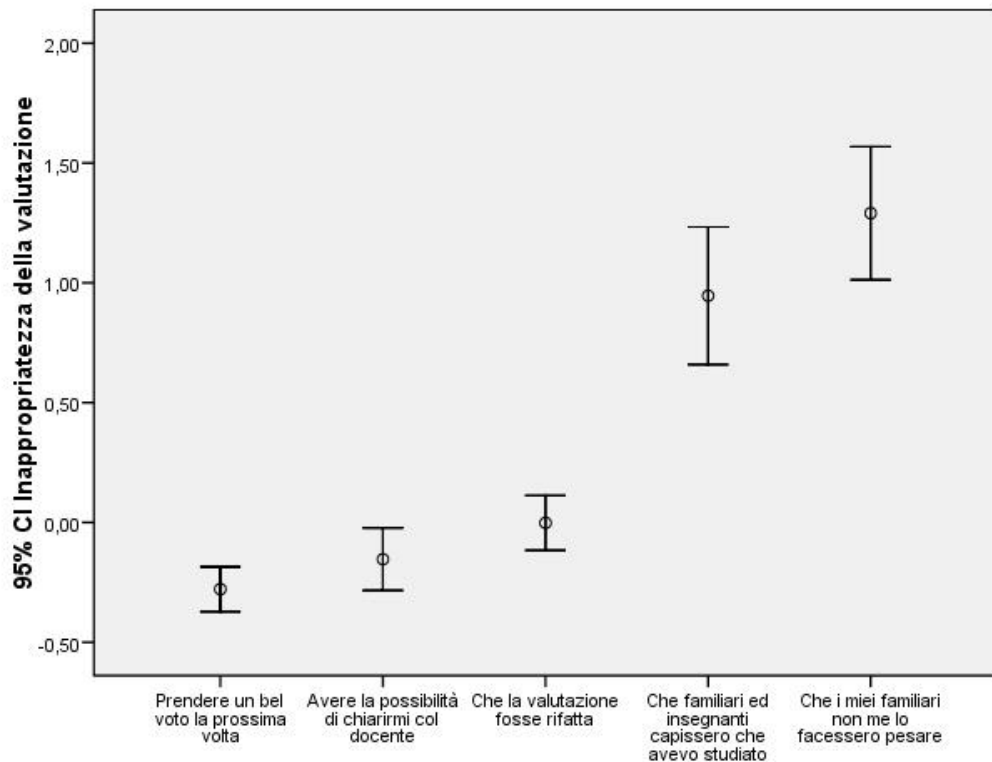


Quali sono le circostanze in cui, secondo te, il docente sbaglia nel mettere il giudizio?

L'adeguatezza di fronte ad un errore del docente si manifesta soprattutto attribuendo alla sua simpatia l'errore, mentre l'inappropriatezza si manifesta con l'aspettativa che il docente valuti l'impegno, poco visibile e poco utile, anziché i risultati, ovvero il profitto conseguito. In altre parole, lo studente con inadeguatezza si aspetta che il docente valuti, al pari di un buon genitore che conosce il

figlio e sa quando si impegna, anche se non raggiunge un buon risultato.

Ecco di nuovo la reazione ad un cattivo voto immeritato: l'inadeguatezza ora si manifesta con una reazione di paura e sottomissione verso i familiari, anziché con il desiderio di fare meglio la prossima volta o di chiarire quello che è successo. Quindi, la dimensione di Inadeguatezza o il suo contrario, Appropriatezza, è un costrutto costituito da una mancanza di razionalità, di un eccesso di affettività incontrollata, di fragilità motivazionale, di atteggiamento di desistenza, bassa resistenza alla frustrazione, di una ricerca di supporto sociale, di un mantenimento di un legame filiale, fondato su presupposti di devozione e di dipendenza.

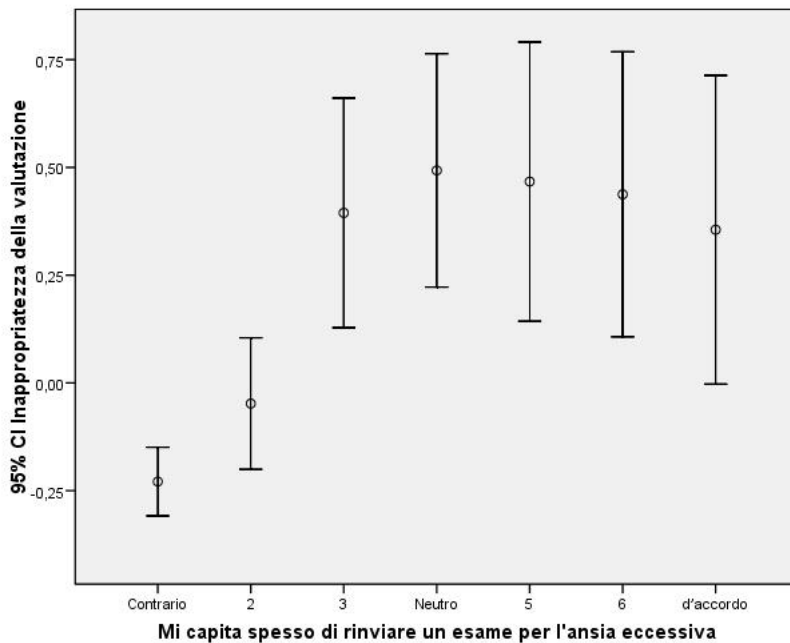
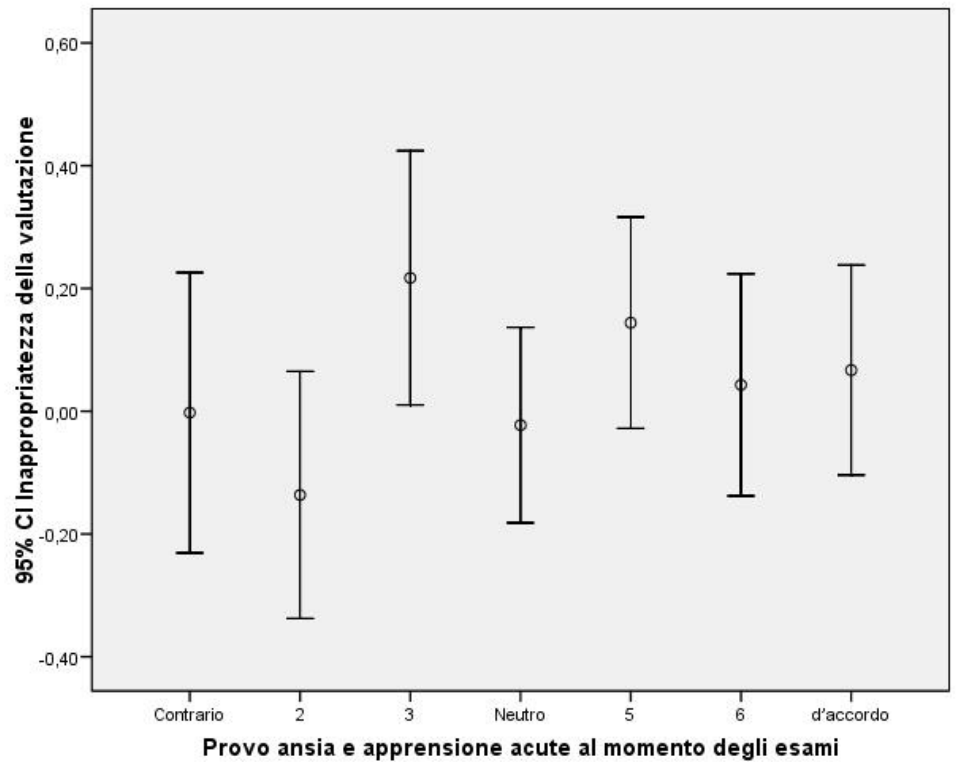


Supponi di prendere un brutto voto immeritato, vorresti soprattutto

## Validazione concorrente e discriminante

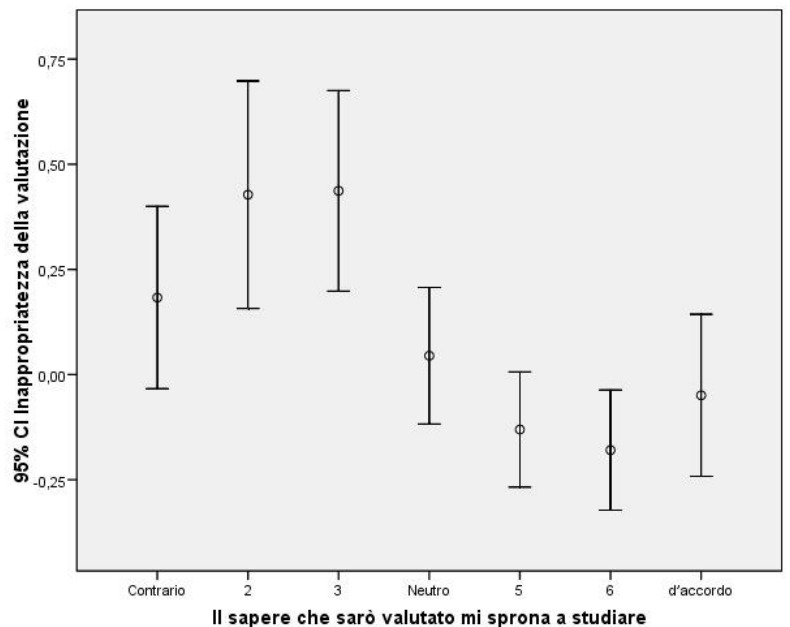
I grafici seguenti, che riportano i punteggi di chi da risposte a domande più semplici, vanno nel senso di questa interpretazione. (Queste frasi non fanno parte della scala e sono delle frasi tipo Likert o altre scale di valutazione.)

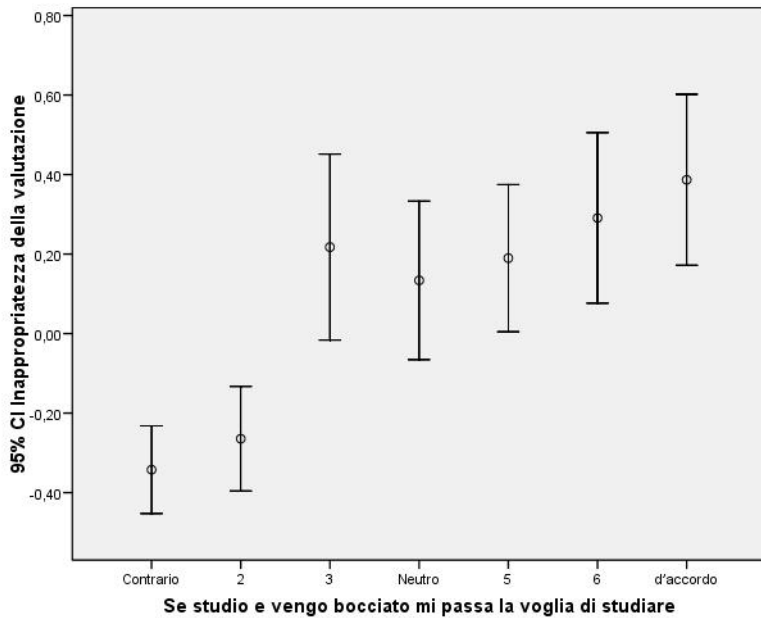
L'inadeguatezza non è ansia ma è un prodotto dell'ansia incontrollata.



L'ansia si manifesta sia negli adeguati sia negli inadeguati, ma in questi ultimi la mancanza di controllo produce effetti destabilizzanti che fanno posporre l'esame

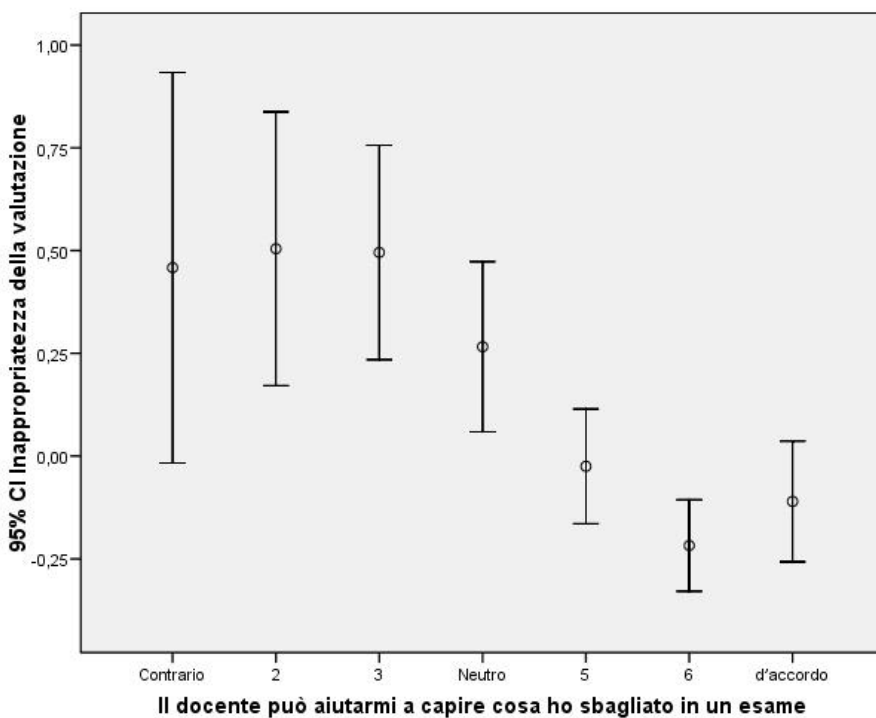
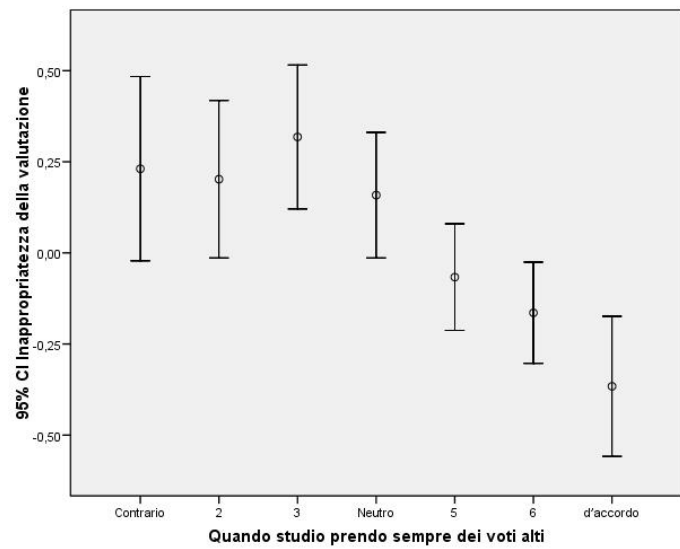
Il sapere di essere valutati è uno sprone per gli adeguati, ma non agisce da stimolo per gli inadeguati.





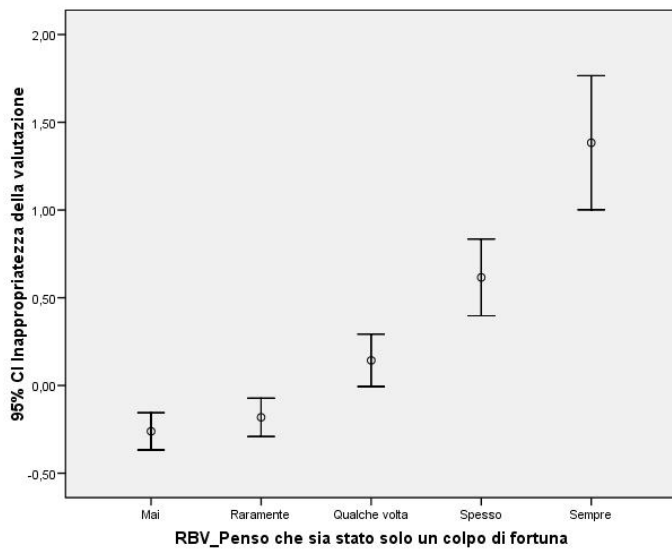
La bocciature ingiusta produce calo intenso di motivazione negli inadeguati

Per lo studente adeguato, lo studio produce voti alti, ma non per quelli inadeguati.



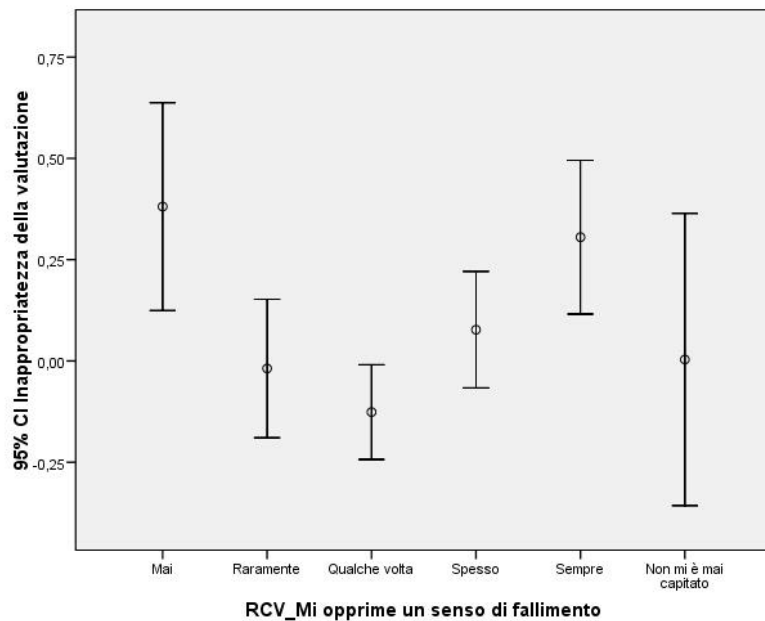
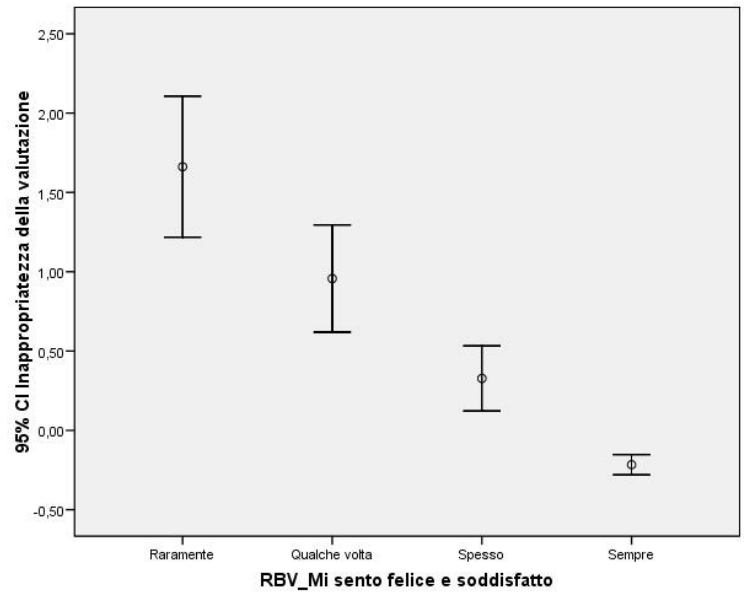
Per l'inadeguato, il docente è impotente nell'aiutare lo studente a capire che cosa ha sbagliato.



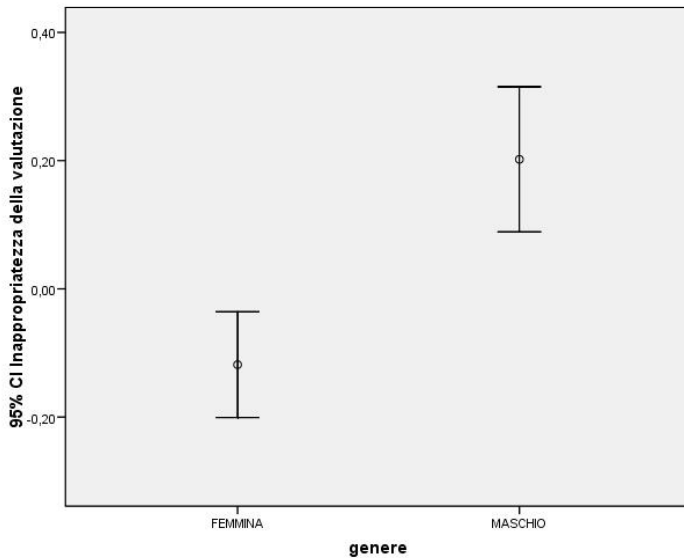


Reazione a un buon voto: Un buon voto per uno studente inadeguato è frutto del caso, mentre non lo è mai per lo studente adeguato

Reazione a un buon voto: Lo studente adeguato è sempre felice e soddisfatto di un buon risultato, mentre l'inadeguato lo è raramente



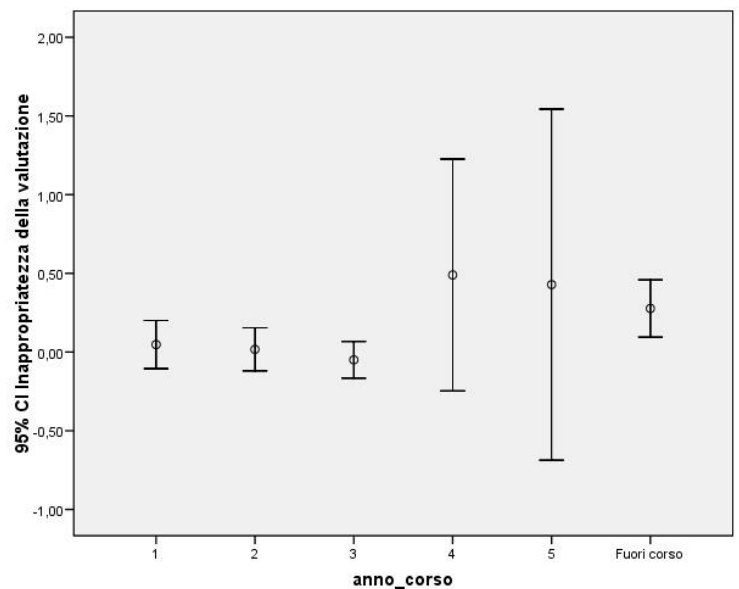
Reazione a un cattivo voto Sentirsi un fallito dopo un brutto voto può essere una reazione non comune, ma per gli inadeguati tale reazione è o molto comune (per un senso generale di bassa efficacia di sé) o molto rara, forse per un processo mentale di accomodamento che rifiuta ogni collegamento fra successo accademico e il sé.



## Altre caratteristiche dell'inadeguatezza

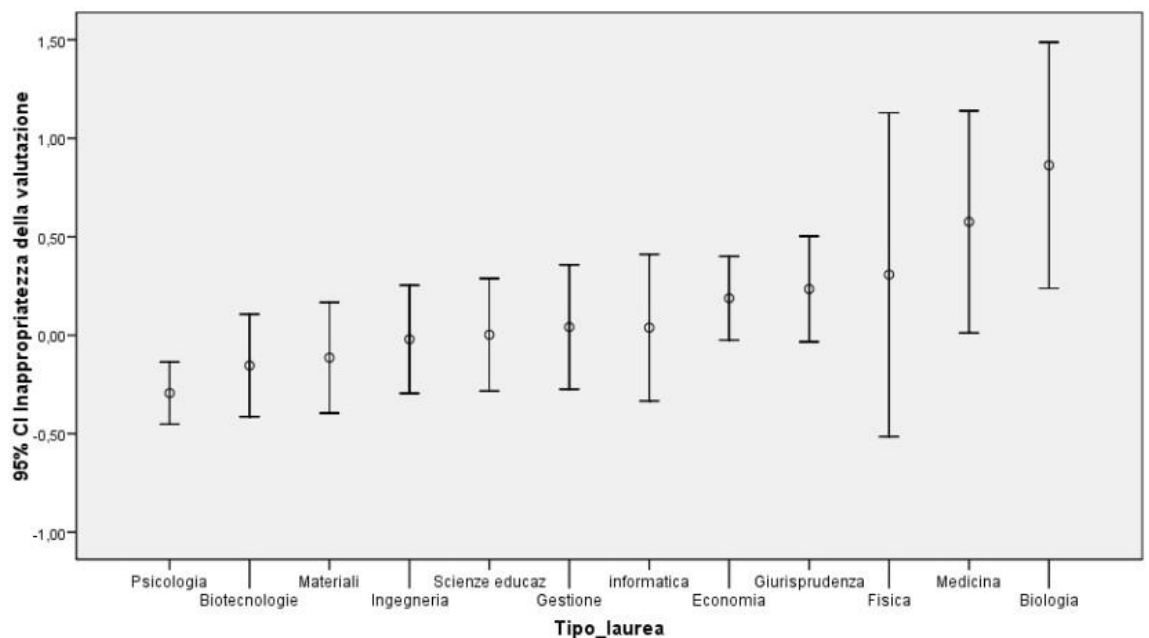
Il genere: i maschi dimostrano di avere valutazioni più inadeguate degli esami, e tale aspetto resta da esplorare

Anni di corso: prevedibilmente, i fuori corso hanno punteggi di minore adeguatezza valutativa rispetto agli studenti in corso. Per gli studenti del 4 o 5 anno non si può dire nulla, vista la scarsa numerosità del campione.



## Tipo di laurea

Nel nostro campione i più adeguati sono gli studenti di Psicologia, e al versante opposto sono quelli di Medicina e Biologia. Non ci sono abbastanza elementi per stabilire se i campioni sono rappresentativi dei vari dipartimenti e pertanto lasciamo aperta la questione, per affrontarla in un successivo studio di approfondimento.



## Conclusioni

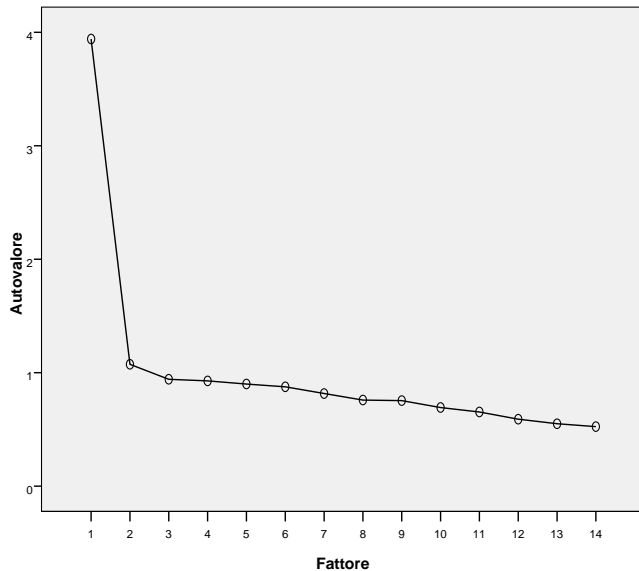
Quindici domande a scelta multipla, concernenti la valutazione affettiva- cognitiva e motivazionale degli esami universitari sono state quantificate con il metodo dei punteggi ottimali. Un item è stato scartato perché non pertinente. Gli altri 14 sono presentano saturazioni elevate su il primo fattore non ruotate e dimostrano l'unifattorialità della scala così costruita. Un'analisi della fedeltà ha confermato che tutti i 14 item contribuiscono ad aumentare il coefficiente alfa di Cronbach e apportano ciascuno una quota di varianza utile alla misurazione. Il costrutto così delineato si può chiamare Adeguatezza della valutazione degli esami universitari: i punteggi alti di Adeguatezza indicano una posizione di equilibrio, razionalità, soddisfazione personale, mancanza di vincoli parentali e sociale pesanti, capacità di continuare lo studio nonostante le difficoltà di applicazione nello studio, motivazione solida allo studio. Al contrario, i punteggi bassi di adeguatezza (o inappropriata) caratterizzano posizione di sfiducia nelle proprie capacità, incapacità di controllare gli effetti destabilizzanti dell'ansia, un locus esterno di responsabilità, una eccessiva dipendenza dal giudizio di altri, una confusione di ruoli del docente con quello di genitore ma anche un vissuto di apatia, indifferenza e mancanza di motivazione generale. Gli studenti fuori corso sono più propensi a mostrare questa inadeguatezza di valutazione.

Il metodo dei punteggi ottimali ha trovato qui una valida applicazione in un campo di difficile definizione teorica - l'adeguatezza valutativa degli esami universitari. Ha dimostrato che alcune risposte possono essere considerate inadeguate se pur all'interno di domande chiuse non modificabili. Il metodo ha delle caratteristiche che le scale Likert non prevedono e può costituire un valido modo alternativo di concepire nuovi concetti psicologici con un impianto psicometrico innovativo e solido.

# Risultati dell'analisi fattoriale e dell'analisi di fedeltà

## Scree-test

Grafico decrescente degli autovalori



MATRICE FATTORIALE NON RUOTATA DELLE VARIABILI QUANTIFICATE	FATTORE 1	CORRELAZIONE DEL TOTALE	ALFA DI CRONBACH SE L'ITEM È
		ITEM CORRETTA	ESCLUSO
TQ1 L'essere valutato	0,440	0,389	0,793
TQ2 Secondo te, la valutazione è più utile	0,479	0,436	0,789
TQ3 Se non passi un esame per il quale ti eri preparato	0,535	0,463	0,786
TQ4 Hai fatto bella figura all'orale ed il docente ti ha dato un giudizio alto Hai qualche preoccupazione?	0,598	0,520	0,781
TQ5 Che cosa è la valutazione secondo te?	0,521	0,448	0,788
TQ6 Quando ottieni una valutazione soddisfacente, desideri che	0,620	0,556	0,777
TQ7 Quando prendi un giudizio positivo vorresti	0,389	0,338	0,796
TQ8 Un bel giudizio meritato	0,588	0,516	0,782
TQ9 Per prendere un giudizio positivo oltre a studiare bisogna anche	0,431	0,397	0,792
TQ10 Prendere un basso voto immeritato	0,466	0,418	0,790
TQ11 Sapresti dire perché alcuni studenti prendono sempre valutazioni basse anche se si sentono preparati?	0,358	0,325	0,797
TQ12 Supponi di prendere un brutto voto immeritato, vorresti soprattutto	0,409	0,365	0,794
TQ13 Prendere un brutto voto meritato	0,469	0,413	0,791
TQ14 Quali sono le circostanze in cui, secondo te, il docente sbaglia nel mettere il giudizio?	0,289	0,265	0,800