



Università degli Studi di Modena e Reggio Emilia
Dipartimento di Economia Politica
CAPP: Centro di Analisi delle Politiche Pubbliche

\\ 535 \\

**Problemi e osservazioni sul questionario
utilizzato per la valutazione della didattica
da parte degli studenti frequentanti**

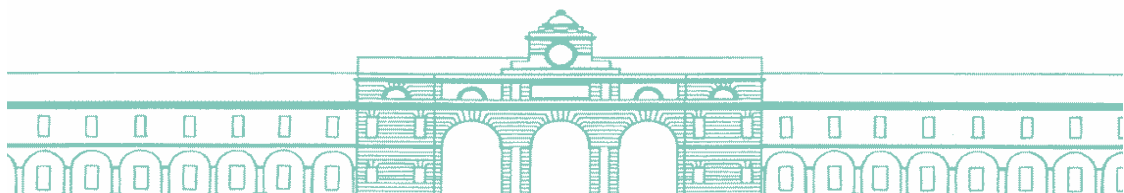
di

Michele Lalla

e-mail: lalla.michele@unimore.it

Materiali di discussione

Università degli Studi di Modena e Reggio Emilia
Dipartimento di Economia Politica
Via Jacopo Berengario 51
41100 Modena (Italia)



Via Jacopo Berengario 51 – 41100 MODENA (Italy) tel. +39-059.2056943 fax +39-059.2056947
e-mail: capp@unimo.it

RINGRAZIAMENTI

Il documento è stato elaborato, nell'ambito delle attività svolte nel Nucleo di Valutazione interna, in base (e come risposta) alle osservazioni critiche provenienti dai Presidi delle Facoltà sul questionario adottato per la valutazione della didattica da parte degli studenti frequentanti, con lo scopo di fornire un documento di riferimento alle molteplici domande inerenti, oltre che al questionario, anche agli obiettivi e a qualche aspetto della procedura dell'indagine. Osservazioni critiche e suggerimenti molto interessanti e puntuali sono stati posti da alcuni docenti del Dipartimento di Scienze Sociali, Cognitive e Quantitative, tra i quali Nicoletta Cavazza, che si ringraziano perché senza i loro preziosi stimoli il documento non sarebbe stato prodotto. Alcune domande sono state formulate direttamente dall'autore, con relativa risposta. Il lavoro è accessibile, via internet, anche dal sito del Nucleo di Valutazione interna, <http://www.pcv.unimo.it>, oltre che dalla collana «Materiali di discussione», che a sua volta è raggiungibile sia dal sito del CAPP (<http://www.capp.unimo.it>, percorso <Pubblicazioni>, <Materiali di discussione>), sia dal sito della Facoltà di Economia (<http://www.economia.unimo.it>, percorso <Dipartimenti e attività di ricerca>, <Dipartimento di Economia Politica>, <Materiali di discussione>).

1. Introduzione

L'attività di valutazione della didattica da parte degli studenti frequentanti condotta e/o coordinata dai Nuclei di Valutazione interna (NdV) degli atenei è, ormai, una attività consolidata e ordinaria. Ciò nonostante, permangono molteplici problemi tecnici e organizzativi che rendono difficoltosa tale attività. Le note critiche sono, poi, numerose e pervase da scetticismo che coinvolge lo strumento o questionario utilizzato per la misurazione delle opinioni (validità, attendibilità, precisione dei risultati), la modalità di somministrazione (tempo o data, ripetizione, sovraccarico, autoselezione, risposte a caso o seriali, acquiescenza), l'elaborazione dei dati, e l'uso dei risultati.

La letteratura sull'argomento è vasta e una ricognizione attenta è stata svolta dall'[Osservatorio \(1998\)](#), ora Comitato Nazionale, per la Valutazione del Sistema Universitario (CNVSU). In questa nota si raccolgono e si discutono le domande più frequenti che sono poste al NdV; in particolare, quelle che si riferiscono allo strumento attualmente utilizzato per la rilevazione dei dati, ossia il questionario. Alcune domande sono seguite dalle risposte che giustificano le scelte, altre sono, invece, senza risposta perché la revisione critica alla soluzione adottata può apportare un miglioramento.

Le argomentazioni contro la valutazione della didattica eseguita dagli studenti assumono spesso che la loro posizione non li metterebbe in grado di valutare con cognizione l'oggetto dello studio, di essere neutrali, e di giudicare con oggettività l'operato del docente; inoltre, si sostiene che non vi sia relazione tra apprendimento e giudizio espresso perché si può imparare ugualmente pur manifestando insoddisfazione per la coartazione cui si è soggetti, per la fatica che si è sostenuta, per la tensione d'esame. Infine, tutte le indagini volte a rilevare la percezione soggettiva degli utenti sul valore di una istituzione e sulla validità del servizio offerto, di cui si è beneficiari, manifestano aspetti suscettibili di critiche alquanto fondate perché i giudizi dei singoli possono essere: vaghi, imprecisi o distorti per finalità non sempre note, ancorati a sistemi di valori diversi che rendono non confrontabili le valutazioni espresse, affetti da forme di reticenza o provocazione o irrisione che li invalidano. La condizione di raccolta delle informazioni e la loro utilizzazione, poi, può essere non appropriata e ostacolare un corretto flusso dagli utenti ai gestori. Gli studenti collaborano tra loro frequentemente durante la compilazione sicché i giudizi non sono indipendenti.

Tali difficoltà sono ineliminabili e costituiscono un elemento «ontologico» e/o fisiologico di ogni indagine sociale, rispetto alla quale si riscontrano limiti e, talvolta, critiche radicali ([Goyder, 1987, cap. 2](#)). L'attività di valutazione della didattica da parte degli studenti frequentanti è conforme al processo di indagine sociale; perciò, è soggetta allo stesso apparato critico. Le osservazioni/obiezioni più frequenti, che si indirizzano alla valutazione della didattica, spesso non tengono conto degli obiettivi che la caratterizzano, delle condizioni concrete di conduzione, dei processi da attivare, degli assunti o «paradigmi» che sottostanno alle scelte intraprese e che «guidano» spesso le risposte alle obiezioni mosse, molte delle quali sono riportate di séguito.

Il *primo assunto* imprescindibile concerne *il valore in sé delle percezioni soggettive degli studenti*, nonostante le possibili imprecisioni e le condizioni di non adeguatezza nell'esprimere un giudizio su ciò che loro non conoscono bene. Gli esiti che emergono dai loro giudizi esprimono sempre istanze, difficoltà, e problemi che l'azione didattica dovrebbe tenere presente, in ogni caso, perché possono fornire elementi di riflessione per rendere più efficace la comunicazione in aula. Sostenere la loro infondatezza equivale soltanto a negare gli ostacoli che intralciano il percorso di

apprendimento dei discenti e che costituiscono, in un certo senso, la *realtà* in cui si opera. L'esperienza sembra avvalorare, infatti, l'affidabilità dei giudizi degli studenti quando riguardano aspetti che li coinvolgono direttamente, come la chiarezza dell'esposizione o il carico di lavoro appropriato (Harvey, Knight, 1997; Hativa, 1998). Nonostante ciò, l'incapacità degli studenti a esprimere un giudizio corretto e ponderato rimane una delle obiezioni più frequenti esplicite o implicite.

Il *secondo assunto* imprescindibile riguarda la *natura censuaria dell'indagine* sicché non può tenere debitamente conto della complessità e degli aspetti multiformi delle questioni indagate, non può essere completa e esaustiva per rispondere alle numerose questioni inerenti alla didattica. Non si tratta di eseguire, infatti, una indagine scientifica volta a scoprire il tessuto di tutte le possibili correlazioni esistenti tra le variabili e come agiscono le indipendenti sulle dipendenti, ma si tratta di una indagine che rileva poche dimensioni essenziali per avere un riscontro «minimo» sull'andamento del processo di insegnamento. Lo strumento di rilevazione (questionario) deve essere agile e semplice per non appesantire il carico dei rispondenti con il relativo logoramento perché autocompilato e anche più volte durante l'anno, quindi, dovrebbe contenere poche domande per non superare più di una facciata di un comune foglio A4. Si dirà che è riduttivo tutto questo e, perciò, inutile. Critiche analoghe si potrebbero rivolgere alla conduzione dei censimenti perché rilevano pochi caratteri essenziali. Si evince, poi, dai risultati finora disponibili, che anche un sistema grezzo e approssimativo, come *pare* l'attuale, è in grado di fornire una valutazione sensata su alcune dimensioni.

Il *terzo assunto*, connesso al secondo, si riferisce agli *obiettivi minimi* o ridotti dell'indagine perché contemplan soltanto la necessità di attivare un processo che ha in sé la potenzialità di rendere più trasparente la funzione della istituzione universitaria: lo studente non ha un ruolo solo passivo perché può esprimere le sue opinioni sull'insegnamento, il docente non è più un «*dominus*» assoluto perché il giudizio dello studente diventa «pubblico» o quantomeno inviato ai rappresentanti dell'organizzazione (Preside e Presidente di corso di studio). Il docente tenderà a porre più attenzione alle sue azioni. Se le sofferenze e le difficoltà che emergono saranno anche risolte, allora la partecipazione e il coinvolgimento dei discenti tenderà a incrementare, instaurando un circolo virtuoso. Finora, è emerso, come già detto, una certa fondatezza del giudizio rilevato, nonostante i numerosi fattori che disturbano il processo di raccolta dei dati.

Il *quarto assunto* tratta il *ritardo nella consegna dei risultati* come intrinseco e/o strutturale perché il processo è complesso: va dalla raccolta dei questionari, alla loro consegna (che non è sempre puntuale) presso l'ufficio del NdV, alla loro lettura ottica, alla elaborazione delle informazioni, quindi alla restituzione degli esiti agli interessati. Per il ritardo, anche fisiologico, nulla si può più fare, se non, forse, migliorare il corso l'anno successivo, a beneficio della coorte successiva. Il processo di intervento in itinere non è, però, antitetico a quello della valutazione, ma si possono attuare entrambi i meccanismi perché per affrontare veramente i problemi che nascono nei corsi e durante gli esami, occorre non adagiarsi sugli allori della valutazione di Ateneo, ma instaurare una prassi che consenta di intervenire tempestivamente, appena il problema si presenta, indipendentemente dall'esito della valutazione. Tutti gli attori devono svolgere il loro ruolo con impegno: il docente deve essere sensibile alle richieste degli studenti, gli organismi rappresentativi degli studenti devono essere attenti alle esigenze dei loro rappresentati, la *Commissione (paritetica) della didattica* deve interessarsi con tutti gli strumenti possibili al buon svolgimento dei corsi. Se si concede, in quest'ultima, ampia facoltà di intervento agli studenti e senza tema di possibilità di ritorsioni, si può

forse attivare un processo di modifica in itinere idoneo alle esigenze degli interessati. Anche per questo tipo di funzione occorre, però, una prassi consolidata e diffusa tra gli studenti stessi; perché nell'ambiente universitario si entra entusiasti e, presto, si comincia a pensare di uscire quanto prima dall'ambiente e non di sollevare i problemi da risolvere; ognuno impara subito, a sue spese, che l'importante è portare in salvo il suo "particolare" e arrivare in fondo al traguardo. D'altronde, anche i rappresentanti devono studiare e agli studenti manca proprio il tempo per studiare; dunque, come si può chiedere loro di seguire l'andamento della didattica, quando non hanno neanche il tempo di fare ciò che devono? La responsabilità non è certo degli studenti, ma di tutti quelli che lavorano dentro l'Ateneo, incluso l'apparato burocratico che, come tutte le burocrazie, lavora *più per complicare* lo svolgimento dell'attività istituzionale, *piuttosto che per semplificare* procedure e regole amministrative. Ovviamente, la «colpa» non è dei singoli attori, ma delle regole esistenti e, raramente, delle abitudini inveterate. Soltanto lo sforzo comune, di tutte le componenti, volto a attuare una didattica efficace e efficiente, a creare un ambiente ideale che risolva i numerosi problemi che gli studenti incontrano durante il loro percorso di studio, può rendere l'Università una istituzione in grado di assolvere completamente i suoi compiti formativi e educativi.

Il *quinto assunto*, connesso sempre al secondo, limita la quantità e la qualità dell'elaborazione dei dati raccolti. Il «ritorno» informativo deve essere semplice e comprensibile a tutti: più l'analisi è sofisticata e più difficile e/o controllabile è la lettura, più elevata è la professionalità dell'operatore, più ampio è il lasso di tempo che occorre per completarla, più impegnativa è questa fase del processo. Le richieste di spiegazioni e le proteste più o meno velate sono già elevate ora; inoltre, la relazione scritta è già abbastanza voluminosa. L'elaborazione più raffinata comporterebbe un aumento di tutti questi fattori con uno spreco inutile di risorse e tempo rispetto agli obiettivi propri della valutazione stessa. Per esempio, le valutazioni rilevate potrebbero essere anche analizzate rispetto a altri caratteri: il genere, il tipo di scuola di provenienza, e forse anche il voto di diploma quale indicatore "implicito" di diverse qualità del rispondente — che, però, non è sempre disponibile e attendibile. Per esempio, accertare come agiscono sul giudizio sia il corso e sia la materia, potrebbe comportare il controllo della risposta dello stesso studente in corsi diversi; ciò richiede un modello statistico un po' più articolato delle tabelle di contingenza sopra utilizzate e una rilevazione più complessa dell'attuale. La conduzione di queste analisi comporta un aumento delle risorse umane, materiali, e temporali; inoltre, esorbita dagli obiettivi da conseguire con una indagine tipica per la valutazione della didattica da parte degli studenti frequentanti. L'analisi e l'esposizione dei risultati possono (e *devono*) limitarsi, a questo stadio e quale compito dei NdV, a strumenti semplici (v. *infra*, §3.1).

Il *sesto assunto* considera l'opportunità/ necessità di avere un *questionario unico* a livello nazionale o con un nucleo minimo di domande «adottate» da tutte le università per consentire la possibilità di confrontare i risultati tra loro. Ne consegue che tutte le obiezioni o critiche alle domande comuni non possono essere prese in considerazione dal NdV locale, ma vanno indirizzate al CNVSU. Il NdV dell'Università di Modena e Reggio Emilia ha adottato solo il nucleo minimo di domande previste dal DOC 9/02 (CNVSU, 2002) sicché le domande non possono essere assolutamente modificate.

La presente nota si occupa prevalentemente del questionario utilizzato per raccogliere le opinioni degli studenti sulla didattica svolta e è strutturata come segue. Il secondo paragrafo descrive sommariamente il disegno dell'indagine e, in particolare, il questionario utilizzato nella Università di Modena e Reggio Emilia, riportato in

«Appendice 1», con una breve nota critica sulla sperimentazione della raccolta dei dati via internet (*on-line*). Il terzo paragrafo contiene le osservazioni sugli aspetti del questionario che competono al NdV con un sottoparagrafo che contempla i suoi compiti di elaborazione dati. Il quarto paragrafo commenta le domande inerenti alla valutazione che, per l'Università di Modena e Reggio Emilia, coincidono con quelle indicate come contenuto minimo di ogni questionario e proposte dal CNVSU nel DOC 9/02 (CNVSU, 2002). Il CNVSU potrebbe, infatti, rivedere il documento (DOC 9/02) con il quale forniva le indicazioni ai NdV e, allora, tali osservazioni e riflessioni potrebbero costituire un certo interesse. Il quinto paragrafo colleziona, infine, alcuni problemi generali emergenti nella valutazione e nel questionario. Il sesto paragrafo illustra brevemente le caratteristiche delle scale di risposta adottate di solito in simili indagini. Il settimo paragrafo conclude brevemente.

Nell'esposizione, le osservazioni sono state indicate con la lettera O, seguita da due cifre: la prima indica la sezione o paragrafo e la seconda indica la numerazione progressiva. La risposta all'osservazione è stata indicata con la lettera R, seguita anch'essa da due cifre in modo analogo alla precedente e, per la versione elettronica, è scritta in blu, come le citazioni bibliografiche. La Contro Obiezione alla risposta, denotata con la sigla CO, è sempre numerata analogamente alla precedente e è scritta in verde. Le domande del questionario sono riportate in colore rosso.

2. Il disegno dell'indagine

In questo paragrafo si descrivono la metodologia (§2.1), il questionario (§2.2), l'organizzazione della rilevazione condotta nell'Università di Modena e Reggio Emilia (§2.3) e una nota critica sulla sperimentazione della raccolta dati via internet (§2.4).

2.1. La metodologia adottata

L'attività di rilevazione è coordinata dall'*Ufficio Pianificazione e Controllo* (UPC) che funge da supporto al NdV. La rilevazione è effettuata su *tutti gli insegnamenti* attivati, come in un censimento, e si sviluppa secondo le seguenti (sei) fasi principali.

- (a) *Censimento della programmazione didattica* delle singole Facoltà (o, in mancanza, dei singoli Corsi di Studio — CdS) e codifica delle variabili «insegnamento» e «docente», mediante la produzione di Tabelle di codici, che vengono restituite dall'UPC alle Facoltà, affinché i codici siano apposti dagli studenti sul questionario. Tale fase, che è necessaria in tutti i casi, è *stata fusa con la successiva* con l'introduzione del sistema «Esse3» che produce i codici degli insegnamenti.
- (b) *Distribuzione dei moduli (in formato PDF)* del questionario, da parte dell'UPC, alle Facoltà e/o ai Presidenti dei CdS, che provvedono alla consegna dei moduli stampati (o del PDF) ai docenti; questi ultimi lo consegnano agli studenti durante la lezione. Ora, ogni docente si collega a «Esse3» e stampa tante copie del questionario relativo alla materia che intende sottoporre a valutazione: la prima sezione del questionario sarà già precompilata (v. *infra*, §2.2). Una scelta che ha pro e contro, ma è efficace perché semplifica l'organizzazione della rilevazione «scaricandola» sul docente.

- (c) *Compilazione del questionario* per insegnamento e docente, in aula. Gli studenti ricevono dal docente il modulo cartaceo e lo compilano in aula a partire dalla seconda sezione, indicando la modalità di risposta preferita con una ics sul, o annerendo, il quadratino disponibile.
- (d) *Raccolta dei questionari compilati* in modo da garantire l'anonimato dei rispondenti. Uno studente è incaricato di ritirare i moduli, inserirli in una busta, e consegnarla presso l'ufficio designato al ritiro. Le Facoltà inviano le buste al Servizio IX Automazione e all'UPC che, tramite lettura ottica dei moduli, generano gli archivi dei dati.
- (e) *Trasformazione e elaborazione delle risposte*, a cura dell'UPC, da formato testo in dati numerici con un programma per SPSS (SPSS, 1997a,b); quindi, elaborazione dei dati per Insegnamento–Docente–Facoltà, dove per ogni domanda si fornisce media, deviazione standard, e numero di rispondenti.
- (f) *Trasmissione dei risultati* elaborati per l'unità di analisi Insegnamento-Docente ai Presidi delle Facoltà che li trasmettono, a loro volta, o direttamente ai docenti o ai Presidenti dei CdS, che li consegnano, infine, ai docenti.

2.2. Lo strumento di rilevazione

Il questionario utilizzato per la rilevazione della valutazione della didattica da parte degli studenti, riportato nell'Appendice 1, è formato da nove sezioni, secondo la proposta del CNVSU contenuta nel DOC 9/02 con qualche marginale adattamento che raggruppa le domande secondo una loro presunta rappresentanza semantica e le dispone nel questionario in modo da ridurre il più possibile l'effetto alone. La discussione è avvenuta nel NdV e nelle Facoltà durante la presentazione dei risultati dell'indagine precedente: un contributo fattivo è stato apportato dal professor Massimo Pilati.

La *prima* sezione concerne i dati di individuazione e riferimento dell'oggetto della valutazione: Facoltà, CdS, denominazione dell'insegnamento, e nome del docente.

La *seconda* sezione richiede le informazioni generali del rispondente: genere, tipo di diploma di maturità, anno di corso di iscrizione, posizione amministrativa di iscrizione, modalità di iscrizione, percentuale di lezioni frequentate del corso valutato, indicazione del motivo se ha frequentato meno del 50% delle lezioni, valutazione del numero medio di frequentanti le lezioni.

La *terza* sezione si occupa dell'«organizzazione dell'insegnamento», ossia l'adeguatezza del carico di studio rispetto ai Crediti Formativi Universitari (CFU) del corso valutato (D01), l'adeguatezza del materiale didattico (D02), l'utilità delle attività didattiche integrative (D03), la chiarezza della modalità di esame (D04).

La *quarta* sezione riguarda la docenza, ossia, il rispetto degli orari previsti (D05), la reperibilità per chiarimenti e spiegazioni (D06), le motivazioni dell'interesse verso la disciplina (D07), la chiarezza e il rigore dell'esposizione (D08).

La *quinta* sezione concerne la valutazione delle «infrastrutture», dove si chiede un giudizio sull'adeguatezza delle aule per le lezioni (D09) e delle aule e delle attrezzature per le attività integrative (D10).

La *sesta* sezione si occupa di raccogliere le opinioni sul possesso di conoscenze preliminari (D11), sull'interesse (D12), e sulla soddisfazione (D13). Le prime due,

assieme alla difficoltà della materia, costituiscono la «triade» che sta alla base dell'apprendimento e, quindi, della soddisfazione.

La *settima* sezione verte sull'organizzazione del CdS, ossia sul carico di studio complessivo (D14) e dell'organizzazione complessiva —orario, esami intermedi e finali— (D15) degli insegnamenti ufficialmente previsti nel periodo di riferimento.

L'*ottava* sezione è stata lasciata vuota per consentire la possibilità di introdurre tre domande libere, che possono essere decise dalla Facoltà, dal CdS, o dal docente.

La *nona* sezione invita a segnalare suggerimenti e difficoltà e osservazioni, espresse in nove proposizioni, per migliorare la qualità dell'attività didattica.

Il CNVSU (2002) richiede l'uniformità nel questionario per consentire confronti a livello nazionale e propone una scala di Likert (1932) a quattro modalità di risposta, ossia senza la posizione centrale, con sintagmi distanti dalle valutazioni scolastiche, famigliari ai rispondenti: ① *decisamente no*, ② *più no che sí*, ③ *più sí che no*, ④ *decisamente sí*. Per l'elaborazione dei dati si associano a tali modalità i valori della «scala decimale»: {2, 5, 7, 10}, rispettivamente, come suggerito da Chiandotto e Gola (2000). Si indicano, poi, media e varianza per la sintesi dei giudizi, ma non sono idonee per tali variabili qualitative ordinate. Per una valutazione empirica del funzionamento e dei limiti di questa scala, si vedano Lalla, Facchinetti, e Mastroleo (2004). Il NdV dell'Università di Modena e Reggio Emilia si è comunque uniformato alla direttiva.

2.3. L'organizzazione della rilevazione

L'obiettivo della scelta del periodo della distribuzione in aula dei questionari, agli studenti frequentanti la lezione, è raggiungere il maggior numero di rispondenti, in grado di esprimere un giudizio consapevole sull'attività didattica. Il NdV ha stabilito che, nelle singole Facoltà, *la rilevazione sia effettuata tra la terzultima e la penultima settimana di lezione* perché in questo periodo il numero degli studenti frequentanti, pur essendo diminuito, è ancora alto e gli studenti possono esprimere un giudizio con maggiore cognizione (Lalla, 2002); ma dove si eseguono quattro rilevazioni l'anno, a volte, si utilizza anche l'ultima settimana di lezione.

2.4. La sperimentazione della rilevazione via internet

Il NdV dell'Università di Modena e Reggio Emilia ha avviato una sperimentazione di raccolta via internet delle opinioni degli studenti sull'attività didattica, utilizzando una procedura all'interno del sistema «Esse3» che dovrebbe condurre, in futuro, all'automazione totale dei processi amministrativi, come la gestione degli esami, dei curricula degli studenti, dei programmi degli insegnamenti, e così via.

L'obiettivo della sperimentazione è ridurre il complesso e faticoso processo di raccolta delle opinioni degli studenti: si sprecherebbe meno carta, si ridurrebbe l'impegno di risorse umane coinvolte nel processo, si consentirebbe un aggiornamento *continuo* del giudizio del docente e una consegna immediata. Per visionare il suo giudizio, il docente deve soltanto collegarsi al sito «Esse3» e aprire il documento che contiene i suoi risultati. Si sono selezionate due Facoltà con caratteristiche specifiche e idonee per gli scopi della sperimentazione: la Facoltà di Bioscienze e Biotecnologie, perché là la frequenza è obbligatoria e è sufficientemente attrezzata per gli scopi; la Facoltà di Scienze della Comunicazione e dell'Economia (Reggio Emilia), perché là si

sta sperimentando la formazione a distanza e i docenti comunicano già ampiamente via internet. A queste si è aggiunto volontariamente un CdS di Medicina e Chirurgia.

Gli studenti delle Facoltà individuate sono obbligati a compilare il questionario perché il docente richiede loro la «ricevuta» attestante la compilazione per ammetterli all'esame. Si è lasciata ugualmente la possibilità al docente di ignorare l'obbligo, ma tutti dovrebbero attenersi a questo. Gli esiti ottenuti nel primo semestre sono lusinghieri: il numero di rispondenti è più alto di quello degli anni precedenti, ma occorre ancora attendere il secondo semestre per un giudizio definitivo. Al fine di evitare che lo studente dichiari di non avere frequentato, dove la frequenza non è obbligatoria, per non compilare il questionario, se ne è predisposto uno per non frequentanti che contiene un numero di domande equivalente a quello per frequentanti. In tal modo si potrà esplorare, con i soliti limiti fisiologici delle indagini, anche l'insieme dei non frequentanti. I primi risultati sono stati appena esplorati al fine di perfezionare il loro questionario.

Le critiche sul tipo di rilevazione sono numerose. Le più rilevanti, qui discusse, riguardano: (1) l'eterogeneità delle risposte perché sono raccolte in momenti diversi del tempo, (2) l'impossibilità di avere un giudizio completo alla fine del corso perché si avrà uno stillicidio di rispondenti che vi si aggiungono, (3) la possibilità che si creino gruppi di studenti che collaborano alla risposta per danneggiare un docente.

Premessa/caveat: ogni scelta, qui come nel séguito, comporta dei vantaggi e degli svantaggi. Bisogna operare, quindi, cercando sempre di equilibrarli il più possibile. È scontato il grande vantaggio economico e funzionale che scaturisce da una diffusa raccolta di valutazioni via internet. Per le critiche si osserva quanto segue.

(1) L'eterogeneità è un danno perché il giudizio sul docente avviene in ritardo e si basa sul ricordo che è fallace (v. *infra*, §5.2) per alcune domande; su altre, però, è più preciso perché lo studente compila quando pensa di sostenere l'esame: ha già una preparazione adeguata e il suo giudizio (per esempio, sul materiale didattico) è più preciso. La compilazione potrebbe avvenire molto precocemente, ma questo si può risolvere con finestre temporali per la compilazione e bloccandola in altri periodi.

(2) Il giudizio che si modifica di continuo è effettivamente un po' frastornante, ma, a regime, prima di iniziare di nuovo l'insegnamento, il docente avrà un giudizio completo, dato da tutti gli studenti che hanno sostenuto l'esame più qualcuno che ha valutato e non ha sostenuto, poi, l'esame. Si perde, comunque, il giudizio di chi ha frequentato, non ha valutato, e non ha sostenuto l'esame o lo sostiene l'anno successivo; ma qui valgono sia la risposta data al punto precedente, sia la premessa/*caveat*.

(3) I gruppi si creano o si possono creare anche nella compilazione in aula. La presenza del docente non riesce a impedire a un gruppo di dare risposte tutte negative.

3. DOMANDE di pertinenza del Nucleo di Valutazione interna

Le domande di pertinenza del NdV riguardano la terza sezione del questionario, dove si rilevano alcuni dati anagrafici, oltre all'organizzazione grafica complessiva.

O.3.1. Aspetto grafico del questionario.

La versione corrente non facilita la lettura e la rapida associazione fra risposte e caselle dove apporre. Le domande sono troppo ravvicinate soprattutto nella parte superiore.

R.3.1. L'osservazione è corretta, ma c'è un vincolo di spazio: tutto il questionario deve essere compreso in un solo foglio e su una sola facciata; pertanto, la compressione delle domande e la riduzione del carattere (che inficiano leggibilità, chiarezza, eleganza, e comprensibilità) sono diventate esigenze irrinunciabili.

CO.3.1. Va da sé che un *layout* (sia orizzontale, sia verticale) basato su due colonne permetterebbe di ottimizzare lo spazio e la leggibilità.

RCO.3.1. Il foglio è già troppo fitto, perciò si suppone che non sia possibile migliorare l'aspetto anche con le due colonne. L'autore (Lalla) ha lavorato sul documento, ma non è riuscito a ottenere un risultato migliore. Si può andare da un tipografo, ma salta tutta l'attuale procedura di rilevazione. Chi vuole, può richiedere il documento originale e provare o migliorarlo. Il NdV è oltremodo felice di avere collaborazioni in tal senso.

O.3.2. Motivo della frequenza inferiore al 50% delle lezioni.

A chi dichiara di aver frequentato meno del 50% delle lezioni, si chiede di specificare il motivo, scegliendo fra tre opzioni specifiche (lavoro, sovrapposizione con altre lezioni, altro). Si ritiene utile aggiungere anche una risposta del tipo: *“La frequenza alle lezioni di questo corso non è necessaria per l'apprendimento della materia”*.

R.3.2. L'osservazione è interessante, ma si ricade ancora nel vincolo di spazio. C'è chi osserva che, presumibilmente, la domanda viene intesa come *“Ritengo che la frequenza alle lezioni di questo corso non sia necessaria per il superamento dell'esame”*. In generale, si potrebbero mettere entrambe le modalità di risposta, ma aumenta semplicemente la necessità di spazio sul foglio, che è proprio quello che manca.

O.3.3. Il numero medio di studenti frequentanti nel periodo di riferimento.

O.3.3a. Che senso ha domandare ciò agli studenti?

R.3.3a. Il NdV ha ritenuto utile avere una stima del numero (medio) di frequentanti del corso per valutare sia l'affollamento, sia il tasso di risposta. Il dato è, pertanto, utile.

CO.3.3a. Il tasso di risposta è oggettivamente dato dal rapporto percentuale tra i questionari compilati e gli iscritti all'anno di corso (per i corsi obbligatori è la soluzione più semplice) o gli studenti che superano l'esame nel corso dell'anno.

RCO.3.3a. Si tratta di due informazioni diverse: la prima esprime il tasso di partecipazione tra i presenti in aula, la seconda esprime il tasso di partecipazione rispetto agli iscritti nell'anno di corso che prevede quell'insegnamento. Tra i rispondenti potrebbero esserci studenti iscritti a anni successivi o addirittura a altre Facoltà.

O.3.3b. Non sarebbe più logico chiederlo al docente?

R.3.3b. Si deve necessariamente accreditare l'onestà della risposta e la capacità di valutare degli studenti, anche se si può presumere che non tutti opereranno con l'attenzione e la precisione richiesta o desiderata. Si potrebbe dire lo stesso dei docenti; infatti, si deve notare che: (1) anche il docente è soggetto allo stesso tipo di errore di valutazione dello studente e, per giunta, in questo caso si avrebbe una sola valutazione; (2) diventa più complicato inserire la risposta nei questionari; (3) è più macchinosa la raccolta dei dati. Si può ancora *obiettare* che il docente ha una visione completa dell'aula, rispetto allo studente, ma egli dovrebbe ugualmente valutare il numero di posti vuoti e la variabilità del numero dei frequentanti nel tempo; in ciò è soggetto allo stesso rischio di errore degli studenti: può sbagliare in più e in meno proprio come tutti.

CO.3.3b. Dati come questo e altri (adeguatezza dell'aula) potrebbero essere rilevati con un questionario che il docente compila contestualmente alla distribuzione/compilazione del questionario degli studenti. Si aggiunga che, conoscendo la capienza dell'aula, il conteggio dei posti occupati/liberi è compito risolvibile senza rischio di errore.

RCO.3.3b. L'obiezione è corretta in sé, ma ignora sia che il docente non conta ogni giorno i posti vuoti e occupati, annotando gli esiti, sia la scansione e le esigenze di una procedura di raccolta dati tramite questionario, finalizzata alla costruzione di una «matrice dei dati». *Le obiezioni devono essere contestualmente mirate, non isolate dal contesto, come una informazione a sé (v. supra, §2.4, il capoverso «premessa/caveat»).*

O.3.3c. Se si ottiene molta variabilità nelle risposte (come sempre avviene), si procede come al solito facendo una media fra di esse, ma così si compie una operazione non legittima, dato che si mettono insieme risposte corrette con risposte sbagliate.

R.3.3c. Risposte vere e risposte sbagliate sono presenti in qualunque tipo di indagine. Le valutazioni di ogni tipo di grandezza sono affette da errori, anche marchiani, ma si assume, in genere che si possa sbagliare in più e in meno con la stessa probabilità: approssimativamente gli errori dovrebbero distribuirsi secondo una normale, ma andrebbe accertato empiricamente. Si può adottare, poi, un indice che ne riduca gli effetti, almeno per i corsi molto affollati (mediana, medie lasche, medie troncate).

O.3.4. Insegnamenti a moduli.

Nel questionario si dovrebbero identificare i diversi moduli, ossia vi dovrebbe essere uno spazio per il docente di ogni modulo.

R.3.4. Gli insegnamenti a moduli rappresentano un problema complesso, specie nella Facoltà di Medicina e Chirurgia, dove uno stesso insegnamento può essere composto da *numerosi moduli*. Diventa, allora, assai complicato predisporre un questionario con tanti moduli, variabili da un insegnamento all'altro e da una Facoltà all'altra. Ogni modulo è svolto da un docente e, dunque, la valutazione è effettuata su quel modulo e su quel docente. Si obietterà che vi sono domande inutili o ridondanti perché già chieste, come quelle inerenti alla adeguatezza della struttura (ma non è detto che le lezioni si svolgano nella stessa struttura: in un insegnamento di anatomia, il modulo sul sistema nervoso potrebbe essere svolto in un'aula diversa dal modulo sul sistema muscolare), o alla organizzazione del CdS, o ai «dati degli studenti». Si deve considerare, però, che per una piccola ripetizione (rispetto al numero di studenti coinvolti), si ottiene una grande semplificazione nella raccolta e nella elaborazione dei dati. Si aumenta, così, il carico di lavoro per gli studenti: è importante anche il loro tempo, non solo quello dei docenti che protestano per il tempo che devono perdere a stamparsi i questionari. Per tale problema non c'è, però, una soluzione ragionevole, completa, e semplice (v. *supra* RCO.3.3b).

Il problema può apparire facilmente risolvibile nella compilazione via internet del questionario perché nella valutazione dei moduli, successivi al primo, si possono presentare già compilate le domande alle quali, *ipoteticamente*, bisogna rispondere una sola volta. Sulla prima difficoltà, in parte si è già detto notando che l'aula può cambiare da un modulo all'altro; infatti, non si può sapere a priori quali sono le domande «fisse», eccetto le D14 e D15: possono cambiare la percentuale di lezioni frequentate, il numero di frequentanti, le conoscenze preliminari (D11), l'interesse (D12), e la soddisfazione (D13). I dati anagrafici sono compilati automaticamente, nella rilevazione via internet. Per evitare che lo studente ripeta le valutazioni delle D14 e D15, bisogna obbligarlo a

valutare tutti i moduli nello stesso momento perché, per la riservatezza, appena un questionario è compilato, è anche «dimenticato» dal sistema. Si può tenere memoria, solo se valuta i moduli uno dopo l'altro. Occorre considerare, poi, che anche il lavoro di programmazione diventa più complesso e impegnativo. Per due domande, si impone un vincolo allo studente e un aumento di lavoro per i responsabili. Allo stato attuale si può soprassedere; in futuro, però, non si può escludere la possibilità di realizzarlo più per l'eleganza della soluzione, che per la convenienza.

O.3.5. Osservazioni e suggerimenti finali.

Meglio una domanda aperta (si usa in altri atenei) da fare avere ai docenti. Si potrebbero mettere alcune (poche) domande aperte nel retro del foglio e togliere la parte finale.

R.3.5. Vi sono diversi Atenei che hanno previsto la possibilità di una domanda a risposta aperta, in cui gli studenti scrivono liberamente le loro osservazioni, da inviare direttamente ai docenti. L'esperienza svolta nella Facoltà di Economia, dal lontano 1989, ha mostrato che tali giudizi erano, sí, spesso interessanti, ma il tasso di risposta era basso; inoltre, nei corsi più difficili vi era solo una mole di insulti e amenità di vario genere. Non si vuole, però, depotenziare la forza dell'osservazione perché si sono riscontrate anche indicazioni interessanti che, tra l'altro, non si sono attuate per varie ragioni. Il NdV ha scelto una strategia di semplificazione perché la struttura, allo stato attuale, non ha le potenzialità per fornire anche questo servizio: si tratta di acquisire lo scritto come fotografia, collezionare le fotografie dei testi, e spedire il documento ai singoli docenti per posta elettronica. Per via cartacea si avrebbe uno spreco enorme di carta con il relativo carico di lavoro! Altre Università lo fanno: per esempio, l'Università di Firenze. *Se il docente è interessato ai giudizi liberi può richiederli fornendo, durante la compilazione del questionario ufficiale, un foglio bianco sul quale richiede il giudizio libero; infine, farli raccogliere da uno studente che li porta in un ufficio a ciò adibito, al fine di garantire l'anonimato.* L'esperienza citata ha mostrato anche che sono più interessanti le risposte libere guidate (questionario semi-strutturato): sul materiale didattico, sull'organizzazione degli esami, sulle parti di programma che si ritengono critiche, e così via; altrimenti, lo studente resta un po' spaesato o il docente non riceve le risposte sui temi ritenuti utili per l'organizzazione del corso.

CO.3.5. La raccolta delle osservazioni degli studenti a metà corso e alla fine del corso non pone alcun problema, ma presuppone l'intenzione del docente di tener conto dei suoi utenti e di migliorare la qualità del suo lavoro didattico.

RCO.3.5. L'Università è costituita dai docenti, che dovrebbero avere cura del proprio lavoro e svolgerlo con senso di responsabilità e attenzione verso gli studenti e la disciplina impartita. Senza il cambiamento radicale delle risorse umane, nell'accezione etimologica «alla radice», non cambierà alcunché e non saranno sufficienti le migliori procedure e i controlli più severi. L'università siamo noi: «orgoglio e dedizione». Senza questi «presupposti» non si va molto lontano.

O.3.6. Scelta del tempo o periodo per la somministrazione del questionario.

Il questionario è somministrato, in genere, alla fine del corso (ultima settimana); ciò introduce un elemento di distorsione nei dati raccolti: tra i rispondenti sono, infatti, sottorappresentati (o addirittura assenti) gli studenti che hanno interrotto la frequenza (causa la loro insoddisfazione per il corso stesso). *Soluzione:* andrebbe previsto un questionario all'inizio del corso e uno alla fine.

R.3.6. Si può rinviare alla relazione per l'anno 1999 (Lalla, 2002), qui riportata e adattata (v. *infra* §5.2). Qui ci si limita a commentare la doppia rilevazione perché se attuata, si otterrebbe subito quasi il raddoppio del costo e del tempo di lavoro. Allo stato attuale è da scartare e contrasta con il secondo, terzo, e quinto assunto. Se fosse istituzionalizzata, infatti, occorrerebbe anche un questionario diverso tra l'inizio e la fine; inoltre, riguarda il CNVSU. Nelle Facoltà con due semestri e in ciascun semestre con due periodi di insegnamento, si potrebbe in parte realizzare, almeno per i corsi con 8CFU (Crediti Formativi Universitari). In tali Facoltà, infatti, si eseguono quattro rilevazioni l'anno. Un corso di 8 CFU che si svolge, per fissare le idee, nel primo semestre si trova a potere eseguire due rilevazioni: una alla fine del primo periodo del primo semestre e una alla fine del secondo periodo del primo semestre; allora, si può anticipare la prima rilevazione verso l'inizio del corso. Con lo stesso questionario si ottiene, forse, una scarsa «resa informativa» e, tuttavia, si ha ugualmente una piccola complicazione nella raccolta e anche nell'elaborazione. Il NdV potrebbe raccogliere correttamente i dati e lasciare agli interessati il documento (*file*) per una elaborazione individuale o da parte del docente stesso o da parte di un responsabile indicato dalla singole Facoltà. Per ora, *si può escludere che tale lavoro lo svolga il NdV.*

CO.3.6. Nulla osta rilevare il dato sui frequentanti in tre momenti diversi del corso: fine prima settimana (inizio), metà corso, penultima settimana. Si avrebbero indicatori interessanti: percentuale di frequentanti sui frequentanti potenziali, variazioni percentuali dall'inizio a metà corso, da metà corso a fine corso, dall'inizio a fine corso.

RCO.3.6. Se ci si limita al dato sui frequentanti si svisciva l'osservazione precedente e non riguarda più il processo che deve seguire il NdV.

O.3.7. Analisi dei dati amministrativi.

Ci si può aspettare che il NdV usi e analizzi dati amministrativi per ricavarne indicatori, come la distribuzione degli iscritti per CFU e voti? Si noti che questi dati sono previsti nel modello di rapporto di autovalutazione della CRUI, ma ancora non ne disponiamo.

R.3.7. Alcuni dati amministrativi sono già riportati nelle relazioni annuali del NdV, che nessuno legge, e a regime riguardano almeno l'insieme dei dati da trasmettere al CNVSU. L'osservazione è corretta e interessante; tuttavia, esistono ardue difficoltà per eseguire tali elaborazioni in modo sistematico. Una possibile soluzione è l'automazione completa di tutti i servizi. Il nostro Ateneo si sta muovendo in tale direzione tramite l'accordo istituito con il CINECA e l'adozione di «Esse3». In prospettiva, si dovrebbe adottare un sistema di *DataWarehouse* che consentirebbe una elaborazione semplificata dei dati amministrativi, ma occorre sempre un investimento in capitale umano e la creazione di procedure standardizzate per ottenere i risultati desiderati.

CO.3.7. No, non ci siamo; per il questionario di autovalutazione (modello CRUI) viene richiesta la compilazione di 77 schede informative, ma non si ritiene opportuno/necessario inviare ai valutatori le informazioni desumibili dai dati in possesso delle segreterie di facoltà. Circa le risorse umane, si valuti il monte ore che l'Università di Modena e Reggio Emilia carica sui docenti coinvolti (40) tra corso di formazione (24 ore) e compilazione (tra le 100 e le 200 ore): siamo su ordini di grandezza che oscillano tra le 4.000 e le 8.000 ore uomo docente.

O.3.8. Titolo di studio con modalità di risposta troppo aggregate.

Per alcune Facoltà, la modalità «Altro» del titolo di studio presenta una percentuale assai elevata, oltre il 30%. Perché non disaggregarlo di più?

R.3.8. Sul questionario cartaceo ci sono problemi di spazio e non sarà facile, quindi, rispondere a questa esigenza, mentre non ci sono problemi per la rilevazione via internet. Purtroppo, nelle Facoltà coinvolte dal dato non c'è neanche la sperimentazione.

3.1. I compiti del Nucleo di Valutazione interna

Osservazioni frequenti riguardano la necessità di eseguire analisi più precise, che tengano conto delle correlazioni esistenti tra le variabili, degli effetti negativi che i caratteri anagrafici o alcune dimensioni hanno sulle altre, delle differenze esistenti tra i CdS, del diverso significato che ha la mancanza di requisiti minimi per lo studio di un insegnamento tra quelli del primo anno e quelli degli anni successivi. O Ancora, di applicare tecniche di analisi longitudinale o di estrapolare alcuni problemi per Facoltà al fine di adeguare le loro specificità, valorizzare i dati tramite confronti tra Facoltà e CdS per migliorare la programmazione, eccetera.

La risposta a tali richieste è negativa e deriva dai paradigmi di riferimento dell'indagine (assunti due, tre, e cinque) perché non ci sono le risorse umane per fornire alle Facoltà le risposte alle loro domande: alcuni approfondimenti esorbitano dagli obiettivi della valutazione della didattica svolta a tappeto. Il NdV può fornire i dati individuali, le singole schede di valutazione degli studenti, alle Facoltà che, con il vincolo del rispetto della riservatezza, possono incaricare il loro personale di eseguire tutte le elaborazioni desiderate. Nulla osta. I dati aggregati, invece, sono già disponibili, elaborati, e commentati; ma solo qualcuno visiona o legge (i commenti). La Facoltà di Lettere e Filosofia ha prodotto una nota di alcune pagine, che rappresenta un saggio esemplare di analisi secondaria; non può essere il NdV a produrre questo, ma sono proprio le Facoltà che, nella piena autonomia e con reale attenzione ai problemi della didattica, si analizzano i dati. Non si può sempre delegare un «altro» a dirci «com'è» e «cosa fare»: per assolvere tali compiti il NdV dovrebbe diventare un «centro studi».

Un altro problema che emerge nell'esame dei dati è come affrontare il caso di un docente che risulta poco chiaro e l'esame non è ben organizzato. Come si può intervenire? Chi può decidere, seppure «cosa», «come» un docente debba insegnare?

Un problema complicato e delicato. Si potrebbero organizzare dei corsi che «insegnino a insegnare» perché, in generale, si è sempre assunto che la conoscenza della materia sia sufficiente di per sé a garantire la capacità di trasmetterla ai discenti. In realtà, insegnare è un'arte e l'esperienza sul campo aiuta a migliorare le capacità organizzative e espositive di ogni docente; tuttavia, cognizioni di pedagogia e di comunicazione associate a un addestramento «controllato» potrebbe contribuire a migliorare la «resa didattica» dei docenti. Anche per questa attività mancano, però, sia i mezzi (risorse finanziarie) e sia i tempi (perché più tempo si dedica alla didattica e meno tempo rimane per la ricerca). In ogni caso, non è chiaro se questo sia un compito del NdV o delle Facoltà, se non delle aree scientifiche più ristrette e meglio definite.

4. DOMANDE del Comitato Nazionale di Valutazione del Sistema Universitario

In questa sezione si esaminano sia alcuni aspetti generali delle domande presenti nel questionario di valutazione, sia le singole domande. L'ordine di esposizione delle domande è quello adottato dal NdV dell'Università degli Studi di Modena e Reggio Emilia, che ha modificato l'ordine proposto dal CNVSU nel DOC 09/02, al fine di porre la valutazione del docente al centro dell'attenzione del rispondente. Per evitare l'effetto alone, infatti, si sono spostate in fondo le domande inerenti all'organizzazione del CdS, che sono state precedute dalle tre domande (triade) inerenti alle conoscenze preliminari, all'interesse, e alla soddisfazione. Rispetto alle domande originali si sono apportate lievi modifiche che saranno illustrate nel séguito, spesso con il parere contrario di chi scrive perché ogni difformità altera inevitabilmente la risposta e inficia la possibilità di confrontare i dati, anche se, come si vedrà, le variazioni sono spesso trascurabili.

4.1. Osservazioni critiche generali sulle domande

Tutte le domande poste hanno la stessa direzione (la risposta diventa più positiva mano a mano che si procede da sinistra a destra della scala). Questo favorisce gli effetti di distorsione: *response set* [«risposte meccanicamente uguali tra loro», (Corbetta, 1999, p.201, p.208) o «serialità»], acquiescenza (tendenza a scegliere le risposte che esprimono accordo, come «Più sí che no»), desiderabilità sociale (scelta della risposta che l'autorità si aspetta dal rispondente). *Il problema potrebbe essere evitato* sostituendo le domande con affermazioni in entrambe le direzioni: positive e negative (come, per esempio, nelle scale di atteggiamenti) rispetto alle quali il rispondente deve dichiararsi più o meno d'accordo (si possono mantenere le quattro modalità di risposta). Dal punto di vista linguistico, è anche più familiare dichiararsi «Abbastanza d'accordo» o «Abbastanza in disaccordo», rispetto a rispondere «Più no che sí» o «Più sí che no».

La valutazione si desume da un insieme di opinioni che ogni individuo esprime. Le domande devono essere formulate, quindi, in modo da permettere al rispondente di esprimere una opinione, cioè una posizione personale variabile che non ha un referente di realtà controllabile in modo oggettivo (cosa che non si può fare attraverso la raccolta delle opinioni). Se esiste un referente di realtà («dato di realtà») non si tratta di una opinione, ma di risposta corretta o sbagliata. Se le risposte a domande di questo tipo variano molto, non indicano un giudizio negativo sul servizio che si offre, ma informano degli errori che fa chi risponde al questionario. Alcune domande del questionario sono formulate in relazione a aspetti che hanno un «referente di realtà» e non possono essere valutati tramite opinione.

L'osservazione contenuta nell'ultimo capoverso ha un fondamento tecnico ineccepibile, eppure la valutazione dei cosiddetti «dati di realtà» da parte degli studenti è ineliminabile perché vi è un conflitto di interessi tra i docenti e gli studenti. I docenti hanno sempre l'interesse a sostenere che le strutture sono adeguate e la didattica svolta per alcuni «dati di realtà» (orario, puntualità) è efficiente e efficace. Si ammetta pure che esistano standard di riferimento e che gli «elementi» valutati siano a essi conformi. La valutazione degli studenti, esprimendo una percezione soggettiva sul dato di realtà, non dovrebbe essere richiesta? Chi valuta, allora, e come questo dato di realtà? Per esempio, un'aula potrebbe essere bella e spaziosa, ma avere difetti di sonorità e/o di visibilità. Chi più dello studente è in grado di valutare? Il docente sostiene sempre che è puntuale. È possibile anche che lo sia e nella maggior parte dei casi è vero. Anche il

docente piú puntuale potrebbe ricevere un giudizio incoerente con il dato di realtà e il docente meno puntuale, ma molto accattivante, avere un giudizio piú favorevole. Perché eliminare la possibilità di esprimere la *percezione soggettiva* dello studente anche sul dato di realtà? L'importante è dargli il giusto peso o *instaurare altri meccanismi efficaci del controllo dei dati di realtà* e le due strategie non si escludono a vicenda. Sul cosiddetto «dato di realtà», infine, si legga il testo in R.4.9a (v. *infra*).

4.2. Osservazioni sulle singole domande

Le parti della *versione originale* delle domande, come proposte nel DOC 9/02, sono indicate *in viola* e riportate *tra parentesi quadre*, mentre tra parentesi graffe c'è la modifica apportata. Se si legge il testo, allora, saltando la parentesi quadra in viola e leggendo la graffa, si ha la domanda presente nel questionario dell'Università di Modena e Reggio Emilia; se si legge il testo saltando la parentesi graffa scritta in rosso e leggendo la quadra scritta in viola, si ha la domanda proposta nel DOC 9/02.

D01 (D09 del DOC 9/02) — «Il carico di studio {di} [richiesto da] questo insegnamento è proporzionato ai crediti assegnati?».

O.4.1a. La presenza di un syllabo (programma del corso lezione per lezione con indicazione dei saperi e dei testi di riferimento) rende la domanda superflua. Se il docente organizza la didattica lezione per lezione non può non definire il carico di lavoro che viene richiesto agli studenti. Naturalmente, ciò presuppone che il docente sappia e voglia redigere un syllabo.

R.4.1a. L'obiezione *presume* che il docente sia l'unico in grado di valutare cosa fare e come farlo bene e proporzionatamente. Esiste una regola aurea? In base a quale criterio, se si vuole rispettare il dettame legislativo sul rapporto tra contenuti e tempi necessari per apprenderli, il docente può decidere il tempo richiesto dalla «quantità» di nozioni impartite? Una indagine condotta nella Facoltà di Economia, sugli insegnamenti dei CdS del vecchio ordinamento, ha mostrato una varietà di distribuzioni dei tempi di studio, necessari per preparare l'esame, dalle quali non era facile trovare un criterio adeguato e ragionevole anche in senso statistico; inoltre, la valutazione dei tempi di studio richiede una indagine impegnativa e comunque difficile e poco attendibile come risultati, se gli intervistati non operano con una dedizione e precisione enormi.

O.4.1b. Siamo alle solite: si chiede allo studente una valutazione che dovrebbe essere richiesta a altri. Non ci dovrebbe essere un coordinatore del CdS in grado di controllare i contenuti e i carichi di lavoro (esigendo la preparazione dei sillabi ai docenti)?

R.4.1b. Il paradigma sotteso da tale obiezione è che sia tecnicamente possibile valutare con precisione il rapporto tra crediti e carico di studio. Ciò sottende, come già detto, una superiorità e gnoseologia del docente che è insostenibile non solo in senso assoluto, ma, soprattutto, in senso relativo perché il carico di studio va commisurato proprio sulla base delle capacità degli studenti che sono ignorate completamente dal docente. Ne consegue che la risposta dello studente è fondamentale. Piú complicata è l'obiezione che considera i non frequentanti, qui non citata, perché non partecipando alle lezioni guadagnano tempo per lo studio, ma potrebbero perderlo per capire argomenti difficili. Ci si astiene, però, da ulteriori approfondimenti perché non sono determinanti per gli obiettivi dell'indagine; inoltre, le due valutazioni non sono incompatibili. Il Presidente

di CdS controlla i contenuti del corso rispetto all'obiettivo formativo e lo studente esprime il suo giudizio sulla pesantezza del corso. O deve solo «obbedir tacendo»?

O.4.1c. Questa domanda andrebbe posta prima dell'esame. Quando il corso è ancora in fase di svolgimento molti studenti non hanno ancora iniziato a prepararsi. La risposta NO è ambigua: può essere NO sia per un corso «leggero», sia per un corso «pesante».

O.4.1d. La domanda può essere riformulata in forma di opinione: «*L'impegno di studio che questo insegnamento ti richiede, rispetto alla media degli insegnamenti con gli stessi crediti, è secondo te: (molto maggiore, un po' maggiore, uguale, un po' minore, molto minore)*». In questo caso si fa riferimento esplicito a una esperienza soggettiva (il tuo impegno di studio) e non a una proporzione fra carico (presumibilmente, in termini di quantità di materiale da studiare) e crediti invariabile per tutti gli studenti di un corso.

R.4.1d. La domanda andrebbe riformulata perché la scala non può essere cambiata per motivi di spazio e uniformità. Si potrebbe chiedere: «*L'impegno di studio che questo insegnamento ha richiesto è molto più alto della media degli insegnamenti con gli stessi crediti?*». I precisi non mancherebbero di formulare osservazioni critiche: chi sostiene gli esami per la prima volta non ha conoscenze sufficientemente adeguate per eseguire confronti, è soggetto agli errori di memoria, il cambiamento della domanda snatura la misura dell'atteggiamento sbilanciandolo in una sola direzione, e così via.

D02 (D10 del DOC 9/02) — «Il materiale didattico (indicato o fornito) è adeguato per lo studio della materia?».

O.4.2.a. I programmi dei corsi devono essere pubblici e visibili da parte di chiunque. I programmi, metodi, e contenuti didattici dovrebbero essere oggetto di analisi da parte di altri docenti dello stesso settore scientifico disciplinare. Valgono, pertanto, le prime due osservazioni precedenti (4.1a, 4.1b).

R.4.2a. Valgono, perciò, le stesse risposte.

O.4.2b. Cosa si intende, inoltre, per materiale didattico? Si intendono sillabi, lucidi, e materiali per esercizi o anche testi?

R.4.2b. Il problema di definizione esiste, è vero; ma, senza bizantinismi, esiste anche l'accezione comune del termine che può includere qualunque tipo di materiale che deve essere visionato per il corso, incluso il programma stesso. Può essere che il programma sia scritto bene e i testi siano scadenti e/o non presenti in commercio, ma il giudizio è complessivo. Il docente interessato a capire i problemi e a provvedere, dopo avere ricevuto un giudizio negativo, può porre in atto diverse azioni: aggiunge le domande nel questionario del NdV (ve ne sono tre libere), conduce interviste informali e formalizzate tra gli studenti, se è il caso, per accertare i termini della questione. Si svolge così, correttamente, il proprio lavoro, che richiede tempo. Tempo che i docenti, purtroppo, non hanno perché finirebbero per svolgere solo didattica; ma una didattica seguita con attenzione richiede un impegno che comporta anche questi compiti di indagine per migliorare la propria prestazione nell'attività formativa.

O.4.2c. Se i libri completi sull'argomento non esistono, almeno per come il docente svolge il corso? Che senso avrebbe la risposta? Il docente si dovrebbe limitare a seguire pedissequamente un libro di testo? Il materiale aggiuntivo dovrebbe essere scritto dal

docente stesso? Il materiale risulterebbe troppo «dispersivo» e il giudizio potrebbe essere negativo, ma in sé, il materiale sparso qua e là potrebbe essere un pregio per la formazione e un difetto per l'efficacia dello studio. Eccetera.

R.4.2c. Le obiezioni/ domande sono molto più insidiose e *vere* nelle lauree magistrali, un po' meno nelle lauree triennali, forse eccetto il terzo anno, ma vale sempre R.4.2b.

O.4.2d. L'adeguatezza del materiale didattico non dovrebbe essere valutata da una persona competente, un docente o più docenti?

R.4.2d. Si vedano anche le risposte R.3.3a e R.3.3.b. Tale rilievo non è cogente perché un materiale perfetto per un esperto della materia potrebbe essere inadeguato per uno studente; perciò, il parere dello studente, seppure derivante da «percezione soggettiva» e non competenza, dovrebbe essere tenuto in grande considerazione. Il materiale didattico deve essere semplice, chiaro, efficace, e in grado di fare apprendere la materia; pertanto, il giudizio dello studente è cruciale e costituisce una vera informazione utile per il docente; semmai, il problema è che solo una valutazione di intensità della qualità non è in grado di suggerire dove sta la lacuna o l'inadeguatezza. Il docente, autonomamente, dovrebbe sondare il parere degli studenti sul materiale didattico, *indipendentemente* dalla valutazione ufficiale eseguita dal NdV; dovrebbe affinare le sue capacità o ne potrebbe discutere con i colleghi della stessa e di altre università.

CO.4.2e. Ripeto che altrove è prassi fornire queste informazioni all'inizio dell'anno per tutti i corsi: nelle università USA *catalogs* e sillabi sono la norma. I sillabi definiscono sempre in modo chiaro le modalità di esame e, nel caso di modalità multiple, il peso percentuale assegnato a ciascuna di esse.

R.CO.4.2e. Altrove è prassi formulare domande simili: ecco una che è tratta da un «*Question Bank*» di Adelaide (Australia, <http://www.flinders.edu.au/teach/SET/>): «*The teaching materials (e.g. handouts, study guides, reading list, etc) were valuable for my understanding of the topic*». Ecco un'altra (aperta) tratta dal questionario di valutazione dell'università del Minnesota (USA, <http://eval.umn.edu/set.html>): «*Please comment about anything in particular that you like or dislike about: ... The reading material*».

D03 (D11 del DOC 9/02) — «Le attività didattiche integrative (*esercitazioni, laboratori, seminari, ecc.*) {risultano} [sono] utili ai fini dell'apprendimento?».

O.4.3. Tale domanda andrebbe posta solo agli studenti che frequentano queste attività.

R.4.3. Si vedano anche le risposte R.3.3a e R.3.3.b. Si deve usare un unico strumento di rilevazione che comprende le possibili situazioni; infatti, è stata prevista la modalità di risposta «NON prevista». Si può ancora obiettare che si riscontrano sul campo risposte e, quindi, valutazioni in casi dove è noto che non ci sono attività integrative. Tale fenomeno, come tanti altri che si osservano sul campo (l'acquiescenza, la desiderabilità, la reattività), è connesso a un tipo di errore fisiologico nella raccolta dei dati, che è stato definito «coercizione alla risposta» valutativa anche quando la modalità corretta sia diversa dalla scala di intensità (Lalla, 2005): è, appunto, il caso in cui occorre rispondere «NON prevista» e, invece, il rispondente continua a esprimere ugualmente una sua valutazione per errore, disattenzione, acquiescenza, «serialità» (*response set*) .

D04 (D03 del DOC 9/02) — «Le modalità di esame sono state definite in modo chiaro?».

O.4.4. Non sarebbe meglio se la risposta la desse una sola persona competente (un docente) sulla base dei sillabi/programmi dei corsi?

R.4.4. Si vedano anche le risposte R.3.3a e R.3.3.b. Se l'informazione, benché esposta correttamente, non ha raggiunto gli interessati è un fatto degno di rilievo e di considerazione per ridurre il tasso di mancato raggiungimento del soggetto ricevente la comunicazione, anche se ciò garantisce il docente da qualunque tipo di contestazione. Se si pensa, poi, che sono studenti frequentanti, una valutazione insufficiente deve fare riflettere. Non esistono sempre risposte al problema e strategie capaci di eliminarlo radicalmente; ma avere indizi di acquisizione delle informazioni è utile e interessante.

CO.4.4. Si ripete con fermezza che altrove è prassi fornire queste informazioni all'inizio dell'anno per tutti i corsi: nelle università USA, *catalogs* e sillabi sono la norma. I sillabi definiscono sempre in modo chiaro le modalità di esame e, nel caso di modalità multiple, il peso percentuale assegnato a ciascuna di esse.

R.CO.4.4. Altrove è prassi formulare domande simili: si citano alcune tratte dal solito «*Question Bank*» di Adelaide (Australia): (a) «*Sufficient information was provided about the exam format*», (b) «*Sufficient guidance was provided for exam preparation*», (c) «*The exam was a fair test of the content of the topic*», (d) «*The exam was a fair test of my knowledge*», (e) «*I received adequate feedback on my exam performance*».

D05 (D04 del DOC 9/02) — «**Gli orari di svolgimento dell'attività didattica sono rispettati?**».

O.4.5a. Si fa riferimento alla pratica del quarto d'ora accademico o a quella delle «lezioni saltate»? In entrambi i casi che senso ha chiedere una cosa del genere a tutti gli studenti? Non basta un bidello (o uno studente)?

R.4.5a. Si può riscontrare, sí, una discrasia tra un dato di realtà e un dato di percezione soggettiva. L'esperienza empirica sul campo e l'analisi delle valutazioni degli studenti della Facoltà di Economia, condotta dal 1989, hanno mostrato che tale giudizio è influenzato dal tipo di (o difficoltà della) materia, dalle qualità del docente (che è poco puntuale, ma affascinante e coinvolgente), dalla mancata puntualità dello studente stesso che non è in grado di valutare e per non sbilanciarsi dà un giudizio approssimativo, ma impreciso. Si può ritenere, tuttavia, che tale domanda abbia un valore di responsabilità verso lo studente, che può giudicare un aspetto importante dell'organizzazione, e di «controllo» sul docente, che sa di essere valutato anche sulla sua puntualità.

O.4.5b. Quale contenuto informativo restituiscono le risposte che si ottengono? Presumibilmente, chi valuta si basa su ricordi più o meno sbiaditi. Per valutare questi aspetti occorre usare strumenti diversi dal questionario rivolto agli utenti.

R.4.5b. Si veda la risposta R.4.5a. Si può aggiungere che il contenuto informativo di questa domanda sia molto basso perché ci si aspetta che tutti i docenti siano responsabili e, quindi, saranno pochissimi quelli che indurranno gli studenti a valutazioni negative. Per avere una precisione accettabile bisogna, forse, introdurre la discutibile e discussa procedura di «timbrare il cartellino» da parte del docente; ma è questione di precisione o di funzionalità? Può accadere che un docente sia spesso assente o in ritardo per ragioni di salute o impegni accademici/istituzionali imprescindibili. Sia nei casi giustificati e sia nei casi non giustificati, in genere, le Presidenze conoscono già la situazione. Il problema è che provvedimenti prendere quando si verificano tali

sgradevoli situazioni, ossia, ritardi e assenze ingiustificate, almeno per quanto è dato dal «dato di realtà», conosciuto dai Presidi e non sempre dagli studenti. Allo stato attuale non vi sono strumenti di pressione adeguati o casi di provvedimenti presi per tali disagi.

D06 (D05 del DOC 9/02) — «Il [personale] docente è effettivamente reperibile per chiarimenti e spiegazioni?»

La modifica introdotta dal NdV altera il significato della domanda perché si passa dall'impressione generale di valutazione globale del personale docente, alla valutazione del singolo che, però, è proprio l'obiettivo della valutazione della didattica da parte degli studenti. Il cambiamento è inopportuno, anche se ragionevole perché è il docente l'oggetto della valutazione della reperibilità; diversamente, finisce tutto in un *calderone*.

O.4.6a. Se si fa riferimento alla reperibilità del docente in orario di ufficio valgono le considerazioni già fatte per gli orari.

R.4.6a. Valgono, perciò, le stesse risposte.

O.4.6b. È ambigua. Significa: “il docente è disponibile durante la lezione a rispondere a domande o richieste di chiarimenti” oppure “il docente è presente durante le ore di ricevimento studenti”? Nel primo caso il termine “reperibile” non è adatto, ma la domanda sarebbe interessante ai fini della valutazione. Nel secondo caso occorre prevedere una modalità di risposta per la maggior parte degli studenti, cioè coloro che non hanno mai cercato il docente nell'orario di ricevimento.

D07 (D07 del DOC 9/02) — «Il docente stimola/motiva l'interesse verso la disciplina?».

O.4.7. Si tratta di una domanda che si presta a differenti interpretazioni che dipendono da differenti punti di vista, sia dello studente, sia del docente. Lo studente può ritenere che il docente non motiva perché svolge un corso facile o perché il corso è difficile e impegnativo e la motivazione si perde nella difficoltà della materia. Il docente può svolgere tutti gli sforzi possibili per motivare, ma gli studenti sono inerti o assenti o non ricettivi perché la materia è «disapprovata», come può accadere, a volte, ai corsi di matematica e statistica nelle Facoltà di Economia.

R.4.7. L'osservazione è corretta, anche se l'ultima parte dell'argomentazione è comune e spesso rima con «ai miei tempi» assai impropriamente o con il decadimento generale delle capacità dei discenti. Su questo si può sgombrare il campo osservando che oggi studia il 60-70% circa dei giovani, quando nel passato studiava il 12-20%, vigeva una forte selezione di classe e di capacità, e gli obiettivi formativi dell'istruzione erano meno sollecitati da fattori sociali. La domanda ha, come si è detto, una sua ragion d'essere che può risultare coerente con gli obiettivi dell'indagine e, quindi, accettabile, nonostante le tante obiezioni, anche non riportate.

D08 (D08 del DOC 9/02) — «Il docente espone gli argomenti in modo chiaro [e rigoroso]?».

L'aggiunta del termine «rigoroso» è emerso con forza nella discussione del NdV anche contro il principio che non si devono modificare le domande per non alterare le

condizioni di somministrazione del questionario e il principio che si deve evitare, in genere, di chiedere due aspetti in una sola domanda al fine di individuare correttamente l'oggetto di riferimento della risposta. Dal punto di vista del docente, si può comprendere l'esigenza di essere valutato sia per la chiarezza e sia per il rigore perché spesso i due aspetti sono in contrasto, ma sarebbe stato meglio escludere il termine «rigore» perché non si sa come lo interpreta il rispondente: c'è chi suppone che si intenda corrispondenza delle spiegazioni al programma, c'è chi suppone che coincida con una esposizione formalizzata, e così via.

D09 (D12 del DOC 9/02) — «Le aule in cui si svolgono le lezioni sono adeguate (*si vede, si sente, si trova posto*)?».

O.4.9a. Il docente non è in grado di rispondere? Che senso ha raccogliere le risposte di N studenti?

R.4.9a. Si veda il §4.1, perché il giudizio dell'aula rientra in un «dato di realtà». Si consideri, in aggiunta, che è lo studente che utilizza «veramente» l'aula per conseguire proprio i fini precisi dell'attività didattica; pertanto, un suo giudizio, seppure inficiato da una percezione erronea, è fondamentale e non è da vituperare in ogni caso: basta dargli il giusto peso. Si nota, infine, che anche questionari più snelli (con poche domande) e utilizzati all'estero, dove la valutazione si esegue da alcuni decenni, tale domanda è presente: non è che si debba fare ciò che altri fanno, ma occorre valutare sempre se ciò che fanno è corretto. In questo caso, il giudizio del discente sull'aula è fondamentale e è corretto perché concerne un problema tecnico di misurazione. Si consideri, per esempio, la valutazione del «si sente». Come si deve procedere per accertare tecnicamente tale qualità di un'aula? Occorrono strumenti tecnici per valutare l'intensità della voce in ogni posto a sedere; ma non basta perché la presenza di persone nell'aula altera la diffusione e, dunque, la misurazione deve essere svolta in un'aula affollata. Anche la variazione del numero di presenti può alterare, inoltre, la diffusione. Si supponga che l'uso della strumentazione tecnica, da sola, non sia sufficiente. Si può procedere, allora, con una valutazione qualitativa di un esperto. A questo punto, si supponga di avere un'aula perfetta oggettivamente; allora, il giudizio dello studente è inutile? In generale, sí; ma se si persegue la precisione bizantina dell'approssimazione alla realtà, allora, potrebbe darsi che gli studenti abbiano un calo uditivo nella banda del parlato per il diffuso ascolto della musica a alto volume o andare in discoteca. Perché non fare esprimere il loro punto di vista? Se fosse molto negativo, si potrebbe procedere a fare i controlli delle aule, a migliorare i microfoni, quando ci sono e a introdurli quando non ci sono. Un discorso analogo vale per il «si vede» che, come per il «si sente», non è un dato esclusivamente oggettivo. Il «si trova posto» rientra nei dati oggettivi e, dunque, si ricade nelle osservazioni precedenti.

O.4.9b. Nella parentesi si invita il rispondente a considerare tre aspetti diversi: «*si vede, si sente, si trova posto*». Se le risposte riportano un giudizio negativo, quale aspetto crea problemi? Non è possibile saperlo.

R.4.9b. L'obiezione è giusta, anche se la formulazione è assai comune nell'ambito della valutazione della didattica. La risposta all'obiezione si enuclea dai paradigmi enunciati nell'introduzione; inoltre, basta andare a controllare lo stato dell'aula per il «si vede» e il «si sente» —anche tramite una osservazione qualitativa—, chiedere al docente se ha avuto problemi con i microfoni o studenti in piedi all'inizio delle, o durante le, lezioni.

D10 (D13 del DOC 9/02) — «I locali e le attrezzature per le attività [didattiche] integrative (*esercitazioni, laboratori, seminari, ecc.*) sono adeguati? (*se non sono previste [attività didattiche integrative] rispondere “non previste”*)»

O.4.10. Il docente non è in grado di rispondere? Che senso ha raccogliere le risposte degli studenti?

R.4.10. L'obiezione è simile alla O.4.9a e la risposta è simile alla R.4.9a; ma vi sono diverse altre risposte che hanno aderenza con questa, come la rilevazione del numero di studenti frequentanti (§3: O.3.3a, O.3.3b, O.3.3.c) o i dati con referenti di realtà (§4.1).

D11 (D06 del DOC 9/02) — «Le conoscenze preliminari [da me] possedute sono risultate sufficienti per la comprensione degli argomenti trattati [in questo insegnamento]?».

O.4.11. Ha senso una domanda del genere in assenza di obblighi di frequenza e di propedeuticità?

R.4.11. Sí, perché in molti CdS non vi sono né propedeuticità obbligatorie, né consigliate; eppure, spesso, non si dà la possibilità, a uno studente del primo anno, di sostenere gli esami del terzo anno: il divieto sarebbe, dunque demenziale, se la domanda fosse insensata in sé. Vi sono insegnamenti che, pur in assenza di propedeuticità, richiedono conoscenze che potrebbero non essere state acquisite o vi potrebbe essere un ampio stato di sofferenza per il tipo di conoscenze richieste da un insegnamento.

D12 (D14 del DOC 9/02) — «{Sei} [Sono] interessato agli argomenti di questo insegnamento?».

Tale domanda suscita, quando l'interesse per la materia è scarso, solo molteplici obiezioni connesse al declino dei tempi, alla scarsa preparazione degli studenti, al giudizio negativo sul docente. Sulla domanda in sé non vi sono critiche di rilievo particolare e la domanda non è sempre presente nei questionari usati in altri paesi. Per molte argomentazioni si può rispondere con la R.4.7. Altre obiezioni riguardano l'interpretazione dei risultati negativi, quando vi è un diffuso disinteresse per la materia.

D13 (D14 del DOC 9/02) — «{Sei} [Sono] complessivamente soddisfatto di come è stato svolto questo insegnamento?».

Anche qui non c'è molto da eccepire. Si noti la piccola variazione introdotta per reggere «meglio» la proposizione interrogativa (anche nella D12), rispetto alla prima persona del presente indicativo che è usata tendenzialmente per le affermative. La ragione della modifica non è, però, così scontata, ma non dovrebbe alterare le risposte degli studenti.

D14 (D01 del DOC 9/02) — «Il carico di lavoro complessivo degli insegnamenti ufficialmente previsti nel periodo di riferimento è [sostenibile] [accettabile]?».

O.4.14. Va da sé che uno studente che non vuole lavorare dirà «No». Uno studente che si dedica a tempo pieno all'università *deve* essere in grado di sostenere il carico di lavoro di 30CFU per semestre: 30CFU non equivalgono a 30×25 ore?

R.4.14. La quantità di 30CFU può corrispondere a più di 7 esami di 4CFU, mentre 30×25 ore corrisponde a 750 ore: 100 ore di studio per insegnamento di 4CFU che equivale a 15 giorni di studio di 6 ore al giorno, senza avere frequentato il corso: anche

una materia facile può diventare impossibile oppure è solo banalizzata. Chi frequenta ha una dispersione enorme di tempi. Il docente deve stare molto attento alla quantità di testi che intende fare studiare e all'ampiezza del programma: gli studenti hanno, semmai, un carico eccessivo. Si sta progredendo nella direzione di adeguamento, anche se vi sono numerosi studenti molto bravi che apprendono programmi ampi e a ritmi serrati. La domanda ha in sé un valore importante per l'organizzazione complessiva. Come risolvere il problema che pone è difficile, ma non si può eliminare l'accertamento di una difficoltà possibile perché si presume che il problema non esista!

D15 (D02 del DOC 9/02) — «L'organizzazione complessiva (*orario, esami intermedi e finali*) degli insegnamenti ufficialmente previsti nel periodo di riferimento (*bimestre, trimestre, semestre, ecc.*) è accettabile?».

O.4.15. Preside e Coordinatore didattico non sono in grado di dare una risposta univoca e attendibile? Che bisogno c'è di chiederlo agli studenti?

R.4.10. L'obiezione è simile O.4.9a e la risposta sarà simile alla R.4.9a; ma vi sono diverse altre risposte che hanno aderenza con questa, come la rilevazione del numero di studenti frequentanti (§3: O.3.3a, O.3.3b, O.3.3.c) e i dati con referenti di realtà (§4.1). Si contraddice sempre il primo assunto della valutazione per smontarla e renderla vana depotenziando le capacità di giudizio degli studenti. Si dà il caso che è lo studente che studia, sostiene gli esami, e paga per una pessima organizzazione: non si è bravi tutti allo stesso modo nell'organizzare; avere un riscontro, quindi, anche positivo, può essere sempre interessante. Bisogna, come sempre, dare ai risultati solo il giusto peso.

4.3. Osservazioni sulle modalità di scelta delle risposte

Le osservazioni sulle modalità di scelta coinvolgono alcuni aspetti teorici non sempre patrimonio comune, sicché le osservazioni sono meno frequenti. *Le più significative critiche sono state riportate nel paragrafo 6.3*, dopo l'illustrazione, seppure sommaria, delle caratteristiche delle scale di risposta, che in questo caso è una scala di Likert, e degli aspetti tecnici coinvolti nella definizione del paniere di scelta delle risposte.

5. Il questionario di valutazione: problemi e contenuti

Le argomentazioni più specifiche contro la valutazione riguardano, come già detto, molteplici questioni. Qui si esaminano le domande, ossia quali dimensioni sottoporre al giudizio, e le modalità di raccolta dei dati, cioè i tempi e le strategie da adottare per ottenere informazioni più accurate.

5.1. La scelta delle dimensioni da sottoporre a giudizio

Ogni processo di indagine deve delineare attentamente come si svolgerà il lavoro o la sua struttura, quali premesse teoriche si assumono, quali ipotesi si devono verificare, quali obiettivi si intendono perseguire. L'organizzazione delle idee e delle proposizioni, su cui impennare le fasi della procedura di rilevazione e di elaborazione, richiede spesso la definizione di tutti i concetti da utilizzare per formulare gli asserti da indagare e le ipotesi da verificare, incluse le teorie di riferimento. Si sviluppa, quindi, un

processo di identificazione che consiste nell'individuazione degli aspetti da esaminare in base al tema, ai soggetti, e agli obiettivi dell'indagine.

Il primo passo consiste nella chiarificazione dei termini, un processo definitorio che consente di eliminare le ambiguità che possono celarsi nelle parole usate per indicare i concetti che si vogliono utilizzare nello svolgimento del lavoro. La scelta dei concetti dipende dall'obiettivo che si deve conseguire con l'attività di valutazione e da cosa si intende per valutazione della didattica. Se l'obiettivo dell'indagine è migliorare l'attività di insegnamento, occorre subito precisare che i tempi per gli interventi, che si possono attuare per migliorare gli aspetti che mostrano deficienze e inefficienze, sono assai lunghi e non possono riguardare gli attori che li hanno denunciati, perché si può solo intervenire da un anno accademico all'altro; infatti, la valutazione è effettuata, per ovvie ragioni, alla fine delle attività (v. *infra*, §5.2). Per migliorare l'insegnamento e risolvere le difficoltà presenti, occorrerebbe potenziare meccanismi, come si chiarirà meglio nel seguito, che facilitino il dialogo tra docenti e discenti e che consentano di intervenire in itinere. Se l'obiettivo della valutazione è instaurare, invece, una prassi che dia un senso di partecipazione —e in alcuni casi, forse, anche di dignità— ai discenti e che induca il docente a una maggiore attenzione, allora la valutazione della didattica acquisita un valore in quanto tale, ossia è determinante e necessario che si faccia, non importa quali e quante dimensioni siano indagate. Tale ipotesi o proposizione è un po' paradigmatica in questa analisi e, pertanto, la descrizione del processo di definizione dei termini e delle dimensioni oggetto di indagine si limiterà solo agli aspetti più rilevati nelle valutazioni della didattica svolte prevalentemente in Italia (Osservatorio, 1998).¹

Le *informazioni personali* dello studente costituiscono un profilo interessante da rilevare con l'indagine; tuttavia, sono numerosi i concetti che possono essere rilevanti per comprendere le loro associazioni con il giudizio espresso, ma in una indagine così diffusa e capillare non si possono porre tante domande; per uno studio strutturale, bisogna ricorrere a indagini campionarie specifiche e con finanziamenti opportuni. I dati più richiesti sono quelli di natura anagrafica: genere, età o anno di nascita, tipo di diploma conseguito, provincia di residenza, anno di corso, voto di diploma, numero di esami sostenuti, media dei voti conseguiti, numero di corsi frequentati nel periodo di lezioni oggetto di valutazione, posizione lavorativa, e così via. Altri aspetti potrebbero interessare, come le condizioni familiari e gli atteggiamenti individuali, ma non si possono rilevare con tale tipo di indagine. Alcuni caratteri elencati presentano problemi di rilevazione non banali. Per esempio, la richiesta del numero di corsi frequentati nel periodo corrente è spesso fraintesa perché i rispondenti tendono a includere anche i corsi del semestre precedente o tutti corsi dall'immatricolazione; la richiesta del voto medio produce dati che sono presumibilmente affetti da errore di memoria o di ignoranza, e presentano un pronunciato effetto telescopico —le distribuzioni del voto medio rilevato nelle indagini sulla valutazione della didattica, condotte presso la Facoltà di Economia nell'Ateneo di Modena e Reggio Emilia, sembrano confermare questa supposizione.

La *logistica* è l'altra dimensione di interesse e concerne la condizione delle aule dove si svolgono le lezioni, gli spazi disponibili per gli studenti quando non vi sono lezioni, le attrezzature per le attività pratiche, i laboratori, la biblioteca,

¹ Il rapporto di ricerca dell'Osservatorio (1988) per la valutazione del sistema universitario, presenta un'ampia rassegna del panorama internazionale sulla valutazione della didattica, e del panorama nazionale sui questionari utilizzati e sulle dimensioni indagate.

l'organizzazione degli orari per l'accesso ai servizi disponibili, l'orario complessivo delle lezioni e dei corsi in parallelo, la presenza dei servizi (telefoni, distributori di bevande, fotocopiatrici, bacheche, sistema di avvisi e informazioni, centralino, segreterie, eccetera).

L'*organizzazione della didattica* e le *qualità del docente* sono i domini di più elevato interesse. La prima riguarda, in genere, il calendario degli esami (che si può ascrivere, però, anche alla logistica), l'ufficializzazione delle modalità dell'esame, l'adeguatezza del materiale didattico, la ripartizione dei tempi tra lezioni e esercitazioni, la presentazione di casi specifici e seminari (eventuali), le modalità di disposizione degli argomenti, i tempi assegnati alle varie sezioni del corso; si possono considerare appartenenti all'organizzazione della didattica, anche le attività di supporto (esercitatori, lettori, tutori, relatori a vario titolo). La seconda concerne, invece, i dati più personali del docente: la chiarezza dell'esposizione a lezione, la disponibilità a fornire spiegazioni a lezione e durante il ricevimento studenti, la puntualità a lezione e al momento del ricevimento studenti, la reperibilità.

5.2. La scelta del tempo di raccolta dei dati

La scelta del periodo di somministrazione del questionario non è agevole perché all'inizio delle lezioni la frequenza è più elevata, ma i frequentanti non sono ancora in grado di esprimere un giudizio. L'ultima settimana di lezione è sconsigliabile perché si ha una forte diminuzione di presenze per l'avvicinarsi degli esami. La *penultima e la terzultima settimana* sono le candidate ideali perché il giudizio può essere già espresso con maggiore cognizione e il numero di frequentanti, seppur diminuito, è ancora alto. Gli studenti che hanno smesso di frequentare non esprimono il loro giudizio, ma l'abbandono della frequenza potrebbe essere proprio la conseguenza di una cattiva organizzazione didattica che sfuggirebbe alla rilevazione perché restano soltanto i migliori, i più determinati, i più motivati, o chi ha trovato congeniale e soddisfacente l'organizzazione didattica adottata: ciò potrebbe distorcere in alto o in «positivo» i risultati. Il CNVSU nel DOC 9/02 suggerisce «*che il questionario sia somministrato tra la metà e i due terzi dell'insegnamento in modo da rendere possibili primi interventi correttivi da parte del docente al quale i risultati dovranno, di norma, essere forniti entro brevissimo tempo dalla compilazione da parte degli studenti*» (CNVSU, 2002, p. 8). Tale suggerimento è assolutamente incompatibile con l'attuale organizzazione, la necessaria riservatezza dei dati, la codifica degli insegnamenti, la centralizzazione delle elaborazioni, e l'opportunità di non appesantire ulteriormente il carico di lavoro dei docenti perché una variazione delle procedure rischierebbe di coinvolgerli ancora di più nel processo. La proposta di una organizzazione diversa esorbita dagli obiettivi di questa nota e presenta comunque problemi organizzativi immani, sebbene risolvibili con una buona organizzazione informatica dei processi amministrativi e la dotazione di lettori ottici in ciascuna Facoltà, se non per ogni Dipartimento.

La valutazione complessiva del corso non può, poi, che essere parziale perché chi risponde potrebbe non essere in grado di tenere conto del carico effettivo di lavoro se seguisse il corso senza studiare con la giusta intensità in quanto i primi esami del suo programma sono altri insegnamenti; inoltre, esclude il giudizio sul processo di valutazione della preparazione dell'esame, ma anch'esso costituisce un elemento importante dell'organizzazione didattica: i docenti possono tenere, infatti, delle ottime lezioni e, per molteplici ragioni, condurre esami inadeguati o incapaci a cogliere con

obiettività la preparazione dei candidati. Per alcuni insegnamenti la qualità consiste anche in una scelta accurata delle letture proposte, che sono spesso effettuate prima dell'esame e non durante il corso sicché è impossibile valutarlo correttamente due o tre settimane prima della fine delle lezioni. D'altronde, se la valutazione della didattica è rilevata successivamente al periodo degli esami, il giudizio può risultare inficiato dall'esito e dagli errori di memoria perché i dati retrospettivi presentano lacune e imprecisioni (Cannell, Miller, Oksenberg, 1981; Sudman, Brandburn, 1973).

La scelta del periodo potrebbe indurre, quindi, una *distorsione da selezione* nelle informazioni raccolte perché tende a escludere sistematicamente alcune unità della popolazione di riferimento che forse sono notevolmente caratterizzate, proprio rispetto agli obiettivi dell'indagine. Per esempio, restano esclusi gli studenti che hanno smesso di frequentare perché l'aula era troppo affollata o per preparare gli esami imminenti. In un corso in cui si prevede una lezione fissata (per settimana) per lo svolgimento degli esercizi si può verificare che vi siano studenti che frequentano solo la teoria o solo le esercitazioni e, pertanto, la scelta dell'ora da riservare alla compilazione del questionario può escludere chi non frequenta quel tipo di lezione.

La consegna del questionario dovrebbe avvenire sempre nelle stesse condizioni, e le informazioni per la compilazione dovrebbero essere uguali per non indurre distorsioni incontrollate; ma attualmente non c'è una prassi consolidata e le procedure differiscono da sede a sede. In alcune Facoltà è il docente che consegna il questionario agli studenti presenti in aula sicché non solo i compilatori non ricevono una corretta presentazione e le stesse istruzioni per la compilazione, ma in diversi corsi il questionario non viene consegnato per molteplici ragioni, non sempre inconsistenti, come nel caso di un supplente esterno non informato. In altre Facoltà è un incaricato che consegna i questionari e li ritira: in tal modo si ha una modalità di somministrazione identica in tutti i corsi.

I questionari compilati e riconsegnati per l'elaborazione possono presentare gravi problemi di distorsione campionaria dovuti essenzialmente all'*autoselezione* dei rispondenti che potrebbe comportare, come si è detto, una *distorsione da non-risposta* nei risultati perché coloro che non compilano i moduli potrebbero appartenere a categorie di soggetti che presentano giudizi tipici sulla valutazione della didattica (Cochran, 1977; Cicchitelli, Herzel, Montanari, 1997); così, chi riceve un giudizio favorevole potrebbe sospettare che abbiano risposto soltanto i più bravi; chi riceve un giudizio sfavorevole potrebbe argomentare a sua discolpa che siano stati soltanto i meno capaci a valutare la sua attività. Non c'è dubbio che meccanismi di rivalsa e di protesta (specie quando i programmi sono vasti, pesanti, e difficoltosi oppure la materia è ostica oppure non è gradita perché vi è un diffuso rifiuto) potrebbero incentivare le risposte più negative (*reattività*). A volte la risposta può essere anche affetta da *leggerezza* con cui si formulano i giudizi, perché l'indagine è considerata un gioco o semplicemente inutile.

Una valutazione più accurata richiederebbe di scandagliare oltre l'immediata percezione soggettiva; ma più le domande sono articolate o più richiedono impegno nella compilazione dei moduli e più aumentano il carico di lavoro, la *disattenzione* del rispondente nell'espressione del giudizio, e le mancate risposte (che sono elementi negativi per ogni indagine). Il ricorso a un intervistatore elimina alcune difficoltà, ma è quasi impensabile usarlo nell'ambito della valutazione della didattica. Anche un buon questionario, poi, non è sempre esente dal fenomeno delle risposte seriali (*response set*) che in questo caso potrebbe essere enfatizzato perché le condizioni e i ruoli degli attori

(risentimento, reattività, stanchezza, noia) potrebbero indurre un aumento della mancata distinzione tra le domande e della propagazione della stessa risposta. Infine, può verificarsi l'effetto «*trainante*» di un solo elemento di giudizio: una materia difficile che richiede uno studio costante durante il corso potrebbe comportare una valutazione poco positiva sulla chiarezza del docente che a sua volta influenza la risposta sugli altri elementi, come la presenza regolare a lezione o all'orario di ricevimento che sono "oggettivamente" misurabili e verificabili. L'esame delle indagini condotte nella Facoltà di Economia (1990-2000) sembra confermarlo: le materie più difficili —come Statistica, Matematica generale, Matematica finanziaria 1, Ragioneria generale, Scienza delle finanze— presentano tendenzialmente una valutazione più bassa nella presenza regolare a lezione del docente o puntualità all'orario di ricevimento; viceversa, i docenti delle materie linguistiche ricevono punteggi tendenzialmente più alti.

Le difficoltà menzionate non possono che generare sfiducia sugli esiti perché non si sa *se e quanto* le valutazioni ottenute siano affidabili e non è chiaro, poi, l'uso che se ne debba fare, ossia le conseguenze di una buona o pessima valutazione. L'impossibilità di distinguere adeguatamente tra l'attendibilità statistica e l'uso amministrativo dei risultati rende controversi e confusi i dibattiti che si concentrano spesso solo nell'esaltazione della novità; ma se non ha conseguenze pratiche, la procedura rischia di diventare soltanto un elemento di visibilità o un rituale e non uno strumento che possa fornire informazioni utili alla valutazione e all'organizzazione delle attività didattiche. Le argomentazioni contro gli esiti ottenuti con tale attività non sono mancate in questi anni di valutazione nelle università italiane e si possono classificare approssimativamente proprio nelle due categorie citate: la capacità dei risultati di descrivere veramente e con esattezza i concetti coinvolti nell'analisi e la possibilità di derivarne indicazioni utili agli aspetti organizzativi.

L'analisi dei giudizi raccolti nella Facoltà di economia (1990-2000) ha mostrato che la valutazione della didattica presenta una correlazione negativa significativa, anche se non sempre elevata, con il livello percepito di difficoltà della materia e una accentuata correlazione positiva con l'interesse per la materia che dipende a sua volta dalla presunta utilità nel lavoro futuro. Le teorie dell'apprendimento sostengono che anche il livello e la qualità delle conoscenze pregresse (prerequisiti) necessari per seguire il corso contribuiscono a determinare il rendimento, la valutazione, e la soddisfazione.² I giudizi così rilevati devono essere tenuti in debito conto per organizzare la didattica, sí, ma occorre sempre la consapevolezza che contengono un intrigo complicato di fattori per cui attribuire importanza assoluta esclusivamente agli esiti della valutazione degli studenti rischia di generare una involuzione che potrebbe peggiorare gli obiettivi dell'insegnamento e impoverire i contenuti delle materie più complesse. Dal lato del docente, si potrebbe incentivare una riduzione dei programmi per favorire un miglioramento del giudizio degli studenti verso la propria materia. Dal

² Alcune argomentazioni contro la valutazione già citate —e altre tipo: "la conoscenza, che costituisce il patrimonio di ogni singolo soggetto, è sempre il risultato di un'attività costruttiva e, pertanto, non può essere acquisita da chi ha un atteggiamento esclusivamente passivo" o "l'apprendimento è una funzione dell'interesse per i contenuti della materia e del grado di conoscenza dei concetti preliminari necessari per seguire lo sviluppo degli argomenti proposti" o "il corso dell'azione formativa può favorire la *viabilità* con azioni di sostegno e di influenza, di stimolo e di contenimento, ma risulta impotente di fronte al rifiuto del discente"— si possono evincere da numerosi testi concernenti le teorie dell'apprendimento, sia specifici e sia affini. Spunti e suggerimenti si trovano, tra gli altri, in [Morelli e Weber \(1996\)](#), [Pontecorvo, Ajello, e Zucchermaglio \(1993\)](#), [Bloom \(1993\)](#), [Bruner \(1996\)](#).

lato del discente, si possono verificare situazioni anche paradossali. Per esempio, nei corsi di laurea con insegnamenti “omogenei”, come matematica (fisica), ci si dovrebbe aspettare un interesse non inferiore alla posizione intermedia (sufficiente) per tutte le materie, anche se un soggetto può prediligere l’analisi matematica (la meccanica) e avversare la geometria (l’ottica). Che pensare se accade che diversi rispondenti dichiarano di non avere alcun interesse per uno o più insegnamenti? A un errore nella scelta del percorso di studi? Viceversa, nei corsi di laurea con insegnamenti “eterogenei”, come economia e biologia, il disinteresse per alcune materie potrebbe essere molto più frequente e motivato, la conoscenza dei prerequisiti potrebbe presentarsi diffusamente lacunosa (i cosiddetti corsi di azzeramento durano, poi, una settimana o al massimo due!). I docenti delle materie di base sarebbero, quindi, ingiustamente penalizzati dal contesto sfavorevole in cui operano. L’analisi di dati protratta nel tempo e in località differenti potrebbe evidenziare le caratteristiche degli insegnamenti, in sé e in relazione ai corsi di laurea, che aiuterebbero a pesare e a apprezzare correttamente la valutazione degli studenti.

Le preoccupazioni dell’uso amministrativo degli esiti delle valutazioni eseguite dagli studenti possono apparire anche comprensibili per le osservazioni precedenti e si suggerisce, perciò, da più parti (Kogan, 1989; Drenth, Van Os, Bernaert, 1989), di utilizzarli assieme ai risultati di valutazioni effettuate da esperti esterni alle istituzioni perché se ci si affida solo al giudizio degli studenti si rischia di generare comportamenti incoerenti con gli obiettivi dell’insegnamento. Infatti, l’utilizzo amministrativo dei giudizi espressi dagli studenti danno loro, sí, incisività e valore incoraggiando i docenti a migliorare l’attività didattica; ma rischiano anche di incentivare una povertà di contenuto stimolando una possibile rincorsa alla *captatio benevolentiae*. Sembra opportuno, quindi, tenere separate le informazioni sui corsi e le persone coinvolte.

C’è una tendenziale unanimità nel considerare quale scopo essenziale della valutazione la raccolta di informazioni sull’attività del docente, nonostante le difficoltà, rispetto all’efficacia dei suoi metodi (Harvey, Knight, 1997; Cave, Hanney, Kogan, Trevet, 1988). Ovviamente, conoscere lo stato di inadeguatezza dei metodi utilizzati non genera automaticamente un loro miglioramento per molteplici motivi: le convinzioni personali dei docenti sulla bontà dei loro metodi, l’ottimizzazione delle risorse perché ogni cambiamento comporta un aggravio dell’impegno organizzativo e delle risorse umane, il conservatorismo che spesso si maschera dietro giudizi apparentemente validi e nasconde talvolta solo la pigrizia, le difficoltà oggettive a spiegare i concetti quando gli studenti hanno una preparazione lacunosa, il possibile disinteresse per la materia presente tra i discenti.

6. La scala di Likert come scala di misura della soddisfazione³

La valutazione della didattica è condotta sulla base di dati raccolti con un questionario contenente domande atte a accertare sia il grado di soddisfazione degli studenti, sia l’efficienza e l’efficacia dell’insegnamento in termini di qualità e di impegno dei docenti. Le domande possono prevedere risposte simili alla scala di Likert, ma la scelta dei sintagmi per le categorie di risposta non è così scontata. Gli studi condotti sul tema

³ Il testo di questo paragrafo si basa essenzialmente su Lalla (2003, §2), che va detto, anche se le autocitazioni non sono gradite, ma necessarie, come per i citati studi condotti sul confronto delle scale.

specifico (Osservatorio, 1998; Chiandotto, Gola, 2000; CNVSU 2002, p. 9) hanno preferito e suggerito l'uso di una scala di Likert (1932) a quattro modalità di risposta: ① *decisamente no*, ② *piú no che sí*, ③ *piú sí che no*, ④ *decisamente sí*; cui ci si riferirà nel séguito con il termine “scala del MIUR”. Di per sé, la scelta non è criticabile; infatti, il dibattito è ancora aperto nella letteratura e da là non si evince quale sia la strategia migliore. In fase di elaborazione dei dati si suggeriscono, tuttavia, procedure che non sono affatto coerenti con la scala adottata, come l'uso della media e della varianza, che sono gli indici piú comuni e immediati per la comprensione, ma richiedono variabili di natura quantitativa. In particolare, gli autori suggeriscono di attribuire un valore numerico uguale a *due* alla prima modalità, *cinque* alla seconda, *sette* alla terza, e *dieci* alla quarta. Tale scelta consente di esprimere i giudizi secondo il sistema di valutazione utilizzato nella scuola italiana (esclusa l'università), che prevede punteggi che vanno da zero a dieci; nel séguito, ci si riferirà a tale sistema con il termine «*scala decimale*». L'equivalenza assunta tra la scala del MIUR e la scala decimale, sebbene sembri ragionevole, è quanto mai arbitraria; richiede, per lo meno, un accertamento empirico tra la popolazione di riferimento per verificare se la sua opinione coincide con il criterio proposto. La scala di riferimento assunta, poi, non è proprio quella decimale perché il punto centrale dei punteggi sembra il *sei* sicché il campo di variazione va da due a dieci, e non da zero a dieci; ma *sei* corrisponde alla sufficienza che è un giudizio positivo, non neutrale!

Scale alternative ve ne sono tante, sempre con pregi e difetti, che devono essere valutati di volta in volta per accertare la loro applicazione su larga scala. Nei termometri (di sentimenti) si possono ottenere, per esempio, valutazioni piú vicine alla natura dei numeri reali sia aumentando il numero delle modalità di scelta, sia lasciando indicare al rispondente il punto o collocazione del suo giudizio sul segmento ancorato e misurando poi in centimetri la distanza del punto indicato dall'origine. La precisione che si ottiene, che è molto elevata in termini numerici, non può garantire la correttezza della *reale* posizione del rispondente perché la sua risposta è data con l'esattezza dell'approssimazione visiva che risulta molto piú bassa degli strumenti che misurano la distanza; inoltre, diverse esperienze empiriche mostrano che la presenza di numerose scelte o di una scala piú fine non migliora la differenziazione delle risposte degli intervistati che forniscono spesso valori riconducibili o riducibili a un numero inferiore, previa analisi adeguata. Hofacker (1984) mostra che da ventuno categorie nominali ci si può ricondurre a otto. Marradi (1992, 1998) sostiene che un'accettabile approssimazione alla cardinalità è ottenuta con scale *autoancoranti*, che riducono l'autonomia semantica delle modalità intermedie; ma sottolinea che i soggetti intervistati, pur avendo una scala da 0 a 100, tendevano a usare cifre tonde, di solito multipli di 10 e, talvolta, multipli di 5. Questa tendenza si è notata in alcune indagini locali; ma il fenomeno di contrazione, nell'applicazione pratica, del numero delle possibilità offerte dalla scala disponibile è noto (Gattullo, 1968). Una struttura piú fine della scala non garantisce, quindi, una migliore precisione e potrebbe aumentare solo la variabilità dei giudizi e il disorientamento di chi deve esprimerlo con tale strumento.

Si è condotto uno studio tra la scala MIUR e una scala alternativa che usa il sistema di valutazione scolastico, basato sulla denominazione verbale dei punteggi o voti. Esso mostra ugualmente varianti o estensioni, diverse da scuola a scuola, o ambiguità sul significato dei termini rispetto al loro valore numerico; per esempio, i termini “discreto” e “mediocre” hanno differenti interpretazioni della loro posizione in un ordinamento (graduatoria). Con qualche costrizione/contrazione della terminologia

usata in pratica, il sistema di votazione espresso in termini linguistici si è ricondotto a una scala tipo Likert a cinque modalità: ① *molto insufficiente*, ② *insufficiente*, ③ *sufficiente*, ④ *buono*, ⑤ *molto buono*; denominata «*scala di voto*»,⁴ e adottata per un certo periodo dal NdV dell'Università di Modena e Reggio Emilia. I risultati sono esposti in [Lalla, Facchinetti, Mastroleo \(2004, §2\)](#) o [Lalla, Facchinetti \(2004, §2\)](#).

La scala di voto non risolve la questione del livello di misura richiesto dai più comuni indici statistici, perché è sempre qualitativa ordinata e il calcolo della media rimane problematico per la singola domanda, ma almeno presenta una struttura che garantisce formalmente la linearità e costituisce una migliore approssimazione alla cardinalità; infine, nonostante la difficoltà teorica, la scala a cinque modalità che include la modalità centrale è largamente utilizzata in pratica come fosse una scala intervalli perché i risultati sono spesso riportati con l'indice media. La scala di voto, poi, elimina di per sé il problema della modalità neutrale o centrale ([Schuman, Presser, 1996](#)) perché la contiene solo implicitamente; infatti, la terza modalità esprime una posizione favorevole e non una neutrale, che è ritenuta il rifugio di chi non vuole esprimersi, degli indolenti, e dei superficiali. La scala di voto non è, quindi, una vera scala di Likert perché la modalità centrale esprime una scelta positiva e non è, perciò, un'autentica modalità centrale.

6.1. Caratteristiche della scala Likert

La scala di Likert è impiegata per misurare in prevalenza atteggiamenti e opinioni mediante l'uso di affermazioni, comune anche a altre scale: Thurstone, Guttman, differenziale semantico e termometro. La tecnica fu ideata nel 1932 dallo psicologo americano Rensis Likert con lo scopo di elaborare uno strumento semplice per misurare atteggiamenti e opinioni ([Likert, 1932](#)).

La scala di Likert richiede la formulazione di una serie di L proposizioni o asserti o affermazioni (*items*) supposte semanticamente collegate agli atteggiamenti oggetto di indagine che devono essere sottoposti a un gruppo di individui e devono essere monotoni, vale a dire formulati in modo unidirezionale rispetto all'oggetto da misurare affinché all'aumentare dell'atteggiamento favorevole del soggetto verso l'oggetto, aumenti il punteggio conseguito in quella proposizione. Pertanto, si devono ideare asserti con un contenuto favorevole al concetto da misurare e asserti con un contenuto contrario, di varia intensità e in modo che i due insiemi abbiano la stessa cardinalità. A ogni individuo viene chiesto di esprimere il suo grado di accordo per ogni proposizione, che prevede cinque possibilità alternative tra loro: completamente d'accordo, d'accordo, incerto, in disaccordo, in completo disaccordo (*strongly agree, agree, uncertain, disagree, strongly disagree*). Per ogni proposizione, le risposte presentano nell'ordine le etichette 5, 4, 3, 2, 1 (oppure 4, 3, 2, 1, 0) con la funzione specifica di mettere in ordine le alternative di risposta. I punteggi attribuiti a ogni soggetto per ogni proposizione coincidono con le etichette per le affermazioni a favore e sono invertite (1, 2, 3, 4, 5 oppure 0, 1, 2, 3, 4) per le affermazioni contro. Il punteggio finale di ciascun soggetto è dato dalla somma di tutti i punteggi parziali corrispondenti a

⁴ Il sistema di valutazione nelle scuole medie è stato pressoché simile alla scala di voto per alcuni anni in cui si usavano voti espressi con le lettere A, B, C, D, E, in senso decrescente. Oggi, la terminologia è un po' diversa: insufficiente, sufficiente, buono, distinto, ottimo. Si usa quindi una sola modalità negativa, per non offendere (*sic!*) il discente. Nelle scuole superiori, invece, si è sempre usata la votazione espressa in decimi; ma i voti superiori a otto sono, in genere, più rari rispetto ai voti inferiori a quattro.

ogni scelta effettuata tra le L proposizioni che costituiscono la scala; in altri termini, “il punteggio individuale sull’intera scala, [può anche essere] costituito dalla [...] somma dei codici numerici attribuiti alle risposte scelte da un individuo ai vari *items* della scala” (Cacciola, Marradi, 1988, pp. 72-73) perché i codici delle categorie di risposta sono espressi, in genere, con numeri naturali.

La procedura presenta celerità di registrazione, di codifica, e di calcolo. Gli assunti sottostanti riguardano: (1) l’unidimensionalità degli atteggiamenti o opinioni oggetto di misura per cui le diverse proposizioni utilizzate nella scala devono riferirsi allo stesso concetto o proprietà; (2) collocabilità dell’atteggiamento su un *continuum* per cui dalle risposte fornite dai singoli soggetti si ottiene un corrispondente valore (numerico) che esprime l’orientamento dell’atteggiamento per ognuno di loro e la posizione in graduatoria è determinata dagli stessi; (3) equidistanza tra le categorie di risposta, per esempio, la distanza tra “completamente d’accordo” e “d’accordo” dovrebbe essere uguale a quella che c’è fra “disaccordo” e “completamente in disaccordo”, così come rispetto alla categoria “incerto” e, infatti, i punteggi attribuiti ne sono una conseguenza. Tuttavia, la procedura di costruzione non garantisce né che la scala misuri una sola proprietà, né che le posizioni percepite siano le stesse per ciascuna categoria e per tutti gli intervistati.

Il numero di cinque categorie di risposta previste dalla scala è arbitrario, anche se pare sia abbastanza efficace perché risulta semplice e riduce la variabilità dello zero e dell’unità di misura individuali; tuttavia, il numero ottimale di categorie è una funzione delle condizioni di misura e dipende dal contenuto specifico (Mattell, Jacoby, 1971; Cox, 1980; Wildt, Mazis, 1978). Nella versione originaria la scala presentava sette possibili modalità di scelta: *strongly agree, mildly agree, agree, uncertain, disagree, mildly disagree, strongly disagree*; in seguito le due categorie introdotte da *mildly* furono abbandonate (Cacciola, Marradi, 1988, p. 67).

Una variante molto discussa consiste nell’eliminare la categoria 3 (“incerto”) per costringere il rispondente a schierarsi, nell’ assunto che i “veri” incerti si distribuiscano in parti uguali tra le categorie dei favorevoli e dei contrari; ma la mancanza dello “zero” nella scala solleva perplessità nell’analisi dei dati perché non c’è più l’equidistanza tra le categorie. In alternativa, si esclude la categoria dal campione, ma i risultati possono essere falsati o distorti (Guy, Norvell, 1977; Ryan, 1980; Garland, 1991). La percentuale di soggetti che scelgono la categoria intermedia tende ovviamente a diminuire all’aumentare del numero di categorie disponibili (Mattell, Jacoby, 1972).

6.2. Vantaggi e svantaggi della scala di Likert

La frequente applicazione della scala di Likert in diversi campi deriva anzitutto dalla semplicità dello strumento, dalla facilità dello stile di registrare le risposte, dalla chiarezza delle categorie di risposta e della scelta che offre. Pertanto, l’intervistatore avrà minori difficoltà a porre le domande e a registrare le risposte. Il soggetto intervistato ha un minimo di possibilità nell’articolare il suo punto di vista perché può esprimere la propria opinione scegliendo tra cinque alternative e risulta, quindi, più flessibile delle scale che prevedono solo una risposta dicotomica (sì/no oppure accordo/disaccordo). Inoltre, si può raccogliere una maggiore quantità di informazioni in minor tempo perché si possono utilizzare le stesse alternative di risposta per più affermazioni. L’ordine delle categorie di risposta è predefinito e stabile perché hanno scarsa autonomia semantica (in altre parole, devono fare riferimento al testo della

domanda e alle altre categorie per essere interpretate); pertanto, è improbabile che gli intervistati scelgano di ordinarle in modo diverso da come sono e ne consegue che si ordinano più facilmente anche le risposte date dai soggetti (Cacciola, Marradi, 1988).

Le difficoltà dello strumento sono diverse e si enumerano di seguito secondo un ordine casuale e non gerarchico.

La *prima* riguarda l'uso di "domande chiuse" che obbligano l'intervistato a scegliere le risposte soltanto entro le limitate alternative previste sicché l'essere indotto a scegliere una modalità che non esprime il proprio reale atteggiamento può generare frustrazione e irritazione con conseguente aumento della probabilità di non risposta; oppure, per reazione, si può scegliere senza riflettere o anche in modo pressoché casuale perché estranei al tema trattato e si agisce soltanto per non mostrare la propria ignoranza o disinformazione sull'argomento. La scelta obbligata potrebbe comportare anche una *deriva* nella graduatoria finale di un soggetto sottoposto a una batteria di domande perché quando le coppie (di domande favorevoli e contrarie al concetto in esame) non hanno la stessa accettabilità sociale, allora si potrebbe produrre uno spostamento più consistente verso l'approvazione o la disapprovazione violando l'assunto della distribuzione equa tra i favorevoli e contrari: una caratteristica che varia facilmente da individuo a individuo, nello spazio, e nel tempo (Orvik, 1972).

La *seconda* si riferisce agli assunti di continuità e equidistanza tra le categorie di risposta perché arbitrarie e non c'è alcuna garanzia che siano considerate uguali e equidistanti da tutti gli intervistati; numerosi studi hanno mostrato: sia la percezione delle etichette con un'ampiezza molto variabile e senza plausibili regolarità (Amisano, Rinaldi, 1988); sia l'effetto "estremità" (*end effect*) che concerne la tendenza di certi soggetti a essere attratti o respinti dalle scelte estreme (Galtung, 1967).

La *terza* concerne l'assunto di linearità che comporta la collocazione di soggetti favorevoli alla proposizione su un lato della scala e dei contrari sul lato opposto; la violazione dell'assunto genera la *curvilinearità*, che emerge quando soggetti con atteggiamenti diversi e opposti danno la stessa risposta; così, può accadere che i *veri* contrari a una certa proposizione siano associati con altri che si dichiarano contrari perché sarebbero ancora *più* favorevoli al concetto espresso dalla proposizione data di quanto consentano le modalità di scelta. Nella stessa classe si hanno, quindi, soggetti che sono in accordo e soggetti che sono in disaccordo perché due soggetti, con opinioni opposte sul tema, effettuano la stessa scelta ottenendo lo stesso punteggio che evidentemente non corrisponde al loro *stato* reale (Guidicini, 1995, p. 98). Il suggerimento di Coombs per evitarla è "scegliere un'affermazione così estrema da rendere improbabile che ci siano individui tanto estremi da respingerla perché non è sufficientemente estrema" (Coombs, 1953, p. 530). Si introducono, così, ancora delle distorsioni perché le posizioni intermedie sono sacrificate a favore di quelle estreme (Marradi, 1984, p. 63).

La *quarta* contempla la possibilità di risposte seriali (*response set*) quando le proposizioni di una scala Likert sono presentate agli intervistati non singolarmente, ma una di seguito all'altra (batteria). Lo stesso schema di risposta facilita la scelta e accelera i tempi di compilazione, ma potrebbe indurre l'intervistato a selezionare meccanicamente sempre la stessa modalità, indipendentemente dal contenuto della domanda: acquiescenza (*acquiescent response set*). Le risposte ottenute possono essere, perciò, inficiate da distorsioni e affette da scarsa attendibilità sicché l'analisi dei dati può condurre a risultati fuorvianti; ma distinguere le risposte date in modo ripetitivo da quelle che effettivamente rispecchiano l'opinione o l'atteggiamento del soggetto non è

semplice. Altre denominazioni sottolineano, talvolta, aspetti diversi: per le alternative dicotomiche (sì/no), si dice anche *yea-saying* o *nay-saying* (Gasperoni, Giovani, 1995); la scelta della prima modalità che si presenta è detta effetto “somaro” (*donkey vote effect*) che sostituisce *yea-saying* (Ray, 1990; Chan, 1991) o anche effetto “primato” (*primacy effect*). Nella valutazione della didattica ciò può avere un’importanza rilevante (Albanese, Prucha, Barnet, Gjerde, 1997). L’intervistatore può rimediare interrompendo la serialità delle risposte (Marradi, 1984, pp. 62-66): (a) sottolineando opportunamente l’indipendenza reciproca delle varie proposizioni della batteria; (b) richiamando “l’attenzione dell’intervistato sull’incongruenza delle risposte”; (c) interrompendo la batteria con l’inserimento domande di forma diversa prese o (c1) dallo stesso questionario, o (c2) “da una lista di domande appositamente concepite per risvegliare l’interesse dell’intervistato”. Tali azioni sono, però, in contrasto con le regole di conduzione dell’intervista. Il punto (a) può influenzare, infatti, la risposta. Il punto (b) potrebbe suscitare reazioni di frustrazione o irritazione (Goode, Hatt, 1952). I punti (c1) e (c2) violano la necessità di mantenere lo stesso ordine nella somministrazione delle domande per confrontare le risposte dei diversi soggetti; (c2) comporta, inoltre, il prolungamento della durata dell’intervista con domande di relativa importanza. Gli intervistatori che appartengono al gruppo di ricerca sono in grado, secondo Pitrone (1986, p. 111), di intervenire con più efficacia perché hanno ragioni e conoscenze “necessarie per improvvisare ulteriori domande che suscitino maggiore interesse nell’intervistato e facciano emergere con più chiarezza la sua opinione”; ma ciò pregiudica comunque il confronto dei dati nel contesto della giustificazione.

La quinta è inerente alla posizione in cui vengono poste le modalità di scelta e la direzione dell’intensità (Chan, 1991). Si tratta di effetti che sono assimilabili alla risposta seriale e potrebbero denominarsi effetti *posizionali* e *direzionali* (*primacy effect*). Nelle indagini che li hanno analizzati, l’entità e il segno della distorsione sono risultati variabili. Nella valutazione della didattica si è visto che gli esiti complessivi sono più elevati (più positivi) quando le categorie con valore più elevato sono poste a sinistra, com’è nella scansione tradizionale della scala di Likert, rispetto a quelli che si ottengono quando le categorie con valore più elevato sono poste a destra. Il numero di categorie può interagire con esso: all’aumentare del numero di categorie diventa più evidente, ma si potrebbe argomentare anche il contrario (Albanese, Prucha, Barnet, Gjerde, 1997). All’aumentare dell’importanza della valutazione attuata dagli studenti, diventa necessario comprendere come funziona il meccanismo di attribuzione del voto.

La sesta difficoltà della scala Likert è la *reazione all’oggetto* che si ha quando l’intervistato non reagisce al significato delle affermazioni, “ma ai personaggi, alle azioni, alle situazioni menzionate dalle affermazioni stesse” (Cacciola, Marradi, 1988, p. 86). Non riesce a separare, quindi, il significato dell’intera frase (affermazione) dagli elementi che la costituiscono (contenuti nell’affermazione) e genera, pertanto, una distorsione nella risposta. Allora, se l’intervistato è favorevole all’oggetto contenuto nella proposizione, si dichiarerà sempre d’accordo anche quando essa esprime un’opinione contraria alla sua (dovrebbe dichiararsi, invece, in disaccordo). Viceversa, se sarà contrario all’oggetto, si dichiarerà sempre in disaccordo anche quando essa esprime un’opinione favorevole alla sua (dovrebbe dichiararsi, invece, d’accordo). Per ridurre gli effetti, Cacciola e Marradi (1988, p. 100) suggeriscono di ricorrere esclusivamente a “affermazioni positive nei confronti del loro oggetto: in tal modo, sia che l’intervistato valuti l’affermazione [significato], sia che egli reagisca all’oggetto [in essa contenuto], la sua risposta sarà la stessa”. Per le affermazioni negative non c’è

modo, invece, di appurare se la disapprovazione derivi dal significato o dall'oggetto (Sapignoli, 1995). Altre distorsioni derivano per lo piú da incomprensioni linguistiche. La reazione all'oggetto "può emergere solo se l'intervistato premette o fa seguire dei commenti alla risposta incasellata nelle categorie previste da Likert; questo spiega perché [tale fenomeno] non è stato notato in occasione dei sondaggi di massa, che non incoraggiano e comunque non registrano in alcun modo i commenti *a latere* degli items Likert" (Cacciola, Marradi, 1988, p. 86).

La *settima* difficoltà della scala Likert è la «mitezza» (*leniency*) o una specie di tendenza centrale che scaturisce in soggetti psicologicamente orientati e che tendono a valutare qualcuno/qualcosa o con modalità troppo elevate o troppo basse o con una riluttanza a scegliere i valori estremi (Albaum, 1997).

6.3. Obiezioni alla scala utilizzata

Le perplessità sulla scala MIUR, utilizzata nel questionario attuale del NdV, è stata manifestata in piú occasioni. Si riporta un giudizio interessante che ne riassume diversi: «la scala è piuttosto inconsueta e a mio giudizio meno incisiva rispetto alla precedente scala di voto e inferiore alla scala decimale. Nell'esperienza di valutazione degli articoli da pubblicare su riviste, della didattica di corsi fuori dell'università, della prestazione del personale non si è mai vista una scala cosí riduttiva. Si rischia di ottenere una valutazione *grigia* che tende a appiattire il giudizio degli studenti e di non discriminare con chiarezza il giudizio o valore dei docenti. Si è sondato un campione di studenti e docenti per verificare la *portata* di questa scala?».

Tali osservazioni pongono tre problemi: (1) i sintagmi utilizzati per definire le modalità di risposta della scala utilizzata e il loro basso numero, (2) la presenza della modalità centrale in termini impliciti, (3) la verifica empirica della scelta effettuata.

Sul *primo* punto, l'obiezione è piú pessimista del necessario perché molti studi hanno mostrato da tempo che anche un numero ridotto di modalità di scelte funziona abbastanza bene (Gattullo, 1968). I sintagmi sono funzionali alle domande poste. Il punteggio sarà riportato comunque su scala decimale attribuendo **2** per «*Decisamente no*», **5** per «*Piú no che sí*», **7** per «*Piú sí che no*», **10** per «*Decisamente sí*».

Sul *secondo*, presenza/ assenza della modalità centrale, l'esperienza empirica non è riuscita ancora a risolvere definitivamente il problema perché controverso e «*bicornio*». Chiandotto e Gola (2000) hanno scelto, dopo un accurato esame delle diverse possibilità, l'eliminazione della modalità centrale nella scala di Likert e una definizione sintagmatica che semplifica la formulazione delle domande perché, come è facile immaginare, lo stile della domanda dipende da come è definita la scelta per il rispondente. La *scala di voto*, utilizzata dal NdV nel precedente questionario, aveva costretto a ricorrere a un tipo di formulazione non sempre immediata.

Sul *terzo* punto, la risposta è positiva solo per caso; infatti, nell'Università di Modena e Reggio Emilia, si è condotta una indagine, ripetuta piú volte, su campioni estesi di studenti della Facoltà di Economia, per accertare problemi, analogie, e differenze delle risposte degli studenti sui vari aspetti di scelta e valutazione delle modalità (Lalla, 2003; Lalla, Facchinetti, 2004; Lalla, Facchinetti, Mastroleo, 2004). In particolare, gli esiti sulle intenzioni di punteggio corrispondenti alle diverse modalità hanno mostrato una peggiore prestazione della scala MIUR nel suo complesso, rispetto alla «scala di voto» che utilizzava il NdV in passato; mentre la corrispondenza con i voti su scala decimale era leggermente inferiore a quella proposta da Chiandotto e Gola

(2000). In base agli assunti dell'indagine, però, rimane sempre una domanda cruciale: «Un risultato che avesse contraddetto le indicazioni del CNVSU, avrebbe dovuto essere preso in considerazione?». Se «Sì», dove sarebbe andata a finire la confrontabilità, necessaria tra i vari atenei? Ne consegue che l'indagine sarebbe stata, ahinoi, inutile; infatti, è stata condotta solo per l'interesse scientifico degli autori che, casualmente, sono dell'Università di Modena e Reggio Emilia.

7. Conclusioni

Si può evincere da quanto precede che la maggior parte delle questioni sono insolute e insolubili in senso assoluto; per ogni aspetto si possono eseguire scelte che presentano pro e contro. L'insieme delle scelte spesso è dettato dai paradigmi e dagli obiettivi che si vogliono conseguire; pertanto, è frutto di compromessi più o meno accettabili.

Si sono discusse soltanto le questioni più spinose e più frequentemente citate, senza approfondirle adeguatamente sicché la panoramica non è esaustiva e, per gli obiettivi della nota, non vi sono particolari conclusioni da proporre. L'*excursus* può terminare, quindi, esaminando due obiezioni fondamentali: (1) la mancanza di conseguenze concrete dell'attività di valutazione della didattica, che la trasformano in un rito vano e inutile, e (2) la mancanza di comunicazione dei risultati individuali agli studenti per motivi di riservatezza e quelli aggregati non vengono discussi pubblicamente con loro.

La mancanza di effetti concreti della valutazione sull'organizzazione dei corsi e sul docente è, allo stadio attuale, pressoché la norma, almeno per tre motivi. *Primo*, non vi sono procedure e prassi consolidate che consentano di passare dai risultati delle valutazioni ai provvedimenti conseguenti sul docente e sull'organizzazione; in tal caso, l'introduzione della pratica valutativa costituirà un'*agorà* di addestramento e di diffusione di un atteggiamento che alla lunga creerà le condizioni per un'abitudine. *Secondo*, i risultati dell'elaborazione sono noti, per ovvie ragioni di sequenze temporali, alla fine dei corsi, quando i danni —di una pessima organizzazione e/o di un discutibile approccio ai programmi e ai modi e tempi di svolgere le lezioni da parte dei docenti— sono stati già tutti «consumati e digeriti». *Terzo*, nonostante le difficoltà note in alcuni insegnamenti, i risultati empirici sono lusinghieri; ciò potrebbe suggerire un effetto di «controllo» del docente, che «plagia», in un certo senso, lo studente con la sua capacità di svolgere ugualmente il suo lavoro in modo soddisfacente per la sua autorevolezza o anche con una velata *captatio benevolentiae*. La Legge n. 370 del 19 ottobre 1999 (G.U. n. 252 del 26 ottobre 1999) istituiva la distribuzione degli incentivi ai docenti, «a condizione che le loro attività didattiche siano valutate positivamente nell'ambito dei programmi di valutazione della didattica adottati dagli atenei» (art. 4); pertanto, ciò costituiva già un primo passo concreto e sperimentale per dare forza e incisività all'azione di valutazione. Questo effetto della legge è stato affossato e non c'è stata un'altra direttiva sull'uso della valutazione degli studenti. Per esempio, si potrebbe ancorare una quota percentuale piccola, 1-2%, dello stipendio al giudizio positivo degli studenti, come premio. Si riacutizzerebbero, però, le critiche al sistema di valutazione perché la perdita dello stipendio «incide» e, forse, richiederebbe anche un giudizio non limitato solo a quello degli studenti. Da questo circolo vizioso non è facile uscire. ma qui si declinano altri commenti o proposte perché esorbitano dagli obiettivi della nota.

La mancanza di pubblicazione degli esiti individuali è la norma. La diffusione dei risultati, riferiti ai singoli docenti per i singoli insegnamenti, è stata riservata, nell'Università di Modena e Reggio Emilia, ai diretti interessati e alle autorità accademiche: ai Presidi di Facoltà e/o ai Presidenti di CdS. Le opinioni, in materia, sono diverse e si possono ricondurre a due categorie contrapposte: la segretezza del dato e la sua pubblicazione *urbi et orbi*. Nel *primo caso* si rispetta il vincolo di riservatezza sancito dalla Legge n. 675/96 (G.U. n. 5 dell'8/01/1997), generando tra gli studenti la percezione di inutilità della valutazione. Nel *secondo caso* si urta la coscienza di chi lavora, magari anche con impegno, e non ottiene i risultati sperati perché opera in condizioni non idonee o insegna una materia «difficile». Concedere a un pubblico vasto, che a volte non è in grado di leggere correttamente i risultati, la possibilità di accedere ai dati può danneggiare ingiustamente l'immagine del docente che, perciò, non vuole esporsi al rischio di essere sotto i «riflettori». A tale resistenza, si può obiettare che il docente svolge un ruolo pubblico e pubblico dovrebbe essere anche il resoconto della sua attività, come pubblico è anche il suo giudizio sullo studente. L'argomentazione pare ineccepibile, ma si possono verificare danni di immagine da parte di una stampa pronta a esaltare i paradossi (un docente noto con esiti negativi) o a spettegolare (quanto ha preso Umberto Eco). Oramai, anche il voto degli studenti, forse per eccesso di riservatezza, è esposto in forma anonima! L'uso di internet facilita i compiti e in alcune Facoltà, per esempio Economia, si comunicano gli esiti degli esami individualmente. Per una ampia diffusione dei risultati sarebbe opportuno un intervento del CNVSU e/o legislativo per «ricomporre» gli opposti pareri.

Riferimenti bibliografici

- Albanese M., Prucha C., Barnet J.H. e Gjerde C.L. (1977). The Effect of Right or Left Placement of the positive Response on Likert-type Scales Used by Medical Students for Rating Instruction, *Academic Medicine*, 72, pp. 627-630.
- Albaum G. (1997). The Likert scale revisited: an alternate version, *Journal of the Market Research Society*, 39(2), pp. 331-348.
- Amisano E., Rinaldi G. (1988). Confronto tra forme diverse di 'chiusura' degli items Likert, in Marradi A. (a cura di), *Costruire il dato. Sulle tecniche di raccolta delle informazioni nelle scienze sociali*, pp. 44-62, Franco Angeli, Milano.
- Bloom B.S. (1976). *Human Characteristics and School Learning*, McGraw-Hill, New York. Tr. it. (1993). *Caratteristiche umane e apprendimento scolastico*, Armando, Roma.
- Bruner J.S. (1996). *The Culture of Education*, Harvard University Press, Cambridge (MA), 1996. Tr. it. (1997). *La cultura dell'educazione. Nuovi orizzonti per la scuola*, Feltrinelli, Milano.
- Cacciola S., Marradi A. (1988). Contributo al dibattito sulle scale Likert basato sull'analisi di interviste registrate, in Marradi A. (a cura di), *Costruire il dato. Sulle tecniche di raccolta delle informazioni nelle scienze sociali*, pp. 63-102, Franco Angeli, Milano.
- Cannell C. F., Miller P. V., Oksenberg L. (1981). *Research on Interviewing Techniques*, in S. Leinhardt (ed.), *Sociological Methodology*, Jossey-Bass, San Francisco (CA), pp. 389-437.
- Cave M., Hanney S., Kogan M., Trevet G. (1988). *The Use of Performance Indicators in Higher Education: A Critical Analysis of Developing Practice*, Jessica Kingsley Publishers, London.
- Chan J.C. (1991). Response-order Effects in Likert-type Scales, *Educational Psychology Measurements*, 51, pp. 531-540.

- Chiandotto B., Gola M.M. (2000). Questionario di base da utilizzare per l'attuazione di un programma per la valutazione della didattica da parte degli studenti, *Rapporto finale del gruppo di ricerca (RdR 1/00)*, MURST, Osservatorio (ora Comitato nazionale) per la valutazione del sistema universitario, Roma (<http://www.cnvsu.it>)
- Cicchitelli G., Herzel A., Montanari G.E. (1997). *Il campionamento statistico*, nuova edizione, il Mulino, Bologna.
- CNVSU (2002). *Proposta di un insieme minimo di domande per la valutazione della didattica da parte degli studenti frequentanti*, CNVSU, Roma. (DOC 09/02)
- Cochran W.G. (1977). *Sampling Techniques*, 3rd edition, Wiley & Sons, New York
- Coombs C.H. (1953). Theory and Method of Social Measurement, in Festiger L., Katz D. (eds.), *Research Methods in the Behavioral Sciences*, pp. 471-535, New York, Dryden.
- Corbetta P. (1999). *Metodologia e tecniche della ricerca sociale*, il Mulino, Bologna.
- Cox E.P. (1980). The Optimal Number of Response Alternatives for a Scale: A Review, *Journal of Marketing Research*, 17, pp. 407-422.
- Drenth P.J.D., Van Os W., Bernaert G.F. (1989). Improvement of Quality of Education: Through Internal Evaluation (AMOS), in Kogan M. (ed.), *Evaluating Higher Education*, Jessica Kingsley Publishers, London, pp. 56-62.
- Galtung J. (1967). *Theory and Methods of Social Research*, Allen & Unwin, London.
- Garland R. (1991). The Mid-point on a Rating Scale: Is it Desirable?, *Marketing Bulletin*, 2, pp. 66-70.
- Gasperoni G., Giovani F. (1995). Come e perché non funzionano le scale Likert con items a polarità semantica invertita, in Marradi A., Gasperoni G. (a cura di), *Costruire il dato 2. Vizi e virtù di alcune tecniche di raccolta delle informazioni*, 2^a edizione, pp. 60-94, Franco Angeli, Milano.
- Gattullo M. (1968). *Didattica e docimologia. Misurazione e valutazione nella scuola*, Armando Editore, Roma.
- Goode W., Hatt P.K. (1952). *Methods in Social Research*, McGraw-Hill, New York. Tr. it., 1971, *Metodologia della ricerca sociale*, il Mulino, Bologna.
- Goyder J. (1987). *The Silent Minority: nonrespondents on sample surveys*, Blackwell, Oxford.
- Guidicini P. (1995). *Questionari Interviste Storie di vita. Come costruire gli strumenti, raccogliere le informazioni ed elaborare i dati*, Franco Angeli, Milano.
- Guy R.F., Norvell M. (1977). The Neutral Point on a Likert Scale, *The Journal of Psychology*, 95, pp. 199-204.
- Harvey L., Knight P. (1997). *Transforming Higher Education*, London, The Society for Research into Higher Education & Open University Press, Buckingham.
- Hativa N. (1998). Lack of clarity in university teaching: A case study, *Higher Education*, 36, pp. 353-381.
- Hofacker C.F. (1984). Categorical Judgment Scaling with Ordinal Assumptions, *Multivariate Behavioral Research*, 19, 1, pp. 91-106.
- Kogan M. (1989). The Evaluation of Higher Education: An Introductory Note, in Kogan M. (ed.), *Evaluating Higher Education*, Jessica Kingsley Publishers, London, pp. 11-25.
- Lalla M. (2002). *La valutazione dell'attività didattica nell'Università di Modena e Reggio Emilia — Anni Accademici 1998-1999, 1999-2000*; (Allegato), in Nucleo di Valutazione Interna (a cura di), *Relazione per l'anno 1999*, Università di Modena e Reggio Emilia, Modena.
- Lalla M. (2003). La scala di Likert per la valutazione della didattica, *Rivista Italiana di Economia Demografia e Statistica*, LV (4), ottobre-dicembre 2001, pp. 149-175.
- Lalla M. (2005). *La valutazione dell'attività didattica — Anni Accademici 2002-2003, 2003-2004*, (Allegato), in Nucleo di Valutazione Interna (a cura di), *Relazione per gli anni 2002 e 2003*, Università di Modena e Reggio Emilia, Modena.
- Lalla M., Facchinetti G. (2004). Measurement and Fuzzy Scales, in *Atti della XLII Riunione Scientifica «Sessioni Plenarie e Specializzate»*, SIS, Bari, 9-11 giugno, pp. 351-362.

- Lalla M., Facchinetti G., Mastroleo G. (2004). Ordinal Scales and Fuzzy Set Systems to Measure Agreement: An Application to the Evaluation of Teaching Activity, *Quality & Quantity*, 38, pp. 577-601.
- Likert R. (1932). *A Technique for the Measurement of Attitudes*, Archives of Psychology, monografia n. 140.
- Marradi A. (1984). *Concetti e metodo per la ricerca sociale*, 3^a edizione (8^a ristampa 1997), La Giuntina, Firenze.
- Marradi A. (1992). *L'analisi monovariata*, Franco Angeli, Milano.
- Marradi A. (1998). Termometri con vincolo di ordinabilità: il «gioco della torre» consente di aggirare la tendenza alla desiderabilità sociale?, *Sociologia e ricerca sociale*, 57, pp. 49-59.
- Matell M.S., Jacoby J. (1971). Is There an Optimal Number of Alternatives for Likert Scale Items? Study 1: Reliability and Validity, *Educational and Psychological Measurement*, 31, pp. 657-674.
- Matell M.S., Jacoby J. (1972). Is There an Optimal Number of Alternatives for Likert Scale Items? Effects of Testing Time and Scale Properties, *Journal of Applied Psychology*, 56 (6), pp. 506-509.
- Morelli U., Weber C. (1996). *Passione e apprendimento. Formazione-intervento: teoria, metodo, esperienze*, Raffaello Cortina Editore, Milano.
- Orvik J.M. (1972). Social Desirability for Individual, his Group, and Society, *Multivariate Behavioral Research*, 7, pp. 3-32.
- Osservatorio per la valutazione del sistema universitario (1998). *Valutazione della didattica da parte degli studenti. Rapporto finale del gruppo di ricerca*, RdR 1/98, Ministero dell'Università e della Ricerca Scientifica e Tecnologica, Roma.
- Pitrone M.C. (1986). *Il Sondaggio*, Franco Angeli, Milano.
- Pontecorvo C., Ajello A.M., Zucchermaglio C. (1993). *Discutendo si impara*, Carocci editore (NIS - Nuova Italia Scientifica), Roma.
- Ray J.J. (1990). Acquiescence and Problems with Forced-choice Scales, *Journal of Social Psychology*, 130 (3), pp. 397-399.
- Ryan M. (1980). The Likert Scale's Midpoint in Communications Research, *Journalism Quarterly*, 57 (2), pp. 305-313.
- Sapignoli M. (1995). L'intervistato reagisce all'intera frase o solo a singole parole?, in Marradi A., Gasperoni G. (a cura di), *Costruire il dato 2. Vizi e virtù di alcune tecniche di raccolta delle informazioni*, 2^a edizione, pp. 100-125, Franco Angeli, Milano.
- Schuman H., Presser S. (1996). *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*; Sage Publications, Thousand Oaks, CA.
- SPSS (1997a). *SPSS® Base 7.5 for Windows® User's Guide*, SPSS, Chicago.
- SPSS (1997b). *SPSS® Base 7.5 for Windows® Application's Guide*, SPSS, Chicago.
- Sudman S., Brandburn N. M. (1973). Effects of Time and Memory Factors on Response in Surveys, *Journal of the American Statistical Association*, 68, pp. 805-815.
- Wildt A.R. e Mazis M.B. (1978). Determinants of Scale Response: Label versus Position, *Journal of Marketing Research*, 15, pp. 261-267.

APPENDICE 1. Questionario del Nucleo di Valutazione di «Modena-Reggio»

L'esempio, che è riportato nella pagina seguente, si riferisce al questionario stampato dall'autore per l'insegnamento di «Statistica» nel corso di studi in Economia Aziendale. Si può notare che la prima sezione è già compilata; infatti, si è scaricato il questionario in formato pdf dal sito <http://www.esse3.casa.unimo.it/Start.do>, dopo il collegamento con *login* che richiede *password* e *username*; si è aggiunto, poi, in fondo al documento del testo in formato pdf.



Facoltà **ECONOMIA "Marco Biagi"** cod. **50**
 Corso di studi dello studente **ECONOMIA AZIENDALE** cod. **50-200**
 Insegnamento **STATISTICA** cod. **MN2-00914**
 Docente **LALLA MICHELE** cod. **014402**

DATI STUDENTE — GENERE: M F MATURITÀ: Classica Scientifica Tecn.-Ind. Tecn.-Comm. Altro

ANNO CORSO: I II III IV V VI TIPO ISCRIZ.: In corso Fuori Corso MODALITÀ: T. Pieno T. Parz.

Percentuale di lezioni frequentate (circa): Meno del 50% 50%-75% Più del 75% Se meno del 50% indicare il motivo: Lavoro Altre lezioni Altro

Numero (medio) di studenti (compreso il compilatore) che nel periodo di riferimento hanno frequentato questo insegnamento

Eccetto i nn. 14 e 15, i quesiti si riferiscono solo all'insegnamento che si sta valutando		← Giudizi negativi		→ Giudizi positivi	
Modalità di risposta	NON prevista	Decisa-mente NO	Più NO che SÌ	Più SÌ che NO	Decisa-mente SÌ
ORGANIZZAZIONE DI QUESTO INSEGNAMENTO					
1. Il carico di studio di questo insegnamento è proporzionato ai crediti assegnati?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Il materiale didattico (indicato o fornito) è adeguato per lo studio della materia?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Le attività didattiche integrative (esercitazioni, laboratori, seminari, ecc.) risultano utili ai fini dell'apprendimento? (se non sono previste, rispondete "non previste")	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Le modalità di esame sono state definite in modo chiaro?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ASPETTI RELATIVI ALLA DOCENZA					
5. Gli orari di svolgimento dell'attività didattica sono rispettati?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Il docente è effettivamente reperibile per chiarimenti e spiegazioni?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Il docente stimola / motiva l'interesse verso la disciplina?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Il docente espone gli argomenti in modo chiaro e rigoroso?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
INFRASTRUTTURE RELATIVE A QUESTO INSEGNAMENTO					
9. Le aule in cui si svolgono le lezioni sono adeguate (si vede, si sente, si trova posto)?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. I locali e le attrezzature per le attività integrative (esercitazioni, laboratori, seminari, ecc.) sono adeguati? (se non sono previste, rispondete "non previste")	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
INFORMAZIONI AGGIUNTIVE E SODDISFAZIONE					
11. Le conoscenze preliminari possedute sono risultate sufficienti per la comprensione degli argomenti trattati in questo insegnamento?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Sei interessato agli argomenti di questo insegnamento (indipendentemente da come è svolto)?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Sei complessivamente soddisfatto di come è stato svolto questo insegnamento?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ORGANIZZAZIONE DEL CORSO DI STUDI					
14. Il carico di studio complessivo degli insegnamenti ufficialmente previsti nel periodo di riferimento (bimestre, trimestre, semestre, ecc.) è sostenibile?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. L'organizzazione complessiva (orario, esami intermedi e finali) degli insegnamenti ufficialmente previsti nel periodo di riferimento (bimestre, trimestre, semestre, ecc.) è accettabile?		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ASPETTI SPECIFICI DI INSEGNAMENTO/ CORSO DI STUDI/ FACOLTÀ					
16. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

OSSERVAZIONI E SUGGERIMENTI

- | | | |
|--|--|---|
| <input type="checkbox"/> Migliorare il coordinamento con altri corsi | <input type="checkbox"/> Alleggerire il carico didattico complessivo | <input type="checkbox"/> Fornire più conoscenze di base |
| <input type="checkbox"/> Migliorare la qualità materiale didattico | <input type="checkbox"/> Eliminare argomenti già trattati in altri corsi | <input type="checkbox"/> Aumentare le ore di esercitazioni |
| <input type="checkbox"/> Fornire in anticipo il materiale didattico | <input type="checkbox"/> Aumentare attività di supporto didattico | <input type="checkbox"/> Inserire prove di esame intermedie |