

Modelli teorici e strumenti di valutazione degli esiti in università

Cristina Coggi • Università di Torino – cristina.coggi@unito.it

Theoretical models and evaluation tools of learning outcomes in Higher Education

Relevant requirement for the University is to define and explain the results, the learning outcomes of students. It's a problem with a relatively recent history, of which we'll discuss some steps. Then we'll propose some elements of reflection about the models that are adopted, on related measurement tools, to finally find new challenges for research.

Keywords: learning outcomes, competences, ranking, higher education, formative assessment, critical thinking.

Definire ed esplicitare gli esiti, vale a dire i risultati di apprendimento degli studenti, è un'esigenza rilevante per l'Università. Si tratta di un problema con una storia relativamente recente, di cui richiameremo qualche passaggio. Proporremo quindi alcuni elementi di riflessione sui modelli che si adottano, sugli strumenti di misura correlati, per individuare infine qualche linea di approfondimento per la ricerca.

Parole chiave: esiti di apprendimento, competenze, classificazione, università, valutazione formativa, pensiero critico.



Modelli teorici e strumenti di valutazione degli esiti in università

1. La valutazione dei *learning outcomes* nelle università americane

In America il problema di valutare la qualità degli esiti universitari si è posto da tempo, a partire dagli anni '80, con il movimento dell'*assessment* (Kuh, Ewell, 2010). Nel decennio successivo si è sviluppata una crescente richiesta alle Università di esplicitare i traguardi raggiunti dagli studenti e dar conto degli stessi. Nel 2006 il *National Report* del Dipartimento dell'Educazione Statunitense ha dichiarato l'*accountability* dell'università pubblica una delle quattro priorità per il futuro delle università (Spellings, 2006). Si ritiene infatti che l'Università sia "responsabile almeno parzialmente del progresso intellettuale degli studenti negli anni universitari" e che questo debba essere documentato, attraverso indagini che comparino i *learning outcomes* in ingresso con quelli in uscita dai percorsi formativi (Liu, 2009).

Nell'aprile 2009 trecentoventuno istituzioni universitarie di 50 Stati hanno aderito al "Sistema volontario di Accountability" (VSA). Lo scopo era quello di valutare il cuore dei risultati formativi nelle università pubbliche e nei college, focalizzandosi su abilità comuni, multidisciplinari, formate a livello universitario, individuandole essenzialmente nelle abilità di comunicazione scritta, nel pensiero critico e nel ragionamento analitico (Liu, 2009).

Per la misura, sono stati dunque selezionati tra i diversi strumenti disponibili tre test, che offrivano garanzie di validità e affidabilità:

- il **CLA** (Collegiate Learning Assessment) del CAE (Council for Aid to Education);
- le **ETS Proficiency Profile**, trasformato poi in MAPP (Measure of Academic Proficiency Profile, Liu, 2011);
- il **CAAP** (Collegiate Assessment of Academic Proficiency) dell'ACT.

I tre test presentano però differenze nel formato degli item, nelle modalità di somministrazione e qualche diversità anche nella valutazione di abilità integrative¹. Date le disparità tra i test e la possibilità per le Università di scegliere liberamente quale utilizzare, si è posto presto il problema della confrontabilità dei risultati degli stessi. La questione è stata ampiamente dibattuta in letteratura. Sono state condotte ricerche per verificare la comparabilità dei dati (Test Validity Study, Shulenberg, Keller, 2010), con esiti ritenuti meritevoli ancora di ulteriori approfondimenti. Secondo i ricercatori del CAE comunque, tutti e tre i test forniscono indicatori affidabili dell'apprendimento degli studenti e sono sensibili al progresso che avviene all'interno di un'università nelle abilità misurate, ma non è chiaro se le scale riferite allo stesso costrutto valutino effettivamente la stessa proprietà (es. capacità critica) e se i loro punteggi di valore aggiunto possano essere sostanzialmente confrontati (Steedle, Kugelmass, Nemeth, 2016).

1 Le abilità in matematica e in lettura sono valutate solo dall'ETS PP e dal CAAP.



2. Valutare i risultati di apprendimento al termine dei curricula universitari europei

Le Università europee sono dal 1999 coinvolte nel processo di Bologna, finalizzato a rendere il più possibile equivalenti e/o comparabili i percorsi formativi accademici e i risultati in uscita. Ricerche europee hanno iniziato a studiare la fattibilità di indagini comparative, simili a quelle americane, allo scopo di accertare il conseguimento di traguardi di competenza nei diversi Paesi (AHELO, OCDE, 2013). A tal fine sono state pianificate le rilevazioni delle cosiddette “competenze generali e trasferibili”, riconducendole ai medesimi costrutti individuati dagli americani: pensiero critico, ragionamento analitico, abilità nella comunicazione scritta e *problem solving*. Quando è stato necessario definire il test per la misura, le scelte tra gli strumenti disponibili con adeguate qualità psicometriche erano ridotte, con pochissimi competitori nel mercato globale (Douglass, 2012).

L'opzione è caduta su uno dei tre test utilizzati dal VSA: il “CLA Higher Education” del CAE, che è stato somministrato insieme ad un test a scelta multipla australiano dell'ACER². Sono state avviate inoltre indagini pilota su specifiche abilità disciplinari e si è provveduto ad approfondire i problemi di misura del valore aggiunto.



3. L'indagine TECO italiana: valutare le competenze di carattere generalista

Lo studio italiano realizzato dall'ANVUR nel 2012 (TECO), sulle “competenze generali e trasferibili” ha adottato il test americano CLA, rinnovato, nella versione CLA+. Lo scopo dell'indagine è stato quello di produrre risultati originali per l'Italia e di analizzare il problema della possibilità di confrontare gli esiti tra Stati (Kostoris, 2014). Si è trattato di un incarico con un ruolo di prestigio per il nostro Paese, in quanto lo studio ha affrontato un problema chiave che è alla base della possibilità di effettuare graduatorie di comparazione internazionale. L'indagine è stata condotta su un campione di 5.900 studenti italiani e di 4.800 studenti americani.

Dal rapporto di ricerca ANVUR 2014 emerge che le distribuzioni dei risultati dei due campioni sono pressoché sovrapponibili, nelle medie e nei quartili (con differenze in alcune sottoscale). Le norme sono analoghe, quindi la comparabilità pare possibile. Si tratta di approfondirne scopi, modalità ed effetti.

4. Riflessioni e sfide

Il breve iter dei tentativi di misura dei *learning outcomes* in Università, che abbiamo sintetizzato, sollecita qualche riflessione.

a. Dipendenza delle scelte europee da quelle americane

L'Europa ha mutuato dall'America le strategie di valutazione degli esiti di apprendimento in Università, con alcuni adattamenti degli strumenti. Occorre interro-

2 Australian Consortium for Educational Research.

garsi però se i traguardi europei debbano essere i medesimi di quelli americani, individuati inizialmente negli USA nel 1991 per le università e i college, e se i costrutti e gli strumenti scelti siano pienamente adeguati oggi al contesto europeo ed italiano in specifico. Più in generale, ci si può domandare se l'università europea debba omologarsi a quella americana o se la sua missione non sia piuttosto quella di concorrere allo sviluppo socio-economico e sociale delle singole nazioni e, in forma più integrata, allo sviluppo della società europea, mantenendo in questa fase storica alcune sue specificità, anche di eccellenza in alcuni campi.

b. Problemi edumetrici da affrontare

Prendendo atto delle scelte effettuate a livello europeo, restano comunque aperti alcuni problemi, per esempio edumetrici. Si tratta di limiti metodologici emersi dallo studio europeo di fattibilità (AHELO, Tremblay, Lalancette, Roseveare, 2012) rispetto alla valutazione dei *learning outcomes* (Wolf, Zahner, Benjamin, 2015) e che erano già stati identificati da ricerche americane (Liu, 2009; Douglass, 2012).



1. Un problema rilevante riguarda il campionamento degli studenti nei vari Paesi, che a livello universitario diventa difficilmente controllabile.
2. Emerge inoltre che sui risultati cognitivi incidono la motivazione e la persistenza degli studenti nel terminare il test, a fronte della scarsa possibilità di riconoscimento dello sforzo erogato dagli stessi.
3. Il diverso formato con cui vengono somministrati i test (con item a scelta multipla o semi-strutturati) può inoltre orientare la misurazione verso costrutti diversi, a dispetto dell'intenzione di rilevare una medesima abilità. Questo rende di fatto complessa l'interpretazione degli esiti.
4. Gli strumenti adottati risultano fortemente saturati nella lingua scritta, in quanto valorizzano abilità di lettura rapida su testi diversi, anche di tipo scientifico, e abilità di argomentazione scritta. Favoriscono così alcune fasce della popolazione e alcune tipologie di percorso universitario, come i curricula scientifici e/o quelli storico-filosofici, in cui è previsto un esercizio sistematico del pensiero critico-analitico su materiale verbale. Questo è confermato peraltro dall'indagine TECO (2014). Diventano quindi problematici eventuali ipotetici confronti tra corsi di studio della medesima università, sulla base dei dati ottenuti con i test in oggetto.
5. Rispetto alla *comparabilità* tra Paesi con il CLA incidono: le caratteristiche specifiche della lingua in cui viene riadattato il test, che possono renderlo più o meno complesso; il tipo di testo scelto (che cambia le graduatorie tra Paesi) e, per la parte semi-strutturata del test, la differenza di severità tra i valutatori, che dipende dalle abitudini professionali di correzione dei diversi Paesi (Zahner, Steedle, 2014).
6. Rispetto all'utilizzo delle misure di *learning outcomes* per l'*accountability* si pone il problema se utilizzare un disegno longitudinale oneroso, cioè svolgere rilevazioni sulla stessa coorte a distanza di anni (con rischio di mortalità sperimentale) o un disegno a coorti incrociate (che suppone però la comparabilità di coorti estratte da popolazioni potenzialmente differenti, quella delle matricole e quella dei laureandi).

c. Questioni etiche

Un ulteriore problema riguarda le prove. Gli items che valutano il pensiero critico spesso coinvolgono anche questioni etiche, che però non vengono prese in considerazione nelle chiavi di correzione³. Occorre però riflettere sul significato implicito di queste scelte che privilegiano la logica formale rispetto al ragionamento etico.

d. Effetti possibili sulla didattica e sull'apprendimento

Impostare in modo accentrato gli strumenti di valutazione degli esiti, secondo un modello di *accountability* guidata da test, porta a definire degli standard da raggiungere: tende quindi a dare orientamenti stringenti alla didattica, riducendone la flessibilità, per favorire la conformità agli standard delle prestazioni e dei programmi dei corsi (Rebora, 2013). L'operazione richiede dunque in generale particolare cautela, anche in considerazione del fatto che comportamenti di adattamento al sistema, portano ad affinare le competenze richieste agli studenti anche su curricoli che non le prevedono come dimensioni rilevanti. Questo si evidenzia, per esempio, in sperimentazioni recenti, condotte negli USA (Anderson, Reid, 2013; Pascarella, Martin, Hanson et alii, 2014) e in diversi altri contesti, sullo sviluppo del pensiero critico in università, tentato anche nei confronti delle professionalità educative con esiti deludenti (Yücel Toy, Ok, 2012).

L'uso poi di test strutturati nella valutazione finale può favorire l'incremento dell'impiego di prove oggettive nei corsi universitari, ancor più a fronte di un numero elevato di iscritti. Tale utilizzo però, secondo la ricerca, incide negativamente sull'abilità di elaborazione degli studenti e favorisce l'apprendimento in superficie (Biggs, 2011).



5. Quesiti di sfondo

Abbiamo evidenziato alcune sfide per l'affinamento della ricerca sull'*accountability* secondo i modelli che sono stati adottati. Potremmo spostare l'attenzione ora sull'uso dei risultati TECO e AHELO e sulle evidenze che se ne possono far emergere, per ricondurre la riflessione sulle scelte preliminari che guidano tali valutazioni.

a. Quali evidenze stiamo raccogliendo in Italia?

Si può osservare che i costrutti misurati dal CLA+ risultano parzialmente riconducibili a quelli accertati dalle prove PISA (*problem-solving* su questioni autentiche). Si può dunque supporre che gli esiti regionali delle due indagini correlino, a meno che le università presentino, ai sensi dell'*accountability*, *effect size* molto diversi tra loro. Se proviamo a correlare, per esempio, gli esiti medi regionali italiani del PISA in matematica (2012) con le medie dei punteggi per regione del CLA+ (indagine TECO

3 Per esempio, nella "Prova Parchi" del CLA+ non si controlla se lo studente adotta criteri egualitari o di privilegio nelle decisioni, ma se i processi di analisi sono adeguatamente esercitati e le scelte sono argomentate in modo stringente e linguisticamente corretto.

2014), si rileva una correlazione elevata ($r=0,81$) e significativa (con $p<0,001$), con una retta di regressione cui si allineano molti esiti del CLA+ (v. fig. 1).

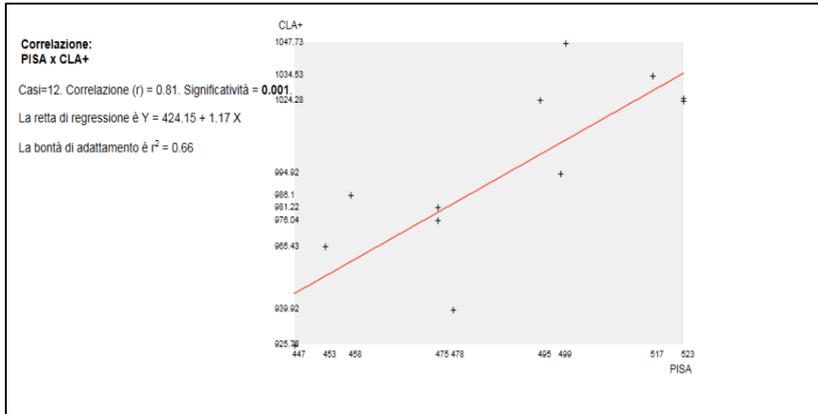


Fig. 1 – Retta di regressione tra i punteggi del PISA- matematica 2012 e il CLA plus.



Le graduatorie regionali che risultano dai dati PISA di matematica non vengono dunque sostanzialmente mutate da anni di Università. Il sistema formativo universitario italiano ha quindi difficoltà a spostare le graduatorie delle secondarie, che a loro volta sono correlate alla stratificazione culturale e sociale⁴, naturalmente con lodevoli eccezioni, rispetto agli ambiti misurati. Si può ipotizzare in alternativa che variando il tipo di competenze misurate come esiti finali dell'Università, si potrebbero ottenere graduatorie differenti, in grado di apprezzare apporti dell'Università non previsti dalle rilevazioni TECO.

b. Quali evidenze stiamo raccogliendo in Europa?

A fronte di uno sforzo così ingente e condiviso tra tanti Paesi e di rilevanza per le politiche, occorre domandarsi su quali aspetti dell'apprendimento degli studenti ci si sta focalizzando. Il sistema di misurazione dei risultati adottato dall'Europa opera infatti scelte esplicite (in analogia con quanto avviene in America), centrandosi in primo luogo sui processi cognitivi e privilegiando in maniera rilevante, oltre al *problem-solving*, soprattutto il pensiero critico. Sia il CLA sia il CLA+ attribuiscono infatti a tale processo un peso singolare, a discapito delle altre operazioni intellettuali. Questo definisce una personalità modale che le università si propongono di formare e che caratterizza i risultati degli studenti in merito ad alcune dimensioni cognitive acquisite, trascurando altri aspetti della formazione. Si presuppone, almeno implicitamente, che il senso critico, misurato con il CLA, caratterizzi futuri professionisti in grado di incidere sul mercato competitivo nazionale e mondiale e sullo sviluppo sociale. In una logica di *accountability*, dovremmo dunque riscontrare che gli esiti degli sforzi formativi delle università, così misurati, caratterizzino dei laureati poi in grado di favorire lo sviluppo economico e il be-

4 Come risulta dai rapporti TECO e PISA.

nessere sociale dei Paesi. Per evidenziare questo risultato occorrerebbe intraprendere una ricerca approfondita, in analogia con gli studi avviati da McClelland⁵, che cercava le connessioni tra il potenziamento della motivazione al successo, ottenuto con l'educazione⁶, e l'indice di sviluppo dei Paesi 25 anni dopo⁷.

Per tentare di esplorare, in forma pilota, nel nostro caso, l'ipotesi di connessione tra esiti universitari e andamenti dell'economia dei Paesi, potremmo provare a correlare i punteggi medi CLA⁸ dei 9 Paesi considerati dall'indagine AHELO 2012 con il PIL nazionale corrispondente⁹. Il tentativo evidenzia un coefficiente di correlazione molto basso ($r=0,16$, ovviamente non significativo con $p=0,68$), che non attesta alcun legame tra le abilità intellettuali rilevate nei laureati con il CLA e il PIL del loro Paese (v. fig. 2).

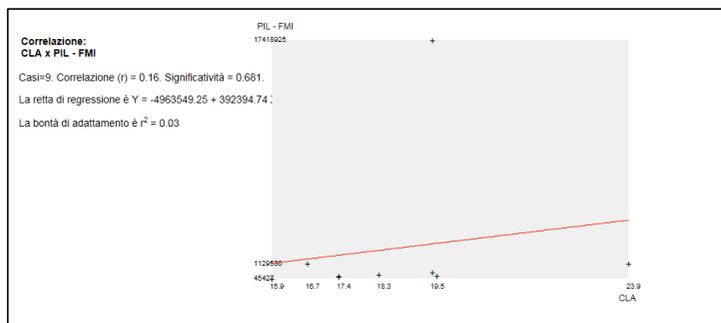


Fig. 2 – Retta di regressione e correlazione tra punteggi medi del CLA e il PIL di 9 Paesi

Se invece supponiamo che i laureati contribuiscano allo sviluppo sociale e umano di un Paese, per verificare tale ipotesi possiamo connettere gli esiti di apprendimento in uscita dalle università con indici come il FIL (o GNH: “Gross national happiness”), finalizzato a misurare “la somma non solo degli esiti economici nei Paesi, ma anche degli impatti ambientali netti, della crescita culturale e spirituale dei cittadini, della salute mentale e fisica e della solidità dei sistemi politici e aziendali”. La correlazione però si conferma non significativa, e per 5 Paesi (v. tab. 1) diventa addirittura negativa, con rho di Spearman = - 0,7 ($p=0,18$).

- 5 Gli studi di McClelland hanno suscitato interesse e di ambito non solo in ambito psicologico, ma anche economico (S.P. Schatz, “Achievement and Economic Growth: A Critique”, *The Quarterly Journal of Economics*, vol. 79, n. 2, 1965, pp. 234-245).
- 6 Con attività di lettura nell'infanzia.
- 7 Calcolato con il fabbisogno di energia elettrica.
- 8 Si ipotizza che le graduatorie ottenute con il CLA siano sufficientemente stabili da poter rappresentare le competenze dei laureati di un Paese per più anni e che quindi le competenze dei laureati 2012 siano indicative di quelle che si sarebbero potute ottenere anni prima, nei laureati provenienti dalle medesime università già entrati nel mondo del lavoro, in grado quindi di contribuire al PIL del Paese.
- 9 Misurato con l’FMI, l’anno successivo (2013-14).

	CLA	FIL (o GNH)
Messico	16	8
USA	17	4
Kuwait	32	7
Colombia	35	5
Korea	41	1

Tab. 1 – Esiti medi del CLA per Paese e l'indice FIL (o GNH) di 5 Paesi

Emerge dunque che la graduatoria per Paese delle abilità intellettuali dei laureati, misurate dal CLA, non consente di prevedere quella definita dall'indice di sviluppo sociale, umano ed economico del Paese stesso.

Siamo dunque chiamati a formare e valorizzare “têtes critiques”, supponendo, ma non riuscendo a dimostrare, che siano utili per il mercato competitivo mondiale, sul modello delle selezioni americane?

Dobbiamo preoccuparci invece di educare “têtes bien faites” e personalità responsabili e creative per sviluppare i contesti sociali e le economie locali e mondiali in modo sostenibile ed equo?



Se optiamo per la seconda ipotesi dobbiamo metterci nella prospettiva di cambiare il modello di valutazione dei risultati di apprendimento, adottando strategie più variate, che difficilmente si connettono con la costruzione di graduatorie tra università, basate su pochi indicatori. Questo però suppone di riflettere sui modelli formativi e di comparazione delle università.

c. La priorità è il ranking internazionale?

Affrontando le questioni in maniera più globale, occorre domandarsi se è essenziale che l'Europa faccia prevalere un modello neoliberista per la valutazione dei risultati del sistema universitario, assumendo un paradigma come quello della “World Class University” (WCU), con il *ranking*. Tale sistema si propone di classificare le università in una graduatoria mondiale sulla base della capacità di promuovere risultati competitivi degli studenti, attrarre docenti e studenti di alto livello, ottenere finanziamenti elevati e realizzare ricerca di qualità (Cheol Shin, Kehm, 2013). Secondo studi condotti in prestigiose università (es. Berkeley), il sistema di valutazione connesso a tale paradigma rischia di non essere adeguatamente informativo e orientativo per consentire alle singole istituzioni, specie se grandi, di individuare metodi e comportamenti organizzativi che possano rendere più produttiva la ricerca, migliore la didattica, più efficienti i servizi erogati e più elevati gli esiti. In questa comparazione inoltre le università europee rischiano di non essere adeguatamente valorizzate (Van der Wende, 2009).

Dunque dobbiamo forse cercare modelli diversi da quelli *ranking* per dar conto degli esiti conseguiti.

5. Modelli alternativi attuali

La missione dell'università è molto ampia e complessa e lo sono anche gli esiti conseguiti. Occorre dunque un modello di rilevazione dei risultati di apprendimento altrettanto articolato e non ridotto a pochi dati per un puro *ranking* competitivo. Si stanno sviluppando dunque a livello internazionale, in diverse università, proposte alternative. Ne riportiamo alcune a titolo esemplificativo.

A. New Flagship model: un modello GLOCAL

L'Università di Berkeley, per esempio, propone un modello GLOCAL, denominato "New Flagship model" (Douglass, 2014). Tale proposta, senza ignorare gli standard internazionali di eccellenza, va nella direzione di espandere più significativamente il ruolo delle Università nelle società di appartenenza, secondo una logica anche di differenziazione. Oggetto della valutazione è la rilevanza regionale e nazionale delle stesse per poi accettare l'influenza e il prestigio globale. Secondo il "New Flagship model (Douglass, 2016)" l'Università è chiamata a darsi una molteplicità di scopi, quali: costituire un ambiente di apprendimento e di ricerca produttivo, creare nuova conoscenza e preservare quella del passato, contribuire a una società più giusta e prospera, favorire l'avanzamento individuale delle capacità umane e svolgere una funzione di monitoraggio costante della società nella quale si trova inserita. L'Università si propone dunque di formare studenti civicamente impegnati, di promuoverne la creatività, la cultura e abilità di alto livello e anche su questi aspetti viene orientata la valutazione.



B. Proposta di ampliamento dei traguardi da valutare e trasformazione degli strumenti

Alcuni consorzi di università americane recentemente stanno andando nella direzione di creare strumenti alternativi per l'apprezzamento dei risultati.

- 1) Ne è un esempio significativo lo strumento proposto dalla rete SERU, promossa dall'Università di Berkeley (Thomson, Douglass, 2009), che ha lo scopo di valorizzare diversi aspetti dell'esperienza accademica degli studenti, attraverso un questionario e scale di autovalutazione. Tale strumento ha evidenziato *face validity* e validità concorrente con i punteggi medi attribuiti dai docenti nei corsi universitari (California GPA) (Douglass, Thomson, Zhao, 2012).
- 2) Allo stesso modo, all'interno del progetto "Rising to the challenge"¹⁰ è stato realizzato e validato dall'American Association of State Colleges and Universities (AAC&U) uno strumento di autovalutazione (*Degrees of Preparation*), che si propone di rilevare negli studenti:
 - le abilità *utili per avere successo in un mondo globalizzato* (valorizzando, per esempio, le esperienze di studio e lavoro all'estero, l'apprendimento delle lingue straniere...);

10 AAC&U, AAC&U, APLU, *Rising to the challenge. Meaningful Assessment of Student Learning*, FIPSE, 2010 (<http://files.eric.ed.gov/fulltext/ED519804.pdf>).

- la *preparazione al lavoro* (es. esperienze di lavoro durante il percorso universitario);
- lo sviluppo del *coinvolgimento civico* (si rilevano, per esempio, le fonti di informazione utilizzate più frequentemente, indicatori di coinvolgimento politico, capacità di lavoro in gruppo, attività di volontariato, pensiero critico e capacità comunicative, senso civico...).

La maggior parte delle scale strutturate risultano valide e capaci di discriminare tra matricole e laureandi.

- 3) All'interno del *Values Project* (Rhodes, 2010) viene privilegiato il giudizio di esperti multipli. Per la valutazione sono stati individuati da più di cento Facoltà 15 apprendimenti considerati irrinunciabili al termine dei curricula universitari e sono state realizzate rubriche per valutare analiticamente il livello raggiunto dallo studente in ciascuno di essi.

Le rubriche hanno coperto tre aree.

- Per le "*abilità cognitive e pratiche*" si rilevano le capacità di: indagare e analizzare; pensare criticamente; creare; comunicare in forma scritta; comunicare oralmente; raccogliere informazioni e comunicarle in termini matematici; individuare, valutare e usare le informazioni; lavorare in gruppo; risolvere problemi; leggere.
- Per rilevare la "*responsabilità personale e sociale*" si valutano: la conoscenza e il coinvolgimento civico (locale e globale); le conoscenze e competenze interculturali; il ragionamento etico; il possesso delle conoscenze fondamentali e delle abilità di apprendimento per la vita.
- Si rileva inoltre l'"*integrazione e il transfer delle conoscenze apprese*".

Le rubriche, una volta validate e riferite a criteri e livelli uniformi, sono state applicate alla valutazione degli e-portfoli degli studenti, che consentono di mostrare adeguatamente e in maniera ampia il percorso universitario e le acquisizioni dei laureandi. Si tratta di strumenti utili per valutare apprendimenti complessi che non potrebbero essere apprezzati da test standardizzati¹¹.

Anche Università di Paesi emergenti come la Cina (Hong Kong) propongono di arricchire la gamma di apprendimenti da valutare, includendo oltre ai traguardi cognitivi anche quelli affettivi, come le trasformazioni di atteggiamenti e valori, lo sviluppo dell'empatia e di comportamenti etici (Tam, 2014).

Si aprono quindi nuove sfide per la ricerca docimologico-didattica sui problemi considerati. Ne prospetteremo alcune.

6. Direzioni di ricerca sulla valutazione degli esiti

Possiamo soffermarci brevemente su cinque ambiti di studio meritevoli di approfondimento.

11 Modelli più articolati di valutazione dei risultati di fine scuola secondaria in vista dell'accesso all'università o al mondo del lavoro si stanno elaborando in letteratura secondo nuovi paradigmi sostenuti da grandi università, come quella di Stanford (L. Darling-Hammond, G. Wilhoit, L. Pittenger, *Accountability for College and Career Readiness: Developing a New Paradigm*, Stanford Center for Opportunity Policy in Education, 2014 (<https://edpolicy.stanford.edu/sites/default/files/publications/accountability-college-and-career-readiness-developing-new-paradigm.pdf>)).



- 1) Alla luce della ricerca internazionale, è possibile pensare in Italia di integrare il modello attuale di valutazione per dar conto meglio della ricchezza degli esiti della formazione universitaria, delle trasformazioni dei curricula e delle istanze provenienti anche dal mondo del lavoro. Si tratta di rilevare, oltre al pensiero critico, dimensioni trasversali aggiuntive, ritenute essenziali come risultati di apprendimento delle università, come il pensiero creativo, le abilità di comunicazione orale, le competenze di cittadinanza attiva di tipo globale, il ragionamento etico...
- 2) Particolare attenzione dovrebbe essere riservata alla creatività. L'Europa e l'Italia soprattutto hanno visto nell'innovazione, non solo tecnologica, la possibilità di sviluppo della propria produzione e cultura. Quindi i curricula formativi dovrebbero dare rilevanza alle *caratteristiche divergenti* del pensiero, considerate una dimensione importante dello sviluppo intellettuale a tutti i livelli scolastici, ma in particolare in università (Orlando, 2012). Per questo occorrerebbe sviluppare coerentemente strumenti in grado di rilevare anche questi aspetti, per dar conto dei progressi ottenuti.
- 3) Rispetto alla rilevazione attuale delle competenze di carattere generalista, rimane aperta la questione delle modalità più adeguate per condurre rilevazioni in alcuni settori specifici, come quello educativo, in cui la ricerca in diversi contesti, da tempo sottolinea che le abilità di pensiero critico, al centro delle rilevazioni, ottengono misure medio-basse. Del resto il CLA+, nella versione americana, non indica, tra le aree di studio per le quali sono previsti i compiti di performance del test, il settore dell'Educazione. Forse i futuri educatori ed insegnanti andrebbero valutati con strumenti adeguatamente riadattati o almeno dovrebbero essere esclusi dai campioni per la misura delle competenze generaliste, nel caso che si continuino ad utilizzare gli strumenti del TECO.
- 4) Attualmente è previsto che la ricerca di valutazione dei *learning outcomes* per l'Italia transiti alla seconda fase, in cui si richiede di individuare per i diversi settori disciplinari le conoscenze irrinunciabili, esito della formazione e quindi della valutazione. A questo proposito emergono alcune questioni. Per curricula come quello degli insegnanti, che prevedono modelli di formazione anche con aspetti più clinici come per i medici, è rilevante centrarsi sulla modalità con cui le conoscenze vengono mobilitate per affrontare i problemi che si pongono nei contesti di lavoro. In questa linea si è sviluppata una corrente di ricerca, detta traslazionale (Montgomery, 2014), e sono stati pubblicati studi applicativi, per esempio dall'Università di Boston, che propongono di adottare per la valutazione degli insegnanti in formazione, un portfolio di evidenze. Si tratta di uno strumento che consente (attraverso studi di caso, strumenti per rilevare gli aspetti etici, rubriche, progetti...), di verificare i progressi in itinere e le competenze conseguite dai futuri docenti (La Velle, 2015).
- 5) Risulta inoltre complesso definire le conoscenze specifiche irrinunciabili di tipo fattuale, concettuale, procedurale, metacognitivo (Anderson, Krathwohl, 2001) per i percorsi con profili d'uscita più variati e/o meno definiti. Un confronto nazionale tra università, con una ricerca connessa, potrebbe portare ad una prima ricognizione delle stesse.
- 6) Preoccupati di valutazione sommativa non dobbiamo perdere di vista il valore formativo che deve avere la valutazione anche in università. Nella logica di connettere i risultati a un miglioramento della didattica e degli esiti degli studenti, occorre approfondire, anche le strategie di valutazione in itinere, in una logica di valutazione regolativa e formativa. Si tratta di ripensare la valutazione legata ai corsi secondo un approccio più attento ai processi attivati e più ri-



spondente ad esigenze di chiarezza e autenticità e quindi di sviluppo di competenze. Occorre come docenti assumere l'impegno di una valutazione trasparente, riferita a criteri espliciti e in grado di offrire feed-back formativi agli studenti, restituendo loro la responsabilità di monitorare i propri processi di acquisizione, in vista di un apprendimento per la vita e della maturazione della *working readiness*. Sarà utile allo scopo approfondire l'uso di scale, questionari, rubriche, schede di autovalutazione e portfoli, anche con il supporto delle tecnologie (in particolare per gli studenti non frequentanti).

7. Conclusione

La valutazione formativa e scelte adeguate per la valutazione sommativa finale possono accompagnare meglio la transizione attuale. Occorre passare da un'università finalizzata a potenziare il capitale culturale e intellettuale per una società della conoscenza applicata, ad un'università che forma competenze e prepara cittadini responsabili per un mondo globalizzato. Questo esige una didattica e una valutazione profondamente rinnovati in Università.



Riferimenti bibliografici

- Anderson, L.W., Krathwohl, D.R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Anderson, P.R., Reid, J.R. (2013). The Effect of Critical Thinking Instruction on Graduates of a College of Business Administration. *Journal of Higher Education Theory and Practice*, 13(3-4), pp. 149-167.
- Biggs, J.B. (1999, fourth edition 2011). *Teaching for Quality Learning at University: What the Student Does*. Maidenhead: McGraw-Hill.
- Cheol Shin, J., Kehm, B.M. (eds) (2013). *Institutionalization of World-Class University in Global Competition*. London: Springer.
- Darling-Hammond, L., Wilhoit, G., Pittenger, L. (2014). *Accountability for College and Career Readiness: Developing a New Paradigm*. Stanford Center for Opportunity Policy in Education. Estratto da: <<https://edpolicy.stanford.edu/sites/default/files/publications/accountability-college-and-career-readiness-developing-new-paradigm.pdf>>.
- Douglass, J.A. (2014). Profiling the flagship University Model: An Exploratory Proposal for Changing the Paradigm From Ranking to Relevancy. *Research & Occasional Paper Series*, CSHE 5(14). Estratto da: <<http://cshe.berkeley.edu/>>.
- Douglass, J.A. (2016). *The New Flagship University: Changing The Paradigm from Global Ranking to National Relevancy*. London: Palgrave Macmillan.
- Douglass, J.A., Thomson, G., Zhao, C.-M. (2012). Searching for the holy grail of learning outcomes. *Research & Occasional Paper Series*, CSHE 3(12). Estratto da: <<http://cshe.berkeley.edu/>>.
- Douglass, J.A., Thomson, G., Zhao, C.-M. (2012). The learning outcomes race: the value of self-reported gains in large research universities. *Higher Education*, 64, pp. 317-335.
- Kostoris, F. (a cura di) (2014). *Le competenze effettive di carattere generalista dei laureati italiani*. Roma: ANVUR. Estratto da: <<http://www.anvur.org/attachments/article/248-/Rapporto%20TECO%202014.pdf>>.
- Kuh, G.D., Ewell, P.T. (2010). The state of learning outcomes assessment in the United States. *Higher Education Management and Policy*, 22(1), pp. 9-28.
- La Velle, L. (2015). Translational research and knowledge mobilization in teacher education: towards a 'clinical', evidence-based profession? *Journal of Education for Teaching*, 41 (5), pp. 460-463.

- Liu, O.L. (2009). Measuring Learning Outcomes in Higher Education. *R & D Connections*, 10, pp. 1-5.
- Liu, O.L. (2011). Measuring value-added in higher education: conditions and caveats – results from using the Measure of Academic Proficiency and Progress (MAPPP). *Assessment & Evaluation in Higher Education*, 36 (1), pp. 81-94.
- McClelland, D.C. (1961). *The Achieving Society*. Princeton: Van Nostrand.
- Montgomery, C. (2014). Translational and transcultural positionality in globalized higher education. *Journal of Education for Teaching*, 40(3), pp. 198-203.
- Orlando, M. (2012). Fostering creativity in higher education. *RHESL*, 5(14), pp. 54-61.
- Pascarella, E.T. et alii (2014). Effects of Diversity Experiences on critical thinking skills over 4 years of College. *Journal of College Student Development*, 55(1), pp. 86-92.
- Rebora, G. (2013). *Nessuno mi può giudicare? L'Università e la valutazione*. Milano: Guerini Associati.
- Rhodes, T.L. (2010). Valid Assessment of Learning in Undergraduate Education. *Rising to the Challenge: Meaningful Assessment of Student Learning*, FIPSE, 2010. Estratto da: <<http://files.eric.ed.gov/fulltext/ED519804.pdf>>.
- Shulenberg, D., Keller, C. (2010). Interpretation of Findings of the Test Validity Study Conducted for the Voluntary System of Accountability. In AAC&U, AASCU, APLU, *Rising to the challenge. Meaningful Assessment of Student Learning*, FIPSE. Estratto da: <<http://files.eric.ed.gov/fulltext/ED519804.pdf>>.
- Soriano de Alencar, E.M.L. (2004). Inventário de Práticas Docentes que Favorecem a Criatividade no Ensino Superior. *Psicoogia: Reflexão e Crítica*, 17(1), pp. 105-110.
- Spellings, M. (2006). *A Test of Leadership: Charting the Future of U.S. Higher Education*. Washington: Department of Education.
- Steedle, J., Kugelmass, H., Nemeth, A. (2010). What do they measure? Comparing Three Learning Outcomes Assessment. *Change*, July/August, 33-37. Estratto da: <www.changemag.org>.
- Tam, M. (2014). Outcomes-based approach to quality assessment and curriculum improvement in higher education. *Quality Assurance in Education*, 22(2), pp. 158-168.
- Thomson, G., Douglass, A. (2009). Decoding learning gains. Measuring Outcomes and the Pivotal Role of the Major and Student Backgrounds. *Research & Occasional Paper Series*, CSHE 5(9). Estratto da: <<http://cshe.berkeley.edu/>>.
- Tremblay, K., Lalancette, D., Roseveare, D. (2012). Assessment of Higher Education Learning Outcomes Feasibility Study Report, OECD. Estratto da: <<http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf>>.
- Wende van der, M.C. (2009). European Responses to Global Competitiveness in Higher Education. In J. Douglass, J. King & I. Feller (eds.) *Globalization's Muse: Universities and Higher Education Systems in a Changing World* (pp. 317-341). Berkeley: Public Policy Press, Institute of Governmental Studies, University of California.
- Wolf, R., Zahner, D., Benjamin, R. (2015). Methodological challenges in international comparative post-secondary assessment programs: lessons learned and the road ahead. *Studies in Higher Education*, 40(3), pp. 471-481.
- Yücel, Toy B., Ok, A. (2012). Incorporating critical thinking in the pedagogical content of a teacher education programme: does it make a difference? *European Journal of Teacher Education*, 35(1), pp. 39-56
- Zahner, D., Steedle, J.T. (2014). Evaluating Performance Task Scoring Compatibility in an International Testing Program. Paper presented at the American Educational Research Association, Philadelphia. Estratto da: <http://cae.org/images/uploads/pdf/Evaluating_Performance_Task_Scoring_Comparability_in_an_International_Testing_Program.pdf>.



