

TECO all'Università: quali usi e funzioni?

Maria Lucia Giovannini • Università di Bologna – marialucia.giovannini@unibo.it

TECO at University: which uses and purposes?

The paper aims to discuss the use of TECO assessment of generic learning outcomes in Italian Universities.

A brief overview of the social, cultural, institutional and political context of higher education and educational assessment is followed by the description of an assessment model use whose specificity lies on a purpose and value oriented approach.

This model emphasizes the relevance of making explicit cultural beliefs and values underlying assessment purposes and uses in compliance with scientific criteria and assessment validity with the aim of developing a democratic debate about uses and misuses of the results.

Keywords: learning assessment; purpose of assessment; use of assessment results; validity; higher education; pilot test TECO.

Il contributo propone, mediante un modello d'uso della valutazione, un'analisi della sperimentazione TECO relativa alla valutazione degli esiti dell'apprendimento di natura generalista all'università. Dopo alcuni brevi riferimenti all'attuale contesto socio-culturale e politico-istituzionale dell'università, viene presentato tale modello in cui gli scopi della valutazione e le connesse scelte valoriali hanno un ruolo guida rispetto agli usi intenzionali dei risultati della valutazione. Sulla base di tale ottica viene discussa l'importanza di esplicitare gli orizzonti culturali e valoriali sottesi a scopi e usi, salvaguardando i criteri del rigore scientifico e della validità valutativa per consentire un democratico confronto tra i diversi utilizzatori circa gli usi appropriati e impropri dei risultati.

Parole chiave: valutazione degli apprendimenti; scopo della valutazione; uso della valutazione; validità; università; test TECO.



studi

TECO all'Università: quali usi e funzioni?

Premessa

Negli a.a. 2013-2014 e 2014-2015 hanno avuto luogo rispettivamente la fase 1 e la fase 2¹ della sperimentazione relativa alla valutazione sugli esiti effettivi dell'apprendimento di natura generalista dei laureandi italiani, svolta dall'ANVUR tramite il test TECO (TEst sulle COmpetenze generaliste). Essa è stata concepita prendendo come riferimento lo studio di fattibilità AHELO (*Assessing Higher Education Learning Outcome*) dell'OCSE e le competenze considerate sono state misurate tramite un adattamento del test CLA+ (*Collegiate Learning Assessment*), costruito e validato scientificamente dal CAE (*Council for Aid to Education*)² (Kostoris Padoa Schioppa, 2012; ANVUR, 2015).

Nello specifico, il test si compone di due moduli dei quali il primo, denominato *Performance Task* o Prova di Prestazione, intende rilevare mediante domande a risposta aperta le competenze relative alle capacità di ragionamento analitico e soluzione di problemi nonché all'efficacia e alla tecnica di scrittura, mentre il secondo, denominato *Selected Response-Questions* o Prova a Risposta Multipla, contiene 20 domande a risposta chiusa che intendono misurare le capacità di lettura critica, di criticare un'argomentazione e di ragionamento scientifico e quantitativo. Per quanto concerne le modalità di somministrazione ai "laureandi" aventi diritto di partecipare alla sperimentazione³, il test è stato compilato on-line e ha avuto una durata complessiva di 90 minuti (60 per il primo modulo e 30 per il secondo). Oltre al livello individuale sulle competenze trasversali oggetto di rilevazione, sono stati raccolti dati relativi ad alcune caratteristiche dei laureandi compresa la riuscita negli studi e alcuni dati di carattere socioeconomico familiare o personale⁴ (*ibidem*).

In relazione agli esiti della valutazione degli apprendimenti rilevati dal test TE-

- 1 Per quanto concerne la fase 2, che ha coinvolto una ventina di sedi universitarie, si ha notizia delle pianificazioni della somministrazione effettuata dagli atenei; i risultati invece non sono stati ancora diffusi. Nella fase 1 le università coinvolte sono state 12.
- 2 Il sistema CLA è nato primariamente per fornire ai docenti e ai responsabili dei programmi di studio elementi di informazione e comparazioni con opportuni *benchmarks*, utili per migliorare i percorsi formativi. Esso comprende inoltre materiali e *performance tasks* che possono essere usati in classe dagli studenti per migliorare le proprie competenze (Kostoris Padoa Schioppa, 2012, p. 51). Esula, quindi, dagli intenti originari del sistema CLA una qualsiasi forma di classificazione (*ranking*).
- 3 Potevano accedere al test studenti del primo ciclo universitario e studenti del ciclo unico a metà percorso con determinate caratteristiche relative all'anno di iscrizione e al numero di crediti acquisiti; per far riferimento a loro in TECO viene utilizzato il termine "laureandi" (Kostoris Padoa Schioppa, 2012, p. 10).
- 4 Per l'interpretazione dei dati sono inoltre state prese in considerazione altre informazioni relative a caratteristiche della sede universitaria e del corpo docente, nonché al tasso di occupazione e di crescita nell'ultimo triennio a livello regionale.

CO, il presente contributo intende porre l'attenzione sulla problematica relativa a chi e a che cosa servono le informazioni raccolte e sugli usi proposti. I potenziali destinatari e utilizzatori dei risultati sono infatti molteplici, ma a seconda del tipo di stakeholder può variare lo scopo per cui viene effettuata la valutazione, come pure la scelta dei *learning outcomes* da rendere oggetto di rilevazione. Nel contempo, la problematica dell'uso dei risultati valutativi non può essere disgiunta a nostro avviso dagli scopi e dalle funzioni assegnate alla valutazione, come pure dal processo valutativo stesso, nonché da scelte valoriali più ampie.

In relazione alla problematica delineata, viene innanzitutto fatto un sintetico riferimento agli elementi di contesto che stanno alla base della sperimentazione del test TECO e viene presentata la cornice teorica sull'uso della valutazione da noi prescelta. Vengono poi riportate le ragioni, formali e sostanziali, presentate nei documenti ufficiali relativi a TECO, e vengono esplicitati gli usi dei risultati in essi indicati. Oltre alla coerenza interna tra tali "dichiarazioni", vengono analizzati gli usi previsti alla luce dei riferimenti teorici precedentemente delineati e in relazione a una funzione formativa della valutazione in una direzione trasformativa tesa a sostenere la riuscita formativa degli studenti e l'innovazione didattica piuttosto che a classificare e a sanzionare i Corsi di studio e le università.



1. Il contesto di riferimento

1.1 *La metafora dei missili balistici e dei missili intelligenti*

Per far riferimento a un contesto sociale caratterizzato da forti trasformazioni, che coinvolgono anche le finalità assegnate all'educazione e ai processi di insegnamento/apprendimento a vari livelli, richiamiamo la metafora dei *missili balistici e intelligenti* proposta da Bauman (2012) in relazione all'ingresso nella "modernità liquida". Nel caso dei missili balistici la direzione, il percorso e il punto di atterraggio sono prestabiliti e calcolati con precisione in base alla conoscenza della posizione del bersaglio da colpire; quando però il bersaglio si sposta in modo irregolare e imprevedibile diventa necessario un missile intelligente, cioè «un missile che possa cambiare la sua direzione durante il volo in base al mutare delle circostanze, che possa localizzare immediatamente gli spostamenti del bersaglio, apprendere da essi tutto ciò che occorre e che può essere desunto dalla direzione e dalla velocità attuali del bersaglio, ed estrapolare dalle informazioni raccolte il punto in cui le loro traiettorie possono incrociarsi. Tali missili intelligenti non possono sospendere, e tanto meno porre fine, alla raccolta e alla elaborazione delle informazioni durante il percorso, perché il loro bersaglio non smette mai di muoversi e di cambiare direzione e velocità e l'operazione di seguire il luogo dell'incontro deve essere costantemente aggiornata e corretta. [...] I missili intelligenti, a differenza dei loro antichi cugini balistici, *apprendono durante il percorso*. Quindi, ciò di cui innanzitutto hanno bisogno di essere provvisti è la *capacità* di imparare, e di imparare rapidamente» (*ibidem*, p. 27).

Un processo di insegnamento/apprendimento volto alla "produzione di missili intelligenti" deve dunque puntare a promuovere negli studenti la formazione delle cornici cognitive e delle predisposizioni che consentano loro di orientarsi autonomamente in situazioni non familiari e in continuo cambiamento, aggiornando e ampliando in modo flessibile conoscenze e abilità possedute. In particolare per i sistemi di istruzione si tratta di sfide impegnative che richiedono trasformazioni anche da parte dei professionisti dell'educazione, investimenti e risorse, ma so-

prattutto la consapevolezza della rilevanza di tali trasformazioni da parte dei decisori politici. Occorre infatti ancorare tale esigenza di migliorare i sistemi di istruzione e formazione e di renderli più efficaci rispetto ai mutamenti radicali dell'attuale contesto a finalità più generali, esplicitando il modello di scuola e di università nonché di società ad essi sotteso. Il miglioramento è rispondente a fini egualitari e democratici o alle esigenze di un mercato del lavoro competitivo e della globalizzazione economica e finanziaria?

1.2 Risultati di apprendimento attesi-acquisiti e qualità della didattica

Un termine sempre più utilizzato nell'ambito formativo è sicuramente quello di competenza. Benché non esista una definizione unanimemente condivisa e manchi un accordo sulla possibilità e modalità di misurarla, a livello internazionale è ormai acquisito l'orientamento a formulare gli obiettivi dei percorsi formativi in termini di competenze attese (Luzzatto, 2011). In seguito all'intesa intergovernativa denominata "processo di Bologna"⁵ non solo è stata rielaborata "l'architettura" dei cicli formativi, ma è stato approvato a Bergen nel 2005 il *Framework for the Qualifications of the European Higher Education Area*, cui ogni singolo Paese deve ricondurre il *Quadro di riferimento nazionale dei titoli di istruzione superiore*.

Un elemento fondamentale di tale *Framework* è costituito dai "descrittori di Dublino", che precisano cinque tipologie di risultati di apprendimento (*learning outcomes*) che gli studenti universitari devono dimostrare di aver acquisito per poter ottenere il titolo corrispondente al Corso di Studio frequentato.

Per assicurare la qualità nelle istituzioni di istruzione superiore che costituisce un aspetto molto importante ai fini della reciproca fiducia tra i Paesi coinvolti nel processo, nel medesimo incontro di Bergen i Ministri dell'istruzione superiore dei Paesi aderenti all'area europea dell'istruzione superiore hanno condiviso il documento *Standards and Guidelines for Quality Assurance in the European Higher Education Area* elaborato dall'ENQA (*European Association Quality Assurance in Higher Education*). L'elemento iniziale e fondamentale precisato da tali linee guida (*European Standards and Guidelines – ESG*) è costituito dall'autovalutazione da parte delle singole istituzioni formative, cui segue il miglioramento che rappresenta l'obiettivo primario della valutazione e per il cui conseguimento devono operare i diretti responsabili dei programmi formativi; solo successivamente si inserisce in tale processo la valutazione esterna, in dialogo con i rapporti di autovalutazione. Gli elementi considerati sono essenzialmente di tipo qualitativo e non sono previsti usi a scopi classificatori (*ranking*).

La coerenza del sistema di valutazione e di assicurazione della qualità degli atenei italiani con gli ESG viene ribadita anche dalla legge 240/2010 e le procedure di accreditamento dei Corsi di studio da parte dell'Agenzia Nazionale di Valutazione (ANVUR) secondo lo schema AVA (Autovalutazione, Valutazione periodica, Accreditamento)⁶ derivano dalle linee guida europee. Tuttavia in tale legge viene delineata nel contempo la necessità di valorizzare la qualità e l'efficienza delle Università e conseguentemente di introdurre «meccanismi premiali nella distribuzione delle risorse pubbliche sulla base di criteri definiti ex ante, anche mediante

5 Cfr. www.processodibologna.it/

6 Cfr. il Decreto Legislativo 19 del 27 gennaio 2012 e ANVUR, 2013.



accreditamento periodico delle università» (art. 5, 1 comma a). In tal modo, a nostro avviso, si cerca di usare la valutazione a scopi classificatori diversamente dall'impostazione delle linee guida europee.

1.3 L'importanza crescente delle competenze chiave/trasversali/generali

Oltre all'affermarsi progressivo del termine competenza/e, sono emerse le espressioni competenze chiave, competenze trasversali, competenze generaliste o generiche. In documenti di diversa natura, così come in letteratura, si ritrovano anche espressioni quali *general academic skills, key competences, cross-curricular competencies, transversal or transferable competences, soft skills, life skills, key skills, core skills, essential skills, necessary skills, generic skills, employability skills, transferable skills*. È evidente che la diversità nel sostantivo (*skill, competence*) e nei diversi aggettivi ad esso affiancati fa emergere un problema di attribuzione/costruzione e riconoscimento dei significati di ciascuna di esse e dei diversi presupposti teorici; non c'è infatti condivisione di quali abilità/competenze considerare, come pure di quanto esse siano indipendenti da uno specifico ambito disciplinare o contesto di attivazione⁷.

L'estendersi dell'uso delle suddette espressioni rispecchia l'importanza crescente attribuita alle competenze in oggetto. In proposito, una tappa importante a livello di linee di indirizzo elaborate dall'Unione Europea è costituita dalla Raccomandazione del Parlamento e del Consiglio 2006/962/CE relativa alle competenze chiave per l'apprendimento permanente (Commissione Europea, 2006), tesa a un rinnovamento dei curricula al fine di far acquisire agli studenti le otto competenze chiave individuate⁸ e renderli in grado di apprendere lungo tutto il corso della vita.

Anche in relazione all'attuazione della strategia *Europa 2020* (Commissione Europea, 2010), viene ancora una volta sottolineata l'importanza cruciale per i sistemi di istruzione e formazione di promuovere, tra l'altro, lo sviluppo di competenze trasversali – comprese quelle connesse alla capacità di apprendere, all'uso delle tecnologie digitali, alla cittadinanza attiva e all'imprenditorialità – rispetto al raggiungimento degli obiettivi relativi a uno sviluppo intelligente, sostenibile e inclusivo, in grado di dotare tutti i cittadini delle conoscenze e capacità utili a incentivare competitività e innovatività dell'economia europea ma anche a favorire la coesione e l'inclusione sociale e l'*empowerment* personale.

Per quanto concerne il ruolo dell'università nella promozione delle competenze

- 7 Il fenomeno del *transfer* di tali abilità/competenze costituisce un aspetto problematico. Gli studi ci dicono che molte di esse, sviluppate in un contesto, difficilmente vengono attivate dal soggetto in altri contesti se il soggetto stesso non viene adeguatamente preparato a farlo. Questo fenomeno pone una seria sfida a tutti coloro che intendono sviluppare in contesti formativi, dunque in qualche modo separati dalla realtà e dalla concretezza della vita quotidiana, competenze che dovranno un domani essere utilizzate al di fuori di essi.
- 8 Le competenze chiave delineate nel quadro proposto, intese come una combinazione di conoscenze, abilità e attitudini necessarie per adattarsi in modo flessibile a un mondo in rapido mutamento e per la realizzazione personale, la cittadinanza attiva, la coesione sociale e l'occupabilità, sono le seguenti otto: 1) comunicazione nella madrelingua; 2) comunicazione nelle lingue straniere; 3) competenza matematica, scientifica e tecnologica; 4) competenza digitale; 5) imparare a imparare; 6) competenze sociali e civiche; 7) spirito di iniziativa e imprenditorialità; 8) consapevolezza ed espressione culturale.



trasversali/per la vita, un riferimento interessante è rappresentato dal progetto *Tuning*⁹, inaugurato nel 2000 e fortemente sostenuto dalla Commissione europea, che propone come obiettivo principale quello di contribuire alla realizzazione del Processo di Bologna attraverso una metodologia utile a progettare, rivedere, sviluppare e valutare curricula e corsi di studio, rendendoli comparabili e trasparenti in termini di risultati di apprendimento attesi dagli studenti. A tal fine, nell'ambito del progetto, un ruolo di vitale importanza viene assegnato non solo alle competenze specifiche delle varie aree disciplinari, ma anche alle competenze "generali/trasferibili", indispensabili in una società in continua trasformazione e in quanto tali comuni a ogni corso di studio. Tali competenze, per la cui individuazione è stato attuato un processo di consultazione a livello europeo coinvolgendo datori di lavoro/imprese, neo-laureati e accademici, sono state classificate in tre gruppi: 1) competenze *strumentali*, che comprendono abilità cognitive (ad es. capacità di analisi e sintesi), metodologiche (ad es. capacità di organizzazione e pianificazione, di presa di decisione e di risoluzione di problemi), tecnologiche (ad es. capacità di utilizzo di dispositivi tecnologici e di gestione delle informazioni) e linguistiche (ad es. capacità di comunicazione scritta e orale nella lingua madre e in una seconda lingua); 2) competenze *interpersonali*, che fanno riferimento ad abilità individuali quali le capacità di espressione delle proprie opinioni/sensazioni, di interazione sociale e collaborazione, di critica e autocritica; 3) competenze *sistemiche*, che richiedono l'acquisizione preliminare dei precedenti due tipi di competenze, riguardano abilità concernenti sistemi integrati e prevedono una combinazione tra conoscenza, sensibilità e comprensione (ad es. capacità di ricerca, di apprendimento, di adattamento a nuove situazioni, di gestione di progetti, di creatività e spirito d'iniziativa).

Anche nei Descrittori di Dublino, ai quali abbiamo fatto riferimento a proposito degli obiettivi di apprendimento, è rintracciabile la presenza di alcune competenze trasversali: autonomia di giudizio, abilità comunicative, capacità di apprendere, conoscenza e capacità di comprensione applicata.

Rispetto alla loro utilizzazione nella progettazione della didattica per *learning outcomes* nei diversi Corsi di laurea di primo e secondo livello delle università italiane, da un'indagine realizzata dal CARED dell'Università di Genova è emersa la consapevolezza da parte dei coordinatori dei Corsi di studio di dover considerare le competenze trasversali oltre agli specifici obiettivi formativi relativi alle discipline accademiche (Luzzatto, Mangano, Moscati & Pieri, 2013). Una successiva ricerca relativa all'a.a. 2014-2015, condotta dall'Università di Genova e dalla Fondazione Giovanni Agnelli, ha rilevato e messo in luce una non completa chiarezza del concetto di tale tipo di competenza da parte di tutti i coordinatori, ma allo stesso tempo la loro consapevolezza di aver bisogno di essere formati per realizzare Corsi di studio che prevedano attività organizzate in maniera tale da poter sviluppare e/o irrobustire le competenze trasversali¹⁰ (Luzzatto & Mangano, 2015)¹¹.

9 Il sito del progetto Tuning è: <http://www.unideusto.org/tuningeu/>. All'Università di Deusto è in atto uno studio di caso sull'apprendimento delle competenze trasversali (*generic competences*) nel quale ha largo spazio la funzione formativa della valutazione (cfr. Villa Sanchez & Poblete Ruiz, 2008).

10 Le competenze trasversali considerate nell'indagine sono otto: risolvere problemi, analizzare e sintetizzare informazioni, formulare giudizi in autonomia, comunicare efficacemente, apprendere in maniera continuativa, lavorare in gruppo, essere intraprendenti, saper organizzare e pianificare.

11 Tale ricerca rientra in un progetto di ampio respiro al quale hanno partecipato anche

L'esigenza che l'università sviluppi le competenze trasversali è stata d'altra parte espressa in modo generalizzato dal mondo imprenditoriale e l'attenzione dei datori di lavoro verso tali competenze viene giustificata sulla base della crescente complessità delle organizzazioni: i laureati assunti risultano carenti nelle *soft skills* – così vengono chiamate nell'ambito lavorativo le competenze trasversali –, ritenute necessarie per un positivo inserimento nelle aziende, mentre risultano preparati negli ambiti disciplinari. È quanto emerge da varie ricerche, tra le quali l'indagine condotta dalla Fondazione Agnelli sui “nuovi” laureati (Fondazione Giovanni Agnelli, 2012) e quella realizzata dal CARED dell'Università di Genova (Luzzatto, Mangano, Moscati & Pieri, 2011; Mangano, 2014).

1.4 Le rilevazioni internazionali e lo studio di fattibilità AHELO

L'importanza attribuita alle competenze “trasversali” dei giovani adulti e degli adulti emerge anche da varie indagini internazionali. L'indagine ALL (*Adult Literacy and Life Skills* – Competenze alfabetiche funzionali e abilità per la vita), realizzata anche in Italia negli anni 2003-2004, è nata all'interno della cornice del progetto DeSeCo (*Definition and Selection of Competencies*)¹² per l'individuazione delle competenze indispensabili a garantire cittadinanza attiva e occupabilità (Galina, 2003, 2006). Poco meno di un decennio più tardi è stata realizzata l'indagine internazionale PIAAC (*Programme for International Assessment of Adult Competencies*), volta a misurare il livello di possesso delle competenze o abilità chiave (*key skills*) nell'elaborazione delle informazioni considerate essenziali per la piena partecipazione di cittadini adulti al mercato del lavoro e alla vita sociale di oggi (Di Francesco, 2013).

Un'altra importante iniziativa è rappresentata dallo studio di fattibilità AHELO (*Assessment of Higher Education Learning Outcomes*), lanciato dall'OCSE nel 2008 e condotto in prospettiva comparata coinvolgendo 249 istituzioni universitarie in 17 Paesi per valutare la possibilità – in termini di fattibilità pratica e scientifica – di sviluppare misure comparative dei risultati di apprendimento degli studenti universitari in grado di consentire un confronto a livello internazionale tra contesti differenti per lingua, cultura e caratteristiche dei sistemi di istruzione superiore (Tremblay, Lalancette & Roseveare, 2012; OECD, 2013a, 2013b). Benché sia stato esplicitato fin dall'inizio l'obiettivo di fondo posto alla base dell'iniziativa, individuato nel fornire alle singole istituzioni universitarie informazioni utili a promuovere il miglioramento dei risultati di apprendimento degli studenti, lo studio

il Consorzio AlmaLaurea e Unioncamere attraverso Excelsior, che hanno indagato – rispettivamente – la percezione che i laureati hanno circa l'acquisizione o meno di competenze trasversali in ambito universitario e l'interesse che i datori di lavoro mostrano rispetto al possesso o meno di tali competenze nei laureati che prevedono di assumere (cfr. http://www.dafist.unige.it/?page_id=1898). Il Consorzio AlmaLaurea ha scelto di inserire, all'interno del questionario di fine corso per i laureandi/neolaureati, alcune domande volte a indagare le loro percezioni in merito all'acquisizione di sette competenze trasversali – risolvere problemi, analizzare e sintetizzare informazioni, formulare giudizi in autonomia, comunicare efficacemente, apprendere in maniera continua, lavorare in gruppo, essere intraprendenti – definite a partire dal più vasto insieme di competenze considerate nel progetto *Tuning* sopra menzionato.

12 Cfr. Rychen & Salganik, 2003, 2007.



AHELO è stato accolto con cautela e scetticismo in alcuni ambienti accademici innescando un vivace dibattito tra gli stakeholder circa la sua natura e le sue finalità. Come si legge nel rapporto pubblicato dall'OCSE nel 2012 (Tremblay et al., 2012), le perplessità e le preoccupazioni espresse si sono concentrate su vari aspetti quali le difficoltà implicate dal confronto tra università e contesti diversi, i rischi associati all'uso di test standardizzati, il potenziale impatto sull'autonomia delle istituzioni e sulla libertà accademica, il focus posto da AHELO sulle competenze generiche, alcune questioni di carattere metodologico e pratico. Tuttavia, l'aspetto rilevato come maggiormente problematico è indubbiamente rappresentato dal rischio di un uso improprio dei dati da parte dei governi e dei responsabili politici, in particolare dalla possibilità che i risultati delle rilevazioni effettuate vengano utilizzati – in un'ottica di mera *accountability* – per elaborare graduatorie delle università e dei corsi di studio e per basare su tali classifiche meccanismi di riallocazione delle risorse imperniati sulla logica dei premi per i migliori e delle sanzioni per i peggiori. A questo proposito viene precisato che in tal modo, piuttosto che favorire il miglioramento della qualità complessiva dei sistemi di istruzione superiore, si finirebbe per rafforzare le situazioni di eccellenza, ovvero le università con prestazioni già elevate, mentre le università poco “performanti” subirebbero dei tagli di risorse che accentuerebbero ulteriormente i loro problemi.

Benché l'OCSE abbia chiarito la propria posizione su tale questione (OECD, 2008), evidenziando la necessità di «”sganciare” i risultati delle valutazioni dalle decisioni inerenti l'uso dei finanziamenti pubblici per evitare ogni effetto perverso» (Tremblay et al., 2012, p. 59), e benché AHELO non sia mai stato concepito come uno strumento per stilare graduatorie, il rischio che i dati prodotti generassero ulteriori classifiche o fossero utilizzati nell'ambito dei sistemi di classificazione esistenti ha avuto un posto di rilievo tra le preoccupazioni espresse dagli stakeholder¹³. Nel terzo dei volumi relativi allo studio di fattibilità AHELO (OECD, 2013b), si fa riferimento alla richiesta espressa da alcuni Paesi di voler AHELO al servizio di obiettivi di politica pubblica più ampi rispetto a quelli di fornire informazioni e feedback utili al miglioramento delle università e dei sistemi di istruzione superiore. Viene pertanto evidenziata la necessità di affrontare tale problema per non compromettere la natura *low-stakes* e formativa di AHELO e viene precisato che una lezione fondamentale tratta da tale studio è l'importanza di stabilire e comunicare chiaramente lo scopo perseguito. AHELO è stato pensato per effettuare una diagnosi a livello delle singole istituzioni e per promuovere al loro interno, alla luce delle evidenze raccolte, il miglioramento della didattica e dei risultati di apprendimento degli studenti. Si ribadisce dunque che esse *non* sono state concepite come: misure della performance delle singole istituzioni universitarie; misure della performance dei sistemi di istruzione superiore dei singoli Paesi; misure della qualità dei docenti o della didattica; misure per la valutazione dei singoli studenti; strumenti per l'*accountability*; strumenti per stilare graduatorie.

Per quanto concerne i possibili sviluppi futuri dell'iniziativa, considerando le preoccupazioni espresse da alcuni stakeholder si afferma che, oltre a ribadire e a sottolineare con maggiore enfasi che AHELO è inteso come uno strumento uti-

13 In una lettera del 2008 agli alti funzionari dell'OCSE, ad esempio, il presidente dell'*American Council on Education* (ACE) ha affermato che anche se «l'intenzione dell'OCSE non è quella di sviluppare un altro sistema di classifica, è altamente probabile che i risultati saranno utilizzati per classificare le università» (Tremblay et al., 2012, p. 57).

lizzabile dalle istituzioni universitarie per sostenere il miglioramento dei risultati di apprendimento degli studenti, può essere utile esaminare la necessità di forme di garanzia e di tutela per assicurare che AHELO rimanga fedele al suo scopo e non venga utilizzato impropriamente.

2. La cornice teorica di riferimento sulla valutazione e gli interrogativi di ricerca

Già da quanto delineato finora è emersa la problematicità, oltre che la complessità, degli usi della valutazione, che costituisce appunto uno dei principali temi sviluppati dai ricercatori in concomitanza con lo sviluppo stesso dei modelli valutativi.

Data l'impostazione del presente contributo, di seguito presentiamo gli elementi fondamentali del modello¹⁴ di *uso della valutazione* degli apprendimenti che assumiamo. Gli usi dei risultati, propri del momento decisionale, non fanno parte della valutazione¹⁵, tuttavia vanno messi in relazione agli scopi intenzionali e alle specificità della valutazione stessa. Infatti, nel processo di progettazione di "impianti valutativi"¹⁶ in ambito educativo/formativo, lo scopo principale attribuito alla valutazione costituisce, a nostro avviso, il "criterio orientatore" di tutti gli altri elementi/momenti della valutazione (partono da quello e a quello ritornano), così come degli usi del processo valutativo e dei suoi risultati¹⁷. A loro volta, gli scopi della valutazione, da individuare a diversi livelli di generalità¹⁸, sono connessi con scelte di fondo più ampie¹⁹. In questo framework, pertanto, la valutazione



- 14 Il termine *modello* viene qui utilizzato per intendere la *cornice concettuale* (rete di concetti e di loro interrelazioni) che intende esplicitare il significato del concetto di *uso della valutazione*. In relazione alla domanda *come è stato elaborato tale modello?* rispondiamo che esso è l'esito di un'integrazione tra modelli precedenti ed esiti di valutazioni realizzate dalla scrivente. In relazione, invece, alla domanda *qual è la funzione/valenza del modello?* rispondiamo che in questa sede esso assume una funzione prevalentemente prescrittiva, ossia il modello intende esprimere ciò che *dovrebbe* essere l'uso della valutazione.
- 15 Il momento decisionale non fa parte della valutazione ma è successivo ad essa (Gattullo, 1986).
- 16 Utilizziamo l'espressione "impianto di valutazione" anziché "disegno" o "progetto di valutazione" perché queste due ultime espressioni vengono usate per far riferimento ai soli aspetti tecnico-procedurali (ad es. Rog, 2005). In molti contesti, nei processi valutativi l'attenzione viene posta prevalentemente sugli aspetti strumentali per cui, spesso, viene perso di vista l'insieme complessivo all'interno del quale si pone lo strumento di misurazione e valutazione; inoltre, le scelte di fondo rimangono in ombra e non vengono esplicitate.
- 17 In letteratura esiste il modello di Patton che, per la valutazione dei programmi/progetti, focalizza la sua attenzione su un altro principio organizzatore, ossia l'uso. Il principio degli "usi" della valutazione da parte dei diversi stakeholder coinvolti costituisce infatti la guida della valutazione (cfr. per esempio Patton, 2008).
- 18 Per esempio, Yarbrough, Shulha, Hopson & Caruthers (2011, p. 82) utilizzano, in ordine decrescente di generalità, questi tre termini: *purpose, goals, objectives*.
- 19 A questo proposito va ricordata, tra i docimologi italiani, la posizione di Gattullo (1978) che esprimeva la sua diffidenza nell'affrontare il discorso sull'uso dei test di profitto in una dimensione esclusivamente tecnica con queste parole: «[L'uso dei test] è comunque

non va identificata soltanto con la sua dimensione tecnica e neppure metodologica con il rischio di cadere nel riduzionismo del «mito del metodo» (Kerlinger, 1960); essa richiama necessariamente, in modo più o meno diretto, gli aspetti valoriali e gli schemi concettuali sottesi²⁰.

In base all'intenzionalità propria delle azioni formali in ambito educativo, gli scopi devono essere ben esplicitati – il che richiede una chiarezza anche dei costrutti cui fa riferimento la valutazione stessa, nonché un legame con finalità educative più generali – e devono essere tenuti presenti nei diversi momenti e fasi della valutazione (cfr. Fig. 1). Di qui la necessità non solo di un impianto valutativo internamente coerente, ma anche di un'esplicitazione e argomentazione trasparente circa i presupposti delle scelte e dell'agire da parte dei valutatori.

Quanto agli usi intenzionali – vale a dire previsti insieme all'impianto valutativo stesso –, occorre che essi svolgano una funzione pertinente, credibile e utile rispetto agli scopi e all'impianto della valutazione, pena il loro scadere in usi non validi e inappropriati (cfr. Fig. 1). Ciò non significa che vadano esclusi usi non previsti dei risultati; vengono richieste però la pertinenza del legame con gli scopi del processo valutativo di cui rappresentano il prodotto e delle evidenze aggiuntive di validità.

A proposito della *validità*, le nostre scelte teoriche la considerano un costrutto “unitario” non riconducibile solo alle caratteristiche della prova di rilevazione – e pertanto ai vari tipi di validità quali per esempio la validità di contenuto o la validità predittiva –, bensì anche all'impianto valutativo e dunque allo scopo perseguito e alle condizioni d'uso dei risultati ottenuti. Né lo strumento utilizzato per raccogliere i dati né i risultati ottenuti possono essere quindi ritenuti validi in modo assoluto: uno strumento che fornisce informazioni valide in relazione a uno scopo può non essere ugualmente valido in relazione ad altri scopi; la responsabilità della validità, inoltre, non ricade soltanto su chi progetta e costruisce lo strumento, ma anche su chi lo utilizza e ne usa i risultati, richiedendo solide argomentazioni a supporto delle decisioni. La definizione di validità in termini di uso è stata espressa da Cronbach (1988) e in particolare da Messick (1989). Per quest'ultimo la validità è un complessivo giudizio valutativo del grado in cui l'evidenza empirica e i presupposti/costrutti teorici «supportano *inferenze e azioni adeguate e appropriate* basate sui risultati valutativi» (p. 13, corsivo nell'originale). È una definizione ripresa più recentemente da Kane, che considera come condizione della validità la plausibilità delle inferenze e degli assunti di riferimento e l'argomentazione utilizzata (Kane, 2006, 2013).

Occorre precisare che nell'ambito dell'*educational measurement* il termine va-

condizionato dalle scelte di fondo sulla funzione formativa e/o selettiva della scuola» e «la consapevolezza della soggettività della valutazione non si accompagna, nella maggior parte dei casi, a quella che i giudizi formulati acquistano un valore oggettivo, perché giuridico, a livello di comunicazione sociale» (p. 246).

- 20 La nostra prospettiva educativa, per esempio, comporta un uso critico della “ragione” (Giovannini, 1995) e un impegno etico tali da poter assicurare a tutti, innanzitutto tramite percorsi intenzionalmente formativi irrinunciabili ed equi, un uguale diritto alla crescita culturale e personale, valorizzando anche, nel contempo, le differenze individuali. Pertanto, anziché rappresentarci il sistema di istruzione (superiore compresa) come entità “economica” che “eroga servizi” a favore di utenti/clienti, seguendo Dewey lo pensiamo invece come occasione privilegiata per mettere in pratica e fare propri i valori pedagogici democratici.



lità è stato usato in modo differenziato; tuttavia – come fa notare Baker (2013) – sin dal 1999²¹ gli esperti di valutazione degli apprendimenti e i costruttori di prove strutturate sono giunti ad accettare una definizione di validità unificata e complessiva in riferimento allo scopo della valutazione. In base al presupposto che i test sono sviluppati per produrre informazioni per un particolare scopo, essa viene ricondotta all'utilità di tali informazioni al fine di prendere le dovute decisioni. Nel caso in cui i risultati siano utilizzati in modo non appropriato rispetto alle funzioni e agli scopi intenzionali, si producono delle conseguenze negative e impreviste sugli studenti e sui docenti ma anche sull'intero sistema d'istruzione.

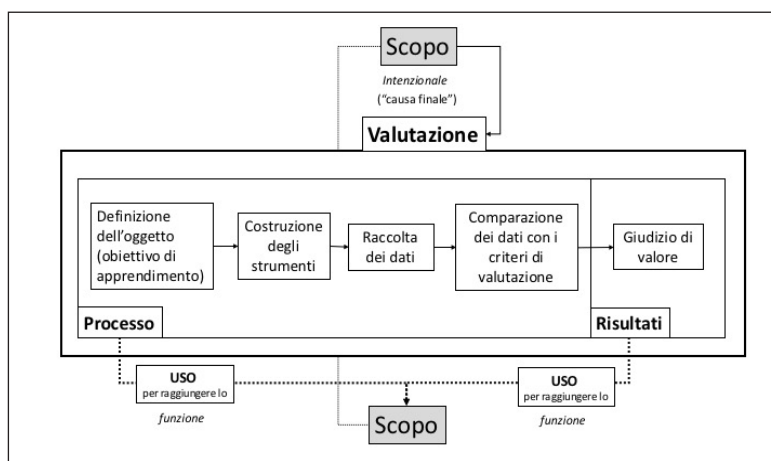


Fig. 1 – Modello d'uso della valutazione

Negli *Standards for educational and psychological testing* del 1999 (AERA, APA & NCME, 1999) viene infatti specificato che quando le applicazioni del test²² si allontanano dagli intenti originali, chi usa i risultati dovrebbe essere in grado di segnalare gli usi inappropriati e di avviare/commissionare studi sulla loro validità formale per raccogliere qualsiasi prova evidente di validità, utile a convalidare le nuove azioni proposte (Standard 1.4, 1999). Indicazioni simili si ritrovano nell'edizione più recente degli Standards (AERA, APA & NCME, 2014). Il capitolo sulla validità inizia proprio con uno standard generale (1.0) che sta alla base dei 25 standard indicati per la validità ed è ritenuto applicabile a tutti i test e a tutti i potenziali utilizzatori: «Dovrebbe essere resa esplicita una chiara articolazione di ogni interpretazione dei risultati destinata a uno specifico uso e dovrebbero essere fornite appropriate evidenze di validità a supporto di ogni interpretazione che si intende fare» (*ibidem*, p. 23). In questo documento si evidenzia chiaramente che è scorretto l'uso della frase “il test è valido”, in quanto

21 Il 1999 è l'anno in cui sono stati pubblicati gli *Standards for educational and psychological testing* (AERA, APA & NCME, 1999). A proposito del rapporto validità e uso cfr. Shepard (2013).

22 Con tale termine, riferito al contesto di istruzione, vengono intese non solo le prove a scelta multipla o a risposta aperta ma anche prove di prestazione, compreso il portfolio (AERA et al. 2014, p. 183).

nessun test consente interpretazioni che siano valide per tutti gli scopi e per tutte le situazioni. Pertanto chi costruisce il test (*test developers*) dovrebbe esporre in modo chiaro gli intenti sulla base dei quali interpretare i risultati e, dunque, usarli. Così pure occorrerebbe delimitare chiaramente la/e popolazione/i cui è destinato il test e descrivere chiaramente il/i costrutto/i che il test intende rilevare e valutare. Ulteriori indicazioni le ritroviamo anche nel capitolo *Educational testing and assessment* della pubblicazione relativa agli Standard del 2014 nel quale gli usi dei risultati delle prove di valutazione dell'apprendimento sono ricondotti agli scopi della valutazione degli apprendimenti formativa, sommativa e certificativa/selettiva, con una sottolineatura dell'importanza di considerare le eventuali conseguenze negative sul piano educativo di usi non appropriati (AERA et al., 2014). Inoltre, viene specificato tra l'altro che chi fa la supervisione in vari momenti della valutazione deve essere in grado di formare in modo efficace gli altri soggetti coinvolti per argomentare loro una spiegazione logica del rapporto tra test utilizzati, scopi perseguiti dai test e interpretazioni dei punteggi dei test per gli usi previsti (Standard 12.4).

Pure nel capitolo *Uses of tests for program evaluation, policy studies and accountability* della medesima pubblicazione viene ripetutamente sottolineata la necessità di fornire informazioni chiare su come si intendono utilizzare i risultati e sui cambiamenti che si pensa di promuovere, insieme alle avvertenze per evitare gli usi inadeguati (Standard 13.7). Per identificare e minimizzare le *conseguenze negative* viene indicato di monitorare l'impatto degli usi (Standard 13.8), mentre vanno integrati i risultati del test con informazioni provenienti da altre fonti se ciò migliora la validità dell'interpretazione complessiva (Standard 13.9).

Il tema dell'uso della valutazione viene affrontato anche nel volume *The Program Evaluation Standards* (Yarbrough et al., 2011) con l'introduzione, tra i criteri di qualità delle valutazioni di programmi/progetti, di quelli relativi all'*utilità*²³. Essi sono stati infatti introdotti allo scopo di incrementare la misura in cui i soggetti portatori di interessi in relazione al programma/progetto valutato trovano sia il processo sia i risultati della valutazione capaci di rispondere ai loro bisogni. In particolare, va sottolineato lo standard di utilità numero otto (U8): «*Preoccupazione per le conseguenze e gli effetti*»²⁴, in cui viene sottolineato che le valutazioni

23 Le altre aree di standard di qualità delle valutazioni sono le seguenti: fattibilità, rispetto, precisione, responsabilità della valutazione.

24 Gli altri sette standard relativi all'Utilità sono i seguenti: *U1 Credibilità del valutatore*. Le valutazioni devono essere effettuate da persone qualificate che stabiliscano e mantengano credibilità nel contesto della valutazione; *U2 Attenzione agli stakeholder*. Le valutazioni devono prestare attenzione a tutta la gamma di individui e gruppi coinvolti nel programma e interessati dalla sua valutazione; *U3 Negoziazione dei fini*. I fini della valutazione devono essere individuati e continuamente negoziati sulla base delle esigenze degli stakeholder; *U4 Esplicitazione dei valori*. Le valutazioni devono chiarire e specificare i valori individuali e culturali alla base di scopi, processi e giudizi; *U5 Pertinenza delle informazioni*. Le informazioni raccolte nel corso della valutazione devono essere al servizio dei bisogni identificati ed emergenti degli stakeholder; *U6 Significatività dei processi e dei prodotti*. Le valutazioni devono elaborare attività, descrizioni e giudizi in modo tale da incoraggiare i partecipanti a riscoprire, reinterpretare o rivedere le loro conoscenze e comportamenti; *U7 Tempestività e adeguatezza nelle comunicazioni e nei rapporti di valutazione*. Le valutazioni devono prestare attenzione alle esigenze di informazione continua del loro diversi interlocutori.



devono promuovere un uso responsabile e adattabile degli esiti e devono prestare attenzione alle conseguenze negative non volute e agli usi impropri²⁵.

Per quanto concerne gli usi della valutazione, in letteratura troviamo diverse tipologie e proposte di classificazione²⁶, tuttavia spesso viene fatto riferimento a un uso: a) *strumentale* (per assumere decisioni o risolvere problemi); b) *concettuale* (per modificare il modo di pensare dei destinatari); c) *per la legittimazione* (*legitimative utilization*) (per supportare una decisione politica già assunta); d) *simbolico* (per dare rilevanza simbolica al soggetto decisore dell'uso della valutazione).

Rispetto a tale tipologia, così come ad altre, preferiamo una semplificazione che risponde anche a una nostra esigenza di chiarezza concettuale. Infatti pensiamo che sia importante, nell'esplicitare il concetto di uso, tenere distinto il "processo pratico" all'interno del quale viene inserita la "valutazione" (gli elementi del suo processo o i suoi risultati) dallo *scopo* che tale processo pratico intende perseguire. Pertanto, come compare nella Figura 2, preferiamo far riferimento all'uso per *conoscere* (specifico dei ricercatori) e all'uso per *prendere decisioni*. Un altro aspetto del modello qui proposto sono i *meccanismi di mediazione*: oltre ai meccanismi di utilizzo (tra i quali consideriamo la *diffusione* degli elementi del processo e degli esiti della valutazione), possiamo farvi rientrare i meccanismi di resistenza e i meccanismi di influenza. Precisiamo che nella Figura 2 viene utilizzata l'espressione "validità d'uso" non tanto per introdurre un nuovo tipo di validità quanto piuttosto, rispetto alla problematica in oggetto, per attirare l'attenzione anche da parte degli utilizzatori sul problema della validità e della pertinenza degli usi rispetto all'impianto valutativo e agli scopi da esso indicati, nonché sulle conseguenze negative.

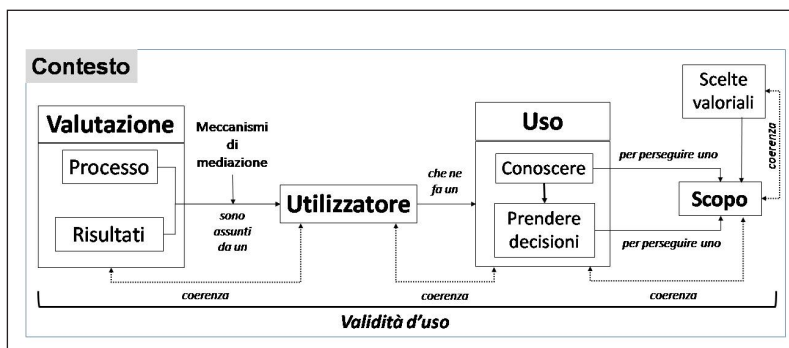


Fig. 2 – Modello d'uso della valutazione – Dettaglio del momento d'uso

25 Va sottolineato che, per rispondere alla finalità di proporsi come strumento da utilizzare in diversi ambiti culturali, gli Standard valorizzano in modo particolare gli aspetti "formali" e/o "processuali" tralasciando di sottolineare gli aspetti valoriali. Questi, infatti, vengono presi in considerazione solo come elementi "portati" dai vari stakeholder, che vanno esplicitati (Standard U3 e U4) ma che non possono far entrare in conflitto i diversi stakeholder poiché è necessario giungere a una negoziazione. In questo aspetto sembra che gli Standard seguano il modello di Patton (*utilization-focused evaluation*) (Patton, 2008).

26 Cfr. per esempio Mark & Henry (2004), Kirkhart K.E. (2000), Johnson K., Greenesid L.O., Toal S.A., King J.A., Lawrenz F., & Volkov B. (2009).

Infine, considerando qui gli usi dei risultati di una valutazione degli apprendimenti, non possiamo non tener presente che nel dibattito scientifico sulla valutazione degli apprendimenti viene sempre più sottolineato che la funzione di valutazione di bilancio (*assessment of learning*) dovrebbe essere preceduta da una valutazione a sostegno dell'apprendimento (*assessment for learning*) e che la valutazione stessa può produrre apprendimento (*assessment as learning*)²⁷.

In relazione ai nostri presupposti teorici e alle indicazioni delineate dagli Standard (in particolare quelli del 1999 e 2014 relativi alla valutazione degli apprendimenti), ci poniamo i seguenti interrogativi sulla sperimentazione TECO: quali sono le indicazioni sugli scopi di tale sperimentazione fornite esplicitamente oppure ricavabili come indizi dalle intenzionalità circa l'uso dei risultati? A chi dovrebbe soprattutto servire? C'è coerenza tra gli scopi/intenti della valutazione e gli usi dichiarati? Quanto è prevalente una funzione formativa nell'uso dei risultati relativi agli apprendimenti degli studenti rispetto a quella classificatoria? In che misura gli usi degli esiti di TECO possono essere trasformativi nel senso di riuscire a sostenere il miglioramento e lo sviluppo degli apprendimenti degli studenti e dei docenti nonché dell'organizzazione della quale essi fanno parte?

Nel prossimo paragrafo cercheremo di rispondere – seppur brevemente – a questi interrogativi facendo diretto riferimento ai documenti dell'ANVUR.



3. Le ragioni della sperimentazione TECO e le dichiarazioni sugli usi dei risultati

La sperimentazione TECO, come si è detto, si è proposta di valutare il livello delle competenze generaliste acquisite dagli studenti universitari del nostro Paese, focalizzando l'analisi sulla capacità di lettura, di analisi critica, di soluzione di problemi nuovi di tipo logico, interpretativo o scientifico-quantitativo, e di comunicazione. Come viene dichiarato nel Rapporto finale relativo al 2014 (ANVUR, 2015), si tratta «di competenze essenziali almeno quanto quelle più legate alle materie specifiche di laurea, che l'Università ha il dovere di porre al centro della sua attività formativa, anche perché necessarie per la migliore adattabilità al mercato del lavoro e alla vita presente e futura delle persone, e dunque fondamentali per accrescere l'occupabilità e l'*empowerment* personali» (pp. 14-15).

Quanto alle ragioni poste alla base della sperimentazione, sono state distinte in «*formali*» e «*sostanziali*» (ANVUR, 2015, p. 11, 17). Le prime sono connesse al sistema di accreditamento iniziale e periodico dei corsi di studio e delle sedi universitarie, in base al quale «i risultati effettivamente raggiunti dagli studenti universitari in termini di competenze, sia specialistiche che generaliste, devono essere non solo confrontati con quelli attesi, ma anche accertati, perché, in prospettiva, vanno considerati ai fini dell'accreditamento e della valutazione periodica» (*ibidem*, p. 17). Le seconde risultano invece «collegate all'interesse da parte dei principali stakeholder (le imprese, le Università, gli studenti e le loro famiglie, il contribuente italiano e la Pubblica Amministrazione) ad avere una sempre migliore qualità didattica nelle nostre Università» (*ibidem*, p. 11). In particolare, in relazione ai vari stakeholder che hanno interesse per gli esiti effettivi dell'apprendimento di natura generalista, vengono precisate le seguenti specifiche esigenze (*ibidem*, p. 18):

27 Cfr. per esempio Assessment Reform Group, 1999; Earl, 2003; OECD, 2005; Black & Wiliam, 2009; Wiliam, 2007.

- per le *imprese*, «individuare laureati di qualunque settore disciplinare, ricchi di quelle competenze trasversali (capacità di analisi critica, di decisione, di comunicazione e altre), che per le aziende sono fondamentali e che nessuno in Italia verifica e certifica»;
- per le *università* stesse, «diagnosticare meglio per accrescere la qualità dell'offerta formativa»;
- per gli *studenti*, con le *famiglie*, «potenziare il loro capitale umano quale fonte di benessere culturale e soddisfazione personale e insieme quale strumento per l'occupazione e l'occupabilità futura in un mercato del lavoro avaro»;
- per i *contribuenti* e la *Pubblica Amministrazione*, «conoscere l'efficacia in termini di risultati delle risorse che essi devolvono all'istruzione universitaria, chiedendo che all'autonomia degli Atenei si accompagnino responsabilità (*accountability*) e la valutazione, tanto più in una fase di crisi profonda».

4. Linee di analisi e proposte

È nostra intenzione, a questo punto, cercare di rileggere la valutazione TECO alla luce del modello d'uso da noi elaborato, presentato precedentemente.

Un primo dato che emerge dall'analisi del testo dell'ANVUR è una duplice sovrapposizione concettuale in relazione al concetto di *stakeholder*: da un lato, infatti, possiamo rilevare la co-presenza in esso di aspetti riconducibili sia al concetto di *utilizzatore* sia a quello di *beneficiario*; dall'altro, si evince che alla domanda “che cosa” dovrebbero utilizzare o “di che cosa” dovrebbero beneficiare gli stakeholder, il riferimento alla valutazione delle competenze, nel testo, si sovrappone a quello del loro sviluppo didattico. Infatti, quanto affermato in relazione agli studenti e alle loro famiglie, per esempio, richiama non tanto l'uso della valutazione quanto piuttosto l'utilità per gli studenti – e per le loro famiglie – dello sviluppo delle competenze trasversali promosse dall'università. In questo caso, studenti e famiglie, di fatto, sono considerati beneficiari delle attività didattiche tese allo sviluppo di competenze trasversali e non utilizzatori dei risultati della valutazione²⁸.

Un secondo dato, che emerge dalla rilettura di TECO nella prospettiva del modello d'uso da noi assunto come riferimento, è il diverso tipo di uso dei risultati della valutazione previsto per gli altri soggetti (ANVUR, università, imprese, contribuenti e pubbliche amministrazioni). L'ANVUR, le università e le imprese sono considerati utilizzatori in quanto inseriscono i risultati di TECO all'interno di processi decisionali tesi a raggiungere specifici scopi, rispettivamente, accreditare e valutare, migliorare la qualità dell'offerta formativa, individuare i laureati in possesso di competenze trasversali. Nel caso, invece, del contribuente e della Pubblica Amministrazione, viene sottolineato l'uso *conoscitivo* dei risultati di TECO la-



28 Anche nel caso in cui tale “beneficio” derivi dal fatto che i risultati siano stati utilizzati dalle Università per rivedere i curricula dei corsi di studi e che questa rivisitazione abbia portato a un miglioramento della didattica e a un incremento nello sviluppo delle competenze trasversali negli studenti, non possiamo, a rigore, fare una lettura che porti a considerare gli studenti come utilizzatori della valutazione; non sono stati loro infatti, ad aver inserito i risultati della valutazione in nuovi processi pratici (conoscenza o processi decisionali) bensì le università.

sciando sullo sfondo – in modo implicito – il loro possibile uso al fine di prendere decisioni (il decidere, per esempio, di non investire più risorse)²⁹.

A partire da queste prime considerazioni, abbiamo proceduto a una ricostruzione di un possibile quadro esaustivo degli utilizzatori e dei possibili usi e scopi ad essi connessi che plausibilmente potrebbe essere associato alla rilevazione TECO allo scopo di farne emergere potenzialità d'uso, per quanto accettabili o meno. In questa direzione, dapprima abbiamo “riletto” come utilizzatori a pieno titolo *almeno* degli *esiti* della valutazione tutti gli stakeholder richiamati dai documenti ANVUR; in secondo luogo, abbiamo ipotizzato ulteriori utilizzatori tra quelli non contemplati dai medesimi documenti. Nel primo caso, ciò ha significato considerare come utilizzatori della valutazione – e non solo come beneficiari – in primo luogo anche studenti e famiglie e in secondo luogo pure i contribuenti e le pubbliche amministrazioni. Inoltre, abbiamo aggiunto come utilizzatori di TECO in processi decisionali almeno i *docenti*³⁰ e i *giornalisti* in quanto espressione dei mezzi di comunicazione sociale.

Successivamente abbiamo individuato alcuni possibili e plausibili usi dei risultati di TECO che studenti, famiglie, giornalisti e docenti potrebbero fare. In modo analogo abbiamo proceduto in relazione ai contribuenti e alle pubbliche amministrazioni, esplicitando a quale scopo potrebbero utilizzare i risultati di TECO in un processo decisionale, oltre a quello di tipo conoscitivo. A questo proposito, risulta plausibile che la situazione sia diversificata tra gli uni e le altre: nel caso del contribuente i dati di TECO potrebbero essere plausibilmente utilizzati per decidere di non votare più – in caso di esiti negativi – a favore del decisore responsabile della situazione dell'Università e/o di non destinare più risorse (5x1000 oppure lasciti o donazioni) all'Università così come nel caso delle Pubbliche Amministrazioni.

Infine, abbiamo anche integrato uno dei possibili scopi d'uso della valutazione per le Università, ossia la diffusione di pratiche didattiche rivelatesi efficaci nel promuovere lo sviluppo delle competenze trasversali.

Sintetizziamo qui questa “rilettura”, indicando in corsivo gli utilizzatori e gli scopi da noi inseriti a integrazione di quanto previsto dai documenti ANVUR:

- ANVUR: accreditare e valutare;
- Università: migliorare la qualità dell'offerta formativa; *formare i docenti e diffondere pratiche didattiche efficaci*;
- imprese: individuare laureati in possesso di competenze trasversali;
- studenti: *scegliere dove iscriversi; autoregolare i processi di apprendimento delle competenze trasversali*;
- docenti: *autoregolare i processi di insegnamento delle competenze trasversali*;
- famiglie degli studenti: *scegliere dove iscrivere i figli; intraprendere azioni di pressione sull'Università per promuovere il miglioramento della didattica*;
- contribuente: *eleggere decisori politici; destinare risorse economiche*;

29 È interessante sottolineare come non vengano inseriti tra gli stakeholder i ricercatori, scelta che sembra rimarcare l'interpretazione eminentemente pratica della valutazione tramite il test TECO. A proposito dell'uso (conoscitivo) dei risultati di TECO da parte di ricercatori, cfr. per esempio Peracchi (2014) e Reale & Zinilli (2014).

30 Inserire i *docenti* tra gli utilizzatori potenziali di TECO significa prendere in considerazione non solo il piano didattico dei curricula dei corsi di studi in capo alla struttura universitaria, quanto le scelte didattiche effettuate a livello di singolo insegnamento in merito al processo di insegnamento/apprendimento.

- Pubblica Amministrazione: *destinare risorse economiche*;
- giornalista/mass media: *divulgare una notizia; influenzare l'opinione pubblica*.

La rilettura in questa prospettiva della rilevazione di TECO ci fornisce uno schema di riferimento per trattare in modo più articolato un aspetto rilevante di ogni modello d'uso: vale a dire le potenziali *distorsioni* legate all'uso dei risultati. Infatti, sebbene il termine valutazione sia spesso proposto come «parola d'ordine, sinonimo di prova, rigore, trasparenza, democrazia» (Damiano, 2011, p. 10), i rischi e pericoli di usi distorti sono stati ampiamente evidenziati in letteratura (per es. AERA *et al.* 1999, 2014; Daniel, 2009; Van Vught, 2009; Hazelkorn, 2011; Olds & Robertson, 2011; Rauhvargers, 2011) e sono stati più volte richiamati anche nei rapporti del progetto AHELO. L'uso distorto dei risultati richiama ed evidenzia una mancanza di *validità* nel sistema di valutazione.

Il concetto di *uso distorto (misuse)* o *inappropriato (inappropriate use)* può riferirsi principalmente a tre dimensioni del modello d'uso da noi assunto come riferimento³¹:

- la prima dimensione riguarda la mancata *coerenza interna* tra i vari elementi del modello d'uso (in questo caso si può parlare anche di *uso incoerente*);
- la seconda dimensione è relativa a un *utilizzatore*, a un *uso* o a uno *scopo* non previsti intenzionalmente (in questo caso si può parlare di *uso non intenzionale*);
- la terza dimensione tocca invece gli assunti teorici e i valori che guidano l'interpretazione dello *scopo* e, di conseguenza, dell'*uso* dei risultati.



In relazione alla prima dimensione della distorsione – la mancanza di coerenza interna tra gli elementi del modello d'uso – un esempio può essere la situazione in cui i risultati di TECO, scaturiti da un campionamento accidentale, vengano utilizzati all'interno di processi decisionali finalizzati a valutare (*scopo*) l'intera università. In questo caso è evidente la mancanza di coerenza pratico-operativa che emerge dal non rispetto del criterio della validità esterna dei risultati di TECO.

Un esempio di uso distorto della seconda dimensione (uso non intenzionale) si può rilevare nel caso in cui il *tipo d'uso* dei risultati, invece di essere quello del “prendere le decisioni”, sia quello – peraltro non preso in esame in dettaglio nel nostro modello – di tipo *simbolico* o di *legittimazione*³².

Infine, l'uso distorto del terzo tipo è quello più complesso da esemplificare in modo sintetico poiché, da un lato, tocca aspetti spesso impliciti in molti documenti e, dall'altro, è di natura *relativa* ossia emerge solo in relazione a soggetti che sostengono differenti assunti teorici e valoriali. In questo caso, infatti, due soggetti con presupposti teorico-valoriali diversi possono ritenere, l'uno, uno specifico uso come appropriato, mentre l'altro inappropriato o distorto. Questo tipo di distorsione può essere esemplificato dal caso in cui lo *scopo* del *valutare* l'università venga interpretato diversamente: da un lato, chi concepisce la valutazione come

31 Ribadiamo che stiamo utilizzando il concetto di *uso distorto* in relazione agli utilizzatori. In letteratura, infatti, troviamo autori che ne parlano anche a proposito dei valutatori. In questo caso, alcuni autori preferiscono parlare di *valutazione distorta* e non di uso distorto.

32 Va, tuttavia, sottolineato che può sussistere il caso in cui vi sia un uso non intenzionale, ma nel contempo coerente al suo interno.

uno *stilare una classifica* in una prospettiva economico-allocaiva; dall'altro, chi la concepisce come un processo con una funzione formativa in una prospettiva pedagogica. In questo caso, per esempio, le affermazioni di Hazelkorn (2011), che sostiene che «molti paesi [...] hanno usato le classifiche per guidare la profonda ristrutturazione dei loro sistemi di istruzione superiore [...] e hanno scelto di premiare i risultati eccellenti e le istituzioni di élite piuttosto che migliorare la capacità e la qualità di tutto il sistema» rivelano come, a partire da un principio di equità distributiva a cui viene associata la funzione della valutazione, sia possibile interpretare come un uso distorto quello di classificare le università.

In questa sede non è possibile sviluppare ulteriormente queste riflessioni. Quanto accennato, tuttavia, ci consente di puntualizzare alcuni aspetti di tipo metodologico che riteniamo possano favorire lo sviluppo del dibattito sulla valutazione. In questo contributo, infatti, abbiamo proposto un approccio metodologico all'analisi del fenomeno valutativo delle rilevazioni degli apprendimenti su larga scala che ha privilegiato una prospettiva contestuale (socio-culturale, istituzionale e politica) dell'uso connesso in modo argomentato agli scopi che prendono forma all'interno di una cornice teorico-valoriale. Questo approccio, tra l'altro, consente di far emergere con chiarezza, per esempio, il fatto che TECO non avendo adottato una prospettiva pedagogica (attenzione ai processi di "apprendimento/insegnamento" implicati in tutte le dimensioni del reale) abbia scelto di non considerare gli studenti come utilizzatori dei risultati per sviluppare i propri processi di auto-regolazione e neppure i docenti per fornire loro un supporto formativo³³. E questo, generalizzando, implica che le scelte valutative sono fortemente influenzate da aspetti di tipo culturale, teorici e valoriali.

Quali indicazioni si possono trarre da questo tipo di considerazioni? Riteniamo che la prima e più rilevante conseguenza sul piano metodologico sia quella di considerare, come criterio di qualità degli impianti di valutazione, l'esplicitazione chiara dei presupposti teorici e degli orizzonti culturali con cui si affronta l'intero processo valutativo sino a comprendere, in base a evidenze argomentate, i potenziali usi in funzione di specifici scopi, dal momento che sono questi ultimi due aspetti che, culturalmente condizionati, orientano le scelte del processo valutativo stesso.

Tale esplicitazione ha a nostro avviso molteplici conseguenze positive. Dapprima il far emergere l'implicito rende possibile effettuare un confronto argomentato sulle scelte sottoponendole al dibattito culturale-politico. In secondo luogo, consente di prevenire scelte che potrebbero indirizzare – anche attraverso il filtro dei documenti a livello europeo – la valutazione verso forme riduttive e non pienamente democratiche³⁴. Infine, consente di richiamare i decisori politici a una complessità del momento decisionale che, sebbene successivo al processo valutativo (Gattullo, 1986), tenga conto non solo di scelte ideologiche o di suggestioni personali ma anche di "elementi" valutativi che, per la loro condivisione all'interno della comunità scientifica, siano riconoscibili dagli altri (credibili) e, in quanto tali, anche apprezzabili per il loro valore.

33 A livello europeo e internazionale è presa ampiamente in considerazione la formazione all'insegnamento universitario; in proposito ci permettiamo di far riferimento a Giovannini (2010).

34 Molto interessante in proposito è il contributo di apertura di Domenici & Lucisano (2011) del dibattito "Valutazione, conoscenza, processi decisionali" proposto dalla rivista *Educational, Cultural and Psychological Studies*, cui ha partecipato anche la scrivente (Giovannini, 2012).

Riferimenti bibliografici

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- ANVUR (2013). *Autovalutazione, valutazione e accreditamento del sistema universitario italiano*. Documento approvato dal Consiglio Direttivo dell'ANVUR, 9 gennaio. Estratto da www.anvur.org
- ANVUR (2014). *Sperimentazione TECO. Valutazione e diagnosi sugli esiti degli apprendimenti effettivi di carattere generalista dei laureandi nelle Università di Napoli Federico II, Lecce, Messina e Cagliari*. Estratto da <http://www.anvur.org/attachments/article/248/R~.pdf>
- ANVUR (2015). *Le competenze effettive di carattere generalista dei laureati italiani 2014*. Roma: Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca.
- Assessment Reform Group (1999). *Assessment for learning. Beyond the black box*. Cambridge: ARG.
- Baker, E. L. (2013). The chimera of validity. *Teachers College Record*, 115(9). Estratto da www.tcrecord.org
- Bauman, Z. (2012). *Conversazioni sull'educazione*. Trento: Erickson.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment Evaluation and Accountability*, 21, pp. 5-31.
- Commissione Europea (2006). Raccomandazione del Parlamento europeo e del Consiglio relativa a competenze chiave per l'apprendimento permanente. Estratto da <http://eur-lex.europa.eu/legal-content/IT/TXT/?uri=celex:32006H0962>
- Commissione Europea (2010). *Europa 2020. Una strategia per una crescita intelligente, sostenibile e inclusiva*. Bruxelles, 3.3.2010, COM(2010)2020.
- Cronbach, L.J. (1988). Five perspectives on validity argument. In H. Wainer & H.I. Braun (Eds.), *Test validity* (pp. 3-17). Hillsdale, NJ: Erlbaum.
- Damiano, E. (2011). Il «senso» della valutazione. Fenomenologia sociale e opzioni epistemologiche. *Education Sciences & Society*, 2(2), 10-39.
- Daniel, J.S. (2009). *Highlights of the UNESCO Global Forum on rankings and accountability: uses and misuses*. Estratto da www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/pdf/RANKINGS/Stamenka-JohnDaniel.pdf
- Di Francesco, G. (a cura di) (2013). *Le competenze per vivere e lavorare oggi. Principali evidenze dall'indagine PIAAC*. Research paper ISFOL n. 9. Estratto da <http://www.isfol.it/pubblicazioni/research-paper/archivio-research-paper/le-competenze-per-vivere-e-lavorare-oggi>
- Domenici, G., & Lucisano, P. (2011). Valutazione, conoscenza, processi decisionali. Dibattito. *Journal of Educational, Cultural and Psychological Studies*, 3, pp. 147-167.
- Earl, L. (2003). *Assessment as learning*. Thousand Oaks, CA: Corwin.
- Fondazione Giovanni Agnelli (2012). *I nuovi laureati. La riforma del 3+2 alla prova del mercato del lavoro*. Roma-Bari: Laterza.
- Gallina, V. (2003). Descrivere e misurare nuove competenze. ALL-letteratismo e abilità per la vita, una ricerca comparativa. *Cadmo*, 1, pp. 103-122.
- Gallina, V. (2006). *Letteratismo e abilità per la vita*. Roma: Armando.
- Gattullo, M. (1978). *Voti, test, schede: ricerche sulla valutazione scolastica*. Firenze: La Nuova Italia.
- Gattullo, M. (1986). Misurare, valutare, decidere. Un modello e tre adattamenti. *La Scuola Se*, 17-18, pp. 55-58.
- Giovannini, M.L. (1995). *La valutazione: ovvero, oltre il giudizio sull'alunno*. Milano: Mondadori.



- Giovannini, M.L. (a cura di) (2010). *Insegnare all'Università. Learning to teach in Higher Education*. Bologna: CLUEB.
- Giovannini, M.L. (2012). Una cultura «critica» della valutazione: un lusso che non possiamo permetterci? *Journal of Educational, Cultural and Psychological Studies*, 6, pp. 207-217.
- Hazelkorn, E. (2011). It's time to move beyond rankings. *Chronicle of Higher Education*, 24, Washington DC. Estratto da <http://chronicle.com/blogs/worldwise/its-time-to-move-beyond-rankings-2/28830>
- Kane, M. (2013), Validation. In R.L. Brennan (Ed.), *Educational measurement* (4th ed.). Westport, CT: Praeger Publishers, pp. 17-64.
- Kane, M. (2013), Validity and fairness in the testing of individuals. In M. Chatterji (Ed.), *Validity and Test Use: An International Dialogue on Educational Assessment, Accountability and Equity*. Bingley, UK: Emerald Group Publishing, pp. 17-53.
- Kerlinger, F.N. (1960). The mythology of educational research: the methods approach. *School and Society*, 88, pp. 149-151.
- Kirkhart, K.E. (2000). Reconceptualizing evaluation use: an integrated theory of influence. *New Directions for Evaluation*, 888, pp. 5-23.
- Kostoris Padoa Schioppa F. (2012). *Ragioni, criteri, e modalità di realizzazione di un test sulle competenze effettive di carattere generalista dei laureandi italiani*. Estratto da http://www.anvur.org/attachments/article/248/criteri_carattere_generalista2.pdf
- Johnson, K., Greenesid, L.O., Toal, S.A., King, J.A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use. A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, 30(3), pp. 377-410.
- Legge 30 dicembre 2010, n. 240, "Norme in materia di organizzazione delle università, dipersonale accademico e reclutamento, nonché delega al Governo per incentivare la qualità e l'efficienza del sistema universitario".
- Luzzatto, G. (2011). La progettazione della didattica universitaria per risultati di apprendimento. In L. Galliani, C. Zaggia & A. Serbati (a cura di), *Apprendere e valutare competenze all'università. Progettazione e sperimentazione di strumenti nelle lauree magistrali* (pp. 33-44). Lecce: Pensa MultiMedia.
- Luzzatto, G., Mangano, S., Moscati, R., & Pieri, M.T. (2011). *Secondo Rapporto - Ricerca su occupazione/occupabilità dei laureati universitari*. Estratto da <http://www.cared.UniGE.it/09062011IRAPPOROTODILAVOROCARED.pdf>
- Luzzatto, G., Mangano, S., Moscati, R., & Pieri, M.T. (2013). *Rapporto sulle risposte a un questionario rivolto ai Responsabili dei Corsi di Laurea e di Laurea Magistrale*. Estratto da <http://www.cared.UniGE.it/Rapportorispostequestionariocompetenze.pdf>
- Luzzatto, G. & Mangano, S. (2015). *Le Competenze Trasversali nelle Università italiane: primi risultati dell'indagine CT3*. Relazione presentata al Workshop Internazionale *Soft skills and their role in employability. New perspectives in teaching, assessment and certification*, Bertinoro (FC), 18-19 Novembre 2015.
- Mangano, S. (2014). *Un confronto tra ricercatori e mondo del lavoro sulle competenze trasversali dei laureati*. Estratto da www.fga.it/uploads/media/S._Mangano_Le_competenze_trasversali_dei_laureati.pdf
- Mark, M.M., & Henry, G.T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), pp. 35-57.
- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (pp. 13-103). New York: American Council on Education/Macmillan.
- OECD (2005). *Formative assessment. Improving learning in Secondary Classrooms*. Paris: OECD Publishing.
- OECD (2008). *Tertiary Education for the Knowledge Society*. Paris: OECD Publishing.
- OECD (2013a). *Assessment of Higher Education Learning Outcomes. Feasibility Study Report. Volume 2 - Data analysis and national experiences*. Paris: OECD publishing.
- OECD (2013b). *Assessment of Higher Education Learning Outcomes. Feasibility Study Report. Volume 3 - Further Insights*. Paris: OECD Publishing.

- Olds, K., & Robertson, S. (2011). On being seduced by the world university rankings. *Inside Higher Education*, 5, Washington, DC. Estratto da http://www.insidehighered.com/blogs/globalhighered/on_being_seduced_by_the_world_university_rankings_2011_12#ixzz2528eObsp
- Patton, M.Q. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Peracchi, F. (2014). *TECO 2013: un primo esame dei risultati*. Università di Roma Tor Vergata. Estratto da http://www.anvur.org/attachments/article/677/Peracchi_RapportoTECO.pdf
- Rauhvargers, A. (2013). *Global university rankings and their impact*. EUA Report on Rankings 2013. Estratto da www.eua.be/Libraries/Publications_homepage_list/Global_University_Rankings_and_Their_Impact.sflb.ash
- Reale, E., & Zinilli A. (2014). I test sulle competenze generaliste dei laureandi italiani: uno strumento per valutare la qualità dell'università? *Rassegna Italiana di Valutazione*, XVIII(59), pp. 78-103.
- Rog, D.J. (2005). *Evaluability assessment: then and now*. Presented at the Joint Meeting of the Canadian Evaluation Society and the American Evaluation Association, Toronto, Ontario, Canada.
- Rychen, D.S., & Salganik, L.H. (Eds.) (2003). *Key competences for a successful life and a well-functioning society*. Gottingen: Hogrefe & Huber Publisher.
- Rychen, D.S., & Salganik, L.H. (2007). *Agire le competenze chiave. Scenari e strategie per il benessere consapevole*. Milano: FrancoAngeli.
- Shepard, L.A. (2013). Validity for what purpose? An afterword. *Teachers College Record*, 115(9).
- Tremblay, K., Lalancette, D., & Roseveare, D. (2012). *Assessment of Higher Education Learning Outcomes. Feasibility Study Report. Volume 1 - Design and implementation*. Paris: OECD Publishing.
- Van Vught, F. (Ed.) (2009). *Mapping the higher education landscape: towards a european classification of higher education*, Vol. 28. Springer Science & Business Media.
- Villa Sanchez, A., & Poblete Ruiz, M. (Eds.) (2008). *Competence-based learning. A proposal for the assessment of generic competences*. Bilbao: University of Deusto. Estratto da <http://www.tucahea.org/doc/Competence-based%20learning%20Alfa%20Project.pdf>
- William, D. (2007). *Assessment for learning. Why what and how*. London: University of London, Institute of Education.
- Yarbrough, D.B., Shulha, L.M., Hopson, R.K., & Caruthers, F.A. (Eds.) (2011). *The Program Evaluation Standards. A guide for evaluators and evaluation users* (3th ed.). Los Angeles: Sage.



