

Towards an assessment of the disciplinary skills of the degree courses L-19. The process of test construction.

Verso una valutazione delle competenze disciplinari dei Corsi di Laurea L-19. La costruzione di un test

Francesca Bracci^a, Alessandra Romano^{b,1}

^a *Università Europea di Roma*, francesca.bracci@unier.it

^b *Università degli Studi di Siena*, alessandra.romano2@unisi.it

Abstract

The article presents the first step of the research process undertaken to construct a test – composed of batteries of closed ended questions – for the assessment of the disciplinary skills of students of degree courses in Education and Training Sciences (L-19) of Italian universities that joined the Teco-D project developed by Anvur. It describes the qualitative phase conducted to define the set of *items* to be submitted to the statistical validation procedure, and which, after the adjustments and calibrations deemed necessary on the basis of the results of the analysis to be carried out, will constitute the test for the assessment of the disciplinary skills provided as part of the Teco-D project.

Keywords: mixed methods; disciplinary skills; assessment.

Abstract

Il contributo presenta il primo *step* del processo di ricerca intrapreso per costruire un test – composto di batterie di domande a risposta chiusa – di valutazione delle competenze disciplinari di studenti e studentesse di corsi di laurea in scienze dell'educazione e della formazione (L-19) di università italiane che hanno aderito al progetto Teco-D sviluppato da Anvur. È descritta la fase qualitativa condotta per definire l'insieme di *item* da sottoporre a procedura di validazione statistica e che, dopo gli aggiustamenti e le calibrazioni ritenuti necessari sulla base dei risultati delle analisi che saranno effettuate, costituirà il test di valutazione delle competenze disciplinari previsto nell'ambito del progetto Teco-D.

Parole chiave: mixed methods; competenze disciplinari; valutazione.

¹ Il paragrafo 1 è scritto da Francesca Bracci. Il paragrafo 2 e le conclusioni sono da attribuire a Alessandra Romano.

1. Il processo di ricerca

Il contributo presenta il primo *step* del processo di ricerca intrapreso per costruire un test – composto di batterie di domande a risposta chiusa – di valutazione delle competenze disciplinari di studenti e studentesse di corsi di laurea in scienze dell’educazione e della formazione (L-19) di università italiane che hanno aderito al progetto Teco-D (Test sulle Competenze) sviluppato da Anvur. È descritta la fase qualitativa condotta per definire l’insieme di *item* da sottoporre a procedura di validazione statistica e che, dopo gli aggiustamenti e le calibrazioni ritenuti necessari sulla base dei risultati delle analisi che saranno effettuate, costituirà il test di valutazione delle competenze disciplinari previsto nell’ambito del progetto Teco-D. Questo segmento di ricerca ha avuto l’obiettivo di facilitare la creazione di un test i cui contenuti siano costruiti, negoziati e condivisi dalle diverse comunità accademiche coinvolte a livello nazionale e i cui risultati possano essere sia utilizzabili per l’autovalutazione di singoli corsi di studio sia confrontabili nell’ambito dello stesso gruppo disciplinare (Galeoto, Rumiatì, Sabella & Sansoni, 2019).

La costruzione di un test di valutazione delle competenze disciplinari caratterizzanti i corsi di laurea L-19 è un processo complesso nel quale, tradizionalmente, gli aspetti misuratori rivestono un ruolo cruciale e trovano nel metodo statistico la loro fondatezza scientifica (Falocci, Gnaldi, Matteucci & Mignani, 2010). È da precisare che tale processo si articola in tre fasi che consistono nella definizione dell’oggetto di misurazione, nella predisposizione dello strumento di rilevazione e nell’analisi dei risultati emersi (Corbetta, 1999). In questo caso, l’oggetto di misurazione comprende le competenze disciplinari legate agli obiettivi formativi specifici del percorso universitario in questione e rilevabili attraverso gli effettivi risultati di apprendimento o *learning outcomes* raggiunti dagli studenti e dalle studentesse in uscita (Castoldi, 2016; Mignani & Ricci, 2005; Gnaldi & Ranalli, 2011). Queste competenze racchiudono reticoli concettuali, scenari di riferimenti, costruiti tra loro eterogenei e multiparadigmatici. Ciò ha sollecitato ricercatori e ricercatrici ad aprire spazi di discussione per condividere come orientarne il processo di traduzione empirica, individuando le componenti di significato e operativizzandole in dimensioni, indicatori e, infine, variabili.

Dato questo quadro, è stato adottato un approccio *mixed method* sequenziale esplorativo che ha utilizzato metodologie qualitative e quantitative (Creswell, 2009). Il disegno di ricerca è articolato in due fasi di cui una qualitativa e una successiva quantitativa basata sulle analisi e sui risultati emersi dalla prima. La sfida è consistita nel comprendere come usare nella seconda fase le informazioni raccolte in quella iniziale (Trincherò, 2013). Questo approccio è considerato particolarmente promettente per sviluppare strumenti che abbiano buone proprietà psicometriche in termini di validità e attendibilità, impiegando, per esempio, porzioni di testo raccolte nello *step* qualitativo per formulare gli *item*, codici per definire le variabili che li raggruppano e temi che raccolgano i codici in scale (Ivankova, Creswell & Stick, 2006; Morse & Niehaus, 2009).

In particolare, la prima fase qualitativa del progetto di ricerca Teco-D sulla valutazione delle competenze disciplinari dei corsi di laurea L-19 – diretta da Paolo Federighi, Francesca Torlone e Loretta Fabbri – si è articolata in tre azioni:

1. Analisi del contenuto e testuale delle Schede Uniche Annuali dei Corsi di Studio L-19, SUA-CdL, per l’anno accademico 2017/2018 di nove università italiane che volontariamente hanno partecipato al progetto.
2. Conduzione di interviste semistrutturate a testimoni privilegiati, quali presidenti e membri dei comitati per la didattica dei CdL-19 degli Atenei coinvolti.

3. Coordinamento di sei gruppi di lavoro costituiti da ricercatori e ricercatrici di pedagogia generale e sociale, didattica, pedagogia speciale e sperimentale che, secondo quanto emerso dall'analisi dei documenti SUA-CdL e dalle interviste, hanno formulato un insieme di item da sottoporre a validazione statistica.

Le azioni 1 e 2 hanno consentito di individuare, mappare e analizzare, da una parte, le diverse articolazioni dei profili professionali che i CdL-19 promuovono per preparare *practitioner* che sappiano far fronte alle sfide e ai problemi che attraversano gli attuali scenari educativi, dall'altra, gli obiettivi formativi finali presenti nelle SUA-CdL, identificando comunaltà e peculiarità caratterizzanti i percorsi universitari esaminati. I processi di analisi del contenuto e testuale delle SUA-CdL e delle interviste – descritti dettagliatamente nell'articolo di Francesca Torlone (2018) – hanno avuto l'obiettivo di sistematizzare i contenuti *core*, cioè gli obiettivi formativi finali espressi in termini di risultati di apprendimento attesi dagli studenti e dalle studentesse al termine del percorso intrapreso. Sono stati identificati sei contenuti *core*:

1. costrutti e teorie per interpretare gli eventi educativi e formativi e sviluppare l'identità professionale;
2. metodologia della ricerca e analisi della domanda formativa nei contesti sociali e organizzativi;
3. modelli di progettazione nei contesti sociali e organizzativi;
4. dinamiche relazionali e situazionali nei contesti educativi e formativi;
5. metodi e tecniche per lo sviluppo e la facilitazione dei processi di apprendimento;
6. *management* delle organizzazioni educative e formative.

Le azioni 1 e 2 sono state condotte dai gruppi di ricerca delle Università degli Studi di Firenze e Siena.

L'azione 3 ha previsto la costituzione di sei gruppi di lavoro formati da ricercatori e ricercatrici di pedagogia generale e sociale, didattica, pedagogia speciale e sperimentale appartenenti alle nove università che hanno aderito al progetto. Il primo compito assegnato loro è stato dettagliare un singolo obiettivo formativo finale, o contenuto *core*, in conoscenze e competenze, declinandole in obiettivi formativi specifici secondo i cinque Descrittori di Dublino. La seconda richiesta è stata quella di precisare per ogni obiettivo formativo specifico i concetti e i problemi considerati fondanti e da cui costruire gli item per la prova di valutazione. La Figura 1 mostra, a scopo esemplificativo, la griglia realizzata da un gruppo di lavoro per il secondo contenuto *core*.

Descrittori di Dublino	Learning outcome	Concetti e problemi <i>core</i>
Conoscenza e capacità di comprensione	Conoscere le metodologie e le tecniche della ricerca sociale in ambito educativo, le metodologie per l'analisi dei bisogni e della domanda di formazione nelle diverse età della vita e nella molteplicità dei contesti sociali e organizzativi.	Concetto di metodologia e metodo; concetto di progetto, disegno e programma di ricerca; strumenti di ricerca qualitativa e quantitativa; strumenti per la rilevazione e l'analisi dei bisogni.
Conoscenza e capacità di comprensione applicata	Saper svolgere ricerche finalizzate all'intervento, anche attraverso l'uso di tecnologie informatiche e <i>database</i> nazionali e internazionali.	Uso dei <i>database</i> , risorse <i>online</i> e <i>open-access</i> ; <i>software</i> e strumenti di calcolo per l'elaborazione dei dati.

Autonomia di giudizio	Saper interpretare e formulare la domanda di formazione. Saper valutare il nesso tra gli obiettivi e i risultati della ricerca. Saper valutare gli strumenti funzionali all'analisi dei bisogni.	Strumenti e metodi di valutazione della coerenza dei risultati attesi con gli obiettivi di ricerca; procedure di analisi della domanda di formazione a livello individuale e organizzativo; scegliere gli strumenti di analisi dei bisogni e della domanda formativa più adatti.
Abilità comunicative	Saper comunicare, argomentare e presentare i risultati della ricerca rispetto ai committenti e ai diversi <i>target</i> .	Strategie di negoziazione con i committenti sulle necessità e le condizioni di realizzazione della ricerca (obiettivi, tempi, costi, qualità); le procedure di stesura e presentazione dei <i>report</i> di ricerca; modalità e strategie per comunicare i risultati della ricerca ai committenti.
Capacità di apprendere	Saper individuare i risultati inattesi della ricerca e i suoi possibili sviluppi sul piano metodologico e dell'impatto. Saper utilizzare metodologie non previste per far fronte a problemi/risultati inattesi.	Strumenti di monitoraggio dei processi di ricerca; criteri e procedure per tradurre l'esperienza di ricerca in casi; individuazione di possibili sviluppi della ricerca.

Figura 1. Metodologia della ricerca e analisi della domanda formativa nei contesti sociali e organizzativi.

A ciascun gruppo, infine, è stato chiesto di formulare, per ogni obiettivo formativo finale, almeno quindici item a risposta multipla per misurare le competenze e le conoscenze disciplinari legate al singolo contenuto *core*. Sono state condivise linee guida metodologiche e indicazioni tecniche per elaborare batterie di domande con quattro alternative di risposta, di cui una corretta e tre distrattori². La scelta di utilizzare questa tipologia di domande è stata indicata da Anvur.

In linea con la letteratura empirica sull'argomento che converge nel riconoscere la natura processuale, situata e complessa delle competenze disciplinari che i CdL-19 consentono di acquisire, sono state sviluppate domande che le rilevassero sia come esperienze di apprendimento sia nelle loro dimensioni prestazionali (Bertagna, 2004; Castoldi, 2016; Pellerey, 2004; Zaggia, 2008). In altre parole, l'attenzione è stata rivolta a costruire batterie di item per valutare sia cosa e come gli studenti e le studentesse in uscita dal CdL-19 hanno appreso sia il grado di padronanza raggiunto nell'affrontare i problemi, i compiti e le sfide caratterizzanti le professioni che si muovono in ambito educativo e formativo.

Sono stati realizzati cento item che, una volta sottoposti a procedura di validazione statistica, costituiranno il test di valutazione delle competenze disciplinari previsto nell'ambito del progetto Teco-D. Il test validato, come suggerito da Anvur, avrà una durata di novanta minuti. La procedura di pre-test sarà condotta al termine del secondo semestre dell'anno accademico 2018/2019 presso le università che hanno aderito al progetto e coinvolgerà gli studenti e le studentesse frequentanti il terzo anno dei CdL-19.

L'insieme di item saranno testati attraverso metodologie legate alla Classical Test Theory

² Nella formulazione di item a scelta multipla è necessario analizzare non solo le risposte corrette, ma le caratteristiche delle opzioni di risposta errate, i *distrattori*. Gli item a scelta multipla contengono alternative di risposta su scala nominale, in quanto le diverse opzioni non sono ordinabili ma sono solo confrontabili tra di loro in termini di diversità. Non è possibile dire che un distrattore sia più corretto di un altro e, allo stesso tempo, la risposta corretta deve possedere le caratteristiche di unicità e univocità.

– che permette un’analisi descrittiva dei risultati – e all’Item Response Theory – che offre un approfondimento sulle loro caratteristiche avvalorandone le proprietà psicometriche (Haladyna, 2004; Haladyn & Rodriguez, 2013; Wilson, 2004). L’analisi dei risultati del pre-test consentirà di modificare le domande e giungere alla definizione di una prova con elevato livello di attendibilità misuratoria. Le indicazioni emerse dallo studio dei dati del pre-test rivestiranno un ruolo cruciale nel ripensare e revisionare, laddove necessario, sia i testi delle domande sia i singoli item sia i distrattori.

2. Il *background* conoscitivo: come valutare le competenze nell’*Higher Education*

La sperimentazione condotta tra il 2012 e il 2013 di Teco, su dodici atenei ha costituito il primo tentativo, in Italia, di valutare il livello delle competenze generaliste e trasversali acquisite da studenti e studentesse in uscita da corsi di studi universitari. Lo scopo di Teco è rilevare le competenze effettive dei laureandi italiani, in particolare le loro capacità di individuare le incoerenze logiche in un ragionamento, analizzare e risolvere problemi mai affrontati in precedenza e comunicare efficacemente il proprio pensiero per iscritto (Anvur, 2015).

La progettazione del percorso di sperimentazione è stata effettuata prendendo come riferimento lo studio di fattibilità AHELO, Assessment of Higher Education Learning Outcomes, sviluppato dall’OECD (Organization for Economic Co-operation and Development) nel 2008 (Tremblay, Lalancette & Roseveare, 2012). AHELO ha coinvolto 249 istituzioni universitarie in diciassette Paesi per valutare la possibilità – in termini di fattibilità pratica e scientifica – di sviluppare misure comparative dei risultati di apprendimento di studenti e studentesse universitari in grado di consentire un confronto a livello internazionale tra contesti differenti per lingua, cultura e caratteristiche dei sistemi di istruzione superiore (OECD, 2013a; 2013b; Tremblay, Lalancette & Roseveare, 2012). AHELO si collocava all’interno di un approccio di tipo *outcome-based*, focalizzato più sui risultati in termini di conoscenze e competenze che gli studenti erano chiamati ad avere al termine del loro percorso di studi, che sulla tipologia di *input* e nozioni disciplinari loro forniti (Tremblay, Lalancette & Roseveare, 2012).

Le competenze considerate sono state misurate tramite un adattamento del test CLA+ (Collegiate Learning Assessment), costruito e validato dal CAE (Council for Aid to Education) (Anvur, 2015; Kostoris Padoa Schioppa, 2012). Potevano accedere al test studenti del primo ciclo universitario, al netto degli iscritti ai Corsi di Studio nelle professioni sanitarie, che nel ciclo triennale avevano superato tutti i crediti formativi (CFU), di base e caratterizzanti e nel ciclo unico almeno 120 CFU di base e caratterizzanti; per far riferimento a loro in Teco è utilizzato il termine *laureandi* (Kostoris Padoa Schioppa, 2012).

Le competenze trasversali misurate nel test Teco – cioè la capacità di ragionare in modo critico per la soluzione di un problema o per prendere una decisione, la capacità di rappresentare e trasmettere un determinato fatto e di apprendere conoscenze nuove legate ad ambiti non necessariamente connessi con le particolarità della disciplina scientifica che è oggetto del proprio percorso formativo – sono quelle capaci di garantire flessibilità e adattamento degli individui ai cambiamenti professionali e personali. Tali competenze sono considerate necessarie per il successo negli studi universitari, e potenziali predittori di successo nel mondo del lavoro e in altri aspetti della vita extrauniversitaria (Giovannini,

2016). Per rispondere bene al test non erano necessarie conoscenze specialistiche in alcun ramo: non esistono risposte giuste o sbagliate, ma solo bene o male argomentate, coerenti o incoerenti, solide o deboli sul piano logico o empirico, descritte con maggiore o minore efficacia e appropriatezza di linguaggio (Anvur, 2015).

Il test si componeva di due moduli: il primo, denominato Performance Task o Prova di Prestazione (PT), intendeva rilevare mediante domande a risposta aperta le competenze relative a capacità di ragionamento analitico e *problem-solving*, efficacia e tecnica di scrittura. Nel PT venivano presentati in un documento centrale uno scenario o un evento o una circostanza di natura realistica, che identificava un tema, corredato da evidenze empiriche aggiuntive, talora incoerenti, spesso con diverso grado di robustezza. I rispondenti erano invitati a rispondere allo scenario stimolo, suggerendo una soluzione o raccomandando l'intervento più adeguato o decidendo tra diverse opzioni che presentassero aspetti più o meno desiderabili, in base alle informazioni fornite.

Le domande in PT erano dirette a valutare tre dimensioni:

1. la capacità di ragionamento analitico e di soluzione di problemi (Analysis and Problem Solving, APS);
2. l'efficacia di scrittura (Writing Effectiveness, WE);
3. la tecnica di scrittura (Writing Mechanics, WM).

Ciascuno dei tre ambiti riceveva un punteggio da 0 a 6: il minimo punteggio assegnabile nel modulo PT poteva essere quindi 0, il massimo 18.

Il secondo modulo, definito Selected Response Questions (SRQ, "domande a risposta chiusa"), conteneva 20 domande a risposta chiusa che misuravano le capacità di lettura critica, di argomentazione critica e ragionamento scientifico e quantitativo. Per rispondere alle 20 domande, studenti e studentesse dovevano scegliere la risposta corretta, scartando tre *distractor*, sulla base delle informazioni riportate o deducibili dalla documentazione fornita. Le domande in SRQ erano dirette a valutare tre dimensioni:

1. la capacità di lettura critica (Critical Reading, CRE);
2. la capacità di criticare un'argomentazione (Critique an Argument, CA);
3. la capacità di ragionamento scientifico e quantitativo (Scientific and Quantitative Reasoning, SQR).

Ogni domanda riceveva un punteggio: 0 se la risposta era errata o mancante, 1 se la risposta era esatta. Il massimo punteggio assegnabile in SRQ era quindi 20. Il test è stato compilato online e ha avuto una durata complessiva di 90 minuti (60 per il primo modulo e 30 per il secondo). Sono stati raccolti dati sul livello delle competenze generali e trasversali oggetto di rilevazione, su alcune caratteristiche dei laureandi compresa la riuscita negli studi e alcuni dati di carattere socioeconomico familiare o personale (Anvur, 2015).

L'attribuzione dei punteggi alla risposta aperta del Teco ha richiesto un'adeguata preparazione dal punto di vista metodologico e una specifica formazione dei docenti che avrebbero svolto questo compito. L'attribuzione dei punteggi alle 20 risposte chiuse è stata invece effettuata direttamente dal CAE, che ha ricevuto tali risposte in forma totalmente anonima.

Le procedure di correzione e valutazione (*scoring*) dei Performance Task hanno richiesto i seguenti passaggi:

- l'Anvur ha proceduto a individuare due *country lead scorer*, che il CAE ha addestrato con formazione di due giorni online e due giorni de visu;
- ciascun Ateneo partecipante ha individuato un suo *lead scorer*;

- i 12 *lead scorer* sono stati addestrati dal CAE a Roma e hanno a loro volta addestrato tutti i valutatori (*scorer*) selezionati dagli atenei partecipanti (circa 110 in totale). La formazione alla correzione del modulo PT, svolta in Italia, è stata accompagnata da griglie di analisi dei contenuti a supporto della valutazione, specificamente redatte con l'ausilio di un Manuale del CAE e di un'ulteriore Guida al Valutatore;
- la valutazione delle risposte aperte dei test è stata effettuata dai valutatori italiani con un metodo *double-blind*. Il punteggio finale nel modulo PT (minimo 0, massimo 18) è stato ottenuto combinando le valutazioni separate fornite da due valutatori indipendenti sui tre diversi ambiti, dopo averle avvalorate e corrette. Il 20% delle domande corrette da ogni SC italiano è stato giudicato anche da un altro SC, e ciascuno di essi ha ricevuto le domande da valutare in modo anonimo e stocastico (non conoscendo il campione di studenti le cui risposte fossero oggetto di valutazione). Per ogni risposta, se lo *scoring* dei due SC coinvolti non era coerente, ovvero se la differenza fra i punteggi assegnati dai due valutatori era maggiore di due per ciascuno dei tre ambiti di valutazione, veniva richiesto il controllo di un terzo valutatore, selezionato da Invalsi (Anvur, 2015).

Il risultato complessivo Teco di ogni studente, da distinguere rispetto ai punteggi finora indicati, è stato calcolato come media aritmetica di opportune trasformazioni dei punteggi, atti a modificare gli *score* (rispettivamente negli intervalli 1-6 e 0-20) in *scale* omogenei. Al termine della prima sperimentazione, il Test è stato sottoposto ad Item Analysis dal CAE (Anvur, 2015), diretta a verificare se nel test vi fossero item troppo semplici, complessi o ambigui, e se gli item del test fossero in grado di dar conto delle differenze conoscitive che caratterizzano studenti diversamente competenti. L'Item Analysis ha confermato la bontà del test. Le analisi dei dati raccolti con le prime sperimentazioni hanno evidenziato l'esistenza di una relazione sistematica e statisticamente significativa tra i punteggi nel test e una serie di caratteristiche osservate degli studenti partecipanti quali l'età, il genere, l'anzianità dal diploma, il voto di diploma, la media dei voti nel corso di laurea, la lingua parlata a casa e la cittadinanza.

Teco, tuttavia, non è il primo progetto di costruzione e sperimentazione di prove di apprendimento. I test sulle competenze generaliste e trasversali erano già regolarmente utilizzati in altre fasce della popolazione, come messo in luce dall'OECD con i test PISA (Programme for International Student Assessment) sui quindicenni (Falocci et al., 2010), relativi alla lettura e alla comprensione di semplici problemi quantitativi, e con quelli PIAAC (Programme for the International Assessment of Adult Competencies) sulla popolazione adulta, dove l'Italia è risultato all'ultimo posto fra i quasi 30 Paesi testati nel grado di *literacy* e penultima nella *numeracy* (Falocci et al., 2010).

Nell'ambito europeo, il progetto Socrates "Diffusion and improvement of mathematical knowledge in Europe" (in breve "Mathematics in Europe"), a cui hanno partecipato il Dipartimento di Matematica dell'Università di Bologna (come capofila) e le Università di Bochum, Cipro, Durham, Parigi 7, ha previsto la costruzione e validazione di un questionario a scelta multipla per valutare le competenze che in matematica avevano gli studenti immatricolati al primo anno dei corsi di studio delle università coinvolte. Sono state sottoposte al test 3441 matricole di corsi di studio i cui curricula prevedevano almeno un corso di matematica obbligatorio (Achilles, 2005).

Lo scopo del test era valutare i livelli di conoscenza matematica degli studenti dei paesi partner e confrontarli per:

1. stimolare la discussione sui programmi di matematica della scuola secondaria superiore in ciascuno dei paesi partecipanti al progetto;
2. dare suggerimenti per l'insegnamento della matematica nel primo anno di università.

Per disegnare un test che prendesse in esame le conoscenze matematiche comuni nei diversi paesi partner, sono stati selezionati i contenuti di alcuni item dal *Syllabus* dell'Unione Matematica Italiana e alcuni item di un test orientativo svolto per gli studenti di ingegneria di Durham. Il primo test era formato da 14 domande focalizzate su sette ambiti fondamentali con due domande per argomento.

Le domande del test sono state formulate secondo questi criteri:

- misurare la capacità di comprensione e le abilità di calcolo;
- cercare di ridurre la possibilità di risposte casuali introducendo l'opzione "nessuna delle risposte suddette è corretta";
- rilevare distorsioni sistematiche nelle risposte fornite;
- fornire alternative di risposta multiple di cui indicare la corretta;
- affrontare i sette argomenti differenti già nelle prime sette domande.

Pur avendo interessato campioni di studenti al primo anno del loro percorso di studio, il progetto "Mathematics in Europe" ha avuto il merito di suggerire proposte migliorative emergenti per l'offerta formativa alla luce della costruzione e formulazione di un primo test di valutazione per competenze disciplinari, quali quelle logico-matematiche.

Nel panorama nazionale, le esperienze reali di validazione preliminare delle prove di apprendimento con esclusione di quelle implementate in un contesto internazionale (PISA, TIMSS - Trends in International Mathematics and Science Study, PIRLS - Progress in International Reading Literacy Study, etc.), sono in fase di decollo (Falocci, et al., 2010). Collocandosi nello scenario degli studi e delle esperienze per la preparazione e la calibrazione delle prove di valutazione delle competenze disciplinari, il progetto di ricerca che Teco-D Pedagogia sta maturando costituisce di fatto un punto di partenza, nonché un protocollo di riferimento essenziale per la procedura di costruzione di una batteria di item, di validazione di un test e di garanzia di risultati della valutazione scientificamente attendibili.

3. Conclusioni

La sfida di costruire un quadro di riferimento per il sapere che è obiettivo di una classe di laurea (L-19), e di avere strumenti per una valutazione dei risultati di apprendimento risponde agli interessi degli studenti e dei docenti, che sono chiamati a orientare il proprio lavoro in termini di competenze acquisite da parte dei *learner*; agli interessi dei Presidenti dei corsi di laurea, al fine di progettare il percorso di studi e modificarlo opportunamente per migliorare i risultati. Può essere utile, infine, per l'*accountability* e l'accreditamento dei corsi di laurea e degli atenei. La portata innovativa di questo processo si situa, da una parte, nello sforzo di tematizzare la complessità del percorso di valutazione, costantemente esposto all'esercizio di una flessibilità tra rigore e condivisione sociale nella scelta degli oggetti di valutazione e della prospettiva metodologica da adottare per analizzarli. Dall'altra parte, si riconosce anche nella strategia di coinvolgere i diversi attori organizzativi e gli *stakeholder* per sostenere i processi di cambiamento e riorganizzazione dell'offerta formativa a partire dai risultati prodotti.

In questo senso, il progetto di coinvolgere gruppi di ricerca di esperti provenienti dalle diverse università per costruire prove per valutare le competenze in uscita dai singoli corsi di studio contribuisce a costruire un sistema di valutazione più ampio. Un sistema che sia in grado di tarare prove valide e affidabili, complete di ancoraggi per i confronti diacronici, di utilizzarle su campioni statistici, rinunciando all'uso dei test come uno strumento di controllo amministrativo, e restituendo le prove e i loro standard alle università per la loro autovalutazione. Il percorso di valutazione che si prospetta, quindi, si caratterizza per alcune opzioni e sensibilità, relative alla particolarità dell'oggetto da valutare e all'ancoraggio a dimensioni di contestualità, che di fatto orientano la valutazione verso una pluralità di obiettivi. Si è, difatti, interessati a ottenere dati ed elementi per valutare gli esiti e l'impatto del percorso di ricerca intrapreso, che vede coinvolti soggetti diversi in processi organizzativi e istituzionali complessi (e in parte inediti), in scambi e transazioni rivolti alla costruzione di accordi significativi attorno al problema di come configurare e mettere in atto un'offerta formativa efficace (Kaneklin, Scaratti & Bruno, 2006). Ciò comporta un'attenzione privilegiata all'utilizzo e all'organizzazione dei processi valutativi, oltre che alla tipologia di prove da costruire e da validare, assumendo le questioni inerenti alla valutazione come una possibile leva di trasformazione dei microsistemi organizzativi dei corsi di laurea universitari.

Bibliografia

- Achilles, R. (2005). La preparazione matematica delle matricole nelle università europee: risultati di un test. In M. Manaresi (ed.), *Matematica e cultura in Europa* (pp. 5-22). Berlino: Springer.
- Anvur. Agenzia Nazionale di Valutazione del sistema Universitario e della Ricerca. (2015). Rapporto TECO 2014. *Le competenze effettive di carattere generalista dei laureati italiani*. <http://www.anvur.it/wp-content/uploads/2014/12/Rapporto%20TECO%202014.pdf> (ver. 15.12.2018).
- Bertagna, G. (2004). *Valutare tutti. Valutare ciascuno. Una prospettiva pedagogica*. Brescia: La Scuola.
- Castoldi, M. (2016). *Valutare e certificare le competenze*. Roma: Carocci.
- Corbetta, P. (1999). *Metodologia e tecniche della ricerca sociale*. Bologna: il Mulino.
- Creswell, J.W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Falocci, N., Gnaldi, M., Matteucci, M., & Mignani, S. (2010). *La Validazione Statistica di test standardizzati di profitto: principali aspetti di metodo e due casi di studio sulla valutazione degli apprendimenti nella scuola primaria*. Invalsi, Working Paper n. 9.
- Galeoto, G., Rumiat, R., Sabella, M., & Sansoni, J. (2019) The use of a dedicated platform to evaluate health-professions university courses. In T. Di Mascio et al. (eds.), *Methodologies and intelligent systems for technology enhanced learning, 8th International Conference. MIS4TEL 2018. Advances in intelligent systems and computing, 804*. Berlin: Springer.
- Giovannini, M.L. (2016). TECO at University: which uses and purposes? *Italian Journal of Educational Research*, 16, 37–57.

- Gnaldi, M., & Ranalli, M.G. (2011). Proprietà degli indicatori di qualità della ricerca scientifica. In L. Fabbris (ed.), *Criteri e indicatori per misurare l'efficacia delle attività universitarie* (pp. 279-298). Padova: Cleup.
- Haladyna, T.M. (2004). *Developing and validating multiple-choice test items*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Haladyna, T.M., & Rodriguez, M.C. (2013). *Developing and validating test items*. New York, NY: Routledge.
- Ivankova, N.V., Creswell J.W., & Stick S.L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field methods*, 18(1), 3–20.
- Kaneklin, C., Scaratti, G., & Bruno, A. (2006). *La formazione universitaria. Pratiche possibili*. Roma: Carocci.
- Kostoris Padoa Schioppa, F. (2012). *Ragioni, criteri, e modalità di realizzazione di un test sulle competenze effettive di carattere generalista dei laureandi italiani*. http://www.anvur.org/attachments/article/248/criteri_carattere_generalista2.pdf (ver. 15.12.2018).
- Mignani, S., & Ricci, R. (2005). Metodi statistici per la misurazione delle competenze: analisi della preparazione matematica delle matricole. In M. Manaresi (ed.), *Matematica e cultura in Europa* (pp. 23-37). Berlino: Springer.
- Morse, J., & Niehaus, L. (2009). *Mixed method design: Principles and procedures*. Walnut Creek, CA: Left Coast Press.
- OECD. Organization for Economic Co-operation and Development (2005). *Formative assessment. Improving learning in Secondary Classrooms*. Paris: OECD Publishing.
- OECD. Organization for Economic Co-operation and Development (2008). *Tertiary Education for the Knowledge Society*. Paris: OECD Publishing.
- OECD. Organization for Economic Co-operation and Development (2013a). *Assessment of Higher Education Learning Outcomes. Feasibility Study Report. Volume 2, Data analysis and national experiences*. Paris: OECD publishing.
- OECD. Organization for Economic Co-operation and Development (2013b). *Assessment of Higher Education Learning Outcomes. Feasibility Study Report. Volume 3, Further Insights*. Paris: OECD Publishing.
- Pellerey, M. (2004). *Le competenze individuali e il Portfolio*. Firenze: La Nuova Italia.
- Torlone, F. (2018). Metodi e strumenti per la definizione dei contenuti core del Corso di laurea L-19. *Form@re - Open Journal per la Formazione in Rete*, 18(3).
- Tremblay, K., Lalancette, D., & Roseveare, D. (2012). *Assessment of Higher Education Learning Outcomes. Feasibility Study Report. Volume 1, Design and implementation*. Paris: OECD Publishing.
- Trincherò, R. (2013). La ricerca e la sua valutazione. Istanze di qualità per la ricerca educativa. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, 3(6), 75–96.
- Wilson, M. (2004). *Constructing measures: An item response modeling approach*. New York, NY: Routledge.

Wright Mills, C. (1959). *L'immaginazione sociologica*. Milano: il Saggiatore.

Zaggia, C. (2008). *L'università delle competenze. Analisi di caso e ricerca applicata sulla progettazione dei corsi di laurea nel Processo di Bologna*. Milano: FrancoAngeli.