

Il dibattito sugli *standard tests* nel mondo educativo statunitense

Un quadro di riferimento

Andrea Mariuzzo

A PARTIRE DALLA STORIA E DALL'EVOLUZIONE DEGLI STANDARDIZED TESTS STATUNITENSI, SI PROPONE UNA RIFLESSIONE SUL CASO ITALIANO, CHE HA VISTO PROPRIO IN QUESTI ANNI LA PROPOSTA DI INTRODUZIONE DI TEST STANDARD PER LA MISURAZIONE DEGLI APPRENDIMENTI.

Ogni anno gli *admission boards* primaverili dei college statunitensi completano le procedure di selezione per i nuovi iscritti, e si ravviva il dibattito sui criteri per selezionare i giovani in uscita dalle *high schools* destinati a ricevere le proposte di iscrizione ai corsi desiderati e, soprattutto, la possibilità di usufruire di agevolazioni economiche e di coperture delle spese, elemento ormai decisivo di fronte alla crescita dei costi per studente. Negli ultimi anni l'attenzione si è concentrata sull'attendibilità dei test a somministrazione nazionale richiesti per offrire una misura in termini quantitativi del "potenziale accademico" dei giovani, ovvero la loro capacità di dominare il ragionamento, il calcolo e la proprietà lessicale e linguistica in contesti nuovi, non necessariamente irrigiditi dai programmi disciplinari tradizionali. A detta dei fautori di tale pratica valutativa, l'affiancamento alla difficoltosa comparazione dei percorsi scolastici e delle attività di formazione integrative, spesso assai disomogenei, di uno strumento uniforme di misura delle attitudini generalmente richieste nei percorsi di formazione superiore, di approfondimento e di specializzazione ha reso le procedure di accesso all'istruzione terziaria maggiormente comprensibili al pubblico dei non addetti ai lavori, e quindi ai loro occhi più legittime e meno arbitrarie.

D'altro canto, una crescente schiera di voci critiche si è concentrata su alcuni aspetti che effettivamente possono mettere in discussione la validità dei criteri di base delle prove di valutazione. Il formato a risposta chiusa o al limite "semi-aperta" dei quesiti rende le prove facilmente *coachable*, ovvero affrontabili con profitto in seguito a una preparazione esclusivamente mirata alle consegne, e non alla creazione di basi culturali effettivamente solide e complete. Di conseguenza, i ragazzi provenienti da am-

bienti familiari più accorti e consapevoli (o fortunati da trovare a scuola un mentore abbastanza competente in materia e capace di andare oltre i suoi doveri) possono svolgere percorsi mirati e raccoglierne i frutti in sede di verifica. Inoltre, la tendenza nella mentalità degli estensori dei test a considerare "standard" alcune abitudini di vita e di cultura e a rivolgersi ai propri destinatari dando per scontate (e quasi connaturate al buon rendimento scolastico) esperienze e contesti "di buona famiglia", finiscono per favorire indirettamente i candidati più avvezzi a simili comportamenti, ovvero quelli provenienti dai gruppi etnici maggioritari e dai ceti medi benestanti e, in termini più sfumati, appartenenti al genere maschile¹.

Di simili discussioni è finora giunta in Italia un'eco disorganica, anche a causa della sostanziale estraneità del mondo scolastico e accademico della penisola, fino a tempi recenti, per simili pratiche di valutazione dell'efficacia del curriculum formativo nello sviluppo delle capacità intellettuali di base. La tradizionale concentrazione sulla veicolazione dei contenuti disciplinari come canale per il consolidamento delle capacità intellettuali di base nell'esperienza didattica, e l'idea che a un sistema educativo nazionale omogeneo corrispondessero pur tra le differenze di successo individuale risultati comunque comparabili, sono però negli ultimi anni messe in discussione dall'importanza assunta nell'opinione pubblica dalle verifiche comparative amministrative dall'OCSE, e dai tentativi di attribuire maggiore centralità alle rileva-

1. Il periodico ritorno del dibattito ha garantito lo stratificarsi di una letteratura copiosa sull'argomento. Mi limito qui a fare riferimento a un'esposizione recente ed equilibrata come R. Zwick, *Rethinking the SAT. The Future of Standardized Testing in University Admissions*, RoutledgeFalmer, London-New York 2004.

zioni dell'Istituto nazionale per la valutazione del sistema educativo². Riprendere qui temi e spunti del dibattito d'oltre Atlantico, però, può risultare incongruo, se non si tiene conto delle profonde differenze di contesto.

Il SAT nel “mercato” educativo statunitense

Riflettere sulla profondità del rapporto che le istituzioni educative americane hanno sviluppato con i test di verifica della preparazione degli studenti aiuta a circoscrivere i termini del dibattito.

Il più antico e comune test standard, il SAT (acronimo originariamente di *Scholastic Aptitude Test*, poi modificato in *Scholastic Assessment Test*) vide le prime sperimentazioni nel 1901, l'anno dopo che alcuni dei maggiori college statunitensi avevano creato il *College Entrance Examination Board* allo scopo di dare nuovi strumenti per migliorare le procedure di ammissione in un momento di crescita accelerata nella domanda di accesso all'istruzione post-secondaria. La somministrazione regolare iniziò nel 1926, ma il punto di svolta per la sua affermazione si ebbe tra gli anni Trenta e Quaranta, quando il presidente di Harvard, James Bryant Conant, divenne un suo aperto sostenitore e promosse la nascita dell'*Educational Testing Service* (ETS), l'agenzia destinata al controllo e all'amministrazione del SAT e alla produzione di altre prove consimili (come ad esempio il *Test of English as a Foreign Language* o TOEFL).

La precocità nell'utilizzo di questo tipo di test e la longevità nel rapporto con le istituzioni educative trovano la propria ragion d'essere nel più caratteristico tratto strutturale del mondo scolastico americano, ovvero il fatto di essersi sviluppato in un contesto in cui, per usare le parole di Martin Trow, già direttore del Center for the Study of Higher Education di Berkeley (CSHE), «il mercato ha preceduto la società e le sue istituzioni di istruzione»³. Un'offerta formativa sviluppatasi, senza evidente supervisione, secondo la pluralità di esigenze che caratterizzava una comunità ampia, complessa e ricca di distinzioni socio-culturali al proprio interno, produsse presto un insieme vario e stratificato di istituti, di programmi di studio e di percorsi didattici. All'inizio del Ventesimo secolo, il SAT si propose come strumento di verifica unitario per offrire a studenti, famiglie e istituti scolastici un parametro comune su cui valutare i risultati di scelte educative tra loro profondamente divergenti. Il successo della tipologia di test amministrata dal College Board si è fondato su una domanda del “mercato” scolastico e accademico, ed è legata alla sua capacità di rispondervi in maniera soddisfacente. Nel corso dei decenni, infatti, il Board che gestisce la prova ha dovuto intervenire più volte e profondamente nella struttura dei quesiti e nel *concept* di base della ve-

rifica, sia per rispondere ai mutevoli orientamenti della propria committenza sia per rintuzzare la concorrenza. Così, da un lato, nel 2001 l'annuncio della possibilità di non prendere più in considerazione i risultati del SAT da parte del presidente della University of California, modello riconosciuto di università pubblica competitiva e vertice di qualità del sistema accademico pubblico del maggiore Stato americano, ha condotto a una complessa indagine di mercato sulle nuove esigenze degli uffici di ammissione dei college, a una minore rigidità nel ricorso alle domande a risposta chiusa in tutti i campi e a un aumento dell'attenzione per le sezioni dei test dedicate alla competenza acquisita nei programmi disciplinari⁴. D'altro canto, dal 1959 il SAT è sottoposto alla competizione con un altro test standard, l'ACT (originariamente acronimo per *American College Testing*)⁵, che ha costantemente guadagnato terreno facendo leva sulle debolezze dell'avversario nella valutazione dell'assimilazione dei contenuti curricolari, e che mettendo in discussione la sua posizione dominante costringe il College Board a migliorare le proprie prestazioni.

Test, classe e merito

Già da questo primo abbozzo di contestualizzazione emerge un paradosso: la “quantificazione” delle competenze assimilate nel percorso di formazione proposta dal SAT, divenuta in anni recenti un idolo polemico dei detrattori del test, era all'origine un punto di forza di simili prove di verifica, perché si proponeva come un comune riferimento per un mondo frammentato, bisognoso di una unità di misura condivisa dei risultati scolastici, di un “vertice” di fatto (a surrogare l'assenza di corpi regolatori di diritto) che fissasse obiettivi universalmente validi e chiaramente identificabili. A questo primo paradosso può aggiungersene un altro. I test standardizzati introdotti con l'esperienza promossa dal College Board, così spesso accusati di dare per scontate nella loro formulazione (e quindi di contribuire a perpetuare) le distinzioni etniche e socio-economiche nordamericane, si sono imposti a tutta prima come strumento di riduzione delle barriere so-

2. Per un'esposizione degli spunti in proposito, rinvio al mio *INVALSI: Spunti per valutare la procedura di valutazione*, in «Termometro Politico», 15 maggio 2014 (<http://www.termometropolitico.it/blog/INVALSI-spunti-valutazione>).

3. M. Trow, *Comparative Perspectives on British and American Higher Education*, in S. Rothblatt - B. Wittrock (eds.), *The European and American University since 1800*, Cambridge University Press, Cambridge 2006, p. 282.

4. Si veda il *paper* del 2004 di R. Zwick - T. Brown - J.C. Sklar, *California and the SAT: Reanalysis of University of California Admissions Data* (<http://www.cshe.berkeley.edu/sites/default/files/shared/publications/docs/ROP.Zwick.8.04.pdf>).

5. Per ulteriori informazioni si veda il sito ufficiale: <http://www.act.org/products/k-12-act-test/>.



ciali e di promozione delle pari opportunità nell'accesso a un'istruzione superiore di qualità.

Come ha ricostruito Jerome Karabel nel suo studio sui criteri di ammissione dei maggiori atenei statunitensi⁶, tra la fine dell'Ottocento e l'inizio del secolo successivo la scelta di sottoporre a selezione le richieste di iscrizione, prima condizionate esclusivamente dalla possibilità di pagare le rette e da semplici prove di competenza disciplinare di base superate da oltre il 90% degli aspiranti, fu dettata da ragioni socialmente conservatrici. La decisione dei vertici di Harvard, Yale e Princeton di valorizzare non tanto la riuscita scolastica quanto il "carattere" degli studenti, metteva a disposizione degli *admission boards* un concetto manipolabile secondo necessità, che infatti divenne rapidamente lo strumento migliore per giustificare l'ingresso nelle università più quotate dei figli degli ex allievi e dei rampolli delle classi privilegiate *wasp*, a discapito, ad esempio, di giovani di origine ebraica. Valutare come titoli di merito elementi del curriculum dei ragazzi come la frequenza ad alcune scuole superiori qualificate, o la partecipazione ad attività extracurricolari di prestigio, ad attività sportive agonistiche o a organismi filantropici di gran nome, significava infatti favorire chi apparteneva ai gruppi generalmente ammessi a tali pratiche di sociabilità e, in definitiva, assumere acriticamente l'ideale di "perfetta umanità" improntato ai valori del tradizionalismo cristiano e del dominio maschile espresso dalla "maggioranza silenziosa" dell'America rurale e benpensante.

La volontà di integrare tra i criteri di scelta e di destinazione dei sostegni finanziari anche i risultati di una verifica delle attitudini e delle competenze per quanto possibile "pura" e asettica emerse gradualmente nella prima metà del Novecento. A rendere possibile tutto ciò fu senz'altro, si è visto, la spinta di un mutamento dei paradigmi ideali, con l'introduzione di valori in senso lato "meritocratici" nel costume accademico da parte di figure come quella di Conant, e del potere dell'esempio dell'università di Harvard da lui diretta in un sistema profondamente condizionato dal continuo confronto tra strutture in concorrenza. A ciò si aggiungeva però la necessità di offrire garanzie di qualità assoluta del corpo studentesco da parte di tutte le università maggiori per ragioni relative al rafforzamento dei propri bilanci: scegliere studenti con rendimento accademico certificato come superiore alla media risultava infatti una pratica vincente per avere maggiori probabilità di attrarre i titolari delle borse di studio federali, in continuo incremento dopo la scelta delle amministrazioni Roosevelt e Truman di sostenere le spese di istruzione dei veterani di guerra⁷; quando poi negli anni Cinquanta le università divennero il principale destina-

Asterischi di Kappa

Francia cattolica

Tutti ricorderanno i "barbuti di Mosul", i fondamentalisti islamici dell'Isis che passarono gli assolati pomeriggi d'inizio primavera 2015 a picconare le lapidi del cimitero cristiano cittadino, ad abbattere campanili e campane, a divellere croci dalle cupole delle chiese, a profanare cappelle e monasteri. In Francia, senza giungere a questi estremi, si assiste, però, a un fondamentalismo laicista dai risultati non dissimili. La statua in onore di Giovanni Paolo II è stata giudicata troppo imponente per essere esposta in luogo pubblico. Se la statua non fosse stata così "vistosa", se l'arco a sormontare il volto rotondo del papa polacco non fosse stato così largo, se quella croce non fosse stata messa in cima, ecco che allora il monumento non avrebbe infranto le rigorose leggi del 1905 sulla separazione tra Stato e religione. Ma visto che l'opera artistica ha una croce e che chiaramente rimanda al culto cattolico, va ridotta in pezzi. Rasa al suolo. Nulla di nuovo, sia chiaro: c'era già stata la candida statuina della Vergine Maria rimossa da un parco pubblico dell'Alta Savoia e la battaglia sui presepi allestiti nei municipi della Vandea. Non si può, stabilirono i tribunali amministrativi: l'esibizione di simboli religiosi in Francia non è ammessa. E pazienza se su quei simboli poggia una bella fetta di storia francese, la vecchia figlia prediletta che ha contemplato i re unti e taumaturghi capaci di guarire i poveri e di sanare le piaghe in nome d'un potere derivante direttamente da Dio. Adesso si litiga perfino sulla possibilità di ricostruire, come era una volta, la guglia di Saint-Denis frantumata da un fulmine in una notte tempestosa, nel pieno Ottocento. C'è chi dice che non ne valga la pena, che costa troppo. Altri ancora, forse, si stanno chiedendo che cosa accadrebbe se un magistrato decidesse che pure la guglia ricostruita "è troppo vistosa".

rio dei finanziamenti pubblici e privati a grandi progetti di ricerca applicata, la possibilità di proporre un ambiente accademico stimolante e di qualità divenne uno dei criteri principali per il successo delle richieste di fondi⁸; più in generale, poi, di fronte a una società sempre più mobile e ricca di opportunità per chiunque desse prova delle competenze più solide e profonde, per tutti gli atenei la

6. J. Karabel, *The Chosen. The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*, Houghton Mifflin, Boston-New York 2005.

7. Sul *Servicemen's Readjustment Act* del 1944, e sulle sue conseguenze, si vedano G.C. Altschuler - S.M. Blumin, *The G.I. Bill. A New Deal for Veterans*, Oxford University Press, Oxford 2009, e K. Frydl, *The G.I. Bill*, Cambridge University Press, Cambridge 2009.

8. Sul tema si veda J. Neusner - N.M.M. Neusner, *The Price of Excellence. Universities in Conflict during the Cold War Era*, Continuum, New York 1995.



Studenti dell'Università di Harvard nel giorno della loro cerimonia di laurea, 29 maggio 2014.

“caccia” a futuri ex allievi destinati con buona certezza a dare buona prova di sé anche senza il sostegno di famiglie illustri alle spalle divenne una prioritaria questione d'immagine. Tutto ciò consentì una ridiscussione almeno parziale dei tradizionali assetti sociali della frequenza universitaria, permettendo a individui privi del *background* tradizionalmente orientato alla prosecuzione degli studi di farsi notare e di entrare nel circuito dei potenziali studenti di università di particolare prestigio. È poi vero, effettivamente, che la concentrazione sulla “misura” delle capacità intellettuali può tramutarsi da strumento di giustizia in arma di consolidamento delle differenze. Lo aveva intuito negli anni Cinquanta Michael Young, nella distopia di *The Rise of Meritocracy*⁹. Nella finzione del romanzo, l'introduzione di verifiche precoci della preparazione e delle capacità degli studenti per indirizzarli verso percorsi di studio adeguati al loro livello e, di conseguenza, a mansioni e ruoli di diverso peso e responsabilità, condusse a tutta prima a un robusto rimescolamento di destini, che però rapidamente si esaurì,

perché i genitori più istruiti iniziarono a curarsi con sempre maggiore preoccupazione della formazione dei figli in vista delle *performance* nelle occasioni di valutazione, e a “congelare” di fatto la mobilità sociale. Ed è in molti casi un fenomeno del genere che si è verificato, con il proliferare di attività parallele all'ordinario percorso scolastico destinate a rafforzare il *test-taking*, appannaggio di giovani provenienti da famiglie in grado di permettersene i costi e, soprattutto, di selezionare con cognizione di causa gli “allenamenti” più produttivi.

Il pericolo è nella rigidità delle pratiche

Questo schematico sguardo storico all'evoluzione degli *standardized tests* negli Stati Uniti permette di cogliere un elemento fondamentale alla radice della loro ragion d'essere, ovvero l'appartenenza a un sistema per sua natura

9. M. Young, *The Rise of Meritocracy*, Thames & Hudson, London 1958.

PROBLEMI PEDAGOGICI E DIDATTICI

flessibile. Prodotto per rispondere a una esigenza di chiarezza e coordinamento tra domanda e offerta per l'accesso al college, il SAT ha accompagnato l'acquisizione di una dimensione pressoché universale dell'istruzione terziaria negli Stati Uniti senza che venisse meno la libertà di tutte le parti in causa nelle loro scelte. Le università, si è visto, ricorrono al SAT e alle prove di verifica derivate da tale esperienza sulla base di scelte che non violano la loro autonomia nelle procedure di selezione della "comunità" che intendono ospitare, per rispondere a precise esigenze; considerano tuttora i risultati dei test standard, così come il rendimento scolastico pregresso di ogni studente, come singoli elementi di un quadro più ampio in cui ancora possono rientrare aspetti come il rendimento atletico a livello agonistico (fondamentale in un mondo accademico in cui la vendita dei diritti sulle manifestazioni sportive rappresenta una voce di bilancio molto importante), la possibilità di contare su lettere di referenza importanti e di preparare autopresentazioni e prove di scrittura brillanti o la possibilità familiare di pagare l'intero costo del posto di studio senza il ricorso a borse o prestiti; sono disponibili a rivedere il l'uso dei risultati del SAT o dell'ACT, e finanche a rinunciarvi, qualora gli effetti non risultino soddisfacenti.

I problemi sorgono quando sulle graduatorie uscite dai test, e sul tentativo di avere un corpo studentesco caratterizzato dal rendimento medio più elevato possibile in queste verifiche, si fonda l'intera politica di ammissioni. Un numero crescente di college emergenti sembra in effetti portare avanti una politica di questo tipo, solo a prima vista sensata. Affidarsi in forma quasi esclusiva ai dati numerici dei test standard nasconde infatti l'incapacità di portare avanti una politica di accessi dotata di una propria identità, e la volontà di sostituirla col semplice *ranking* dato dalla verifica.

Appare quindi tanto più grave che la politica federale negli ultimi anni abbia apertamente promosso un utilizzo rigido dei risultati delle verifiche standardizzate della preparazione degli studenti anche per l'elaborazione di scelte istituzionali di portata ben più ampia. È ad esempio del 2001 il *No child left behind Act*, provvedimento quadro per l'accreditamento di qualità delle scuole primarie e secondarie, fondato sul monitoraggio continuo delle *performance* degli studenti e in particolare del miglioramento dei ragazzi problematici, disabili, provenienti da contesti sociali svantaggiati, e sull'utilizzo di tali risultati per l'erogazione dei fondi ai progetti di recupero e inserimento scolastico, la redistribuzione nel sistema degli insegnanti, il controllo sulla qualità delle istituzioni scolastiche da parte delle famiglie degli studenti. Gli effetti dell'uso obbligatorio come indicatore esclusivo di valori

pensati come puramente informativi in un contesto di valutazioni e di decisioni assai più complesso sta ormai mostrando i suoi effetti in termini di impoverimento delle esperienze didattiche e di *coaching* in vista dei test, e sembrano imporre un ripensamento urgente¹⁰.

Ed è proprio questa "cultura" del test standardizzato di misurazione della preparazione, basata sull'attribuzione indebita di un carattere normativo a risultati progettati come supplemento informativo in un'elaborazione politica complessa, a essere proposta negli ultimi anni anche in Italia. Nel nostro Paese, più ancora che negli USA di *No child left behind*, i test INValSI o il Program for International Student Assessment dell'OCSE si stanno trasformando nel discorso pubblico in surrogati dell'individuazione di un progetto complessivo su ragioni e obiettivi di fondo del servizio scolastico. La rilevazione della diffusione delle capacità intellettuali di base, dei suoi squilibri e delle eventuali carenze, non è colta tanto come la base imprescindibile con cui individuare eventuali criticità e sanarle rendendo più significativo l'impatto della trasmissione didattica e più efficaci le "reti di recupero" degli studenti in difficoltà, quanto come certificazione di un lavoro mal fatto e processo di individuazione dei colpevoli, prossime vittime dell'ennesima (annunciata) "bonifica" dei "fannulloni" che minano il rendimento di tutti i pubblici servizi¹¹. Tutto, questo, peraltro, senza nemmeno considerare che un sistema scolastico amministrato come quello italiano avrebbe a disposizione strumenti di controllo anche più complessi (ispezioni, verifica dei programmi, mutamento nelle pratiche per l'accreditamento dei titoli ecc.) di quelli disponibili nel disorganico arcipelago americano in cui il SAT, "padre" di tutti i test, ha visto a suo tempo la luce.

Andrea Mariuzzo
Scuola Normale Superiore, Pisa

10. Per ulteriori informazioni rinvio al recente T. De Mauro, *Test in fermento*, in «Internazionale», 30 gennaio 2015, p. 23.

11. Oltre ai riferimenti già citati, sulla natura e sui possibili usi "virtuosi" delle indagini comparative sulle *performance* rinvio all'intervento di un educatore di esperienza internazionale: L. Ravaglia, *La scuola non serve a lavorare, ma ad imparare. Ricordiamocelo*, in «Gli Stati Generali», 2 novembre 2014 (<http://www.glistatigenerali.com/scuola/gli-ideali-riformatori-della-scuola-visti-da-bruxelles/>).