



Inclusione e culture valutative

Cristiano Corsini

*Professore associato di Pedagogia sperimentale
Università d'Annunzio di Chieti-Pescara / cristiano.corsini@unich.it*

Nella genesi della selvicoltura scientifica prussiana la misurazione assume un ruolo rilevante. Come riporta Scott (1998), considerato il valore economico del legname, tra Settecento e Ottocento il monitoraggio delle foreste si focalizza sui soli alberi delle specie più redditizie. Raccolti i dati, nuove foreste sono impiantate sulla base del modello pensato per monitorarle, in modo da agevolare il processo di misurazione, ma quel che ne consegue è una riduzione della biodiversità e un'augmentata vulnerabilità nei confronti dei parassiti. Nella prima parte di questo contributo vengono restituiti gli elementi essenziali di una cultura valutativa che, in campo educativo, agisce come fattore di inclusione, e lo fa tenendo a mente il carattere performativo delle misure sfuggito ai burocrati prussiani. Nella seconda parte questa cultura valutativa viene confrontata con quella che, attualmente, impronta le scelte politiche su istruzione e ricerca.

1. Una cultura inclusiva della valutazione

La riflessione scientifica sulla valutazione nel corso dell'ultimo secolo ha posto in aperto dialogo *assessment* (accertamento educativo) ed *evaluation* (il più generale processo valutativo). Schematizzando, partendo dalla Teoria della Valutazione di Dewey (1939), si procede con la *formative assessment* di Scriven (1967) per giungere all'*authentic assessment* di Wiggins (1998), all'*assessment for learning* di Stiggins (2002), Wiliam (2000, 2011) e altri, all'*inclusive assessment* di Watkins (2007) e all'*assessment as learning* di Earl (2013). Questi *assessment* educativi fanno espliciti riferimenti a visioni complementari dell'*evaluation* e queste, a loro volta, non sono prive

di richiami a prassi valutative educative. Quella tra *assessment* ed *evaluation* è una dialettica che, accentuando la funzione formativa e di sviluppo della valutazione, facilita l'emersione di una cultura valutativa inclusiva basata su due fondamentali processi.

Il primo processo è l'integrazione tra insegnamento e valutazione: la valutazione è formativa (Scriven, 1967) se dà forma all'insegnamento, se chi insegna è disposto a impiegare il giudizio, mettendo in discussione quanto svolto, per regolare la successiva didattica (Black, William, 1998; Greenstein, 2010). Un secondo processo, che si fonda sul primo, si gioca tra valutazione e apprendimento e conferisce a chi apprende responsabilità e potere nel processo valutativo in modo tale da assicurargli maggiori possibilità di apprendimento significativo. Da questo punto di vista, la valutazione per l'apprendimento (*assessment for learning*) e la valutazione come apprendimento (*assessment as learning*) delineano un'evoluzione della valutazione formativa alla luce delle teorie dell'apprendimento. Alcuni item presenti nell'Indice per l'inclusione, con le relative questioni-guida (aiutare tutti gli alunni a riflettere sui prodotti, a comprendere le ragioni della valutazione), rappresentano un'ottima definizione operativa di questo processo, che integra affidabilità ed equità dando voce a una pluralità di soggetti e di intelligenze.

Per analizzare i legami tra valutazione e inclusione è utile partire dalle prerogative della cultura valutativa sopra delineata, seguendo uno schema già tracciato in altre sedi (Corsini, 2015; Corsini, 2018). Secondo tale schema, la valutazione può essere definita un processo con tre prerogative ben riconoscibili: essa è (I) un giudizio di valore, emesso sulla (II) distanza rilevata tra una situazione auspicata e una effettivamente riscontrata e finalizzato (III) all'assunzione di decisioni volte alla riduzione di tale distanza. Come di seguito argomentato, queste tre prerogative, già implicite nella teoria della valutazione deweyana, rappresentano il collante che consente di inserire le diverse riflessioni su *assessment* ed *evaluation* in un'unica cultura valutativa in grado di incentivare dinamiche inclusive.

La prima prerogativa, la valutazione come giudizio di valore, statuisce la natura inevitabilmente soggettiva di ogni valutazione. Si tratta di un aspetto fondamentale nell'avvio e nello sviluppo degli studi docimologici – incentrati sull'analisi dell'inaffidabilità delle valutazioni scolastiche – che in seguito si è arricchito di riflessioni relative all'inqiuità valutativa, con riferimento al ruolo che in essa giocano stereotipi e pregiudizi di genere, di classe, etnici o sulla disabilità. Tuttavia la stessa docimologia, piuttosto che

proporre un'improbabile rimozione della dimensione soggettiva nella valutazione, si pone il problema del suo controllo, ovvero dell'impiego di forme intersoggettive di accertamento: ineliminabile è la soggettività valutativa, non certo l'arbitrio. Questo aspetto della valutazione è intimamente connesso con l'idea di valutazione come processo di indagine: i valori e le conseguenti valutazioni sorgono, per Dewey (1939), dove ve ne è motivo, ovvero dove avvertiamo mancanze, e si sviluppano allo scopo di ipotizzare soluzioni possibili. Il giudizio valutativo è, sempre, una forma di gestione del potere, inteso come possibilità di incidere sui soggetti: ma, come ricorda Barbier (1989), tale potere può essere gestito autocraticamente, escludendo gli stessi individui valutati (col paradosso di oggettivare dei soggetti in apprendimento), oppure condiviso, includendo chi apprende nel processo valutativo fino alla coelaborazione del giudizio. In sintesi, questa prima prerogativa si contrappone a ogni imposizione dall'alto o dall'esterno di finalità e strumenti valutativi e ribadisce la necessità di un confronto o di una negoziazione con individui e contesti valutati (Guba, Lincoln, 1989).

La seconda prerogativa richiede al giudizio valutativo di esprimersi sulla distanza rilevata tra una situazione auspicata e una effettivamente riscontrata, e consente di delineare il ruolo della misurazione nel processo valutativo. Il ricorso alla misura, se inteso come processo che consente di collegare concetti astratti a indicatori empirici (Lucisano, 1989, 2018) – e dunque come operazione di confronto – è caratteristico di ogni compiuta valutazione. Per Dewey (1939) vi è sempre qualche osservazione del risultato ottenuto in paragone e in contrasto con quello atteso, in modo che il paragone getti luce sull'adeguatezza delle cose adoperate come mezzi. Si tratta di porre dunque la misurazione in una posizione intermedia, tra due atti valutativi. Nel 1955, sulla scorta di Dewey, Visalberghi reclama un rapporto dialettico tra misurazione e valutazione nel processo educativo (“la misurazione nasce nella valutazione e in essa confluisce”, p. 21), richiamandosi all'abito di equilibrio e discrezione che un atteggiamento privo di senso della misura rischia di smarrire. Questo significa che i dati non sono ipostatizzati, ma assunti all'interno delle ipotesi interpretative che li producono e che orientano le azioni seguenti. Di qui la necessità di descrivere i pieni, non i vuoti, ovvero di ricorrere a strumenti in grado di fornire un accertamento criteriale e rendere conto dei progressi di ciascuno in termini di padronanza. I lavori sulla motivazione all'apprendimento e sugli obiettivi di padronanza rilevano da decenni (Ames, 1992) come una valu-

tazione gestita totalmente dall'alto, che scoraggia la collaborazione e incentra il giudizio sul confronto con le prestazioni altrui, orienti verso obiettivi di prestazione. Al contrario, orienta verso la padronanza un giudizio espresso sulla posizione di chi apprende all'interno del proprio percorso di apprendimento, una distanza colmabile se si esplicitano obiettivi intermedi in termini di comportamenti osservabili. Una valutazione formativa si fa dunque inclusiva se riconosce e valorizza le differenze senza ricondurle a distanza da una norma¹, dando voce a diverse intelligenze e diverse abilità. Diverse misure e misure diverse conferiscono affidabilità ed equità alla valutazione: riconoscere e valutare differenze (Ianes, 2005) significa conferire loro valore (Tomlinson, Moon, 2013; Dovigo, 2007), e consente a tutti di esprimere e sviluppare compiutamente il proprio apprendimento, rendendo possibile coltivare differenti interessi e diverse intelligenze.

La terza prerogativa determina il perché della valutazione: ridurre la distanza accertata attraverso la misurazione. A rendere efficace la valutazione non è l'impiego di certi strumenti, ma l'uso delle informazioni raccolte attraverso di essi, e la funzione attribuita al processo risulta determinante da questo punto di vista. Spesso la valutazione tende a essere percepita come mera incombenza burocratica o come strumento di erogazione di premi o punizioni (Lucisano, Corsini, 2015; Corsini, 2015). La letteratura di riferimento relativa agli atteggiamenti verso la valutazione da parte dei docenti (Brown, 2011; Davis, Neizel, 2011) attribuisce però efficacia educativa alle sole concezioni che assegnano alla valutazione finalità di miglioramento dei processi di insegnamento e apprendimento (Hattie, Timperley, 2007). Da questo punto di vista, una cultura della valutazione inclusiva si basa sull'impiego proattivo dell'errore (che non è mai stigma) e su un feedback quanto più possibile analitico (Gentile, 2018).

2. La cultura della valutazione egemonica

Da un ventennio si sta affermando nelle politiche valutative dell'istruzione e della ricerca, di scuole e università, una cultura della valutazione che, come verrà evidenziato, tradisce le tre prerogative sopra delineate. Questa

1 Come nel *dynamic assessment*, cfr. Aquario, 2015.

cultura della valutazione inizia a prendere forma in concomitanza con l'autonomia scolastica e universitaria. Si tratta di un passaggio non irrilevante: a tale cultura si tende infatti ad assegnare il compito di tenere sotto controllo le prestazioni di studentesse e studenti, di scuole e università, in modo da rendere conto dell'efficienza degli investimenti in istruzione e in vista di un miglioramento che dovrebbe scaturire da meccanismi competitivi. In altri termini è una cultura della valutazione che si descrive come *accountability* e accentua la funzione sommativa della valutazione. È utile provare a inquadrare tale cultura valutativa a partire da due posizioni espresse nel 2008. La prima sembrerebbe un *lapsus*. Nel rapporto sull'indagine Pirls (Invalsi, 2008, p. 99) si parla di “sviluppo di una cultura della valutazione (*teach to test*) con l'introduzione nelle scuole di prove di verifica strutturate (quesiti a scelta multipla) e semistrutturate (quesiti con stimolo chiuso e risposta aperta)”. Sebbene la definizione di una cultura non possa evitare riferimenti a valori, credenze e finalità, nel passaggio citato ci si limita a riportare strumenti che garantiscono oggettività, un requisito che, come vedremo, svolge un ruolo di rilievo nella retorica con la quale questa cultura della valutazione si racconta. Un secondo aspetto caratterizzante di questa cultura valutativa è l'induttivismo ingenuo rappresentato dalle parole che il ministro dell'Economia Tremonti (2008) spende per sancire il passaggio dai giudizi ai voti nella scuola primaria:

i numeri sono una cosa precisa, i giudizi sono spesso confusi. Ci sarà del resto una ragione perché tutti i fenomeni significativi sono misurati con i numeri [...]. I numeri sono insieme precisi e semplici. Il messaggio che trasmettono è un messaggio diretto. Se gli stessi fenomeni [...] fossero espressi non con numeri ma attraverso frasi complesse con finalità descrittive, il messaggio resterebbe impreciso (p. 37).

È utile collocare i tratti essenziali di questa cultura della valutazione egemonica nelle tre prerogative che definiscono la cultura della valutazione inclusiva. Quanto alla prima prerogativa (la valutazione come giudizio di valore), si riscontra una sostanziale negazione del carattere soggettivo della valutazione: l'*accountability* chiede riferimenti tali da rassicurare l'opinione pubblica sul carattere rigoroso del controllo. A scuole e università vengono pertanto imposti tempi, modalità e strumenti valutativi concepiti per dare luogo a misure oggettive. Questo si traduce in un rapporto gerarchico che tradisce quanto evidenziato in precedenza in relazione al

carattere negoziato proprio di una cultura valutativa inclusiva e rischia di riprodurre, come verrà di seguito evidenziato, effetti deleteri sulla qualità dei processi oggetto di misura. Siamo così in presenza della costruzione di una verità, intesa come sistema di obblighi (Foucault, 2014), che porta gli individui a interiorizzare meccanismi che li depotenziano. Secondo Borrelli (2015) e Pinto (2012), tali dinamiche sono più pervasive se numericamente espresse.

È dunque la seconda prerogativa, ovvero la misurazione, ad assumere una posizione dominante. In questa nuova cultura valutativa la quantificazione si fa valutazione e le misure, strappate alle ipotesi che le hanno prodotte, vengono reificate: da mezzi, diventano fini. Si tratta di un evidente tradimento di quella riflessione che afferma, già con Dewey, il carattere intermedio e strumentale delle misure, il loro adombrare soluzioni possibili a problemi autenticamente percepiti. La validità è ridotta ad affidabilità, laddove il ricorso massiccio a strumenti normativi fa sì che la distanza tra realtà accertata e aspettative venga percepita come lontananza da valori medi (o mediani), e dunque come deficit. La stessa eterogeneità è avvertita come problema.

Di conseguenza, passando alla terza prerogativa (il fine della valutazione), si può rilevare come questa cultura valutativa, incentrata su dispositivi premiali o reputazionali, comporti effetti deleteri sulla qualità della didattica e sui processi d'inclusione. Tra i primi si ricorda il *teaching to the test* (dedicare sforzi, spazi e tempi didattici nell'addestrare studentesse e studenti alla forma del test) e il *cheating* (imbrogliare), mentre lo *skimming* (selezionare per il test solo i soggetti migliori, allontanando quelli ritenuti peggiori o richiedendo per essi particolari certificazioni) è un chiaro esempio del secondo genere di distorsioni. Non desta stupore – all'interno di un sistema che risponde a una cultura della valutazione che concepisce la diversità come problema – l'esistenza di rapporti di autovalutazione che associano la selezione e l'assenza di eterogeneità alla qualità di apprendimento e insegnamento. La posta in palio legata alle misure (reputazione di dirigenti, scuole e docenti, possibilità di attrarre iscrizioni e fondi) comporta la tendenza a rafforzare meccanismi selettivi, in ingresso o in itinere, in modo tale – ad esempio – da consentire alla scuola di ottenere posizioni migliori nella classifica Eduscopio, incentrata sul rendimento di ex studenti al primo anno di università. Elementi non ritenuti inclini ad affrontare un buon percorso accademico potrebbero essere scoraggiati dal proseguire gli studi in quello stesso istituto. In questa cultura valutativa, includere non conviene.

Considerazioni conclusive

Nel *narrative assessment* la produzione – da parte di chi valuta – di narrazioni relative ai progressi evidenziati da chi apprende (*learning stories*) consente di porre l'accento su quel che studentesse e studenti imparano progressivamente a fare piuttosto che su obiettivi stabiliti in termini di distanza da prestazioni medie. La distanza rilevata non è più concepita come deficit: “*the teacher increasingly sees the learner, not the disability*” (Dunn, 2004, p. 126). Si tratta di un processo inclusivo, perché impegna chi insegna (e valuta) in una rimodulazione del contesto e in un riorientamento del proprio punto di vista che consentono, come ogni autentica prassi narrativa (Batini, Bertolucci, De Carlo, 2017), di conferire senso all'esperienza pregressa e presente per orientare quella futura. Questo riferimento temporale consente un ultimo richiamo a Dewey, utile a tracciare la distanza tra la cultura della valutazione inclusiva che si è andata delineando nei decenni e quella che recentemente ha acquisito egemonia. Nella “Teoria della valutazione” emerge con chiarezza come il tempo di “appropriate valutazioni” sia il futuro, dato che i giudizi di valore sono relativi a cose che ancora devono esser poste in essere. Una valutazione è efficace se comporta un impegno futuro, una modifica dell'agire ancorata ai giudizi di valore emessi sulle misure raccolte. Ma, diversamente dalle indicazioni fornite in ottica formativa, i premi e le punizioni erogati dai sistemi di *accountability* tendono a disincentivare l'utilizzo delle misure come elementi di sviluppo e miglioramento. Quanti accademici impiegano il punteggio ottenuto alla VQR per migliorare la qualità della propria ricerca, piuttosto che per verificare la possibilità di entrare o permanere in collegi dottorali, o per misurare il contributo apportato alla quota di finanziamenti riservata al proprio dipartimento? La differenza con l'impatto di referaggi che impegnano gli stessi accademici a modificare articoli inviati a riviste è evidente e, in ultima analisi, attesta tutta la distanza tra valutazione e *accountability*.

Non merita forse di essere messa in discussione la stessa inclusione dell'*accountability* tra i sistemi di valutazione? E non solo in base ai tradimenti che una cultura valutativa incentrata su funzioni rendicontative evidenzia rispetto alle prerogative di una cultura valutativa inclusiva, ma anche prendendo in considerazione quanto la sua tendenza a separare dati e valori incida sulla dissociazione tra sguardo sul passato, impegno nel presente e cambiamento futuro. Eppure Cronbach, citatissimo da quanti si occupano di *assessment* e misure come padre dell'Alpha (indicatore di affidabilità), nel

1980, in un testo orientato a riformare l'*evaluation*, invita a distinguere nettamente l'*accountability*, che “implica il guardarsi indietro per distribuire meriti o colpe”, dalla valutazione, “che serve per comprendere eventi e processi e indirizzare attività future”. Un invito lasciato cadere nel vuoto. Colpevolmente: perché ne risulta un sistema in cui la valutazione viene costretta sul letto di Procuste di un'*accountability* che non solo compromette la validità delle misure (Corsini, Scierri, Scionti, 2018; Losito, 2016), ma pregiudica anche la possibilità dei giudizi di valore di incidere positivamente sui processi educativi e di promuovere prassi inclusive.

Bibliografia

- Ames C. (1992). Classroom: Goals, structures and student motivation. *Journal of Educational Psychology*, 84: 261-271.
- Amrein-Beardsley A. (2014). *Rethinking Value-Added Models in Education*. New York: Routledge.
- Aquario D. (2015). *Valutare senza escludere*. Parma: Junior.
- Barbier J.M. (1989). *La valutazione nel processo formativo*. Torino: Loescher.
- Batini F., Bertolucci M., De Carlo E. (2017). Fight Dispersion Through Education: The Results of the First Cycle of the NoOut Project. *Mind, Brain, and Education*, 11 (4): 201-212.
- Black P., William D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2): 139-148.
- Booth T., Ainscow M. (2014). *Nuovo Index per l'inclusione*. Roma: Carocci.
- Borrelli D. (2015). *Contro l'ideologia della valutazione. L'ANVUR e l'arte della rotamazione dell'università*. Sesto San Giovanni: Jouvence.
- Brown G.T.L. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters* 3: 45-65.
- Canevaro A. (2007). *L'integrazione scolastica degli alunni con disabilità*. Trento: Erickson.
- Catalfamo A. (ed.) (2017). *Cultura inclusiva nella scuola e progettazioni curricolari*. Trento: Erickson.
- Corsini C. (2013). La validità di contenuto delle prove INVALSI di comprensione della lettura. *Giornale Italiano della Ricerca Educativa*, 10: 46-61.
- Corsini C. (2015). *Valutare scuole e docenti*. Roma: Nuova Cultura.
- Corsini C. (2018). Sull'utilità e il danno di misurazione e valutazione in educazione. In C. Corsini (ed.), *Rileggere Visalberghi* (pp. 13-28). Roma: Nuova Cultura.
- Corsini C., Losito B. (2013). Le rilevazioni INVALSI: a che cosa servono? *Cadmo*, 2: 55-76.

- Corsini C., Scierri I.D.M., Scionti A. (2018). La validità delle prove INVALSI di comprensione della lettura. In A.M. Notti (ed.), *La funzione educativa della valutazione* (pp. 335-352). San Cesario di Lecce: Pensa.
- Corsini C., Zanazzi S. (2015). Valutare scuola e università: approccio emergente, interventi e criticità. *I problemi della pedagogia*, 2: 305-334.
- Cronbach L.J. (1980). Our Ninety-five thesis. In L.J. Cronbach (ed.), *Toward Re-form of Program Evaluation* (pp. 25-34). San Francisco: Jossey-Bass.
- Davis D.S., Neitzel C. (2011). A self-regulated learning perspective on middle grades classroom assessment. *Journal of Educational Research*, 104: 202-215.
- Dewey J. (1938). *Logic, the Theory of Inquiry*. New York: Holt and Co.
- Dewey J. (1939). *Theory of Valuation*. Chicago: University press.
- Dovigo F. (2007). *Fare differenze*. Trento: Erickson.
- Dunn L. (2004). Developmental Assessment and Learning Stories in Inclusive Early Intervention Programmes: Two Constructs in One Context. *NZ Research in Early Childhood Education*, 7: 119-133.
- Earl L.M. (2013). *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning*. Thousand Oaks: Corwin Press.
- Foucault M. (2014). *Subjectivité et vérité. Cours au Collège de France 1980-81, édition établie sous la direction de Francois Ewald et Alessandro Fontana par Frederic Gros*. Paris: Seuil/Gallimard (trad. it *Soggettività e verità*, Feltrinelli, Milano 2017).
- Gentile M. (2018). Valutare meglio, imparare di più. In C. Corsini (ed.), *Rileggere Visalberghi* (pp. 111-140). Roma: Nuova Cultura.
- Giovannini M. L. (2014). Il dibattito su valutazione ed evidenze: per un processo valutativo credibile e trasparente. *Journal of Educational, Cultural and Psychological Studies*, 1-9: 101-126.
- Greenstein L. (2010). *What Teachers Really Need to Know About Formative Assessment*. Alexandria: Association for Supervision & Curriculum Development.
- Guba E.G., Lincoln Y.S. (1989). *Fourth Generation Evaluation*. Newbury Park: Sage.
- Hattie J., Timperley H. (2007). The power of feedback. In *Review of Educational Research*, 77(1): 81-112.
- Ianes D. (2005). *Bisogni educativi speciali e inclusione*, Trento: Erickson.
- Ianes D. (2014). *L'evoluzione dell'insegnante di sostegno. Verso una didattica inclusiva*. Trento: Erickson.
- Invalsi (2008). *La ricerca IEA-PIRLS 2006*. Milano: FrancoAngeli.
- Losito B. (2016). Misurare e valutare per il miglioramento delle scuole. *Giornale italiano di psicologia*, 3: 483-488.
- Lucisano P. (1989). Misurare come e perché. In P. Lucisano (ed.), *Lettura e comprensione* (pp. 43-68). Torino: Loescher.
- Lucisano P. (2016). La misura delle misure e la validità educativa. *Giornale italiano della ricerca educativa*, IX, 16: 59-69.

- Lucisano P. (2018). L'uomo è misura di tutte le cose (non è sempre necessario prendere i voti). In C. Corsini (ed.), *Rileggere Visalberghi* (pp. 29-44). Roma: Nuova Cultura.
- Pinto V. (2012). *Valutare e punire*. Napoli: Cronopio.
- Ravitch D. (2010). *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. New York: Basic Books.
- Scott J. (1998). *Seeing Like a State: How Some Schemes To Improve The Human Condition Have Failed*. Binghampton: Vail-Ballou Press.
- Scriven M. (1967). The methodology of evaluation. In R.W. Tyler, R.M. Gagne, M. Scriven (eds.), *Perspectives of curriculum evaluation* (Vol. I, pp. 39-83). Chicago: Rand McNally.
- Slee R. (2013). How do we make inclusive education happen when exclusion is a political predisposition? *International Journal of Inclusive Education*, 17, 8: 895-907.
- Stiggins R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*. 83(10): 758-765.
- Stufflebeam D. (2001). *Evaluation models. New Directions in Evaluations*. San Francisco: Jossey Bass.
- Tomlinson C.A., Moon T. (2013). Differentiation and classroom assessment. In J.H. MacMillan (ed.), *Sage Handbook of research and classroom assessment* (pp. 415-430). Los Angeles: Sage.
- Tremonti G. (2008). Il passato e il buon senso. *Il Corriere della Sera*, 22 agosto 2008.
- Visalberghi A. (1955). *Misurazione e valutazione nel processo educativo*. Milano: Edizioni di Comunità.
- Watkins A. (2007). *Assessment in inclusive settings: key issues for policy and practice*, Odense: European Agency for Development in Special Needs.
- Watkins C. (2001). *Learning about Learning enhances performance*. London: Institute of Education, University of London.
- Wiggins G. (1993). *Assessing Student Performance: Exploring the Purpose and Limits of Testing*. San Francisco: Jossey-Bass.
- Wiggins G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.
- William D. (2000). *Integrating formative and summative functions of assessment*. Paper presented to Working Group 10 of the International Congress on Mathematics Education.
- William D. (2010). Standardized test and school accountability. *Educational Psychologist*, 45(2): pp. 107-122.
- William D. (2011). What is assessment for learning. *Studies in Educational Evaluation*, 37: 3-14.