

## LE RUBRICHE DI VALUTAZIONE

GRUPPO DI RICERCA INVALSI - RESP. DONATELLA POLIANDRI

**Presentazione dello strumento** - Nel settore dell'istruzione, una rubrica è un insieme di criteri e standard collegati ad obiettivi di apprendimento, che viene utilizzata per la valutazione autentica delle prestazioni di uno studente. Una rubrica per essere tale deve contenere i criteri di prestazione e una descrizione delle prestazioni buone e meno buone per ciascun criterio considerato. In molti casi le rubriche di valutazione sono utilizzate per delineare criteri uniformi per la classificazione di compiti complessi. Poiché i criteri di valutazione sono pubblici, ciò permette a chi le utilizza (valutatori e valutati) di giudicarne i criteri. Una rubrica di valutazione può anche fornire una base per l'auto-valutazione, la riflessione e revisione tra pari. Essa mira a una valutazione accurata e equa, a comunicare aspettative di qualità, a favorire la comprensione, a indicare un modo di procedere per un successivo miglioramento. Questa integrazione di prestazioni e *feedback* definisce la rubrica come uno strumento per la valutazione formativa.

Le rubriche di valutazione sono usate non solo per la valutazione delle prestazioni dei singoli, ma anche per la valutazione delle scuole (cfr. quelle usate dagli Ispettori scolastici inglesi, Office for Standards in Education, Children's Services and Skills, 2014). L'INVALSI ha adottato le rubriche di valutazione per valutare la qualità delle azioni realizzate dalle scuole nei progetti sperimentali PON Valutazione e Miglioramento (2010), VSQ (2011), VALES e Valutazione e Miglioramento nuova edizione (VM) (2013).

**Struttura** - Per i progetti VALES e VM sono state elaborate 11 distinte *Rubriche di valutazione*, corrispondenti alle aree oggetto di valutazione, presenti nel quadro di riferimento per i percorsi valutativi delle scuole (INVALSI, 2012). All'interno di ogni area sono presenti: a) il criterio di prestazione; b) la *check list*; c) uno spazio per appunti e uno per indicare i documenti esaminati a scuola; d) la scala di valutazione; e) uno spazio per descrivere le principali evidenze emerse, i punti di forza e i punti di debolezza riscontrati. I criteri di prestazione o *standard* sono definiti come proposizioni (es. per l'area Inclusione: "Strategie di gestione delle diverse forme di diversità, adeguamento dei processi di insegnamento e apprendimento ai bisogni formativi di ciascun allievo"). Ogni criterio di prestazione viene esplorato attraverso una prospettiva "multifocale": infatti all'interno della *check list* sono individuati degli *indicatori*, basati su precisi elementi o evidenze empiriche, afferenti a diverse dimensioni. Per ciascun criterio di prestazione sono predisposte delle scale di valutazione a quattro livelli (inadeguato, accettabile, buono, eccellente). Ogni livello è descritto in modo analitico in relazione al criterio di prestazione corrispondente. La descrizione del livello contiene gli elementi o le evidenze empiriche da considerare per esprimere un giudizio.

**Modalità di compilazione** - Dopo avere svolto individualmente le interviste alle diverse componenti scolastiche il Team compila congiuntamente le *check list*. A questo livello avviene una prima negoziazione tra valutatori in merito alle pratiche, alle modalità organizzative e all'efficacia delle azioni messe in campo dalla scuola. Successivamente viene assegnato un livello per ogni area dei processi e per i risultati. Il livello scelto scaturisce dal confronto tra i valutatori (triangolazione) che sono portatori dei punti di vista specifici delle diverse componenti scolastiche incontrate (Dirigenza e staff, docenti, studenti e famiglie). I valutatori selezionano il livello che descrive meglio la situazione della scuola. Nel caso in cui tali descrizioni non riflettano in maniera puntuale la situazione incontrata, i Team possono decidere di far valere maggiormente un aspetto e di ridurre la rilevanza di un altro nell'espressione del giudizio, basandosi sulla propria esperienza di osservazione degli ambiti considerati. Oltre all'assegnazione del livello il Team, per ogni area di processo e per i risultati, compila un giudizio descrittivo, in cui sono brevemente presentate alla scuola le principali evidenze emerse e i punti di forza e di debolezza riscontrati.

**Partecipanti** - Nell'a.s. 2013-2014, le scuole valutate tramite le *Rubriche di valutazione* sono state nel complesso 721 (288 scuole VALES e 433 scuole VM). Le rubriche sono state utilizzate da Team composti da due valutatori formati all'utilizzo degli strumenti. I valutatori sono stati nel complesso 430, distinti in 4 profili professionali: A1 (dirigente scolastico), A2 (insegnante), B1 (esperto in ambito socio-psico-

pedagogico), B2 (esperto di valutazione delle organizzazioni). I Team erano composti da due tipologie di coppie, A1-B1 e A2-B2. Il disegno della ricerca ha previsto la rotazione sia dei Team, sia dei valutatori all'interno dei Team stessi. Il 90% delle scuole è stato valutato da Team diversi.

**Attendibilità/validità** – Le *Rubriche di valutazione* sono strumenti che si collocano nell'ambito della ricerca qualitativa, pertanto l'attendibilità dipende essenzialmente dalla descrizione esplicita delle procedure di osservazione, e quindi dalla replicabilità e ripercorribilità dell'intero itinerario di processo e dalla successiva ispezionabilità della base empirica (Gobo, 2011). L'attendibilità è anche legata fortemente alla triangolazione delle tecniche di rilevazione utilizzate durante la visita dei Team a scuola. La validità definita come 'interna' è di natura soprattutto semantica, poiché indica il grado in cui uno strumento contiene tutte le caratteristiche che definiscono il concetto che si intende misurare (Kirk e Miller, 1986) e se sono state effettuate alcune analisi per verificarne l'attendibilità e la validità. E' inoltre legata al concetto di 'credibilità' (Guba e Lincoln, 1995; Merriam, 1998): nei diversi progetti in cui le *Rubriche di valutazione* sono state utilizzate, i valutatori che hanno compiuto le visite a scuola sono sempre stati almeno due; ciò ha permesso la triangolazione dei ricercatori per confermare e irrobustire i risultati che di volta in volta sono emersi, al fine di costruire una comprensione olistica e una plausibile spiegazione del fenomeno oggetto di studio attraverso un accordo intersoggettivo. Infine, attraverso la restituzione dei Rapporti di valutazione alle scuole è stato possibile attuare il *member check*: le interpretazioni sono state ripresentate alle scuole da cui sono stati raccolti i dati, chiedendo loro di riconoscerne la plausibilità. Nello specifico si presentano alcune analisi effettuate sul funzionamento delle Rubriche.

- Coerenza interna: le *Rubriche di valutazione* presentano una buona correlazione tra il numero di check inseriti nella *check list* e il livello assegnato nella scala di valutazione (valori compresi tra 0,4 e 0,7), ne deriva che la variabilità nei giudizi spiegata dal numero dei *check* oscilla tra il 20% e il 50% ( $R^2$ ).
- Attendibilità dello strumento nell'utilizzo da parte di soggetti diversi: sono stati posti a confronto i giudizi espressi dalle due tipologie di coppie di valutatori A1-B1 e A2-B2. Dalle analisi non emergono differenze significative tra i giudizi attribuiti dalle due tipologie di coppie (Test del Chi quadrato).
- I livelli che i valutatori esterni hanno assegnato con le *Rubriche di valutazione* presentano un buon accordo con i livelli che le scuole VALES si sono auto assegnate nel Rapporto di autovalutazione (RAV) (coefficienti di correlazione con valori compresi tra 0,2 e 0,7). Sebbene i Team di valutazione conoscessero i RAV, non erano però tenuti a rispecchiarne i giudizi. I valutatori esterni in media hanno confermato i giudizi espressi dalle scuole nel 54% dei casi, hanno invece espresso giudizi superiori a quelli espressi dalle scuole nel 27% dei casi, e solo nel 18% dei casi in media hanno espresso giudizi inferiori.

### Bibliografia minima

- Comoglio, M. (2002). La valutazione autentica. *Orientamenti Pedagogici*, 49 (1), pp. 93-112.
- Gobo, G. (2001). *Descrivere il mondo: teoria e pratica del metodo etnografico in sociologia*. Roma: Carocci.
- INVALSI (2012). *Valutare le scuole: le logiche generali del progetto VALES*. Disponibile su [http://www.invalsi.it/invalsi/ri/vales/documenti/Logiche\\_gen\\_progetto\\_VALEs.pdf](http://www.invalsi.it/invalsi/ri/vales/documenti/Logiche_gen_progetto_VALEs.pdf) [Data di accesso: giugno 2014]
- Kirk, J & Miller, M.L. (1986). *Reliability and validity in qualitative research*. Newbury Park, CA: Sage.
- Lincoln, YS. & Guba, EG. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage.
- Merriam, S.B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.
- Office for Standards in Education, Children's Services and Skills (2014). *School Inspection Handbook. Handbook for inspecting schools in England under section 5 of the Education Act 2005*. Disponibile su <http://www.ofsted.gov.uk/resources/school-inspection-handbook> [Data di accesso: giugno 2014].
- Wiggins, G. (a cura di) (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco, CA: Jossey Bass.

