

Assessment of students' digital competences: from the measurement scale to threshold levels. A systematic literature review

La valutazione delle competenze digitali degli studenti: dalla scala di misurazione ai livelli soglia. Revisione sistematica della letteratura

Marco Giganti

Università degli Studi di Bergamo / Dipartimento di Scienze Umane e Sociali

OPEN ACCESS

Double blind peer review

Citation: Giganti, M. (2024). Assessment of students' digital competences: from the measurement scale to threshold levels. A systematic literature review. *Italian Journal of Educational Research*, 33, 30-46.
<https://doi.org/10.7346/sird-022024-p30>

Corresponding Author: Marco Giganti
Email: marco.giganti@unibg.it

Copyright: © 2024 Author(s). This is an open access, peer-reviewed article published by Pensa Multimedia and distributed under the terms of the Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. IJEDuR is the official journal of Italian Society of Educational Research (www.sird.it).

Received: September 05, 2024
Accepted: November 25, 2024
Published: December 20, 2024

Pensa MultiMedia / ISSN 2038-9744
<https://doi10.7346/sird-022024-p30>

Abstract

This paper aims to present a systematic literature review (SLR) on the methodologies and tools utilized for assessing students' digital competencies, with particular focus on the definition of threshold levels to describe outcomes achieved in standardized tests. The SLR constitutes the initial phase of a research project that seeks to integrate a pedagogical approach into a predominantly technical and psychometric process. The research questions address both the methodology for assessing digital competencies and the procedure for creating and describing threshold levels. A comprehensive search was conducted across three article repositories (ERIC, Scopus, and Web of Science) and several institutional websites (such as UNESCO, the European Union, and IEA), resulting in the retrieval of 5751 articles. Applying inclusion and exclusion criteria (based on students' grade levels, assessment methods, etc.), 128 articles were selected for further examination. However, the analysis of the 21 most relevant articles revealed an absence of pedagogical considerations, indicating a potential area for further exploration. The delineation of competency levels, as determined by standardized test outcomes, has significant implications for schools, students, and their families, providing critical feedback that can inform the teaching-learning process, educational redesign, and decision-making processes.

Keywords: Digital Competences, Standardized Assessment, Threshold Levels, Pedagogical Approach, Systematic Literature Review.

Riassunto

L'articolo presenta una revisione sistematica della letteratura (SLR) sui metodi e gli strumenti per valutare le competenze digitali degli studenti, in particolare sulla costruzione di livelli soglia volti a descrivere i risultati ottenuti in test standardizzati. La SLR è il primo passo di un progetto di ricerca il cui scopo è l'applicazione di un approccio pedagogico in un processo eminentemente tecnico e psicometrico. Le domande di ricerca sono legate alla metodologia di valutazione delle competenze digitali e alla procedura di definizione e descrizione dei livelli di soglia. Tre repository di articoli (ERIC, Scopus e Web of Science) e alcuni siti istituzionali (UNESCO, Unione Europea, IEA ecc) sono stati interrogati, ottenendo 5751 articoli. Attraverso criteri di inclusione/esclusione (classe degli studenti, metodo di valutazione, ecc.) sono stati recuperati 128 articoli. I 21 articoli analizzati non offrono riflessioni pedagogiche sistematiche e questo apre ulteriori prospettive di riflessione. Descrivere i livelli di competenza, come risultato di test standardizzati, ha un impatto sulle scuole, gli studenti e le loro famiglie perché offre un feedback importante che potrebbe essere impiegato nel processo di insegnamento-apprendimento, nella riprogettazione dell'istruzione e nei processi decisionali.

Parole chiave: Competenze Digitali, Valutazione Standardizzata, Livelli Soglia, Revisione Sistematica della Letteratura.

1. Introduzione

Il presente contributo si inserisce nel quadro degli studi nazionali e internazionali volti a definire e valutare le competenze digitali, con particolare attenzione a quelle degli studenti ai diversi gradi scolastici.

Le competenze digitali sono in cima all'agenda politica europea, con l'obiettivo di migliorarle per la trasformazione digitale. L'*European Skills Agenda* (European Commission, 2020a) promuove le competenze digitali e sostiene gli obiettivi del *Digital Education Action Plan 2021-2027* (European Commission, 2021) per lo sviluppo di un sistema di istruzione digitale ad alte prestazioni. Il *Digital Compass 2030* (European Commission, 2020b) e l'*European Pillar of Social Rights Action Plan* (European Commission, 2022a) fissano gli obiettivi di raggiungere almeno l'80% della popolazione con competenze digitali di base e 20 milioni di specialisti delle tecnologie dell'informazione e della comunicazione entro il 2030. In Italia la normativa vigente (PNRR) prevede che il Paese si doti di un sistema di certificazione/attestazione delle competenze digitali dal 2025.

È perciò necessario circoscrivere cosa si intenda per competenze digitali per poterle misurare. In questa prospettiva è stato avviato un progetto di ricerca nazionale per definire un modello prototipale per attestare le competenze digitali; riferimento sono gli studenti della II secondaria di secondo grado ma con la prospettiva di osservare l'evoluzione delle competenze digitali dal termine della secondaria di primo grado fino alla fine della secondaria di secondo grado.

Il quadro di riferimento assunto dai partner di progetto e dalla revisione sistematica qui proposta è DigComp 2.2 (*Digital Competence Framework for Citizens 2.2.*; Vuorikari et al., 2022), sviluppato dalla Commissione Europea per descrivere e valutare le competenze digitali dei cittadini dai 16 anni. Dal 2013 a oggi, DigComp ha trovato applicazione nel contesto dell'occupazione, dell'istruzione, della formazione e dell'apprendimento permanente; è stato adottato a livello europeo per costruire il *Digital Skills Indicator* (European Commission, 2022b) e per monitorare il *Digital Economy and Society Index*.

Nello specifico, il progetto di ricerca in atto, della cui prima fase si intende dare conto nel presente contributo, tra altri affronta il tema della definizione dei livelli di competenza digitale e delle soglie di adeguatezza corrispondenti ai diversi gradi scolastici.

Nelle indagini su larga scala in campo educativo, molte volte le variabili in questione sono rappresentate da abilità, conoscenze o competenze avute in una fase del percorso scolastico o in una determinata fascia di età, costruiti non direttamente osservabili ma definiti in base a un quadro teorico di riferimento e operazionalizzati per somministrare prove standardizzate.

Un esito comunicato come punteggio numerico, però, non fornisce informazioni circa ciò che gli studenti con uno specifico punteggio conoscono e sono in grado di fare rispetto al dominio indagato; ciò è un limite per chi è interessato a interpretare gli esiti di una rilevazione e a ricavarne informazioni per interventi o pratiche didattiche.

L'attribuzione di un livello esplicitamente descritto permette a studenti, famiglie e insegnanti di avere un *feedback* significativo, integrabile dallo studente alla propria percezione di competenza e utile agli insegnanti per la didattica. Molte indagini nazionali e internazionali affiancano a un esito in termini di punteggio una descrizione del livello corrispondente; analogamente fa il progetto in questione.

Il contributo presenta gli esiti di una revisione sistematica della letteratura; molteplici sono le motivazioni che giustificano tale studio: la mancanza di revisioni sul tema specifico dell'approccio pedagogico alla costruzione di scale a livelli descrittive i punteggi ottenuti nelle prove standardizzate, la necessità di offrire ai partner della ricerca un quadro il più esaustivo possibile circa i metodi e gli strumenti di valutazione delle competenze digitali e di definizione di una scala a livelli.

2. Framework teorico

In questa sezione sono riportati: una breve panoramica della concettualizzazione e dei *framework* esistenti sulle competenze digitali, una breve presentazione del *framework* DigComp 2.2 (Vuorikari et al., 2022), descrizione dei *competence based* test, descrizione di metodi e strumenti per la definizione delle scale a livelli. Inoltre, è fornita una motivazione per i criteri selezionati per documentare l'affidabilità e la definizione di un argomento di validità.

2.1 Panoramica della concettualizzazione e dei framework sulle competenze digitali

Per Siddiq, Hatlevik, Olsen, Throndsen, e Scherer (2016) esistono molteplici possibilità per descrivere i contenuti e le modalità con le quali gli studenti imparano e impiegano il digitale. Termini come competenza digitale (Calvani et al., 2012; Ferrari, 2013; Hatlevik et al., 2015), alfabetizzazione digitale (Aviram & Eshet-Alkalai, 2006; Bawden, 2001; Sefton-Green, et al., 2009), alfabetizzazione informatica (Fraillon, et al., 2013), competenze internet (Kuhlemeier & Hemker, 2007; Van Deursen & Van Dijk, 2009), alfabetizzazione ICT (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2012; Erstad, 2006, 2010), alfabetizzazione mediatica (Erstad, 2010), alfabetizzazione ai nuovi media (Lee, et al., 2015), alfabetizzazione di rete (Combes, 2009), fluidità tecnologica (Bunz et al., 2007) e competenze del XXI secolo (Binkley et al., 2012; Griffin et al., 2012; Partnership for 21st Century Skills [P21], 2012) sono spesso usati per descrivere il vasto ambito delle competenze digitali.

Per Siddiq et al. (2016) questi termini condividono un approccio che combina una prospettiva di dominio (digitale, ICT ecc.) con una prospettiva di conoscenza (alfabetizzazione, competenza, abilità). Sebbene a prima vista questi termini possano sembrare diversi, un'analisi più approfondita rivela molte somiglianze. Per esempio, i concetti di competenza digitale, alfabetizzazione ICT e competenze del XXI secolo enfatizzano il recupero e l'elaborazione delle informazioni, la produzione di contenuti, l'uso responsabile ed etico dell'ICT e gli aspetti comunicativi. In continuità con il *framework* DigComp, in questo contributo è impiegato il termine competenza digitale quale definizione comprensiva anche di altre; inoltre, si noti che le competenze del XXI secolo includono generalmente una gamma più ampia di competenze, con quella digitale spesso vista come parte integrante.

Tra altre, vi sono tre definizioni di competenza digitale che per diffusione e impiego in svariati studi è significativo riportare (Tab. 1).

Autore	Definizione
<i>Ministerial Council on Education, Employment, Training and Youth Affairs</i> (MCEETYA, 2005, p. 2)	The ability of individuals to use ICT appropriately to access, manage, integrate and evaluate information, develop new understandings, and communicate with others in order to participate effectively in society.
<i>Educational Testing Service</i> (ETS, 2007, p. 2)	Using digital technology, communications tools, and/or networks to access, manage, integrate, evaluate, and create information in order to function in a knowledge society.
<i>Framework DigComp</i> (Ferrari, 2013, p. 2)	The confident, critical and creative use of ICT to achieve goals related to work, employability, learning, leisure, inclusion and/or participation in society

Tab. 1: Principali definizioni di competenza digitale

Inoltre, Voogt e Roblin (2012) hanno condotto una revisione di 178 pubblicazioni dalle quali hanno estratto otto diverse tipologie di *framework* volti a descrivere le misure dei risultati delle capacità degli studenti di impiegare la tecnologia (Siddiq et al., 2016, p. 61): quattro sono nazionali (Acara, 2012; Calvani, et al., 2008; International Society for Technology in Education [ISTE], 2007; Korea Education and Research Information Service [KERIS], 2013; Norwegian Directorate for Education and Training [NDET], 2012) e quattro internazionali (Balanskat & Gertsch, 2010; Binkley et al., 2012; Ferrari, 2013; Fraillon, Ainley, Schulz, Friedman, & Gebhardt, 2014). Questi sono stati sintetizzati e sistematizzati dal *Joint Research Institute for Perspective Technological Studies* (JRC-IPTS) e sono confluiti nel *framework* DigComp.

2.2 Il Digital Competence Framework for Citizens (DigComp)

Come affermato da Siddiq et al. (2016), DigComp¹ è basato su una revisione completa di 15 *framework* di alfabetizzazione digitale, condotta con diversi strumenti tra cui: (a) una mappatura concettuale della competenza digitale (Ala-Mutka, 2011); (b) una raccolta e analisi di vari *framework* di competenza digitale tramite studi di caso (Ferrari, 2013); e (c) una sintesi delle opinioni sui componenti della competenza digitale da parte di *stakeholder* ed esperti rilevanti (Janssen et al., 2013).

Il presente contributo assume il *framework* DigComp come modello per categorizzare le dimensioni dell'alfabetizzazione digitale per diverse ragioni. In primo luogo, DigComp è fondato su una revisione approfondita e rigorosa dei *framework* esistenti. In secondo luogo, rappresenta uno dei tentativi più completi di delineare l'alfabetizzazione digitale e i suoi componenti specifici. In terzo luogo, pur essendo un *framework* generale applicabile a tutte le fasce d'età, affronta specificamente la questione dell'istruzione fornendo esempi di come ciascuna competenza possa essere applicata all'apprendimento. Inoltre, offre descrizioni dettagliate di ogni competenza, rendendolo adatto a varie applicazioni, come lo sviluppo di indicatori e test, e fornendo un riferimento equilibrato per l'analisi degli studi sulla valutazione dell'alfabetizzazione digitale (Siddiq et al., 2016). Infine, è il *framework* adottato dal progetto di riferimento.

Il *framework* DigComp definisce la competenza digitale specificando cinque livelli con una risoluzione concettuale crescente. Il primo livello identifica cinque aree: informazione, comunicazione, creazione di contenuti, sicurezza e *problem solving*. Il secondo introduce le competenze specifiche all'interno di ciascuna di queste cinque aree. Il terzo delinea un numero discreto di livelli, mentre il quarto delinea esempi di conoscenze, abilità e atteggiamenti rilevanti. Il quinto e ultimo fornisce un'elaborazione contestuale illustrando l'applicabilità di ciascuna competenza per vari scopi. La tabella 2 offre una panoramica delle aree e delle competenze correlate.

Aree di competenza	Competenze
<i>Alfabetizzazione su informazioni e dati</i>	Navigare, ricercare e filtrare dati, informazioni e contenuti digitali
	Valutare dati, informazioni e contenuti digitali
	Gestire dati, informazioni e contenuti digitali
<i>Comunicazione e collaborazione</i>	Interagire con gli altri attraverso le tecnologie
	Condividere informazioni attraverso le tecnologie digitali
	Esercitare la cittadinanza attraverso le tecnologie digitali
	Collaborare attraverso le tecnologie digitali
	Netiquette
	Gestire l'identità digitale
<i>Creazione di contenuti digitali</i>	Sviluppare contenuti digitali
	Integrare e rielaborare contenuti digitali
	Copyright e licenze
	Programmazione
<i>Sicurezza</i>	Proteggere i dispositivi
	Proteggere i dati personali e la privacy
	Proteggere la salute e il benessere
	Proteggere l'ambiente
<i>Risolvere i problemi</i>	Risolvere problemi tecnici
	Individuare bisogni e risposte tecnologiche
	Utilizzare in modo creativo le tecnologie digitali
	Individuare i divari di competenze digitali

Tab. 2: Aree e competenze del framework DigComp 2.2 (Vuorikari et al., 2022)

1 Il lavoro di elaborazione del *framework* è iniziato nel 2005 e la sua prima edizione risale al 2013; nel 2016 è stata pubblicata la seconda edizione con aggiornamenti della terminologia, del modello concettuale e con la proposta di esempi di implementazione. La versione 2.1 è stata pubblicata nel 2017 con una modifica ai livelli di padronanza che da 3 diventano 8. L'ultima versione è del 2022 (2.2) con l'integrazione di 250 esempi di conoscenze e atteggiamenti, il riferimento all'intelligenza artificiale e al lavoro ibrido/remoto.

2.3 Test basati sulla conoscenza e sulla competenza

Le revisioni sistematiche della letteratura condotte da Godaert, Aesaert, Voogt, & van Braak (2022) e Mattar, Ramos, & Margarida (2022) notano che la maggior parte degli strumenti costruiti per la rilevazione delle competenze digitali è di natura *self-report*; vi è perciò un numero inferiore di quelli volti a misurarle in modo standardizzato. Madaus e O'Dwyer (1999) e Darling-Hammond e Adamson (2010) descrivono tipologie di test finalizzati a elicitar l'abilità degli studenti nel generare risposte o creare prodotti. Mediante compiti realistici, tali strumenti permettono di misurare la capacità degli studenti di pensare e ragionare (Siddiq et al., 2016). Nelle medesime revisioni gli autori rilevano una prevalenza di valutazioni *computer-based* le quali permettono una maggiore interattività, una diversa struttura del test e degli item e la possibilità di dare automaticamente un punteggio rispetto a quelle *paper-based*. Inoltre, gli autori rilevano quattro categorie che descrivono la struttura dei compiti o degli item per la valutazione (tab. 3).

Categoria	Descrizione	Esempio
<i>MC</i>	a) Test che consistono prevalentemente in formati di risposta vincolati. b) Le informazioni fornite sono in gran parte statiche.	Test standardizzati a scelta multipla con determinate opzioni di risposta
<i>MC-INT</i>	a) Test che consistono prevalentemente in formati di risposta vincolati. b) Le informazioni fornite sono in larga misura dinamiche.	Gli stimoli del compito/item si basano sull'interattività o sulla necessità di interagire con gli ambienti di test per trovare la risposta giusta che è data come risposta selezionata.
<i>INT</i>	a) Test che in gran parte consistono di elementi che richiedono al compilatore di interagire con l'ambiente di test per fornire una risposta. b) Le informazioni sono fornite dinamicamente al compilatore.	Compiti in cui i compilatori devono impostare i valori di un certo numero di variabili di un dato sistema (per esempio, un telecomando artificiale), osservare i risultati e trarre conclusioni sulle relazioni tra le variabili.
<i>AUTH</i>	a) Prove in cui al compilatore sono forniti compiti che rappresentano una situazione pienamente autentica. b) Non è richiesta alcuna interazione con un ambiente di prova tipico.	Test in cui il compilatore riceve compiti da un amministratore e le sue azioni sono osservate e segnate durante la risoluzione dei compiti su un dispositivo digitale e senza interagire con un particolare ambiente di test computerizzato.

Tab. 3: Categorie per descrivere la tipologia di compiti/item (Siddiq, 2016, p. 63; trad. it. dello scrivente)

I test a scelta multipla sono considerati da alcuni autori (Stecher, 2010) incapaci di riflettere le azioni tipiche della vita reale ma Kane, Crooks, & Cohen (1999) asseriscono che anche scegliere tra diverse alternative può essere considerata una *performance* e perciò misurare una competenza.

2.4 Metodi e strumenti per la definizione di scale a livelli

Secondo Desimoni (2017), nel contesto internazionale si possono individuare diverse tipologie di livelli che differiscono sia per le modalità di identificazione sia per la loro concettualizzazione. Tra queste, emergono principalmente l'approccio dei livelli basati su *standard* e l'approccio dei livelli di competenza descrittivi delineati nelle scale di competenza descrittive o metriche di apprendimento. Entrambi i metodi comportano la divisione del *continuum* della variabile oggetto di studio in segmenti che rappresentano livelli di apprendimento o competenza, determinati da punteggi soglia che permettono di classificare i partecipanti in base alla loro *performance*.

Secondo l'autrice, nell'approccio basato sugli *standard*, si parte dalla descrizione di ciò che uno studente dovrebbe sapere e saper fare riguardo al dominio in esame, in una determinata fase del percorso scolastico o dello sviluppo di una specifica competenza (*content standards*). Questa descrizione è suddivisa in categorie ordinate, i livelli di performance (*performance levels*), delimitati da *standard* di performance (*performance standards*), che concettualmente corrispondono al livello minimo di abilità, conoscenze o competenze che un rispondente deve possedere per essere classificato in un certo livello. L'applicazione pratica di questi

standard a un test specifico sono i punteggi soglia (*cut-scores*), che permettono di determinare il passaggio tra diversi livelli e quindi la categorizzazione dei rispondenti. I livelli sono formalmente definiti tramite etichette (*performance level labels*; PLLs), per esempio «livello base», «livello intermedio», «livello avanzato», e sono accompagnati da descrittori che indicano in termini qualitativi ciò che ci si aspetta che un rispondente a quel livello conosca e sappia fare (*performance level descriptors*, PLDs). La definizione e descrizione di queste categorie sono solitamente curate da una commissione di esperti della materia, con possibile approvazione da parte di comitati esterni, e dovrebbero essere effettuate prima delle procedure conosciute come *standard setting* (Cizek & Bunch, 2007), mirate a tradurre operativamente le categorie in punteggi soglia (o *standard*) di transizione tra un livello e i livelli adiacenti, oppure come primo passo degli *standard setting* stessi da parte dei valutatori coinvolti.

Un secondo approccio rilevato dall'autrice per l'identificazione e la descrizione dei livelli di apprendimento o competenza nelle indagini su larga scala è quello delle metriche di apprendimento o scale di competenza descrittive (Turner, 2014), ossia scale di rilevazione di caratteristiche latenti riportate sia in termini numerici, come punteggi di competenza, sia in descrizioni di cosa comporti essere a quella posizione sul *continuum*. Nelle scale di competenza descrittive (*descriptive proficiency scales*; DPS), la linea continua della variabile latente indagata è concepita come rappresentazione di un costrutto latente graduabile, anche se non osservabile, che si rifà al concetto di apprendimento come una variabile in continua evoluzione, ipotizzando che le abilità, conoscenze e competenze in un certo punto della scala includano quelle dei punti precedenti del *continuum* (Turner, 2014). Desimoni afferma che le descrizioni delle DPS non riguardano ogni singolo punto della scala ma identificano livelli (*proficiency levels*) in cui il *continuum* è suddiviso, descrivendo i gradi di abilità, conoscenze e/o competenza (Masters & Forster, 1996; Turner, 2014). La crescente adozione di questo approccio alla costruzione dei livelli è associata alla diffusione dei modelli e metodi della *Item Response Theory* e del modello di Rasch (1960, 1980), grazie alla caratteristica fondamentale di questi modelli, che è la possibilità di esprimere sia la distribuzione della stima dell'abilità degli studenti sia la difficoltà degli item sulla stessa scala, rappresentante il *continuum* del tratto latente. Pertanto, osservando la posizione degli item sul *continuum* dell'abilità latente, è possibile determinare con quale probabilità uno studente a un determinato punto della scala supererà ciascun item e proprio sulla mappatura delle posizioni degli item sono basate le descrizioni dei livelli delle DPS.

3. Metodo

3.1 Metodologia della revisione sistematica

Green et al. (2011) definiscono la revisione sistematica come segue: « a systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made».

Tra altri, Gough, Oliver, & Thomas (2012) e Petticrew e Roberts (2006) propongono alcune procedure predefinite: formulazione delle domande di ricerca, ricerca e *screening* di studi secondo criteri di inclusione/esclusione, descrizione delle caratteristiche dello studio, valutazione e sintesi. Nello specifico, per il presente lavoro sono state adottate le procedure definite dal *PRISMA 2020 Statement* (Page et al., 2021).

Le domande di ricerca alle quali si intende trovare risposta sono:

- RQ1: quali sono e quali caratteristiche hanno i metodi e gli strumenti di valutazione standardizzata delle competenze digitali degli studenti?
- RQ2: quali sono e quali caratteristiche hanno le tecniche psicometriche e gli approcci pedagogici sottesi al passaggio dalla scala di misurazione ai livelli soglia?

3.2 Procedure di ricerca

Sono state individuate tre banche dati citazionali: ERIC, Scopus e Web of Science. La prima, *Education Resources Information Center*, è di natura eminentemente educativa e pedagogica ed è gestita dall'*Institute of Education Sciences* (IES) del Dipartimento U.S.A. dell'Istruzione; la seconda e la terza sono di proprietà rispettivamente delle società Elsevier e Clarivate. Sono stati inoltre consultati i siti istituzionali e le relative banche dati dell'UNESCO, dell'OCSE, della Commissione Europea, del Ministero italiano dell'Istruzione e del Merito e dell'Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione (INVALSI). La consultazione di piattaforme specifiche contenenti la letteratura grigia (OpenGrey² e OpenDOAR - Directory of Open-Access Repositories³) è stata condotta senza portare alcun risultato significativo.

La stringa di ricerca è stata individuata a seguito di una riflessione approfondita sul tema e sulle domande dello studio. Sono state indagate le competenze digitali degli studenti e la valutazione standardizzata; sono stati impiegati gli operatori booleani OR prima dei sinonimi e AND prima di differenti termini di ricerca, è stato applicato un troncamento ad alcune parole per intercettare le diverse forme delle parole (come *competence* vs. *competences* o *competencies*, è stata troncata in *comp**). La stringa risultante è "student*" AND "digital comp* OR skill*" AND "standardized assessment*". Inoltre sono stati aggiunti manualmente i seguenti filtri: decimo grado, scuola secondaria (escludendo *higher education*, formazione medica, scuole mediche, sviluppo professionale, studenti universitari), sono state escluse alcune aree disciplinari (ingegneria, medicina, salute, infermieristica, odontoiatria, neurologia, immunologia, matematica, *business*, ecologia, fisica, energia, scienze dei materiali, scienze della terra, agraria, farmacia, chimica e veterinaria) e alcune lingue (russo, portoghese, tedesco, estone, cinese, persiano, norvegese, cinese, turco, svedese, sloveno, coreano, croato e afrikaans). La ricerca è stata ristretta agli anni dal 2013 (anno di pubblicazione del DigComp) al 2024, con alcune eccezioni per testi significativi per la letteratura successiva; sono stati usati termini in inglese ma sono stati inclusi studi anche in altre lingue (italiano, francese e spagnolo) che avessero corrispondenza nel titolo, nelle parole chiave e nell'abstract. La ricerca è stata condotta nei mesi di marzo e aprile 2024.

3.3 Criteri di inclusione/esclusione

I seguenti criteri di prim'ordine per includere gli studi sono stati stabiliti *a priori*:

1. Definizioni, concettualizzazioni, misure e variabili chiave che richiama esplicitamente o implicitamente quelle del DigComp; la tipologia di competenze (sono escluse le computazionali, intanto, in quanto non presenti nel *framework* di riferimento);
2. la popolazione di riferimento e il rispettivo campionamento (decimo grado e campionamento probabilistico).

Poiché la RQ1 fa riferimento alla tipologia di valutazione standardizzata delle competenze digitali degli studenti, sono stati definiti i seguenti criteri:

3. la standardizzazione del test di rilevazione;
4. le tipologie di studio quantitative e/o qualitative che fanno riferimento alla teoria classica dei test, all'*Item Response Theory/Modello* di Rash, ai panel di esperti per la validazione degli item e l'attribuzione di un livello di difficoltà.

Per rispondere alla RQ2 (tecniche psicometriche e approcci pedagogici alla definizione dei livelli soglia) sono stati definiti i seguenti criteri:

2 <http://www.opengrey.eu>

3 <http://v2.sherpa.ac.uk/opensoar/>

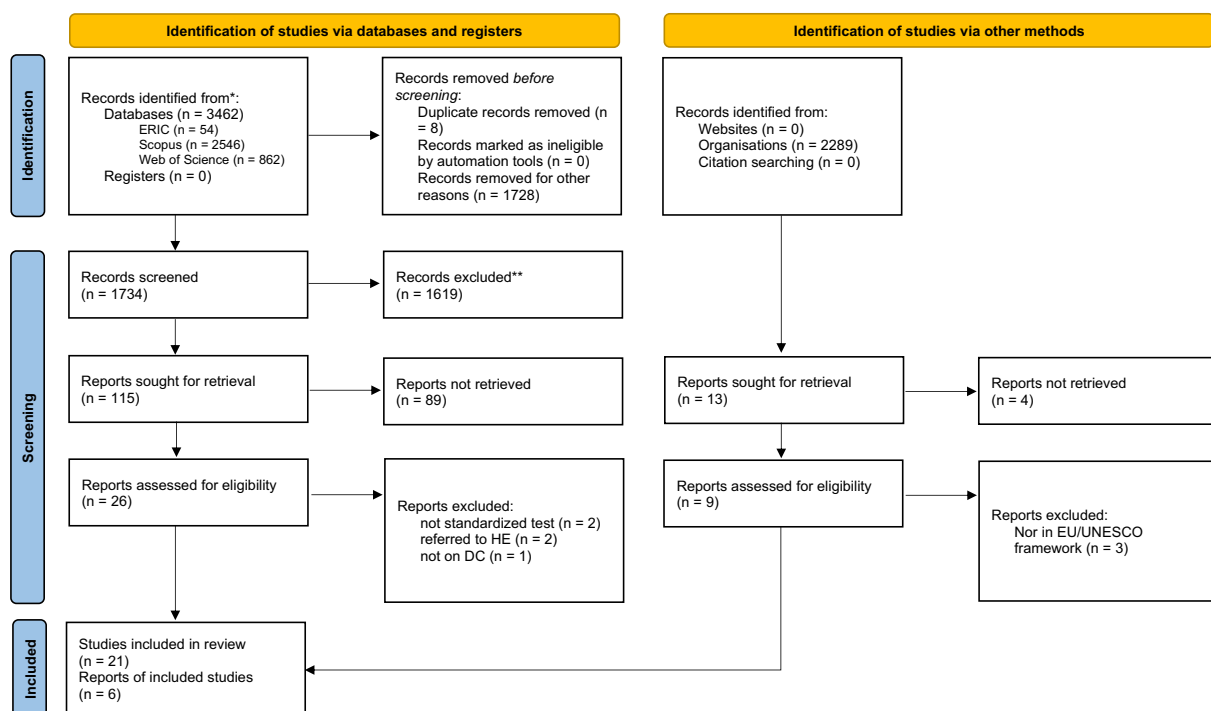
5. riferimento ai livelli di competenza degli esiti dei test, procedure di definizione;
6. la qualità dei dati (valori statisticamente significativi, controllo dei *bias*).

Sono stati inoltre aggiunti criteri di secondo ordine:

7. l'anno di pubblicazione dello studio (sono incluse ricerche precedenti alla pubblicazione del DigComp per la loro significatività per gli studi successivi),
8. la lingua (sono inclusi testi inglesi, italiani, francesi e spagnoli),
9. il processo di revisione tra pari (sono esclusi studi che non riportano la dicitura *peer review*).

3.4 Processi di ricerca e screening

Come mostrato nella fig. 1, la ricerca iniziale ha prodotto complessivamente 5.931 risultati: 3462 provenienti dai database citazionali (54 da ERIC, 2546 da Scopus, 862 da Web of Science) e 2289 dalle banche dati delle organizzazioni menzionate. Dopo aver applicato i filtri manuali sopra riportati sono stati rimossi 1728 articoli. Sono perciò stati scansionati 1734 titoli e abstract e, in seguito ad attenta lettura, ne sono stati esclusi 1619 e recuperati 115. La stessa operazione è stata condotta per le pubblicazioni delle Organizzazioni delle quali 13 sono state selezionate. In seguito a una prima lettura completa sono stati eletti 26 articoli e 9 pubblicazioni. Dopo una seconda lettura sono stati ridotti a 21 (poiché 2 non descrivevano test standardizzati, 2 erano riferiti all'*Higher Education* e 1 non era sulle competenze digitali) e 6 (perché le pubblicazioni non facevano riferimento diretto/indiretto al *framework* DigComp). 21 sono perciò gli articoli inclusi nella revisione mentre le 6 pubblicazioni sono state impiegate per definire il quadro teorico del presente lavoro.



*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/register).

**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: <http://www.prisma-statement.org/>

Fig. 1: Diagramma Prisma

3.5 Codifica ed estrazione dei dati

Ciascuno dei 21 articoli è stato codificato secondo uno schema definito dall'autore (<https://drive.google.com/file/d/112QYqVL3Ue8E1lGNrk4LYExDoEIQW5oL/view?usp=sharing>) sulla base della letteratura in materia di SLR (PRISMA Statement; Page et al., 2021). Ulteriori proprietà sono state indagate, quali le caratteristiche del test (*framework* di riferimento, formato/progettazione dell'item) e del campione (per es., paese, dimensione del campione, età o grado, ordine di scuola), dimensioni della competenza digitale misurata, prove di affidabilità e la definizione di un argomento di validità.

3.6 Indicatori di affidabilità e validità (proprietà psicometriche)

Sulla scorta di quanto realizzato da Siddiq et al. (2016) nella revisione sistematica degli strumenti di rilevazione delle competenze digitali negli studenti della scuola primaria e secondaria, sono stati estratti dagli articoli e sistematizzati alcuni indicatori di affidabilità e validità (https://drive.google.com/file/d/1CY5KTUbj7OXgVQsaWUETtXi09qknddqM/view?usp=drive_link). Tale azione è finalizzata a integrare la risposta alla seconda domanda di ricerca circa le tecniche psicometriche di costruzione delle scale di misurazione e dei livelli soglia. Gli autori propongono due categorie principali entro le quali raggruppare questi indicatori: prove di affidabilità e argomenti di validità.

In aggiunta ai criteri proposti è stata rilevata la modalità di individuazione e di concettualizzazione delle tipologie di livelli; nello specifico, se sia stato impiegato l'approccio *standard-referenced* o i *descriptive proficiency levels* delineati nelle *descriptive proficiency scales* o *learning metrics* (Desimoni, 2017).

3.7 Approcci pedagogici

La seconda domanda di ricerca prevede anche l'individuazione dei principali approcci pedagogici sottesi alla costruzione di scale a livelli. In letteratura non sono stati rinvenuti studi di questa natura; perciò, nella fase di estrazione dei dati sono stati considerati i riferimenti alle ricadute sui processi di insegnamento-apprendimento, sulla riprogettazione didattica, sulla valutazione *del* e *per* l'apprendimento, sull'impiego di *feedback* per gli studenti e le loro famiglie, sulle indicazioni per i *policy maker*.

4. Risultati

In questo paragrafo sono presentati i principali risultati della revisione sistematica, nello specifico sono mostrati quelli rispondenti alle domande di ricerca.

4.1 RQ1: quali sono e quali caratteristiche hanno i metodi e gli strumenti di valutazione standardizzata delle competenze digitali degli studenti?

La tabella 4 riporta sinteticamente alcune caratteristiche dei 21 studi inclusi nella revisione sistematica.

Prevalentemente sono stati condotti nel periodo pandemico e post-pandemico (16 studi). L'area geografica predominante è quella europea, in particolare spagnola (n=9) e centro-nord continentale (n=5); vi sono ricerche anche dell'area asiatica (n=3) e sudamericana (n=2). Non vi sono studi in lingua italiana o condotti in Italia.

Il *framework* di riferimento che ricorre maggiormente è il DigComp (13 studi), seguito dal *21st Century Skills* (n=2) e da quello UNESCO (n=1); gli altri studi si fondano su quadri di riferimento esito di ricerche o di studi nazionali (n=5).

Molti studi indagano tutte le aree del *framework* (n=5), altri solo alcune (*problem solving* n=3, sicurezza n=1, *literacy* n=2, responsabilità digitale n=2).

La maggior parte dei lavori è stata condotta con approcci quantitativi (n=17), 2 qualitativi e 2 mixed

method. Per quanto riguarda i primi, il disegno prevalente è la validazione psicometrica del test mentre per i secondi si tratta di validazione del contenuto, per i terzi entrambi. Alcuni studi sono condotti facendo riferimento all'*Item Response Theory* e al modello di Rash ($n=3$) e descrivono la realizzazione di *Item bank* e forme di test ($n=2$) e/o lo studio del funzionamento differenziale degli item ($n=2$); altri mediante metodologie più tradizionali quali l'analisi fattoriale esplorativa ($n=1$) e/o confermativa ($n=2$) per verificare la corrispondenza tra i fattori latenti e il riferimento teorico sul quale sono stati costruiti gli strumenti.

Composto da un test oggettivo di 16 item e da una scala di atteggiamenti di 17 item, strumento ricorrente è ECODIES il quale è oggetto di studio da parte di numerose ricerche (E1, E2, S2, S9, S12).

Autori	Anno	Area geografica	Framework di riferimento	Area del framework	Tipologia di studio
<i>Cabezas-González, Casillas-Martín, & García-Valcárcel Muñoz-Repiso (E1)</i>	2022	Europa (Spagna)	DigComp	Problem solving	Quantitativo (con design cross-sectional)
<i>García-Valcárcel Muñoz-Repiso, Casillas Martín, & Basilotta Gómez-Pablos (E2)</i>	2020	Europa (Spagna)	DigComp	Tutte	Qualitativo (validazione del contenuto)
<i>Claro et al. (WOS1)</i>	2012	Sud America (Cile)	21 st Century Skills	ND	Quantitativo (validazione psicometrica)
<i>Hernández-Martín, Martín-del-Pozo, & Iglesias-Rodríguez (WOS2)</i>	2021	Europa (Spagna)	DigComp	Sicurezza	Quantitativo
<i>Pedaste, Kallas, & Baucal (S1)</i>	2023	Europa (Estonia)	DigComp	Tutte	Quantitativo (validazione psicometrica)
<i>Cabezas-González, Casillas-Martín, & García-Valcárcel Muñoz-Repiso (S2)</i>	2023	Europa (Spagna)	DigComp	Problem solving	Quantitativo (con modelli di mediazione)
<i>Ackermans et al. (S3)</i>	2023	Europa (Olanda)	ECC-ICT	Basic, media, information e CT skills	Mixed method
<i>Rocha, Muñoz-Repiso, & Costa (S4)</i>	2023	Europa (Spagna)	DigComp	ND	Mixed method
<i>Turpo-Gebera et al. (S5)</i>	2023	Sud America (Perù)	UNESCO	ND	Quantitativo (design non sperimentale e descrittivo)
<i>Fernandes, Matos, & Festas (S6)</i>	2022	Europa (Portogallo)	DigComp	Tutte	Quantitativo (design non sperimentale)
<i>Baeza-González, Lázaro-Cantabrana, & Sanromà-Giménez (S7)</i>	2022	Europa (Spagna)	DigComp	Tutte	Quantitativo
<i>Pan, Reichert, de la Torre, & Law (S8)</i>	2022	Asia (Hong Kong)	ND	ND	Quantitativo
<i>Fernández-Bringas, Sandoval-Arteta, Suárez-Guerrero, & Mercado (S9)</i>	2022	Europa (Spagna)	DigComp	Alfabetizzazione su informazione e dati	Quantitativo (validazione psicometrica)
<i>Cabezas-González, Casillas-Martín, García-Valcárcel Muñoz-Repiso & Basilotta-Gómez-Pablos (S10)</i>	2021	Europa (Spagna)	DigComp	Problem solving	Qualitativo (validazione del contenuto)
<i>Reichert et al. (S11)</i>	2020	Asia (Hong Kong)	DigComp	ND	Quantitativo (studio di secondo livello)
<i>Casillas-Martín, Cabezas-González, García-Valcárcel Muñoz-Repiso (S12)</i>	2020	Europa (Spagna)	DigComp	Tutte	Quantitativo (validazione psicometrica)
<i>Kaarakainen, Kivinen, & Vainio (S13)</i>	2018	Europa (Finlandia)	DigComp	ND	Quantitativo (validazione psicometrica)

<i>Li & Ranieri (S14)</i>	2010	Asia (Cina)	Calvani et al. (2008)	Tecnologica, cognitiva ed etica	Quantitativo
<i>Seufert, Stanoevska-Slabeva, & Guggemos (S15)</i>	2020	Europa (paesi tedescofoni)	Stanoevska-Slabeva et al. (2015)	Bisogni, fonti, strategie di accesso e ricerca, valutazione, uso, presentazione, processo e riflessione sui risultati circa le informazioni	Quantitativo (validazione psicometrica)
<i>Hatlevik & Radtke (S16)</i>	2017	Europa (Norvegia)	Curriculum nazionale	Responsabilità digitale	Quantitativo (definizione dei livelli di competenza)
<i>Siddiq, Gochyyev, & Wilson (S17)</i>	2017	Europa (Norvegia)	21 st Century Skills	Alfabetizzazione all'informazione, ICT e responsabilità personale e sociale	Quantitativo
<i>Legenda. E: database ERIC; WOS: database Web of Science; S: database Scopus; ND: non disponibile.</i>					

Tab. 4: Descrizione sintetica degli studi inclusi nella systematic review

La tabella 5 approfondisce la popolazione di riferimento dello studio e la tipologia di campionamento. Sono indicati lo studio, l'età dei partecipanti, il grado, la numerosità per ordine scolastico o per esperti coinvolti e la tipologia di campionamento. Non vi sono studi esclusivamente rivolti al decimo grado, come da progetto; perciò, sono stati inclusi gli studi su un *range* di età che fosse riconducibile a esso. In alcuni studi non è dichiarata l'età e in altri il grado; il numero di persone componenti il campione varia da 175 (S17) a 3847 (E1); non tutti gli studi dichiarano la tipologia di campionamento (4 di natura probabilistica/casuale; 2 non probabilistici di convenienza).

Studio	Età	Grado	Primaria	Secondaria	Esperti	Campionamento
<i>E1</i>	Da 12 a 14		4175	3847	77	Casuale
<i>E2</i>	15 e 16					
<i>WOS1</i>	15			1185		Probabilistico stratificato
<i>WOS2</i>	Da 11 a 13			595		Casuale
<i>S1</i>	15 e 16	9		836		
<i>S2</i>	Da 12 a 14			772		
<i>S3</i>	Da 7 a 12	Da 2 a 6		575		
<i>S4</i>		Da 7 a 9		810		
<i>S5</i>		Da 9 a 11		1250		Non probabilistico di natura incidentale, basato sull'accessibilità dei soggetti
<i>S6</i>		4, 6 e 9		400		
<i>S7</i>			187			
<i>S8</i>		7 e 9				
<i>S9</i>	Da 15 a 17			201		Non probabilistico di convenienza
<i>S10</i>	Da 12 a 14			595		Probabilistico stratificato
<i>S11</i>		5 e 8	1320	1302		
<i>S12</i>						
<i>S13</i>	Da 12 a 22	Da 7 a 9		3159		
<i>S14</i>		9		350		Casuale
<i>S15</i>		Da 9 a 13		432		
<i>S16</i>		7		1026		
<i>S17</i>	15	9		175		

Tab. 5: Popolazione e campione

4.2 RQ2: quali sono e quali caratteristiche hanno le tecniche psicometriche e gli approcci pedagogici sottesi al passaggio dalla scala di misurazione ai livelli soglia?

La figura 2 mostra la tipologia e la definizione dei livelli di competenza così come riportati negli studi. Prevalentemente, anche se con talune differenze, i livelli definiti sono 3; ciò che varia sono le etichette attribuite e le modalità di individuazione (n. *logit*, punteggi ottenuti, percentuali di risposte corrette). Due studi fanno esplicito riferimento ai livelli definiti dal DigComp.

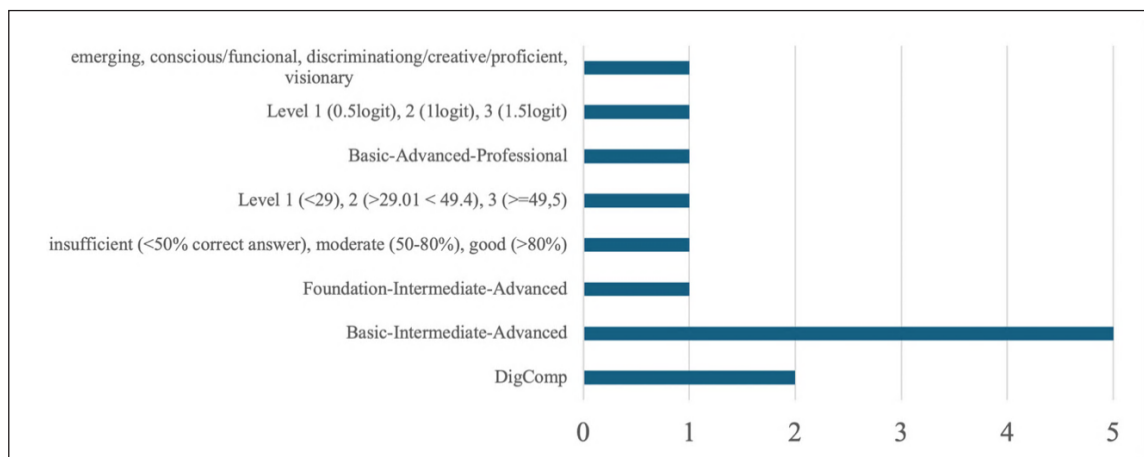


Fig. 2: Livelli di competenza

In https://drive.google.com/file/d/13qm4S9wvtOpUMLxJCM-cptLAIKMKSCtS/view?usp=drive_link sono riportate le evidenze di affidabilità e gli argomenti di validità. Per quanto riguarda i primi, 14 studi descrivono coefficienti di affidabilità (per es. di Cronbach, affidabilità inter-valutatore ecc.) e 15 lo sviluppo dei livelli di *performance* impiegando specifiche e riconosciute metodologie. A tal proposito e per rispondere alla seconda domanda di ricerca sono state estratte le seguenti informazioni: 5 studi fanno riferimento agli approcci *standard referenced* (E1, E2, S2, S10, S12), 10 a quelli *descriptive proficiency levels* (WOS2, S1, S3, S6, S7, S8, S13, S15, S16, 17). Tra questi ultimi, 2 citano esplicitamente le metodologie Angoff (S7, S16), 1 Bookmark (S16) e 1 PISA (S15). 19 studi riportano la difficoltà degli item e 13 anche la discriminatività; 16 offrono informazioni circa la modalità con cui sono stati dati i punteggi agli item.

Per quanto riguarda gli argomenti di validità, 20 studi fanno riferimento esplicito a un *framework* teorico e indicano la modalità con la quale hanno impiegato il test; 6 riportano di aver condotto uno studio pilota prima di quello principale e ne indicano le caratteristiche e gli esiti; 11 offrono anche informazioni qualitative per le strategie di valutazione (panel di esperti, interviste, protocolli *think-aloud* e Delphi ecc.); la comunicazione del contenuto del test è avvenuta principalmente esplicitando il *framework* e le competenze indagate (n=16) e solo in alcuni casi (n=5) sono stati dati limitati esempi degli item; le strategie di campionamento sono state descritte solo in 8 casi; la correttezza del test (n=3), la dimensionalità (n=2), le relazioni predette con altre variabili (n=7) sono state precisate in pochi studi; infine, per quanto riguarda la progettazione degli item/compiti, 5 test sono costruiti con stimoli e format di risposta interattivi, 6 a risposta multipla, 2 con entrambi e 3 in ambiente completamente e autenticamente digitale.

Di grande rilevanza è la completa assenza di riferimenti ad approcci e finalità di natura pedagogica alla definizione dei livelli di prestazione.

5. Discussione

La revisione sistematica è finalizzata a descrivere i metodi e gli strumenti che la letteratura scientifica offre in merito alla valutazione standardizzata delle competenze digitali con particolare attenzione alla costruzione delle scale a livelli. Oltre a un angolo visuale descrittivo, volto a mettere a disposizione dati utili al progetto, tale studio intende avviare una riflessione di natura pedagogica.

5.1 Valutazione standardizzata delle competenze digitali: tra tecnica e ricadute sui processi formativi (RQ1)

La revisione condotta segnala alcune questioni di fondo. I test standardizzati per la valutazione delle competenze digitali sono in numero inferiore rispetto a quelli *self-report*; in fase di *screening* dei titoli e degli abstract tale aspetto è parso evidente. Ciò può essere spiegato dalla complessità e dalla tecnicità richieste dalla costruzione di strumenti validi e affidabili. La distribuzione per regioni geografiche non è uniforme, vi è la prevalenza di studi di area ispanofona e centro-nordestina, probabilmente a motivo dell'adozione del DigComp quale principale *framework* di riferimento. A questo proposito, è interessante notare che tale quadro teorico sembra raggiungere l'obiettivo prefissato di dare la possibilità a studiosi ed enti di ricerca di sviluppare strumenti di valutazione. Il suo influsso è osservabile anche nella scelta delle aree specifiche di competenza poiché anche gli strumenti definiti su altri quadri, in ultima istanza, possono essere ricondotti a quelle del DigComp. Vi è, inoltre, una prevalenza di studi di natura quantitativa (psicometrica) rispetto a quelli *mixed-method* o qualitativa; seppur ovvio, poiché oggetto dello studio era la valutazione standardizzata, è un risultato degno di nota che richiama l'attenzione sulla docimologia come branca pedagogica. Altro elemento interessante in termini di ricerca è la prevalenza di test basati su item a risposta multipla, talora inseriti in un ambiente interattivo, a fronte di una minor presenza di test in ambienti autentici; ciò richiede ulteriori riflessioni circa l'orientamento pedagogico alla valutazione delle competenze digitali. Quanto detto si estende anche all'angolo visuale degli studi inclusi nella revisione sistematica: un approccio pedagogico al tema sembra non essere adeguatamente sviluppato. Nello specifico non è dato congruo spazio alle ricadute sui processi educativi e di istruzione e alla funzione formativa che anche la valutazione standardizzata può avere.

5.2 Descrivere i livelli per spiegare i punteggi (RQ2)

L'interesse dal quale scaturiscono il progetto di ricerca e questa fase di studio è la descrizione esplicita dei livelli di padronanza delle competenze digitali degli studenti del decimo grado scolastico. La revisione sistematica ha offerto numerose informazioni circa le tecniche psicometriche sottese a questo processo. In prima istanza sono osservabili riferimenti espliciti e impliciti al DigComp, prevalentemente nelle etichette assegnate (base, intermedio, avanzato) e seppur pochi studi comprendono anche il livello "altamente specializzato". Nella maggior parte degli studi inclusi nella revisione sono state descritte in modo approfondito le tecniche principali (Angoff, Bookmark ecc.) e vi è esplicito riferimento all'*Item Response Theory* e al modello di Rash ma anche in questo caso la dimensione pedagogica non trova adeguato spazio. In nessun articolo sono sviluppate riflessioni circa la descrizione dei livelli per spiegare i punteggi ottenuti al test, legata alla possibilità per alunni, famiglie e docenti di avere un *feedback* significativo, integrabile dallo studente alla propria percezione di competenza e utile agli insegnanti per la didattica.

6. Conclusioni

I risultati della revisione sistematica e la relativa discussione hanno evidenziato che la maggior parte degli studi è di natura quantitativa e che l'approccio principale al tema è prevalentemente psicometrico. Ciò può essere spiegato in parte dall'impiego di taluni termini di ricerca (valutazione, standardizzazione ecc), altresì rivela che queste parole sono considerate semanticamente più affini a questi angoli visuali. Pur richiedendo ulteriori approfondimenti, vi potrebbe essere una questione di indicizzazione nelle più diffuse banche dati citazionali delle riviste, e quindi dagli articoli, con taglio pedagogico. Ancorché limitati e parziali, tali esiti richiedono una riflessione alla comunità scientifica educativa poiché, come segnalato dalla letteratura pedagogica in tema di valutazione e delle conseguenti ricadute sul processo di insegnamento-apprendimento, tale questione non può essere ignorata. Ulteriore elemento di riflessione è la tipologia di test emersi dalla revisione; solo alcuni intendono valutare le competenze digitali in ambienti autentici a fronte di molti altri che impiegano item a risposta multipla, seppur inseriti in contesti digitalmente interattivi. Ciò richiede e apre ulteriori prospettive di riflessione sui metodi e gli strumenti di valutazione delle competenze (digitali).

Riferimenti bibliografici

- Australian Curriculum, Assessment and Reporting Authority (2012). *National assessment program - ICT literacy years 6 & 10 report 2011*. Sidney: Australian Curriculum, Assessment and Reporting Authority. Retrieved July 5, 2024, from http://www.nap.edu.au/verve/_resources/nap_ictl_2011_public_report_final.pdf.
- Ackermans, K., Bakker, M., Gorissen, P., van Loon, A. -M., Kral, M., & Camp, G. (2023). Development and validation of a test for measuring primary school students' effective use of ICT: The ECC-ICT test. *Journal of Computer Assisted Learning*, 1-13. <https://doi.org/10.1111/jcal.12924>.
- Ala-Mutka, K. (2011). *Mapping digital Competence: Towards a conceptual understanding*. Seville. European Commission, Joint Research Centre. Institute for Prospective Technological Studies. Retrieved July 5, 2024, from <http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=4699>.
- Aviram, R., & Eshet-Alkalai, Y. (2006). Towards a theory of digital Literacy: Three scenarios for the next steps. *European Journal of Open Distance Learning*. Retrieved July 5, 2024, from http://www.eurodl.org/materials/contrib/2006/Aharon_Aviram.htm.
- Baeza-González, A., Lázaro-Cantabrana, J. L., & Sanromà-Giménez, M. (2022). Evaluación de la competencia digital del alumnado de ciclo superior de primaria en Cataluña [Assessment of the digital competence of pupils in upper primary school in Catalonia]. *Pixel-Bit. Revista de Medios y Educación*, 64, 265-298. <https://doi.org/10.12795/pixelbit.93927>.
- Balanskat, A., & Gertsch, C. A. (2010). *Digital skills working Group. Review of national curricula and assessing digital competence for students and teachers: Findings from 7 countries*. Brussels: European Schoolnet.
- Bawden, D. (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, 57, 218-259.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., et al. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 17-66). Amsterdam: Springer.
- Bunz, U., Curry, C., & Voon, W. (2007). Perceived versus actual computer-email-web fluency. *Computers in Human Behavior*, 23, 2321-2344. <http://dx.doi.org/10.1016/j.chb.2006.03.008>.
- Cabezas-González, M., Casillas-Martín, S., & García-Valcárcel Muñoz-Repiso, A. (2022). Mediation Models Predicting the Level of Digital Competence of 12–14 Year-Old Schoolchildren in the Area of Digital Problem Solving. *Journal of New Approaches in Educational Research*, 11(2), 168-185. doi:10.7821/naer.2022.7.789.
- Cabezas-González, M., Casillas-Martín, S., & García-Valcárcel Muñoz-Repiso, A. (2023). Theoretical Models Explaining the Level of Digital Competence in Students. *Computers*, 12, 100. <https://doi.org/10.3390/computers12050100>.
- Cabezas-González, M., Casillas-Martín, S., García-Valcárcel Muñoz-Repiso, A. & Basilotta-Gómez-Pablos, V. (2021). Psycho-Technical Study of a Digital Competence Assessment Tool for Problem-Solving Skills of Compulsory Education Students. *Revista Electrónica Educare (Educare Electronic Journal)*, 25(3), 1-21.
- Calvani, A., Cartelli, A., Fini, A., & Ranieri, M. (2008). Models and instruments for assessing digital competence at school. *Journal of E-learning and Knowledge Society*, 4, 183-193.
- Calvani, A., Fini, A., Ranieri, M., & Picci, P. (2012). Are young generations in secondary school digitally competent? A study on Italian teenagers. *Computers & Education*, 58, 797-807. <http://dx.doi.org/10.1016/j.compedu.2011.10.004>.
- Casillas-Martín, S., Cabezas-González, M., & García-Valcárcel Muñoz-Repiso A. (2020). Psychometric analysis of a test to assess the digital competence of compulsory education students. *Relieve*, 26(2), art. 2. <http://doi.org/10.7203/relieve.26.2.17611>.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard Setting. A guide to Establishing and Evaluating Performance Standards on Tests*. SAGE Publications.
- Claro, M., Preiss, D. D., San Martín, E., Jara, I., Hinostroza, J. E., Valenzuela, S. et al. (2012). Assessment of 21st century ICT skills in Chile: Test design and results from high school level students. *Computers & education*, 59(3), 1042-1053. <https://doi.org/10.1016/j.compedu.2012.04.004>.
- Combes, B. (2009). Generation Y: Are they really digital natives or more like refugees? *Synergy*, 7, 31-40.
- Darling-Hammond, L., & Adamson F. (2010). *Beyond basic skills: The role of performance assessment in achieving 21st century standards of learning*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education. Retrieved July 5, 2024, from <https://scale.stanford.edu/system/files/beyondbasic-skills-role-performance-assessment-achieving-21st-century-standards-learning.pdf>.
- Desimoni, M. (2017). *I livelli per la descrizione degli esiti delle prove INVALSI. Le rilevazioni degli apprendimenti A.S. 2017-2018*. Retrieved July 5, 2024, from https://www.researchgate.net/profile/Marta-Desimoni/publication/326416324_I_LIVELLI_PER_LA_DESCRIZIONE_DEGLI_ESITI DELLE_PROVE_INVALSI/links/5b4c716fa6fdccadaecf739f/I-LIVELLI-PER-LA-DESCRIZIONE-DEGLI-ESITI-DELLE-PROVE-INVALSI.pdf.

- Erstad, O. (2006). A new direction? Digital literacy, student participation and curriculum reform in Norway. *Education and Information Technologies*, 11, 415-429. <http://dx.doi.org/10.1007/s10639-006-9008-2>.
- Erstad, O. (2010). Educating the digital generation. *Nordic Journal of Digital Literacy*, 5, 56-71. Retrieved July 5, 2024, from <http://www.idunn.no/ts/dk/2010/01/art05>.
- Educational Testing Service (2007). *Digital transformation: A framework for ICT literacy. A report of the international ICT literacy panel*. Princeton: Center for Global Assessment. Retrieved July 5, 2024, from https://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ictreport.pdf.
- European Commission (2020a). *European Skills Agenda*. Retrieved July 5, 2024, from <https://ec.europa.eu/social/main.jsp?catId=1223&langId=en>.
- European Commission (2020b). *Digital Compass 2030*. Retrieved July 5, 2024, from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_it.
- European Commission (2021). *Digital Education Action Plan 2021-2027*. Retrieved July 5, 2024, from <https://education.ec.europa.eu/it/focus-topics/digital-education/action-plan>.
- European Commission (2022a). *European Pillar of Social Rights Action Plan*. Retrieved July 5, 2024, from <https://ec.europa.eu/social/main.jsp?catId=1607&langId=en>.
- European Commission (2022b). *Digital Skills Indicator*. Retrieved July 5, 2024, from <https://digital-skills-jobs.europa.eu/en/inspiration/resources/digital-skills-indicator-20-measuring-digital-skills-across-eu>.
- Fernandes, P., Matos, A. P. M., & Festas, I. (2022). Methodology for the evaluation of media literacy in children and youth. Recommendations and trends. *Profesional de la información*, 31(6), e310616. <https://doi.org/10.3145/epi.2022.nov.16>.
- Fernández-Bringas, T., Sandoval-Arteta, F., Suárez-Guerrero, C., & Mercado, G. O. (2022). Psychometric validation of the information area of digital competence in high school students in Perú. *Educational Process: International Journal*, 11(4), 53-68.
- Ferrari, A. (2013). DIGCOMP: A framework for developing and understanding digital competence in Europe. In Y. Punie, B.N. Brecko (Eds.), *JRC scientific and policy reports*. Luxembourg: Publications Office of the European Union. <http://dx.doi.org/10.2788/52966>.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). *Preparing for life in a digital age. The IEA international computer and information literacy study, international report*. IEA: Springer Open.
- Fraillon, J., Schulz, W., & Ainley, J. (2013). *International computer and information literacy Study: Assessment framework*. Amsterdam: IEA.
- García-Valcárcel Muñoz-Repiso, A., Casillas Martín, S., & Basilotta Gómez-Pablos, V. (2020). Validation of an Indicator Model (INCODIES) for Assessing Student Digital Competence in Basic Education. *Journal of New Approaches in Educational Research*, 9(1), 110-125. doi: 10.7821/naer.2020.1.459.
- Godaert, E., Aesaert, K., Voogt, J., & van Braak, J. (2022). Assessment of students' digital competences in primary school: a systematic review. *Education and Information Technologies*, 27. 10.1007/s10639-022-11020-9.
- Gough, D., Oliver, S., & Thomas, J. (2012). *An introduction to systematic reviews*. London: SAGE.
- Green, S., Higgins, J. P. T., Alderson, P., Clarke, M., Mulrow, C. D., & Oxman, A. D. (2011). What is a systematic review? In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (version 5.1). The Cochrane collaboration (chapter 1). Retrieved July 5, 2024, from <https://training.cochrane.org/handbook/archive/v5.1/>.
- Griffin, P., McGaw, B., & Care, E. (2012). *Assessment and teaching of 21st century skills*. Melbourne, Australia: Springer. <http://dx.doi.org/10.1007/978-94-007-2324-5>.
- Hatlevik, O. E. & Radtke, I. (2017). Standard Setting in a Formative Assessment of Digital Responsibility Among Norwegian Eighth Graders. In S. Blömeke & J.-E. Gustafsson (ed). *Standard Setting in Education. The Nordic Countries in an International Perspective*. Springer International.
- Hatlevik, O. E., Gudmundsdottir, G., & Loi, M. (2015). Digital diversity among upper secondary students: A multilevel analysis of the relationship between cultural capital, self-efficacy, strategic use of information and digital competence. *Computers & Education*, 81, 345-353. <http://dx.doi.org/10.1016/j.compedu.2014.10.019>.
- Hernández-Martín, A., Martín-del-Pozo, M., & Iglesias-Rodríguez, A. (2021). Pre-adolescents' digital competences in the area of safety. Does frequency of social media use mean safer and more knowledgeable digital usage?. *Education and Information Technologies*, 26, 1043-1067 <https://doi.org/10.1007/s10639-020-10302-4>
- International Society for Technology in Education (2007). *The national educational technology standards and performance indicators for students*. Retrieved July 5, 2024, from <http://www.iste.org/standards/ISTE-standards/standards-for-students>.
- Janssen, J., Stoyanov, S., Ferrari, A., Punie, Y., Pannekeet, K., & Sloep, P. (2013). Experts' views on digital competence: Commonalities and differences. *Computers & Education*, 68, 473-481. <http://dx.doi.org/10.1016/j.compedu.2013.06.008>.

- Karakainen, M. -T., Kivinen, O., & Vainio, T. (2018). Performance-based testing for ICT skills assessing: a case study of students and teachers' ICT skills in Finnish schools. *Univ Access Inf Soc*, 17, 349-360 <https://doi.org/10.1007/s10209-017-0553-9>.
- Kane, M. T., Crooks, T. J., & Cohen, A. S. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, 18, 5-17.
- Korea Education and Research Information Service (2013). *White paper on ICT in education Korea*. The Ministry of Education and Korea Education & Research Information Service. Retrieved July 5, 2024, from http://english.keris.or.kr/whitepaper/WhitePaper_eng_2013.pdf.
- Kuhlemeier, H., & Hemker, B. (2007). The impact of computer use at home on students' Internet skills. *Computers & Education*, 49, 460-480.
- Lee, L., Chen, D. -T., Li, J. -Y., & Lin, T. -B. (2015). Understanding new media literacy: The development of a measuring instrument. *Computers & Education*, 85, 84-93. <http://dx.doi.org/10.1016/j.compedu.2015-02.006.0360e1315>.
- Li, Y. & Ranieri, M. (2010). Are 'digital natives' really digitally competent? A study on Chinese teenagers. *British Journal of Educational Technology*, 41(6), 1029-1042.
- Madaus, G. F. & O'Dwyer, L. M. (1999). A short history of performance assessment: Lessons learned. *Phi Delta Kappan*, 80, 688-695.
- Masters, G. N. & Forster, M. (1996). *Developmental Assessment*. Camberwell, Australia: Australian Council for Educational Research.
- Mattar, J., Ramos, D., & Margarida, L. (2022). DigComp-Based Digital competence Assessment Tools: Literature Review and Instrument Analysis. *Education and Information Technologies*, 27, 1-25. [10.1007/s10639-022-11034-3](https://doi.org/10.1007/s10639-022-11034-3).
- Ministerial Council on Education, Employment, Training and Youth Affairs (2005). *National assessment program information and communication technology literacy 2005, years 6 and 10: An assessment domain for ICT literacy*. Carlton: Curriculum Corporation. Retrieved July 5, 2024, from https://www.iste.org/docs/pdfs/australia_ict_assessment.pdf?sfvrsn=2.
- Norwegian Directorate for Education and Training (2012). *Framework for basic skills*. Retrieved July 5, 2024, from http://www.udir.no/PageFiles/66463/FREAMEWORK_FOR_BASIC_SKILLS.pdf?epslanguage=no.
- Partnership for 21st Century Skills (2012). *Learn for the 21st century. A report and mile guide for 21st century skills. Partnership for 21st Century Skills*. Retrieved July 5, 2024, from http://www.p21.org/storage/documents/P21_Report.pdf.
- Page, M. J., McKenzie, J. E, Bossuyt, P. M., Boutron, I., Hoffmann, T. C, Mulrow, C. D. et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372(71). doi:10.1136/bmj.n71
- Pan, Q., Reichert, F., de la Torre, J., & Law, N. (2022). Measuring Digital Literacy during the COVID-19 Pandemic: Experiences with Remote Assessment in Hong Kong. *Educational Measurement: Issues and Practice*, 41(1), 46-50.
- Pedaste, M., Kallas, K., & Baucal, A. (2023). Digital competence test for learning in schools: Development of items and scales. *Computers & Education*, 203, 104830.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences. A practical guide*. Malden, USA: Blackwell.
- Rasch, G. (1960, 1980). *Probabilistic models for some intelligence and attainment tests* (Copenhagen, Danish Institute for Educational Research), expanded edition (1980) with foreword and afterword by B.D. Wright. Chicago: The University of Chicago Press.
- Reichert, F., Zhang, J., Law, N. W. Y. et al. (2020). Exploring the structure of digital literacy competence assessed using authentic software applications. *Education Tech Research Dev*, 68, 2991-3013. <https://doi.org/10.1007/s11423-020-09825-x>.
- Rocha, M. T. D., Muñoz-Repiso, A. G. -V., & Costa, E. (2023). ICT Skills – Study concerning students from seventh to ninth grade in Viana do Castelo district. *Journal of Information Systems Engineering and Management*, 8(2), 21014. <https://doi.org/10.55267/iadt.07.13229>.
- Sefton-Green, J., Nixon, H., & Erstad, O. (2009). Reviewing approaches and perspectives on “digital literacy”. *Pedagogies: an International Journal*, 4, 107-125. <http://dx.doi.org/10.1080/15544800902741556>.
- Seufert, S., Stanoevska-Slabeva, K., & Guggemos, J. (2020). Assessing subjective and objective information literacy at upper secondary schools – an empirical study in four German-speaking countries. *Int. J. Learning Technology*, 15(1), 82-103.
- Siddiq, F., Gochyyev, P., & Wilson, M. (2017). Learning in Digital Networks - ICT literacy: A novel assessment of students' 21st century skills. *Computers & Education*, 109, 11-37.
- Siddiq, F., Hatlevik, O. V., Olsen, R. V., Throndsen, I., & Scherer, R. (2016). Taking a future perspective by learning from the past - A systematic review of assessment instruments that aim to measure primary and secondary school

- students' ICT literacy. *Educational Research Review*, 19, 58-84.
- Stecher, B. (2010). *Performance assessment in an era of standards-based educational accountability*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education. Retrieved July 5, 2024, from <https://scale.stanford.edu/system/files/performance-assessment-era-standards-basededucational-accountability.pdf>.
- Turner, R. (2014). *Described proficiency scales and learning metrics*. Assessment GEMs no.4. Melbourne, Australia: Australian Council for Educational Research.
- Turpo-Gebera, O., Zea-Urviola, M., Huamaní-Portilla, F., Girón-Pizarro, M., Pérez-Zea, A., & Aguaded-Gómez, I. (2023). Media and information literacy in secondary students: Diagnosis and assessment. *Journal of Technology and Science Education*, 13(2), 514-531. <https://doi.org/10.3926/jotse.1746>.
- Van Deursen, A. J. A. M. & Van Dijk, J. A. G. M. (2009). Using the Internet: Skill related problems in users' online behavior. *Interacting with Computers*, 21, 393-402.
- Voogt, J. & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. *Journal of Curriculum Studies*, 44, 299-332.
- Vuorikari, R., Kluzer, S., & Punie, Y. (2022) *DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes, EUR 31006 EN*. Luxembourg: Publications Office of the European Union. doi:10.2760/490274, JRC128415.