XXXIX Seminario Nazionale di Ricerca
in Didattica della Matematica Giovanni Prodi
Rimini, 23-26 Marzo 2023

Valutazioni su larga scala e ricerche in didattica della matematica: esempi di studi condotti in Italia.

A cura di:
Federica Ferretti (*Università di Ferrara*), **Giorgio Bolondi** (*Libera Università di Bolzano*), **Alessandro Gambini** (*Sapienza Università di Roma*), **Chiara Giberti** (*Università di Bergamo*), **Alice Lemmo** (*Università dell'Aquila*), **Francesca Martignone** (*Università del Piemonte Orientale*), **Camilla Spagnolo** (*Libera Università di Bolzano*)

Questa relazione è stata preparata per i partecipanti al Seminario Nazionale.
È un documento ancora in bozza che sarà rivisto anche tenendo conto dei contributi dei controrelatori e dei partecipanti.

Si prega di non diffonderlo.

Indice

Premessa

In questo seminario presentiamo alcune delle ricerche condotte dal gruppo di ricerca negli ultimi dieci anni. Queste ricerche si inseriscono in un più ampio programma di studio e lavoro sull'utilizzo, la conoscenza e l'impatto dei materiali provenienti dal Sistema Nazionale di Valutazione - e nello specifico dalle *Prove INVALSI* di Matematica - in una prospettiva di ricerca didattica e di ricaduta sul contesto scolastico italiano.

L'impatto delle valutazioni standardizzate è tradizionalmente un impatto con dinamica *top-down*. Il programma di studio e lavoro si muove invece nella prospettiva di utilizzare in un'ottica di ricerca questi materiali, per innescare un processo *bottom-up* di conoscenza e miglioramento del sistema scolastico italiano.

La presentazione non seguirà lo sviluppo cronologico delle ricerche, ma cercherà piuttosto di presentare una fotografia, ad oggi, della logica secondo la quale questo approccio è andato prendendo forma in questi anni.

La maggior parte delle ricerche presentate hanno utilizzato *mixed-methods*, con piani sperimentali in cui le fasi qualitative e quelle quantitative si sono articolate in modi diversi, a seconda dei problemi affrontati. Non era possibile presentare tutti i dettagli tecnici delle ricerche. Si è preferito quindi, di caso in caso, dare maggiore spazio alla presentazione dei metodi, o all'esposizione dei dati utilizzati, o al commento dei risultati ottenuti. Speriamo in questo modo di aver dato una panoramica abbastanza esauriente dei problemi, dei metodi, dei piani sperimentali, del tipo di risultati raggiunti in questi anni.

La presentazione seguirà questo percorso.

Nel Capitolo 1 inquadriamo brevemente la nostra posizione nel dibattito sull'uso delle valutazioni su larga scala nella ricerca didattica.

Nel Capitolo 2 presentiamo tre esempi di ricerche che supportano la legittimità e la validità dell'utilizzo delle prove INVALSI per lo studio e la comprensione del contesto scolastico italiano, e in particolare per l'interpretazione delle dinamiche di insegnamento-apprendimento attraverso la lettura dei dati con le lenti fornite dalla ricerca in didattica della matematica.

Il Capitolo 3 espone alcuni fatti tecnici relativi alle prove INVALSI, aspetti psicometrici e statistici che hanno rilevanza nelle ricerche presentate.

Il Capitolo 4 presenta due esempi di ricerche in cui la triangolazione realizzata tra dati psicometrici, osservazioni qualitative e costrutti teorici ha permesso di ottenere interessanti risultati sulle manifestazioni del contratto didattico, e "insights" sul fenomeno del *gender gap* in Matematica.

Nel Capitolo 5 viene presentata la ricerca a partire dalla quale ha preso forma il nostro approccio. Questa ricerca ha messo in luce la complessità interpretativa dei fenomeni evidenziati dai risultati

dalle prove INVALSI, e ha motivato lo sviluppo di metodologie *mixed-methods* adeguate, nonché di strumenti di analisi dei dati, sia di quelli quantitativi che di quelli qualitativi.
Nel Capitolo 6 presentiamo, con il relativo inquadramento teorico, alcuni esempi di ricerche condotte sul tema della formazione insegnanti. In esse le prove INVALSI, lette con le lenti fornite dalla didattica della Matematica, diventano un elemento dinamico per lo sviluppo professionale degli insegnanti, e strumento per la riflessione critica sul proprio insegnamento. La valutazione, infatti, è un elemento cruciale nella costituzione dell'identità dell'insegnante, e quindi si è rivelato importante indagare sia come inserire nei percorsi formativi l'elemento "di rottura" portato dalle prove INVALSI, sia in generale il tema delle convinzioni verso le prove INVALSI e il significato dei risultati delle stesse. Un'attenzione specifica è dedicata allo sviluppo di strumenti specifici destinati a favorire questa riflessione.

Tutti i materiali e i dati relativi alle prove INVALSI utilizzati sono di pubblico dominio e accessibili liberamente a tutti i ricercatori.

Il gruppo di ricerca ringrazia sentitamente tutti i docenti e ricercatori che in questi anni hanno contribuito, su vari piani, allo sviluppo delle presenti ricerche. In particolare, si ringraziano Ferdinando Arzarello, Simone Banchelli, Marta Desimoni, Franca Ferri, Rossella Garuti, Cristina Lasorsa, Andrea Maffia, Nicoletta Nolli, Aurelia Orlandoni, Stefania Pozio, George Santi, Ketty Savioli, Ira Vannini e Matteo Viale.

Glossario condiviso

Alcuni termini e locuzioni utilizzati in questo seminario sono stati e vengono tuttora utilizzati con significati talvolta variabili a seconda del contesto, talvolta in sovrapposizione. Per chiarezza espositiva, esplicitiamo in questo breve glossario l'accezione nella quale utilizziamo alcune di queste espressioni.

Prova INVALSI. Ci riferiamo con questa espressione all'insieme di prove che l'INVALSI ha somministrato sul campo, nelle scuole italiane, a partire dal 2008 e fino ad oggi. Anche in precedenza l'INVALSI aveva realizzato piani per la valutazione delle scuole e degli allievi, in particolare con i *Progetti Pilota* (PP1, PP2, PP3, PP4) in diversi gradi scolastici; solo dal 2008 però le prove hanno acquisito caratteri di regolarità e sistematicità e i risultati relativi sono stati rilasciati in maniera completa. Le presenti ricerche si sono pertanto basate sulle prove realizzate dal 2008 in poi, ed è anche questo il significato comunemente condiviso all'interno delle scuole italiane.

Valutazione standardizzata. Con il termine *valutazione standardizzata* (*standardized assessment*) si intende qualunque forma di valutazione nella quale tutti i soggetti coinvolti rispondono alle stesse domande o a domande estratte da una medesima *banca di quesiti*. In questo senso, anche le valutazioni basate su compiti scritti, in classe, sono valutazione standardizzate.

Valutazione su larga-scala. Le valutazioni su larga scala sono valutazioni disegnate per raccogliere informazioni su popolazioni di media e grande dimensione, con risultati anche aggregabili per gruppi. Le prove INVALSI sono valutazioni standardizzate su larga scala. Spesso, nel linguaggio comune, i due termini sono utilizzati come sinonimi.

Prove cartacee e CBT. Le Prove INVALSI sono somministrate attraverso due supporti: cartaceo e digitale. Nel primo caso, tutti gli studenti ricevono i medesimi quesiti su un fascicolo cartaceo (in cui eventualmente può ruotare l'ordine dei quesiti o delle opzioni di risposta a un quesito); nel secondo caso, che chiameremo CBT (*Computer Based Testing*), ogni studente riceve una "forma" (dall'inglese *form*, per riferirsi ad un modulo costituito da un insieme di quesiti) contenente quesiti estratti da una banca di quesiti appositamente costruita per ciascun grado scolastico. Le prove per la primaria sono prove cartacee. Le prove per la scuola secondaria sono cartacee fino al 2017, CBT dal 2018.

Prova. In questo testo con il termine *Prova* ci riferiamo all'insieme delle domande poste ad un allievo nel corso di una valutazione INVALSI. Nel caso delle prove cartacee (per la scuola primaria dal 2009 ad oggi; per la scuola secondaria fino al 2017) all'insieme dei quesiti presenti in un fascicolo; nel caso delle prove CBT (per la scuola secondaria dal 2018 ad oggi) all'insieme dei quesiti di una forma e più in generale a quelli della banca di quesiti da cui sono tratti.

Quesito. Con il termine *quesito* ci si riferisce in questo testo ad un singolo task proposto in una prova INVALSI. I quesiti sono numerati progressivamente all'interno del fascicolo o della forma, e questa numerazione è quella attraverso la quale il quesito viene identificato in questo testo. Nel caso di quesiti ruotati nei fascicoli, adottiamo la numerazione presente nelle *Guide alla lettura* rilasciate dell'INVALSI (che si riferisce al fascicolo convenzionalmente indicato con 1), e riportata nel database *Gestinv*.

Item. Un quesito può contenere diverse richieste e quindi essere costituito da diversi *item*, usualmente indicati, all'interno del quesito, progressivamente con lettere. Avremo quindi, ad esempio nel quesito D20 della prova di grado 5 dell'anno scolastico 2021/22, un item D20a e un item D20b che afferiscono ad uno stimolo comune presentato nel quesito D20.

Opzione di risposta. Nel caso di quesiti a risposta chiusa, con *opzioni di risposta* ci riferiamo a una delle opzioni proposte allo studente. Nel caso delle prove di grado 2 normalmente sono presenti 3 opzioni di risposta, negli altri gradi sono presenti 4 opzioni di risposta. Utilizziamo il termine *opzioni* anche nel caso di domande dicotomiche, in cui la risposta è del tipo Vero/Falso oppure Sì/No.

Distrattore. Con il termine *distrattore* ci riferiamo a qualunque opzione di risposta sbagliata a un quesito a risposta chiusa.

Grado. Seguendo la prassi internazionale, indichiamo con *grado* di una prova l'anno di scolarità in cui viene proposta. Così il grado 2 corrisponde alla classe seconda primaria, il grado 5 alla classe quinta primaria, il grado 6 alla classe prima secondaria di primo grado, il grado 8 alla classe terza secondaria di primo grado (in cui la prova fa parte della valutazione conclusiva del primo ciclo di istruzione), il grado 10 alla classe seconda della scuola secondaria di secondo grado e il grado 13 alla classe quinta secondaria di secondo grado.

Abilità. Quando facciamo riferimento a dati psicometrici, con il termine *abilità* ci riferiamo al *carattere latente* (*latent trait*) misurato dalla prova nel suo insieme. Alti livelli di abilità corrispondono quindi a studenti che hanno risposto correttamente a molte domande della prova, bassi livelli di abilità a studenti che hanno risposto correttamente a poche domande. Gli studenti sono spesso aggregati per percentili rispetto all'abilità.

Scala di abilità. È la scala su cui il modello psicometrico utilizzato riporta l'abilità dello studente. La scala utilizzata dall'INVALSI è centrata in 0, che quindi corrisponde all'abilità mediana. Nel caso del modello di Rasch, la difficoltà dei quesiti e l'abilità degli studenti sono collocate sulla stessa scala.

1. Gli studi condotti dal gruppo di ricerca

Le nostre ricerche si inseriscono all'interno dell'ampio dibattito epistemologico e didattico su come integrare i risultati, i metodi, i quadri teorici e i risultati delle prove su larga scala - che sono progettate per avere un impatto a livello di sistema – nell'azione di insegnanti, formatori e ricercatori su scala locale, a livello di classe e istituzioni scolastiche. Gli studi che presenteremo si riferiscono tutti al contesto italiano.

Le valutazioni su larga scala hanno un forte impatto sistemico e sono state anche contestate per le loro possibili implicazioni pedagogiche, educative e politiche (si vedano, ad esempio, Carnoy, 2015; Cochran-Smith, 2001; Di Martino & Signorini, 2019). Tutto questo ha fatto sì che le valutazioni standardizzate su larga scala siano diventate, in diverse direzioni, oggetto di ricerche in didattica della matematica (De Lange, 2007; Meinck, Neuschmidt & Taneva, 2017; Suurtamm et al., 2016). All'interno di questo dibattito, la discussione verte sulla prospettiva con cui queste valutazioni/rilevazioni vengono utilizzate, rientrando in una problematica più ampia che coinvolge la valutazione sommativa in generale. Qualsiasi tipo di prova può essere utilizzata come strumento o come fine didattico. Noi abbracciamo la prospettiva per cui le prove, i dati che esse permettono di raccogliere e la loro elaborazione possano essere uno strumento utilizzabile per scopi didattici, educativi e di ricerca (Doig, 2006).

In parallelo allo sviluppo delle più importanti esperienze internazionali (OCSE-PISA, IEA-TIMSS, IEA-PIRLS), negli ultimi decenni quasi la totalità dei paesi ha introdotto rilevazioni nazionali su larga scala degli apprendimenti nella scuola primaria e nella scuola secondaria (Eurydice, 2016; Looney, 2011; Morris, 2011).

In Italia, le valutazioni su larga scala nazionali sono gestite dall'INVALSI (www.invalsi.it), l'ente di ricerca che, tra le varie funzioni che gli sono affidate dalla legge, effettua verifiche periodiche e sistematiche sulle conoscenze e abilità degli studenti e sulla qualità complessiva dell'offerta formativa delle istituzioni scolastiche.

La dichiarata consistenza del quadro teorico delle prove INVALSI di matematica con le Indicazioni Curricolari Ministeriali[1] (coerentemente con quanto messo in evidenza da Looney, 2011) e con i principali risultati di ricerca in didattica della matematica, come anche le caratteristiche del disegno

[1] Indicazioni Nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione https://www.miur.gov.it/documents/20182/51310/DM+254_2012.pdf/1f967360-0ca6-48fb-95e9-c15d49f18831?version=1.0&t=1480418494262, Assi Culturali https://archivio.pubblica.istruzione.it/normativa/2007/allegati/all1_dm139new.pdf , Indicazioni Nazionali per i Licei https://www.indire.it/lucabas/lkmw_file/licei2010/indicazioni_nuovo_impaginato/_decreto_indicazioni_nazionali.pdf, Linee Guida per Istituti Tecnici e Professionali https://www.indire.it/lucabas/lkmw_file/nuovi_tecnici/INDIC/_LINEE_GUIDA_TECNICI_.pdf https://nuoviprofessionali.indire.it/linee-guida-prof/

delle rilevazioni e le modalità di restituzione dei dati (INVALSI, 2018), fanno sì che i macro-fenomeni evidenziati in sede di rilevazione nazionale possano fornire informazioni utili e diventare strumenti interpretativi di molteplici aspetti dei processi di insegnamento-apprendimento della matematica specifici del contesto italiano.

Analisi qualitative e quantitative sui testi delle prove e sui risultati delle rilevazioni INVALSI sono state utilizzate in ricerche in didattica della matematica sul contesto italiano da parte di diversi ricercatori italiani[2].

Senza voler quindi presentare in questo seminario una rassegna esaustiva delle ricerche condotte dal nostro gruppo di ricerca, proponiamo una presentazione di alcune di esse che in qualche modo può delineare la mappa concettuale della nostra riflessione. I risultati ottenuti si collegano dialetticamente ad alcuni dei più importanti e condivisi risultati di ricerca in didattica della matematica, e al tempo stesso portano all'attenzione della comunità dei ricercatori nuovi fenomeni e nuove caratteristiche di fenomeni già conosciuti.

L’approccio di ricerca che accomuna i nostri studi è quindi l’uso delle valutazioni su larga scala come strumento di riflessione, analisi e ricerca sui processi di apprendimento e insegnamento degli studenti italiani, anche in un’ottica verticale (Ferretti, 2020a; Ferretti & Gambini, 2017; Gambini, Banchelli & Nolli, 2020).

Decliniamo ora brevemente le componenti dell’approccio di ricerca che sta alla base delle ricerche che saranno oggetto del seminario.

Il primo principio comune è una visione fenomenologica dell'interazione del ricercatore coi processi di insegnamento-apprendimento. Nei nostri studi partiamo da una osservazione ecologica di quello che viene percepito e rilevato sul campo dai diversi attori del sistema (ricercatori, insegnanti, allievi, INVALSI, commentatori, …), tenendo presente che comunque questa percezione e questa rilevazione sono a loro volta inquadrate e condizionate da convinzioni, teorie, ideologie. Ma è proprio questa dialettica tra fenomenologia e teorie che permette l’evoluzione della comprensione dei fenomeni stessi da un lato, e delle teorie dall’altro.

---

[2] In questo seminario presenteremo alcuni dei risultati del nostro gruppo di ricerca, ma naturalmente sono molti gli altri lavori prodotti o da membri di questo gruppo in collaborazione con altri ricercatori e insegnanti-ricercatori o da altri ricercatori. Per fare alcuni esempi, senza nessuna pretesa di esaustività, si vedano ad esempio Banchelli, Garuti e Nolli (2021), Barana e Marchisio, 2020; 2021, Bolondi, Ferretti, e Spagnolo (2021), Carotenuto, Di Martino e Lemmi, 2021, Ferrara, Ferrari e Savioli (2021), Ferrara, Gilardi e Savioli (2021), Ferrara e Pozio (2021), Ferri e Martignone (2018), Gambini, Spagnolo, Capone, Ferretti e Casalvieri (2022), Garuti, Lasorsa e Pozio (2017), Garuti e Martignone (2015), Garuti e Martignone (2019), Graziani e Babini (2018), Spagnolo, Capone, Ferretti e Gambini (2021), Spagnolo, Capone e Gambini (2021).

Tutte le nostre ricerche si basano inoltre su una visione formativa della valutazione in matematica, abbracciando la definizione di valutazione formativa delineata all'interno del Progetto Europeo *LLP-Comenius FAMT&L* (Ferretti, Michael-Chrysanthou & Vannini, 2018).

Le valutazioni di classe e le valutazioni su larga scala hanno, a un primo livello di lettura, obiettivi e scopi diversi. Le valutazioni su larga scala sono valutazioni di sistema e hanno come obiettivo il monitoraggio di un sistema educativo e, nel caso italiano, la verifica dell'implementazione di programmi ministeriali. Peraltro, le informazioni restituite da queste valutazioni, quando coerenti con il sistema educativo, possono integrarsi nelle prassi scolastiche in un'ottica formativa (Looney, 2011; Bolondi, Ferretti & Spagnuolo, 2016). La coerenza di queste valutazioni con il contesto educativo in cui sono inseriti i processi di insegnamento-apprendimento oggetto delle nostre ricerche è un fatto essenziale, che permette tra l'altro di ritrovare nelle prove INVALSI esempi di quesiti in linea con gli obiettivi dichiarati nelle Indicazioni Ministeriali. Infatti, il quadro di riferimento delle prove INVALSI di matematica si richiama esplicitamente alle Indicazioni Ministeriali e agli altri riferimenti normativi. Inoltre, il quadro di riferimento definisce l'organizzazione delle prove sull'articolazione delle Indicazioni stesse, e il processo di costruzione delle prove, in tutti i suoi passaggi, è agganciato alle Indicazioni, quindi al *curriculum intended* della scuola italiana nel suo complesso (INVALSI, 2018). Questo fatto è cruciale e spiega il perché i nostri studi si possano basare solo sulle valutazioni INVALSI e non su qualsiasi altro tipo di valutazione su larga scala.

Questa dichiarata coerenza delle prove (*curriculum assessed*) con il curriculum *intended* va peraltro poi affiancata da una analisi della coerenza delle prove stesse con le pratiche effettive, che fanno parte del curriculum *implemented*, presenti nelle scuole italiane.

Nelle ricerche si è quindi sempre considerato come elemento essenziale sia della progettazione dei piani sperimentali, sia dell'interpretazione dei risultati, la rete di relazioni tra questi tre elementi: le Indicazioni nazionali, le prove effettivamente mandate sul campo dall'INVALSI, le pratiche dichiarate o rilevate nel contesto.

Dal punto di vista metodologico, nel senso di tecniche di raccolta e interpretazione di dati supportati dal sistema di principi, i nostri studi hanno utilizzato *mixed methods*, inserendo i materiali delle rilevazioni INVALSI in disegni sia di tipo *explanatory* sia di tipo *exploratory* (Creswell & Plano-Clark, 2017). Una mixed-method research consiste essenzialmente in due o più componenti o trend di ricerca (Teddlie & Tashakkori, 2009; Schoonenboom & Johnson, 2017) provenienti da prospettive metodologiche diverse - nella maggior parte dei casi sia qualitative sia quantitative. Si possono *combinare* le componenti tra loro, se l'obiettivo della ricerca è quello di unire e legare reciprocamente i risultati di ciascuna o *integrare* le componenti tra loro, se l'obiettivo è la convalida reciproca dei risultati della ricerca.

In tutti in nostri studi, la prima fase consiste nell'*integrazione* delle due componenti. Infatti, le ricerche che presenteremo nascono da una prima fase comune di individuazione di quelli che abbiamo definito *macro-fenomeni* emergenti dai risultati delle rilevazioni INVALSI.

Un secondo principio alla base del nostro approccio è il fatto che la dinamica dei processi di insegnamento e apprendimento è quella di un sistema altamente complesso, che all'interno di questa metafora può essere considerato *caotico*. Non sono normalmente identificabili relazioni deterministiche di causa-effetto, né su scala locale (ad esempio, il risultato dell'azione didattica del singolo insegnante), né su scala sistemica (ad esempio, i risultati delle scelte dei decisori politici). Come si usa dire in questi casi, *una farfalla che batte le ali in Brasile può scatenare un uragano in Texas*. Come in tutti i sistemi di questo tipo, però, si manifestano dei macro-fenomeni emergenti a livello sistemico (*c'è un uragano in Texas*) che è importante rilevare e, quando possibile, misurare, ed è poi possibile descrivere traiettorie tipiche di singoli elementi del sistema. La combinazione e integrazione di questo tipo di osservazioni ed analisi, da realizzare appunto con una integrazione e combinazione di metodi, permette di iniziare a comprendere le dinamiche globali e descrivere possibili traiettorie locali.

Il macro-fenomeno è quindi un fatto complesso, la cui rilevazione e prima interpretazione si può ottenere con l'integrazione di due componenti, secondo Johnson e Onwuegbuzie (2004), una qualitativa e una quantitativa, integrate e complementari. Le informazioni quantitative provengono dalle valutazioni INVALSI: le informazioni della componente quantitativa del macro-fenomeno sono quindi ottenute da un quadro di rilevazione designato e effettuato da Ente di ricerca nazionale. Da qui la necessità, nei nostri studi, della condivisione del quadro teorico su cui sono costruiti i quesiti, le modalità di raccolta dei dati, le analisi effettuate e le modalità di restituzione dei dati, sia a livello campionario nazionale, sia a livello locale alle singole istituzioni scolastiche.

La fase qualitativa dell'individuazione del macro-fenomeno è condotta dai ricercatori, che con le lenti interpretative fornite da teorie e risultati della ricerca analizzano e interpretano i risultati INVALSI delineando il macro-fenomeno.

A partire da questa prima fase comune, ciascuna ricerca è caratterizzata da specifiche domande e da conseguenti impianti metodologici finalizzati alla loro risposta.

Le altre fasi delle ricerche a volte sono di tipo quantitativo, con l'utilizzo di dati provenienti da impianti sperimentali ad hoc, a volte di tipo qualitativo, e a volte di tipo *mixed-method*.

Durante il seminario, presenteremo alcuni degli studi che, in linea con questo approccio di ricerca, hanno permesso una caratterizzazione, quantificazione e misurazione della portata di diversi macro-fenomeni (come, ad esempio, nelle ricerche presentate in Bolondi, Ferretti & Santi, 2019; Ferretti, 2020b; Ferretti & Giberti, 2021; Ferretti, Santi & Bolondi, 2022) o hanno messo in luce nuove evidenze e interpretazioni legate ai fenomeni stessi (fenomeni nuovi come l'effetto "Età della Terra", Ferretti & Bolondi, 2019).

Il gruppo di ricerca si è occupato anche dello studio di come questi macro-fenomeni possono costituire un ingrediente della formazione degli insegnanti, con quali strumenti e con quali modelli. Gli esempi di ricerche con cui concluderemo il seminario riguardano precisamente l'utilizzo di uno strumento di documentazione sulle prove nato per la formazione degli insegnanti e poi divenuto strumento di ricerca, e un modello di formazione basato sull'analisi e discussione delle prove e dei loro risultati.

2. I risultati delle Prove INVALSI nel contesto italiano: alcuni esempi

Come primo elemento di questo quadro, presenteremo tre esempi di ricerche che, in direzioni diverse, mostrano come le prove INVALSI si interfacciano da un lato con i risultati della ricerca didattica, e dall'altro con il contesto socio-scolastico italiano. Queste ricerche, oltre all'interesse intrinseco, supportano quindi la considerazione delle informazioni restituite dalle prove INVALSI come un elemento utilizzabile criticamente per ricerche sulla scuola italiana. Si tratta innanzitutto della ricerca illustrata nel paragrafo seguente, in cui si sono studiati alcuni aspetti predittivi delle prove INVALSI del grado 13 in relazione alla carriera universitaria (Gambini, Desimoni & Ferretti, 2022). Nell'esporre le ricerche faremo riferimento ad alcuni strumenti psicometrici (in particolare i *distractor plot*), che saranno presentati in dettaglio nel Capitolo 3.

2.1 Il potere predittivo delle prove INVALSI di matematica per il successo in ambito accademico

La validità e la coerenza delle prove INVALSI nel sistema educativo italiano è messa in luce anche dai risultati di una ricerca, di durata triennale, nata dagli studi condotti all'interno del progetto di ricerca "Misurazione diacronico-longitudinale dei livelli di apprendimento degli studenti: relazione tra le prove INVALSI per la scuola secondaria di secondo grado (II e V anno) e il successivo percorso all'Università degli studenti, in particolare dal punto di visto dei contenuti i legami tra le prove e i test di ammissione all'Università e la possibile correlazione tra i risultati che gli studenti ottengono, i risultati dei test e la carriera universitaria" (Gambini, 2018). La ricerca mette in relazione i risultati dei pre-test INVALSI di matematica di grado 13, i voti degli Esami di Stato conclusivi del secondo ciclo d'istruzione e i progressi in ambito matematico della carriera universitaria di studenti dell'Università di Bologna. Come mostrato dalla letteratura internazionale (Di Martino, Gregorio & Iannone, 2022), la transizione tra la scuola secondaria e l'università è una questione problematica per diversi studenti, e, in particolare, la matematica gioca un ruolo cruciale sotto diversi punti di vista. Ciò accade anche nel contesto italiano (si veda, ad esempio, Di Martino & Gregorio, 2019). Le ricerche di Camara (2013) hanno messo in luce il problema dell'identificazione di indici attendibili e validi di preparazione degli studenti della scuola secondaria in vista dell'accesso all' università, nonché la necessità di validi strumenti per prevedere il successo accademico degli studenti del primo anno. Il nostro studio (Gambini, Desimoni & Ferretti, 2022) si inserisce in questa linea di ricerca e prende in esame le relazioni tra i risultati ottenuti nei pre-test INVALSI di matematica per il grado 13 dell'a.s. 2017/18, i voti finali dell'esame di Stato e il successo in matematica in riferimento alla carriera universitaria dei primi due anni di 177 laureandi dell'Università di Bologna.

Gli studenti del campione risultano così suddivisi: 56 di Scienze Statistiche; 59 di Scienze della Formazione; 43 di Farmacia; 19 di Ingegneria. In totale hanno partecipato 65 maschi e 112 femmine; l'età media alla fase 1, autunno 2016 è 19 anni con *Deviazione standard età* = 1,49.

Il 75,7% degli studenti del campione ha conseguito il diploma in un liceo (63 licei scientifici; 71 in altri licei), il 20,4% in un istituto tecnico e il 3,9% in un istituto professionale. Questa distribuzione è simile a quella delle matricole in Italia nel 2016, quindi, il campione scelto è rappresentativo della situazione attuale (MIUR, 2017).

La ricerca è stata svolta in due fasi: la prima fase ha avuto luogo nell'autunno 2016, anno in cui gli studenti coinvolti erano appena immatricolati, mentre la fase di follow-up risale all'autunno del 2018, inizio del terzo anno di università dei partecipanti.

Nella prima fase, gli studenti hanno sostenuto la prova contenente alcuni quesiti del pre-test INVALSI di matematica per il grado 13 del 2017 (per brevità espositiva chiameremo la prova "pre-test INVALSI") all'inizio del loro primo anno di università. Le informazioni su sesso, età, voto finale della scuola superiore (HS – *High School final grade*) e percorso di ogni studente sono state raccolte tramite un questionario.

Dopo due anni, nella fase di follow-up, sono stati raccolti per 186 studenti del campione iniziale, i dati sulla media dei voti universitari (GPA – Grade Point Average), il numero di crediti accademici acquisiti e il punteggio individuale nell'esame di matematica.

La nostra prima ipotesi di ricerca (H1) afferma che i voti finali HS avrebbero predetto il GPA universitario (H1a) e il numero di crediti universitari acquisiti (H1b).

La nostra seconda ipotesi di ricerca (H2) afferma che il pre-test INVALSI di matematica avrebbe anche predetto il rendimento universitario di uno studente in termini di GPA (H2a) e di crediti acquisiti (H2b). Ci aspettavamo inoltre che i risultati della prova con i quesiti del pre-test INVALSI e i voti finali del HS potessero fornire, in modo indipendente tra loro, informazioni sui successivi risultati universitari, e in particolare che il voto finale del HS e il rendimento nel pre-test INVALSI di matematica avrebbero svolto ruoli complementari nella previsione del rendimento universitario di uno studente.

Infine, abbiamo ipotizzato che il pre-test INVALSI di matematica potesse prevedere l'esito degli esami universitari di matematica del primo anno (H3a) mentre, al contrario, si è ipotizzata una correlazione più debole tra il voto finale HS e il primo esame universitario di matematica (H3b).

In effetti, con poche eccezioni, i nostri risultati mostrano che i risultati ottenuti nella prova con i quesiti del pre-test INVALSI abbiano un potere predittivo per quanto riguarda il superamento e i risultati ottenuti negli esami di matematica del primo anno più forte rispetto a quello dei voti dell'Esame di Stato.

La Tabella 1 mostra i coefficienti di correlazione tra le variabili analizzate sui sotto-campioni relativi ai diversi corsi universitari analizzati. Nella Tabella 2 vengono mostrati i coefficienti di correlazione delle variabili analizzate sul campione complessivo. Come mostrato in Tabella 1, per quanto riguarda l'ipotesi H1a, sono state rilevate correlazioni positive medio-grandi, statisticamente significative nei sotto-campioni di Farmacia ($p = 0{,}006$), Ingegneria ($p = 0{,}001$) e Scienze statistiche ($p < 0{,}001$). Una correlazione significativa e positiva (anche se con un $p = 0{,}039$) è emersa anche per gli studenti di Scienze della Formazione Primaria.
È stata trovata una correlazione positiva di medie dimensioni statisticamente significativa ($p < 0{,}001$) tra il voto finale HS e la proporzione di crediti acquisiti (sul totale) su tutto il campione (Tabella 2); una correlazione positiva significativa, di medie dimensioni, è emersa anche nel sotto-campione di Scienze statistiche ($p = 0{,}002$) e di Farmacia ($p = 0{,}015$) come si evince dalla Tabella 1. Non è invece emersa alcuna correlazione significativa tra il voto finale HS e la percentuale di crediti acquisiti (sul totale) nei sotto-campioni Ingegneria ($p = 0{,}08$) e Scienze della Formazione Primaria ($p = 0.14$). I risultati del campione totale (Tabella 2) sono quindi coerenti con l'ipotesi H1b, sebbene sia emersa una certa variabilità tra i corsi.
La Tabella 2 indica una correlazione significativa medio-grande ($p < 0{,}001$) con i risultati degli esami di matematica del primo anno nel campione complessivo (Tabella 2) ma anche nei vari sotto-campioni: in Scienze della formazione ($p = 0{,}005$), Farmacia ($p = 0{,}010$) e una correlazione significativa di grandi dimensioni ($p < 0{,}001$) nel sottocampione di Scienze statistiche. Tutte le correlazioni sono positive. Un'altra correlazione positiva di dimensione medio-grande è emersa anche nel piccolo sotto-campione Ingegneria, tuttavia, quest'ultima non ha raggiunto la significatività statistica ($p = 0{,}14$). È emersa anche una correlazione di dimensione medio-grande statisticamente significativa ($p < 0{,}001$) tra il pre-test INVALSI e il GPA ottenuto in due anni (Tabella 2).
Correlazioni significative di medie dimensioni sono state trovate per gli studenti di Farmacia ($p = 0{,}026$) e Scienze della Formazione Primaria ($p = 0{,}005$); una correlazione forte e statisticamente significativa ($p < 0{,}001$) è emersa anche nel sotto-campione di Scienze statistiche (Tabella 1). Tutte queste correlazioni sono positive. In linea con la letteratura, i nostri risultati confermano il ruolo dei voti dell'Esame di Stato come predittori del "successo accademico" nell'istruzione universitaria (in termini di crediti acquisiti e media dei voti universitari) anche nel contesto italiano. La ricerca mostra anche che i risultati ottenuti nei pre-test INVALSI migliorano il potere predittivo dei voti dell'Esame di Stato non solo ai fini della previsione del successo nei corsi di matematica del primo anno, ma anche quando si tratta di criteri più generali di successo universitario.
Lo studio svolto è esplorativo ed ha alcune limitazioni dovute al campione, all'analisi di solo quattro corsi universitari e al fatto che i quesiti del pre-test INVALSI che costituiscono la prova fossero quelli

della campagna di pre-test primo anno del disegno della rilevazione per il grado 13, ancora in una fase di assestamento. Ciononostante, nella nostra ricerca longitudinale, i voti finali del HS e il pre-test INVALSI di matematica si sono confermati strumenti idonei per la previsione del rendimento universitario. Considerare entrambi gli aspetti può consentire una migliore previsione degli esiti nell'insegnamento universitario piuttosto che considerare solo uno dei due predittori. La complementarità nella previsione potrebbe riflettere la complementarità negli scopi della valutazione: gli esiti dei pre-test INVALSI di matematica del grado 13 potrebbero essere utilizzati come misurazione della preparazione per i corsi universitari o come test di ingresso, e possono fornire dati idonei per la progettazione di piani d'azione finalizzati al miglioramento dei processi di apprendimento degli studenti universitari e all'orientamento in ingresso.

Tabella 1. Coefficienti di correlazione tra le variabili analizzate sui sotto-campioni

			Descriptive statistics						Correlation-Spearman's Rho				
Course		Variable	N	M	Median	SD	Skew	Kurt.	1.00	2.00	3.00	4.00	5.00
Pharmacy	1	Age	43	19.09	19.00	0.37	4.27	19.08	1.00				
	2	HS final grade	43	86.95	87.00	9.31	−0.34	−0.68	−.15	1.00			
	3	Maths test	43	20.74	22.00	5.26	−0.67	0.38	−.19	.13	1.00		
	4	U-Maths exam	43	27.70	30.00	4.26	−1.30	0.26	−.28	.19	.39**	1.00	
	5	U-GPA	43	26.40	26.75	2.20	−0.51	−0.39	−.25	.41**	.34*	.67**	1.00
	6	U-Academic credits	43	0.85	0.95	0.21	−1.87	3.18	−.23	.37*	.10	.46**	.73**
	7	Gender – male	43	0.44									

			Descriptive statistics						Correlation-Spearman's Rho				
Course		Variable	N	M	Median	SD	Skew	Kurt.	1.00	2.00	3.00	4.00	5.00
Engineering	1	Age	19	19.32	19.00	0.58	1.77	2.54	1.00				
	2	HS final grade	19	82.00	82.00	12.86	−0.19	−1.06	−.24	1.00			
	3	Maths test	19	16.32	16.00	3.73	0.43	−0.32	−.33	.20	1.00		
	4	U-Maths exam	19	23.26	23.00	4.53	0.11	−0.89	−.14	.48*	.40	1.00	
	5	U-GPA	19	23.95	23.65	2.60	0.59	−0.55	−.29	.69**	.36	.76**	1.00
	6	U-Academic credits	19	0.61	0.57	0.24	−0.03	−0.80	−.04	.43	.49*	.48*	.69**
	7	Gender – male	19	0.74									
Education Sciences	1	Age	59	19.75	19.00	2.40	4.04	16.96	1.00				
	2	HS final grade	59	79.14	78.00	9.66	0.45	−0.28	.22	1.00			
	3	Maths test	59	9.81	9.00	3.91	0.23	−0.21	−.02	−.08	1.00		
	4	U-Maths exam	59	22.80	23.00	3.67	0.18	−1.03	.17	.02	.36**	1.00	
	5	U-GPA	59	26.15	26.18	1.07	0.42	0.39	−.02	.27*	.36**	.45**	1.00
	6	U-Academic credits	59	0.84	0.87	0.15	−1.49	3.29	.05	.18	.03	.02	−.03
	7	Gender – male	59	0.08									
Statistical Sciences	1	Age	56	19.21	19.00	0.80	1.78	4.37	1.00				
	2	HS final grade	56	84.32	83.00	10.30	0.02	−1.13	−.06	1.00			
	3	Maths test	56	19.71	21.00	6.25	−0.46	−0.48	.10	.19	1.00		
	4	U-Maths exam	56	22.20	21.50	4.08	0.49	−1.14	.16	.19	.59**	1.00	
	5	U-GPA	56	26.41	26.50	2.32	−0.17	−0.77	.01	.60**	.62**	.63**	1.00
	6	U-Academic credits	56	0.86	0.93	0.18	−1.35	0.97	−.03	.41**	.42**	.25	.63**
	7	Gender – male	56	0.46									

Tabella 2. Coefficienti di correlazione tra le variabili analizzate sul campione

		Age	HS final grade	Standardized maths test	U-Maths exam	U-GPA	U-Academic credits
Correlation	Age	1.00					
	HS-final grade	.02	1.00				
	Standardized maths test	−.04	.08	1.00			
	U-Maths exam	.05	.17*	.46**	1.00		
	U-GPA	−.06	.44**	.43**	.61**	1.00	
	U-Academic credits	−.07	.33**	.26**	.25**	.45**	1.00
Descriptive statistics	Mean (raw score)	19.37	82.98	16.30	23.85	26.05	0.83
	Median (raw score)	19.00	82.00	16.00	24.00	26.22	0.90
	SD (raw score)	1.50	10.53	6.92	4.58	2.11	0.20
	Skewness (raw score)	6.07	0.01	0.03	0.10	−0.33	−1.37
	Kurtoses (raw score)	43.78	−0.92	−0.96	−1.35	−0.09	1.41
	Skewness (stand. score)		0.01	0.03	−0.09	0.01	
	Kurtoses (stand. score)		−0.92	−0.96	−0.78	−0.38	

Note: GPA = Grade Point Average; HS = High School; stand. score = standardized scores within-groups; * $p < .05$; ** $< .01$

2.2 Il *loss of meaning* dei simboli algebrici: quantificare un macro-fenomeno

La ricerca in didattica della matematica ha i suoi *solid findings* (EMS, 2011). Una prima lista di “solid findings” è stata proposta, assieme a una possibile definizione, in una serie di uscite della Newsletter della European Mathematical Society. Questa lista include, tra gli altri, il Contratto Didattico, le Norme Socio Matematiche, la Linearità e, su un argomento che riguarda questa ricerca, i Modelli e la Modellizzazione (Niss, 2012). Questi risultati sono ottenuti e validati con paradigmi e metodologie di ricerca condivisi, con la prevalenza di un approccio qualitativo (Hart, Smith, Swars & Smith, 2009). La discussione su cosa possa essere considerato un “solid finding” si è sviluppata all’interno di *communities* di ricercatori, come CERME e altre (si veda, ad esempio, la sintesi proposta da Bosch, Dreyfus, Primi & Shiel, 2017 e Dreyfus, 2017). In particolare, un punto di vista psicometrico-quantitativo è stato proposto, in un caso particolare, da Primi (2017). Per molti fenomeni didattici sono disponibili descrizioni dettagliate di casi e quadri teorici generali, e queste descrizioni e questi quadri sono indubbiamente utili anche agli insegnanti quando affrontano nelle proprie classi le situazioni relative e cercano di interpretare le azioni dei propri allievi. Quello che manca, spesso, è una quantificazione del fenomeno all’interno di uno specifico contesto (nazionale o su diversa scala), quantificazione che potrebbe aiutare gli insegnanti nella propria azione, e i ricercatori nel comprendere l’articolazione e l'estensione del fenomeno. Avere dati di diffusione di un fenomeno in uno specifico contesto di sistema può essere uno strumento per collegare i risultati della ricerca con le pratiche sul campo degli insegnanti.

Una ricerca in cui i risultati delle prove INVALSI hanno permesso una quantificazione di un fenomeno che può essere considerato un *solid finding* è presentata da Bolondi e Ferretti (2021) e riguarda la *perdita di significato dei simboli algebrici*, in particolare nelle situazioni di modellizzazione.

L'apprendimento dell'algebra è una questione centrale nell'educazione matematica. Si sviluppa in un nuovo scenario cognitivo in cui il pensiero procedurale e il pensiero relazionale sono strettamente interfacciati. Ha la sua origine in processi di *modellazione* (per esempio, di situazioni del mondo reale) e di *generalizzazione* (ad esempio, di relazioni aritmetiche): l'algebra può essere vista come un punto di svolta nel percorso educativo dello studente perché comporta nuove forme di pensiero e organizzazione concettuale; segna il passaggio dal puro pensiero aritmetico alla matematica avanzata. Il pensiero algebrico è caratterizzato dall'introduzione di un linguaggio simbolico formale in cui il rapporto tra *senso* e *significato* è cruciale (Arcavi, 1994; Capraro & Joffrion, 2006; Kieran, 1992, 2007). L'introduzione del linguaggio algebrico è peraltro viziata da una difficoltà intrinseca: gli oggetti di base del discorso algebrico sono necessari in un preciso momento del curriculum, e devono essere presentati agli studenti di una ben determinata età, ma tali oggetti possono essere adeguatamente definiti con rigore matematico solo in un contesto formale che non è accessibile agli studenti di quell'età (Bolondi, Ferretti & Maffia, 2020). L'introduzione del linguaggio formale implica importanti difficoltà cognitive per lo studente: un esito molto frequente dell'algebra scolastica è un uso meccanico di segni che sono privi di significato. Il principale problema educativo è quindi quello di prevenire un tale uso manipolativo delle lettere, e inserire l'algebra nel curriculum matematico come un vero e proprio strumento per nuove forme di pensiero e generalizzazione.

Arzarello, Bazzini e Chiappini (2001) descrivono gli ostacoli e le difficoltà nell'apprendimento dell'algebra, delineando come l'uso del formalismo algebrico provochi una dissonanza tra il pensiero procedurale indotto dall'uso del linguaggio naturale e il pensiero relazionale trasmesso da quello simbolico. Il risultato è una perdita di controllo semantico delle operazioni algebriche, che si traducono in una trasformazione senza significato dei segni.

La ricerca che presentiamo mira a *quantificare* questa perdita di significato nell'uso dei simboli algebrici. Abbiamo individuato due quesiti somministrati dall'INVALSI in grado 8 (Figura 1) e grado 10 (Figura 2), che descrivono chiaramente ciò che accade. Abbiamo analizzato i dati relativi alle risposte fornite da campioni di studenti di grado 8 e grado 10, composti da 28.361 studenti di 1.312 classi e 50.838 studenti di 2.302 classi rappresentativi di 527.318 studenti, rispettivamente, rappresentativi delle rispettive popolazioni nazionali.

D17. **La formula $L = L_0 + K \times P$ esprime la lunghezza L di una molla al variare del peso P applicato. L_0 rappresenta la lunghezza in centimetri "a riposo" della molla; K indica di quanto si allunga in centimetri la molla quando le si applica una unità di peso.**
Quale delle formule elencate si adatta meglio alla seguente descrizione:
"*È una molla molto corta e molto dura (cioè molto resistente alla trazione)*"?

- ☐ A. $L = 10 + 0{,}5 \times P$
- ☐ B. $L = 10 + 7 \times P$
- ☐ C. $L = 80 + 0{,}5 \times P$
- ☐ D. $L = 80 + 7 \times P$

Figura 1. Quesito 17, Prova INVALSI di matematica Grado 8, 2011, www.gestinv.it

D24. **La formula $l = l_0 + k \cdot P$ esprime la lunghezza l di una molla al variare del peso P applicato. l_0 rappresenta la lunghezza in centimetri "a riposo" della molla; k indica di quanto si allunga in centimetri la molla quando si applica una unità di peso.**
Quale delle formule elencate si adatta meglio alla seguente descrizione:
"*È una molla molto lunga e molto resistente alla trazione*"?

- ☐ A. $l = 15 + 0{,}5 \cdot P$
- ☐ B. $l = 75 + 7 \cdot P$
- ☐ C. $l = 70 + 0{,}01 \cdot P$
- ☐ D. $l = 60 + 6 \cdot P$

Figura 2. Quesito 24, Prova INVALSI di matematica Grado 10, 2011, www.gestinv.it

L'analisi dei risultati mostra che il controllo semantico delle espressioni letterali, che sono qui utilizzate come strumento di modellizzazione di una situazione, diminuisce tra i due gradi, nonostante i due anni di scuola che intercorrono tra di essi siano dedicati proprio allo sviluppo del linguaggio algebrico: il valore dell'indice di difficoltà misurato dal modello all'interno dei test salta da -0,42 a 0,53 (per una definizione tecnica di questi parametri psicometrici rimandiamo al successivo Capitolo 3). In termini di percentuali, il numero delle risposte corrette diminuisce dal 58,3% al 38,1%. La percentuale di risposte mancanti (che di solito sono molto poche in un item a scelta multipla, in un test in cui non c'è penalizzazione per le risposte sbagliate come in questo caso) aumenta dal 4% all'11%, e anche questo può essere visto come una prova della difficoltà degli studenti nel dare un significato al compito. Entrambi gli item hanno buone caratteristiche psicometriche: c'è un buon adattamento con il modello ed entrambi gli item discriminano tra gli studenti con diversi livelli di abilità.

Un interessante fatto messo in luce dalla distribuzione delle risposte corrette lungo la scala di abilità è che questi "studenti perduti" sono studenti che raggiungono nella prova alti livelli di abilità. Senza entrare qui nel dettaglio dei dati restituiti, basti dire che nell'ottavo decile della scala delle abilità, la probabilità di risposta corretta crolla dall'80% al 50% quando si passa dal grado 8 al grado 10; la probabilità di risposta corretta nel nono decile diminuisce dall'85% a circa il 60%. Altre evidenze interessanti possono essere tratte dalla distribuzione delle scelte "sbagliate": tutto comunque mostra che la quota di studenti, anche di alta abilità, che “perdono di vista” il significato dei simboli è molto alta.

Questa brevissima sintesi dei risultati quantitativi dei dati INVALSI mostra che è possibile *quantificare*, nel contesto italiano e nel momento specifico del passaggio tra il primo e il secondo ciclo di istruzione, l’entità di questa *perdita di significato* dei simboli algebrici. Un fatto individuale, la cui traiettoria nel singolo allievo, in termini di cause ed effetti nel percorso, è impossibile da predeterminare o individuare, ma la cui consistenza emerge a livello sistemico come *macro-fenomeno*. Aggiungiamo, per completare il quadro delle informazioni che le prove INVALSI possono fornire sul tema del controllo semantico in una modellizzazione algebrica, che in Pozio e Bolondi (2019), è studiato attraverso protocolli di studenti il caso di un quesito INVALSI in ambito geometrico.

2.3 Manipolare le espressioni algebriche: interpretare un macro-fenomeno con lenti teoriche

La seconda ricerca che presentiamo (oggetto di pubblicazione in Ferretti, Santi & Bolondi, 2022) è volta a interpretare con specifiche lenti teoriche alcune criticità messe in luce dai risultati delle prove INVALSI. Come vedremo, è un esempio di risultato di ricerca che si presta ad essere utilizzata nelle prime fasi del nostro modello di formazione. Anche per questo esempio abbiamo scelto un caso relativo all’apprendimento dell’algebra.

Per indagare le difficoltà degli studenti italiani in questo ambito abbiamo utilizzato come strumento di ricerca il database *Gestinv* (www.gestinv.it), di cui parleremo diffusamente nel Capitolo 6, e in particolare i *distractor plot* che Gestinv raccoglie, che presenteremo nel dettaglio nel prossimo Capitolo 3. Il database ci ha permesso di effettuare un'analisi dei dati quantitativi relativi ai quesiti INVALSI di algebra di grado 10 che hanno avuto percentuali di risposta corrette molto basse a livello nazionale e i cui contenuti sono vicini a quelli presenti nel curriculum standard di tutti gli indirizzi delle scuole superiori italiane durante i primi due anni di scuola secondaria. Si tratta quindi di situazioni di convergenza tra il *curriculum assessed* (le prove INVALSI), quello *intended* (previsto nelle Indicazioni curricolari) e quello *implemented* (effettivamente implementato nelle scuole).

Nonostante questa convergenza, alcuni risultati sorprendono, facendo risultare molto difficili quesiti che, a priori, non sembrano esserlo. Si tratta quindi di un caso in cui una approfondita analisi teorica "fine" delle caratteristiche degli item e soprattutto dei distrattori proposti può risultare utile, anche per le possibili applicazioni.

Tra tutti i quesiti restituiti dal database, sono risultati particolarmente interessanti due quesiti riguardanti le operazioni di trattamento (nel senso di Duval, 2006) di disuguaglianze.

Il primo quesito è quello rappresentato in Figura 3, somministrato nella prova INVALSI di grado 10, somministrato a circa 560 mila studenti di grado 10 nella prova INVALSI di matematica del 2013.

D19. Nell'insieme dei numeri reali, la disequazione $x^2 > 0$ è verificata

A. ☐ per ogni $x \neq 0$

B. ☐ per ogni x

C. ☐ solo per ogni $x < 0$

D. ☐ solo per ogni $x > 0$

Figura 3. Quesito 19, Prova INVALSI di matematica Grado 10, 2013 www.gestinv.it

Come mostrato in Figura 4, ha risposto correttamente al quesito meno della metà degli studenti del campione nazionale costituito da 38.533 studenti. Il fatto che l'opzione C abbia ottenuto le percentuali di scelta più basse era prevedibile in quanto, da un punto di vista semiotico c'è una forte asimmetria tra il segno di disuguaglianza nella domanda e nella risposta.

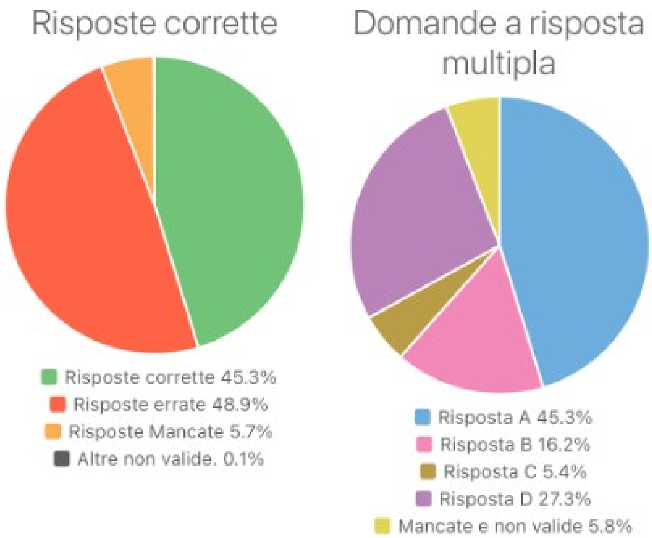


Figura 4. Percentuali nazionali, Quesito 19, prova INVALSI di matematica Grado10, 2013 www.gestinv.it

Le Figure 5 e 8 riportano i *distractor plot* dei quesiti che stiamo analizzando. Questo strumento sarà presentato in dettaglio nel prossimo Capitolo 3; per comprendere il ruolo che gioca nella nostra analisi del fenomeno basti qui sapere che questi grafici riportano la distribuzione delle scelte delle varie opzioni di risposta (le linee tratteggiate, che riportano il dato empirico rilevato) in funzione dell'abilità dello studente, sul complesso della prova; gli studenti sono distribuiti per decili.
Nella Figura 5, l'opzione A (la risposta corretta), rappresentata dalla linea verde con punti vuoti, è la più scelta solo da studenti posizionati oltre il quinto decile. Invece, il distrattore D tra tutte le opzioni errate, è il più scelto fino al penultimo decile e ha una probabilità di scelta più alta di quella della risposta corretta fino al 4° decile.

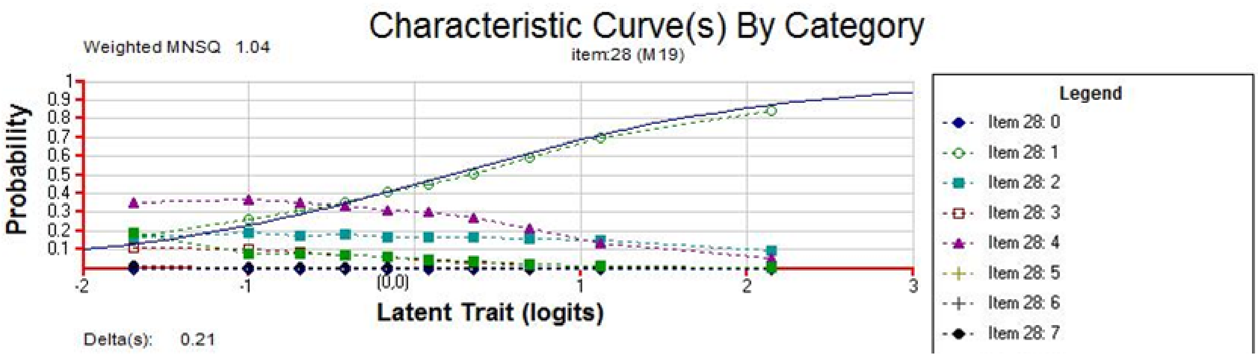


Figura 5. Curva caratteristica, Quesito 19, Prova INVALSI di matematica Grado 10, 2013 www.gestinv.it

Il secondo quesito preso in esame (Figura 6) fa parte della Prova INVALSI di matematica 2015 di grado 10 ed è stato somministrato a circa 550.000 studenti italiani e i risultati si riferiscono a un campione di 27.207 studenti.

D2. **Nell'insieme dei numeri reali la disequazione $x^2 + 1 \geq 0$ è verificata**

A. ☐ solo per $x \geq 0$

B. ☐ solo per $x \geq -1$

C. ☐ per ogni x

D. ☐ per nessun x

Figura 6. Quesito 2, Prova INVALSI di matematica Grado 10, 2013 www.gestinv.it

La risposta corretta è scelta da meno del 40% degli studenti. I distrattori A e B sono le due opzioni errate più scelte, con, rispettivamente, il 19,3% e il 23,4% di scelte (Figura 7).

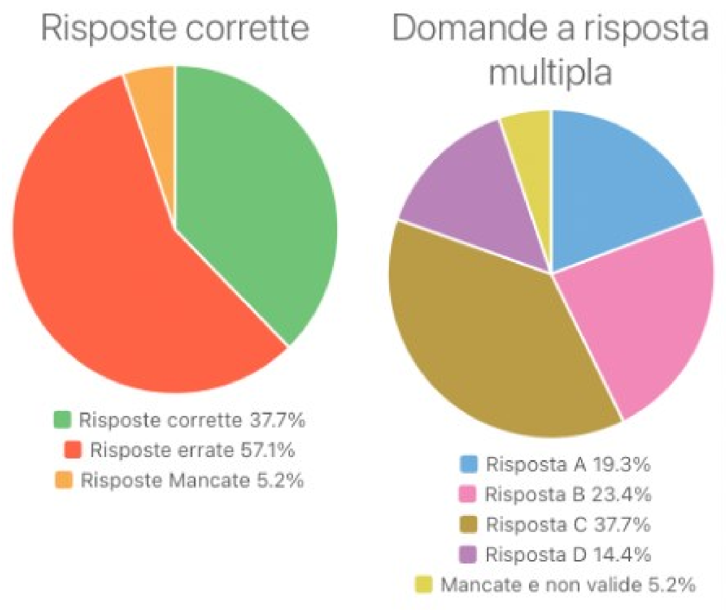


Figura 7. Percentuali nazionali, Quesito 2, Prova INVALSI di matematica Grado 10, 2013 www.gestinv.it

Come si evince dal grafico della curva caratteristica, fino a quasi il quinto decile, i distrattori A e B hanno una maggiore probabilità di essere scelti anche rispetto alla risposta corretta e rimangono tali fra le opzioni sbagliate fino al 7° decile (Figura 8).

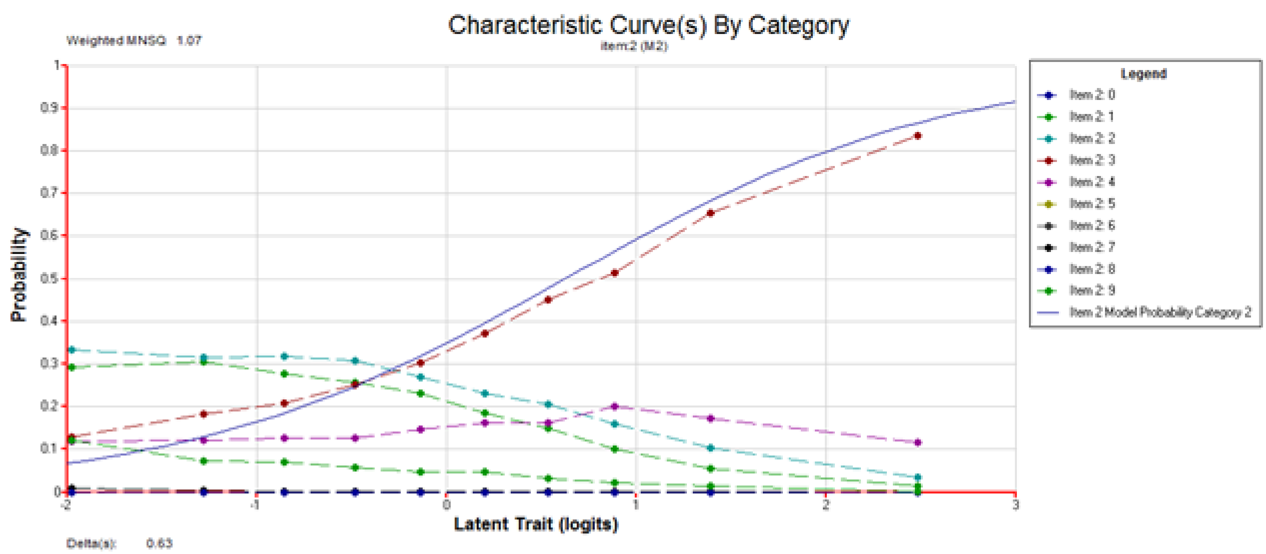


Figura 8. Curva caratteristica, Quesito 2, Prova INVALSI di matematica Grado 10, 2013 www.gestinv.it

I risultati evidenziano difficoltà diffuse nell'assegnare il significato corretto alle disuguaglianze espresse nel linguaggio simbolico.

In particolare, i due quesiti sono caratterizzati dal fatto che la soluzione non richiede necessariamente operazioni di trattamento esplicite.

Per comprendere l'evidente difficoltà emersa a livello di sistema di fronte a una consegna che a priori era possibile considerare standard siamo partiti utilizzando la Teoria della Reificazione della Sfard

(1991). Questa teoria considera i concetti matematici caratterizzati da una doppia natura, di processo e oggetto: si tratta di due aspetti complementari dell'oggetto matematico. Per avere una lente teorica che permetta di caratterizzare il lato-processo e il lato-oggetto abbiamo utilizzato la Teoria dell'Oggettivazione di Radford (2006, 2021) e l'approccio strutturale di Duval (1993, 2017).
Lo studio, basato sull'analisi delle principali possibili strategie per la risoluzione dei quesiti, mostra una possibile difficoltà nell'oggettivazione, nel senso di Radford, del significato corretto delle disuguaglianze. Il fragile controllo semantico nella comprensione delle disuguaglianze e il significato non radicato in un adeguato processo di oggettivazione si evincono anche dalle difficoltà nella gestione sia di trattamenti, che secondo gli studi Duval (2017) dovrebbero invece essere supportati dalla sintassi del registro algebrico, sia di conversioni che, in questi casi, renderebbero le soluzioni banali. Infatti, ad esempio, all'interno di un registro figurale, il significato sarebbe radicato in attività senso motoria in cui la variabile x si muove sull'asse x dallo 0 in poi nel primo quesito (Figura 3) e dall'1 in poi nel secondo (Figura 6) rendendo la soluzione intuitiva nel senso di Fischbein.

In dettaglio, si rileva quindi una generale distanza, nel momento della pratica, tra il significato culturale e il significato personale degli studenti.
A partire dal macro-fenomeno emerso abbiamo quindi condotto un'analisi a priori delle opzioni per inquadrare le possibili strategie messe in atto dagli studenti italiani e per avere un quadro interpretativo di risultati così diffusamente fallimentari.
Nell'analisi a priori abbiamo delineato sostanzialmente tre strategie che portano alla soluzione corretta o errata:

- svolgere calcoli algebrici, ovvero applicare la procedura per risolvere le disuguaglianze;
- non svolgere nessun calcolo e ricorrere ad argomentazioni in linguaggio naturale o interpretando grafici;
- senza ricorso a nessun calcolo, la soluzione è reificata (nel senso della Sfard) nell'espressione algebrica dell'equazione.

La prima strategia, ovvero il ricorso ai calcoli algebrici, è efficacemente interpretata da un approccio strutturale e funzionale: gli studenti eseguono un trattamento nel registro simbolico. Idealmente da un punto di vista strutturale, il significato dovrebbe provenire dal sistema semiotico: ad esempio, il passaggio da $x^2 + 1 > 0$ a $x^2 > -1$ (nel quesito in Figura 6) dovrebbe essere una trasformazione semiotica di trattamento. In termini di Duval, ci aspettiamo che le due espressioni algebriche abbiano sensi diversi ma la stessa denotazione. Questo non è il caso della maggioranza del campione nazionale che ha scelto le opzioni A o B; infatti, chi sceglie le prime due opzioni sbagliate non riconosce che le espressioni in A e B non sono equivalenti alla disuguaglianza presente del testo del quesito. Le due rappresentazioni simboliche non hanno alcun riferimento all'oggetto matematico (la disuguaglianza originale) e gli studenti cadono e sono bloccati nel paradosso cognitivo: avviene l'identificazione della rappresentazione con l'oggetto stesso e viene compromessa la relazione significato-significante (Duval, 2017). Gli studenti collegano la disuguaglianza originale con uno dei due distrattori in base al loro *significato personale*, ricorrendo forse ad analogie, somiglianze o simmetrie tra le due rappresentazioni simboliche. Siamo quindi di fronte a una distanza tra significato personale e significato culturale nel senso di Radford (2021). Probabilmente gli studenti hanno eseguito un calcolo mentale approssimativo e rapido senza eseguire tutti i passaggi della soluzione. Questo è un ulteriore importante elemento che dimostra un fragile controllo semantico nella comprensione delle disuguaglianze. Questo risultato, dal punto di vista educativo, è particolarmente significativo se si tiene conto del fatto che gli studenti italiani sono fortemente "addestrati" nel calcolo simbolico (Bolondi & Ferretti, 2019) e non sono così spesso esposti alle attività algebriche di problem solving e modellizzazione. Ci si aspetterebbe che gli studenti riconoscano un tipo familiare di esercizio e che possono facilmente svolgere ricorrendo a procedure che hanno "allenato" più volte. Il loro controllo sintattico si scontra (e prevale) con il loro scarso controllo semantico.

Chi adotta la seconda strategia, porterebbe eseguire una conversione in linguaggio naturale o interpretare la disuguaglianza in un registro grafico (immaginando x^2 che si muove su una retta di numeri o disegnando una parabola), senza la necessità del ricorso ad alcun calcolo simbolico. Gli studenti potrebbero quindi o far ricorso al linguaggio naturale per svolgere un'argomentazione o interpretare la disuguaglianza in senso grafico. È molto difficile che gli studenti di grado 10 abbiano effettivamente eseguito conversioni perché in Italia la geometria analitica viene introdotta nel grado 11 e le attività di argomentazione sono molto rare nelle pratiche didattiche tradizionali in campo algebrico e più comuni in geometria.

L'adozione della terza strategia ipotizzata deriva da una generalizzazione simbolica riuscita. Lo studente *oggettifica* la disuguaglianza ricorrendo alla designazione funzionale del linguaggio simbolico. Assumiamo che nelle generalizzazioni simboliche il significato è completamente *de-soggettivato,* ma sostenuto dalle oggettivazioni a livello di generalizzazioni più basse caratterizzate da aspetti *embodied* e da un uso indessicale dei simboli. Il ricorso a questa strategia richiede una buona competenza matematica e, nella maggior parte dei casi, necessita di una profonda ingegneria didattica che permette di esporre gli studenti a idonee attività riflessive mediate che coinvolgono anche le generalizzazioni simboliche.

I risultati nazionali INVALSI mostrano che, in entrambi i casi, meno della metà degli studenti del campione ha scelto la risposta corretta. I dati quantitativi non ci consentono di dedurre se gli studenti eseguono effettivamente il calcolo o scelgono un altro tipo di strategia. Il dato interessante è che oltre il 40% sceglie le opzioni errate B o D nel primo quesito e oltre il 30% sceglie le opzioni A o D nel secondo quesito. Questo è un risultato statisticamente rilevante che ci dice che gli studenti non hanno oggettivato il significato corretto delle disuguaglianze.

Per comprendere veramente la portata didattica di questo comportamento, è necessario considerare l'impatto che un quesito a scelta multipla ha sulla strategia di risoluzione. Infatti, la forma dei distrattori può amplificare la perdita del significato ed evidenziare la natura del comportamento pseudo-strutturale in termini di dialettica tra uso della semiotica in senso strutturale-funzionale e come mediatore delle pratiche matematiche. In particolare, la struttura del quesito mette in luce la rottura del significato dei segni come relazione significato-significante e come oggettivazione nei diversi livelli di generalizzazione (Radford, 2003). In un compito scolastico comune, che richiede esplicitamente la soluzione della disuguaglianza $x^2 + 1 > 0$, lo studente potrebbe attivare il meccanismo di risoluzione formale. La relazione con l'insegnante, le pratiche matematiche riconosciute, una presentazione familiare del compito e le clausole del contratto didattico potrebbero compensare e nascondere la mancanza di un significato matematico appropriato quando si tratta di linguaggio algebrico. Una prova "esterna", a cui gli studenti devono rispondere in sede di valutazione standardizzata nazionale, cambia lo scenario emotivo e cognitivo e può "portare allo scoperto" difficoltà e ostacoli che altrimenti possono rimanere nascosti nelle usuali dinamiche di classe.

3. Le analisi statistiche su cui si basano alcune delle nostre ricerche

3. 1 Il modello di Rasch

Le ricerche presentate in questo Seminario Nazionale prendono in considerazione il quadro teorico e metodologico proposto dall'Istituto INVALSI nei due principali documenti relativi alle prove che escono con cadenza annuale in riferimento alle rilevazioni: il *Rapporto dei Risultati* e il *Rapporto Tecnico*. Le metodologie quantitative di analisi dei risultati delle prove utilizzate nelle ricerche sono pertanto pienamente coerenti con le metodologie adottate dall'Istituto. La scelta dei modelli statistici utilizzati viene ampiamente illustrata all'interno dei Rapporti Tecnici (si veda, ad esempio INVALSI, 2017) e nel corso delle ricerche abbiamo più volte verificato la coerenza e l'adeguatezza, dal punto di vista didattico, di questi modelli (si veda, per esempio, Bolondi & Cascella, 2020). Vi sono infatti due principali approcci che possono essere adottati per analizzare rilevazioni su larga scala: la *Teoria Classica dei test* (CTT) e l'*Item Response Theory* (IRT) (Barbaranelli & Natali, 2005) ed entrambi vengono presi in considerazione per l'analisi dei dati raccolti dalle prove INVALSI all'interno del rapporto tecnico. La CTT fornisce importanti strumenti statistici per lo studio delle prove e permette una prima analisi relativa alla difficoltà dei quesiti o degli item che li compongono (intesa come percentuale di risposte corrette), discriminatività degli item (ovvero capacità di ogni item di distinguere studenti con livelli diversi di abilità) e coerenza interna della prova (in termini di Alpha di Cronbach) (Barbaranelli & Natali, 2005). Nonostante ciò, le analisi e le restituzioni dei risultati delle prove INVALSI sono basate principalmente sulla più moderna IRT (Rasch, 1960) che permette di superare le principali limitazioni della CTT tra cui per esempio la allora dipendenza tra la stima dell'abilità dei soggetti e la difficoltà dei quesiti della prova. Il modello di Rasch è il più semplice tra i modelli afferenti alla IRT, si tratta infatti di un modello logistico a un parametro che opera una stima congiunta tra il parametro di difficoltà relativo a ogni quesito della prova e il parametro di abilità attribuito a ogni studente.

Il modello di Rasch consente di calcolare la probabilità di rispondere correttamente a un determinato item, in funzione dell'abilità dello studente nella prova e delle caratteristiche psicometriche dell'item stesso (in particolare la difficoltà dell'item). Consideriamo quindi la matrice dicotomica di risposta a una prova con k item a cui hanno risposto N studenti; il modello di Rasch a partire da questa matrice opera una stima congiunta di due tipologie di parametri:

- a ogni studente *j* viene associato un parametro reale θ_j che rappresenta l'abilità del soggetto rispetto al tratto latente indagato dal test;
- a ogni item *i* viene associato un parametro reale β_i che rappresenta la difficoltà dell'item.

La probabilità che lo studente j risponda correttamente all'item i dipende quindi dall'abilità del soggetto θ_j e dalla difficoltà dell'item β_i, in particolare:

$$P(X_{ji} = 1) = \frac{e^{\theta_j - \beta_i}}{1 + e^{\theta_j - \beta_i}}$$

Per ogni item i di difficoltà β_i, la funzione:

$$\varphi(\theta_j - \beta_i) = \frac{e^{\theta_j - \beta_i}}{1 + e^{\theta_j - \beta_i}}$$

permette di esprimere la probabilità di rispondere correttamente in funzione del livello di abilità dello studente θ_j. La relazione tra abilità del soggetto e probabilità di rispondere correttamente all'item può essere rappresentata graficamente attraverso la curva teorica ipotizzata dal modello, detta anche *curva caratteristica dell'item* (Item Characteristic Curve – ICC), si tratta di una funzione logistica monotona crescente (si riporta un esempio in Figura 9).

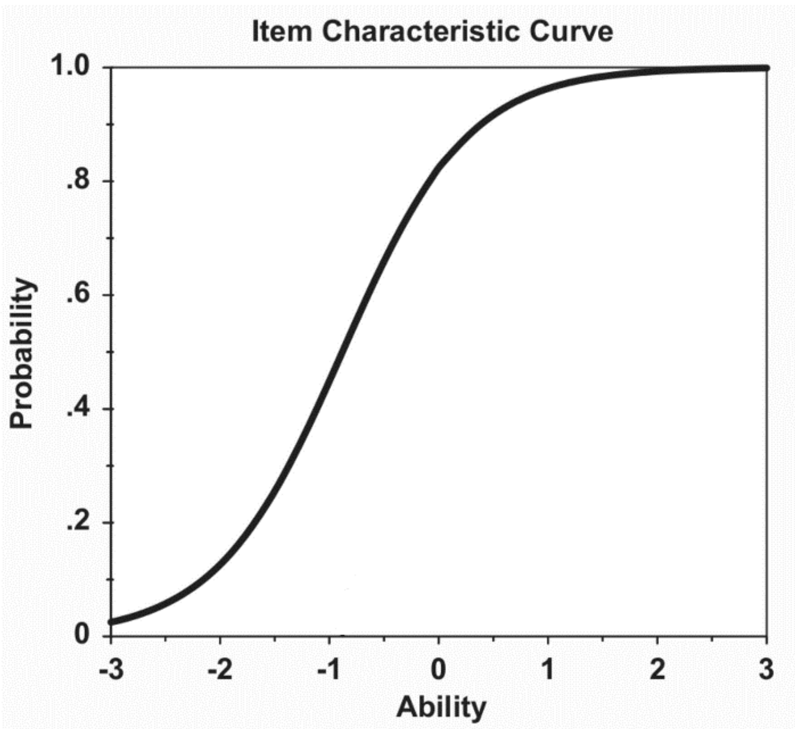


Figura 9. Esempio di *curva caratteristica* di un item (Item Characteristic Curve -ICC)

Nel modello di Rasch la risposta di uno studente è governata dall'interazione di due componenti: l'abilità del rispondente e la difficoltà dell'item. Non si parlerà quindi più di uno studente "bravo" in assoluto o di un item "difficile" in assoluto. La definizione della difficoltà di un item nella IRT è quindi molto diversa dall'interpretazione della difficoltà nella CTT (percentuale di soggetti nel campione che risponde correttamente) che risultava essere fortemente legata alle caratteristiche del campione a cui era stato somministrato il test.

Inoltre si osserva che un individuo con abilità θ_j ha una probabilità pari al 50% di rispondere correttamente a un item con livello di difficoltà $\beta_i = \theta_j$, infatti:

$$P(X_{ji} = 1) = \frac{e^0}{1+e^0} = \frac{1}{2} = 0{,}5$$

La difficoltà di un item nel modello di Rasch viene quindi interpretata come quel punto nella scala di abilità in cui la probabilità di rispondere correttamente è pari al 50% (Barbaranelli & Natali, 2005).

Assunzioni del modello di Rasch

Il modello di Rasch può essere però applicato solo nel caso in cui siano verificate alcune assunzioni, che permettono l'applicazione del metodo e la stima dei parametri (Hambleton, Swaminathan & Rogers, 1991):

- *Unidimensionalità* del test: l'abilità misurata dai diversi item del test deve essere relativa ad un unico tratto latente. La dimensione dominante deve essere quindi unica e questa condizione può essere verificata attraverso una analisi fattoriale esplorativa.

- *Indipendenza Locale*: mantenendo costante il livello di abilità θ degli studenti, le risposte agli item sono eventi indipendenti.

- *Monotonicità*: per ogni item, la probabilità di rispondere correttamente deve crescere monotonicamente all'aumentare del livello di abilità dei rispondenti.

Indipendenza locale e unidimensionalità sono fattori strettamente legati e, avendo una conferma attraverso l'analisi dei dati empirici della unidimensionalità del test, allora anche l'ipotesi di indipendenza locale risulta essere verificata.
La monotonicità può essere verificata attraverso la rappresentazione grafica dei dati empirici, in particolare si può rappresentare la probabilità di risposte corrette a ogni item in funzione del punteggio ottenuto sull'intera prova (punteggio di Rasch o punteggio grezzo).

Le analisi relative a queste assunzioni sono effettuate per ogni prova direttamente dall'Istituto INVALSI e i risultati sono riportati nel rapporto Tecnico INVALSI relativo all'annualità della prova.

3.2 Distractor Plot

Una volta applicato il modello di Rasch, informazioni molto utili riguardanti i singoli item possono essere estrapolate osservando specifici grafici chiamati distractor plot. Nello stesso grafico in cui viene rappresentata la curva teorica ICC relativa alla risposta corretta (linea continua), vengono rappresentati anche i dati empirici relativi alla risposta corretta e alle altre opzioni di risposta. In questo modo è possibile osservare quanto la curva empirica della risposta corretta sia coerente con la curva teorica e, inoltre, si può analizzare l'andamento di ogni distrattore (inteso come risposta non corretta) in funzione del livello di abilità degli studenti.

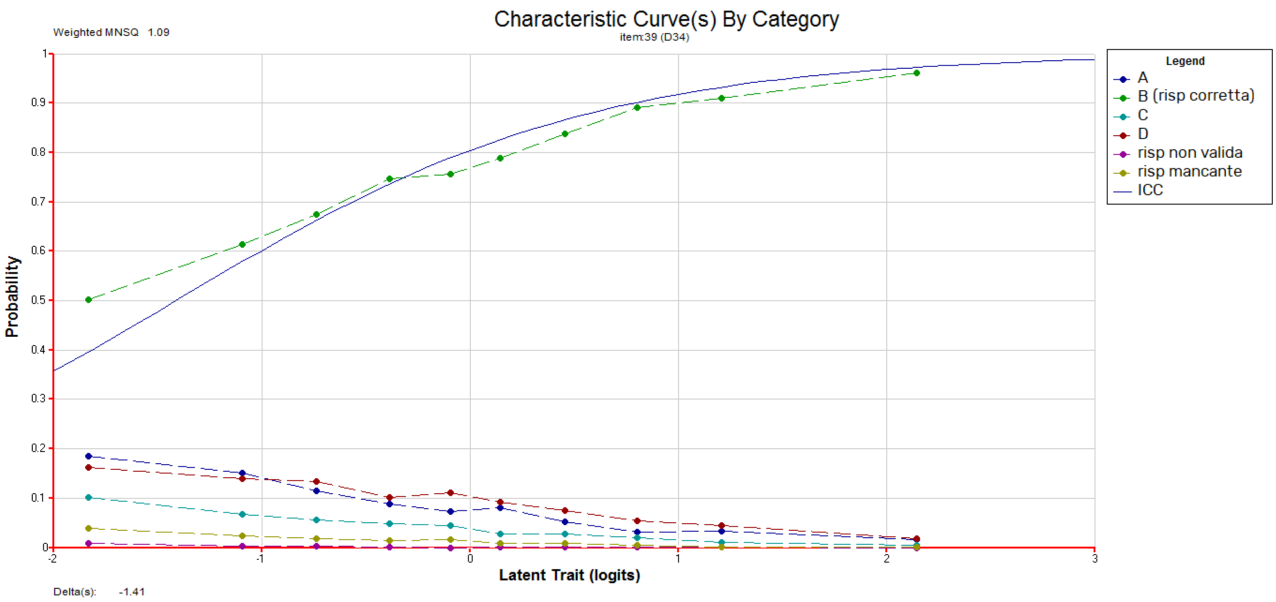


Figura 10. Esempio di *distractor plot* di un item

Il grafico (Figura 10), in ascissa, riporta il punteggio di Rasch in termini di abilità degli studenti θ sull'intera prova e la linea continua rappresenta la curva teorica del modello che esprime la probabilità di rispondere correttamente in funzione del livello di abilità. Le spezzate tratteggiate rappresentano i dati empirici ricavati dalla somministrazione dell'item: per ogni decile (corrispondenti ai pallini), relativo al livello di abilità sull'intera prova, è stata riportata la probabilità di scelta di ciascuna opzione di risposta.

In questo item, il confronto tra l'andamento della risposta corretta empirica e la curva teorica risulta essere accettabile (come confermato anche dall'indice *weighted* = 1,09) anche se il modello tende a sottostimare gli studenti con livelli di abilità bassi. I distrattori più attrattivi risultano essere l'opzione A e l'opzione D che hanno una significativa probabilità di essere scelti principalmente nei livelli bassi. Anche l'altro distrattore comunque mostra un buon funzionamento (risulta infatti attrattivo e coerente con il tratto latente). Infine si può notare che solo pochi studenti non rispondono all'item, quasi tutti appartenenti al decile più basso in termini di abilità.

In generale, al crescere dell'abilità dello studente nella prova si dovrebbe osservare una probabilità maggiore che egli fornisca una risposta corretta mentre le opzioni di risposta errate, i distrattori, dovrebbero mostrare un andamento decrescente al crescere dell'abilità degli studenti. Esistono però casi in cui l'andamento delle opzioni di risposta errate non seguono questo principio: in alcuni quesiti può succedere che un'opzione di risposta errata risulti particolarmente attrattiva per studenti con livelli di abilità medi o medio-alti. Ed è proprio l'analisi dell'andamento di queste curve, che abbiamo definito "andamento a pancia" o "humped performance trend", che sarà alla base delle ricerche presentate nel Capitolo 4. Risulta fondamentale sottolineare quanto questo non vada a ledere l'assunzione di monotonicità, infatti l'andamento strettamente crescente della risposta corretta è sempre verificato e, se si osservano le risposte errate e mancanti nella loro totalità, è confermato anche per queste un andamento strettamente decrescente; gli andamenti non monotoni possono però essere osservati in relazione a singoli distrattori e possono essere interpretati attraverso le lenti teoriche della didattica della matematica.

4. Due esempi di macro-fenomeni emersi dall'analisi delle curve empiriche: distrattori con andamento "a pancia"

In questa sezione verranno analizzati due quesiti tratti da diversi livelli scolastici (dalla scuola primaria alla scuola secondaria di secondo grado) in cui si presenta un macro-fenomeno interessante da interpretare sul piano didattico, Si tratta di item che mostrano buone proprietà misuratorie (INVALSI, 2017; Barbaranelli & Natali, 2005) e in cui almeno un distrattore evidenzia un andamento "a pancia".

L'analisi di questi quesiti è stata alla base di diverse ricerche (Bolondi, Ferretti & Giberti, 2018, Ferretti, Giberti, & Lemmo, 2018; Ferretti, Giberti & Lemmo, 2020; Ferretti & Giberti, 2021; Giberti, 2018) condotte con l'obiettivo di indagare caratteristiche e peculiarità dei macro-fenomeni emersi dai quesiti i cui distractor plot presentano questo particolare andamento.

Nel primo esempio mostriamo come il peculiare andamento della curva empirica di un determinato distrattore evidenzi un macro-fenomeno inquadrabile con alcuni dei costrutti più noti e diffusi in letteratura, nel caso specifico riconducibili all'idea di *contratto didattico* nel senso di Brousseau (1986). Nel secondo esempio, invece, mettiamo in risalto come lo stesso andamento permetta di sviluppare spunti di riflessioni per quanto riguarda il legame tra il contratto didattico e il gender gap.

4.1 Un primo esempio: andamento "a pancia" e contratto didattico

Il primo quesito con andamento "a pancia" è tratto da una prova INVALSI somministrata in quinta primaria (grado 5) nel 2015. Il quesito appartiene all'ambito "Relazioni e Funzioni" e richiede di determinare la soluzione di un problema descritto attraverso un testo scritto. Esso può essere identificato come *problema verbale di matematica* (word problem): un compito presentato tramite un testo scritto in forma verbale eventualmente integrato attraverso il simbolismo matematico (Gerofsky, 1996). In questo caso, alcune informazioni numeriche sono fornite nel registro numerico-simbolico, altre nel registro verbale nel flusso del discorso di descrizione della situazione.

D7. Francesca prepara per il gatto due pasti al giorno utilizzando cibo in scatoletta.

Con il contenuto di una scatoletta Francesca prepara 3 pasti per il gatto.

Francesca ha comprato 8 scatolette di cibo per gatti. Per quanti giorni al massimo le bastano?

A. ☐ 24

B. ☐ 16

C. ☐ 8

D. ☐ 12

Figura 11. Quesito 7, Prova INVALSI di matematica Grado 5, 2015, www.gestinv.it

Si tratta di un quesito a risposta multipla con un'unica risposta corretta che corrisponde all'opzione D, scelta solo dal 30% degli studenti del campione nazionale (Figura 12). Per individuare la risposta corretta al quesito, è necessario considerare le informazioni presenti rappresentate attraverso diversi registri semiotici. Ad esempio, il numero di scatolette di cibo per gatti e il numero di pasti che prepara con ogni scatoletta sono espressi in forma numerica, rispettivamente "8" e "3", mentre il numero di pasti al giorno che Francesca dà al gatto sono indicati in linguaggio naturale, "due".

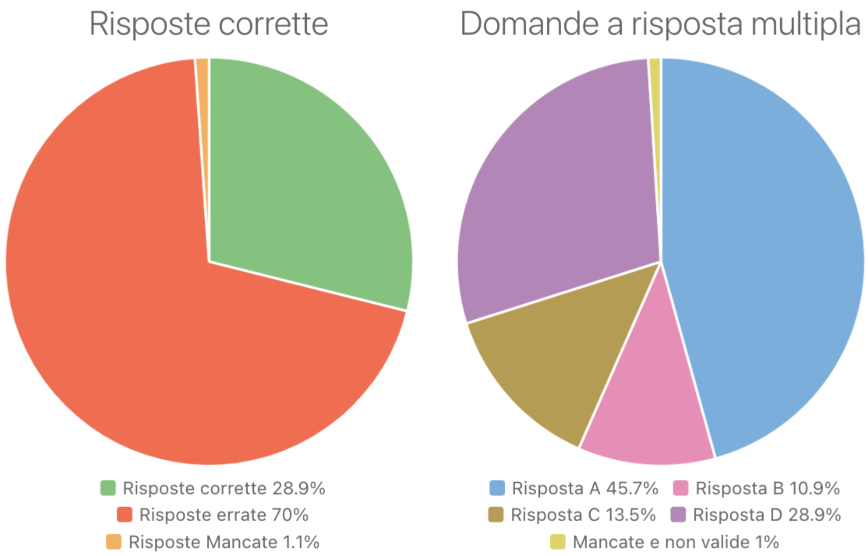


Figura 12. Percentuali nazionali, Quesito 7, Prova INVALSI di matematica Grado 5, 2015 www.gestinv.it

Come mostrato in Figura 12, si tratta di un quesito che può essere considerato difficile (delta=1,10), l'andamento della risposta corretta empirica rispetto alla curva teorica risulta essere più che accettabile (weighted = 1,04) e il grado di discriminazione è buono (0,35).

IRT

item:10 (D7)

Cases for this item 22030 Item-Rest Cor. 0.30 Item-Total Cor. 0.35

Item Threshold(s): 1.10 Weighted MNSQ 1.04

Item Delta(s): 1.10

Label	Score	Count	% of tot	Pt Bis	t	sig. p	PV1Avg:1	PV1 SD:1
1	0	10058	45,66	-0,07	-9,98	0,000	-0,110	0,910
2	0	2407	10,93	-0,1	-15,22	0,000	-0,310	0,950
3	0	2969	13,48	-0,19	-28,09	0,000	-0,490	0,960
4	1	6363	28,88	0,3	46,42	0,000	0,540	1,070
7	0	22	0,1	-0,02	-3,14	0,002	-0,740	1,210
9	0	211	0,96	-0,06	-8,97	0,000	-0,630	0,870

Figura 13. Analisi IRT, Quesito 7, Prova INVALSI di matematica, Grado 5, 2015, www.gestinv.it

Il 70% degli studenti del campione nazionale ha scelto come risposta un'opzione non corretta (Figura 13); in particolare, circa un quarto degli studenti del campione nazionale ha scelto l'opzione B (16) o l'opzione C (8). Guardando il distractor plot per entrambe queste due opzioni (Figura 14), si può notare che le curve hanno un andamento decrescente: all'aumentare dell'abilità degli studenti nella prova, diminuisce la probabilità di scelta di entrambi i distrattori. La probabilità di scegliere questi distrattori è quindi maggiore per gli studenti di questi primi decili.

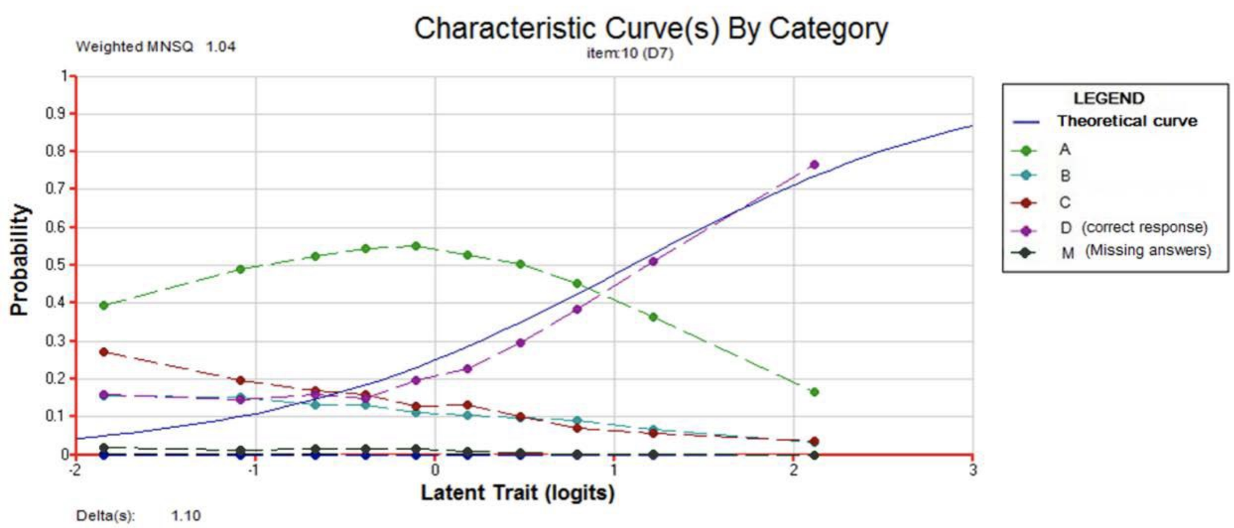


Figura 14. Curva caratteristica, Quesito 7, Prova INVALSI di matematica Grado 5 - 2015, www.gestinv.it

Sempre in riferimento al grafico (Figura 14), si può notare che la curva riferita all'opzione di risposta A presenta un andamento differente dalle curve delle altre opzioni di risposta errate. L'opzione A è stata scelta da oltre il 45% degli studenti del campione nazionale ed è risultata l'opzione di risposta più attrattiva fino a livelli di abilità molto alti. Il grafico mostra infatti una probabilità di scelta di quasi il 20% nel decile più alto, e di quasi il 40% nel nono decile. La semplice analisi delle percentuali complessive di risposta non avrebbe permesso di inferire il grado di attrattività dell'opzione rispetto al livello di abilità degli studenti; il grafico delle curve riferite alle diverse opzioni di risposta, mette in luce la relazione tra l'attrattività delle singole opzioni di risposta con le abilità degli studenti nella prova. Nel caso dell'opzione A si nota che la risposta è stata particolarmente attrattiva soprattutto per gli studenti che hanno mostrato un livello di abilità medio in riferimento alla prova. Entrando nel dettaglio, per gli studenti che si collocano nel primo decile, la probabilità di scegliere l'opzione A è circa il 40% e, al posto di decrescere, questa probabilità si alza nei decili successivi e raggiunge il suo massimo (corrispondente a una probabilità circa pari al 55%) in corrispondenza del quinto decile della distribuzione, vale a dire per gli studenti di livello mediano (corrispondente al valore 0 della scala). Per livelli di abilità superiori a 0, la curva relativa all'opzione A comincia a decrescere e si interseca con la curva della risposta esatta solo tra l'ottavo e il nono decile. In altri termini, la scelta di questo distrattore è massima per gli studenti sulla mediana della distribuzione di abilità.

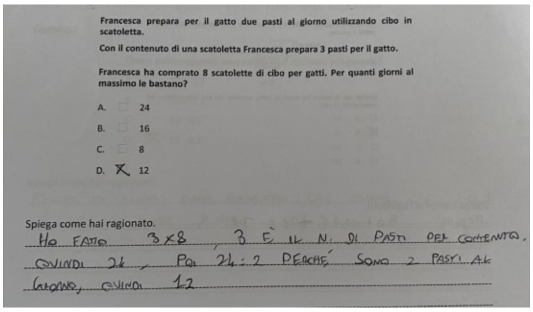
Per poter analizzare i procedimenti risolutivi messi in atto dagli studenti, è stata svolta una ricerca di tipo qualitativo in cui è stato somministrato il quesito così come presente nella prova INVALSI ed è stato richiesto di spiegare il ragionamento effettuato.

Alla ricerca hanno partecipato 58 studenti della scuola secondaria di I grado dell'Istituto Comprensivo Paolo Stefanelli (Roma). Il campione è composto da 20 studenti di grado 6, (classe prima), 18 di grado 7 (classe seconda) e 20 di grado 8 (classe terza). Il focus di questa ricerca è incentrato sull'analisi delle strategie utilizzate e delle motivazioni sottostanti le scelte delle varie opzioni: proprio per questo motivo abbiamo scelto un campione variegato nei gradi e non limitato al grado specifico in cui era stato somministrata la domanda in questione.

Trattandosi di un quesito rivolto alle classi quinte di scuola primaria, si è riscontrato un auspicabile aumento delle percentuali di risposte corrette rispetto a quelle ottenute nel campione nazionale: la metà o più degli studenti degli studenti che hanno preso parte alla ricerca ha infatti fornito la risposta corretta.

La maggior parte degli studenti che ha risposto correttamente, indipendentemente dal grado scolastico, ha motivato fornendo una spiegazione che metteva in relazione il numero di pasti totali possibili con 8 scatolette (3x8) e il numero di pasti al giorno (Figura 15a). Due studenti, uno di grado

7 e uno di grado 8, hanno rappresentato la situazione attraverso schemi (Figura 15b); infine, tre studenti di grado 8 hanno impostato una proporzione (Figura 15c).
La stessa strategia è stata quindi implementata attraverso il ricorso a diverse forme di matematizzazione della situazione.

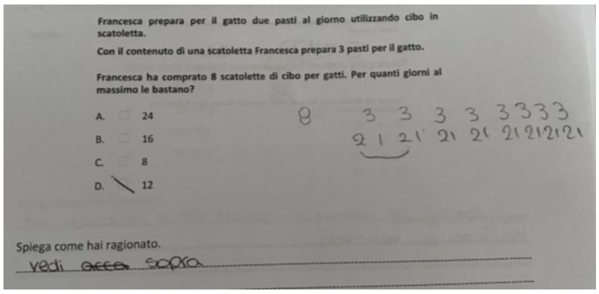


Francesca prepara per il gatto due pasti al giorno utilizzando cibo in scatoletta.
Con il contenuto di una scatoletta Francesca prepara 3 pasti per il gatto.
Francesca ha comprato 8 scatolette di cibo per gatti. Per quanti giorni al massimo le bastano?

A. 24
B. 16
C. 8
D. X 12

Spiega come hai ragionato.
HO FATTO 3x8, 3 È IL N. DI PASTI PER CONTENUTO, QUINDI 24, POI 24:2 PERCHÉ SONO 2 PASTI AL GIORNO, QUINDI 12

Figura 15a



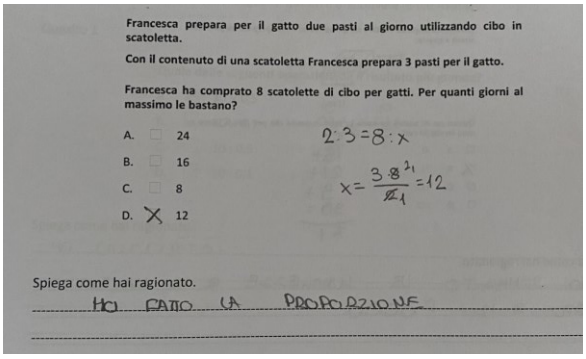
Francesca prepara per il gatto due pasti al giorno utilizzando cibo in scatoletta.
Con il contenuto di una scatoletta Francesca prepara 3 pasti per il gatto.
Francesca ha comprato 8 scatolette di cibo per gatti. Per quanti giorni al massimo le bastano?

A. 24
B. 16
C. 8
D. \ 12

8 3 3 3 3 3 3 3 3
21 21 21 21 21 21 21 21

Spiega come hai ragionato.
vedi sopra

Figura 15b



Francesca prepara per il gatto due pasti al giorno utilizzando cibo in scatoletta.
Con il contenuto di una scatoletta Francesca prepara 3 pasti per il gatto.
Francesca ha comprato 8 scatolette di cibo per gatti. Per quanti giorni al massimo le bastano?

A. 24
B. 16
C. 8
D. X 12

$2:3=8:x$

$x=\frac{3\cdot 8}{2}=12$

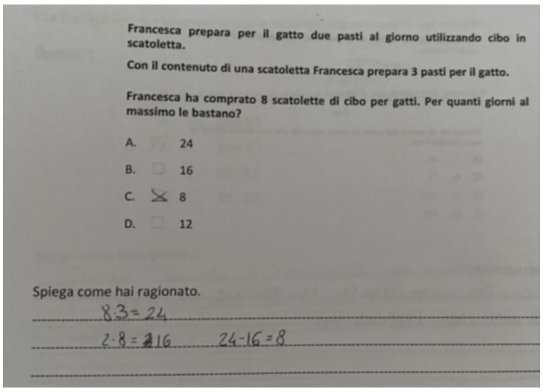
Spiega come hai ragionato.
HO FATTO LA PROPORZIONE

Figura 15c

Figura 15. Esempi di protocolli - risposte corrette

Solo uno studente di grado 8 ha scelto l'opzione B (16) motivando con l'affermazione "ho fatto i calcoli". Questo genere di risposta non ci permette alcuna supposizione in merito alla strategia utilizzata per scegliere questa opzione di risposta. Però, nella sua motivazione lo studente esplicita di aver svolto una qualche forma di calcolo; questo potrebbe suggerire che si sia concentrato esclusivamente su due dei dati presenti nel testo, 2 e 8, e successivamente abbia cercato tra le opzioni di risposta possibili un risultato ottenibile combinando i due dati (8x2=16). Questo genere di comportamento potrebbe essere associabile ad una lettura selettiva del testo (Zan, 2016). In questo caso, il testo non presenta particolari parole che possano suggerire l'operazione da utilizzare ma il contesto suggerito dall'affermazione "due pasti al giorno" potrebbe aver guidato lo studente nella scelta della moltiplicazione.
Cinque studenti di grado 8 hanno scelto l'opzione di risposta C (8). Due hanno moltiplicato il numero di scatolette sia per i pasti giornalieri che per il numero di pasti possibili con ciascuna scatoletta e

successivamente ne ha fatto la differenza (Figura 16a); uno studente ha invece cercato di operare con i numeri in modo da evitare un risultato decimale per poi tornare al valore iniziale (Figura 16b). Questi ultimi tre casi presentati potrebbero essere interpretati con la clausola di *esigenza di giustificazione formale* (D'Amore & Sandri, 1998, D'Amore, 1999) in quanto sembra che gli studenti manipolino i numeri presentati nel testo allo scopo di identificare una soluzione corrispondente ad una delle opzioni di risposta presentate o un numero naturale. In nessuno di questi casi, emerge una motivazione che esuli dai semplici calcoli o dalla ricerca di una soluzione presente tra le opzioni di risposta. Solo uno studente propone una motivazione che collega la strategia scelta al testo del problema (Figura 16a). Dalla sua motivazione, sembra aver supposto che una scatoletta fosse sufficiente per fornire esclusivamente i pasti giornalieri. In questo caso, al contrario dei precedenti, potremmo pensare anche a un riferimento all'esperienza personale, alla "conoscenza enciclopedica" degli studenti (Zan, 2016) per cui è ragionevole che una scatoletta sia dimensionata per preparare i pasti di un'intera giornata o di due giornate.



Francesca prepara per il gatto due pasti al giorno utilizzando cibo in scatoletta.

Con il contenuto di una scatoletta Francesca prepara 3 pasti per il gatto.

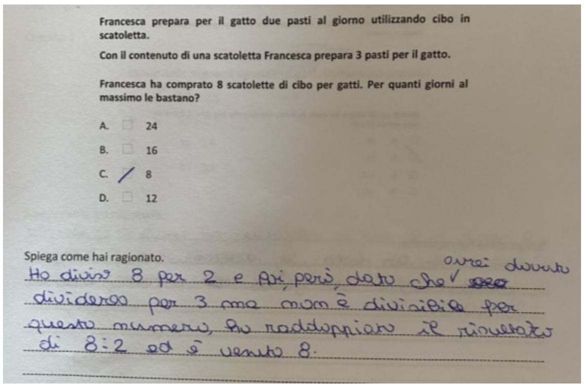
Francesca ha comprato 8 scatolette di cibo per gatti. Per quanti giorni al massimo le bastano?

A. 24
B. 16
C. 8
D. 12

Spiega come hai ragionato.

8·3=24
2·8=16 24-16=8

Figura 16a



Francesca prepara per il gatto due pasti al giorno utilizzando cibo in scatoletta.

Con il contenuto di una scatoletta Francesca prepara 3 pasti per il gatto.

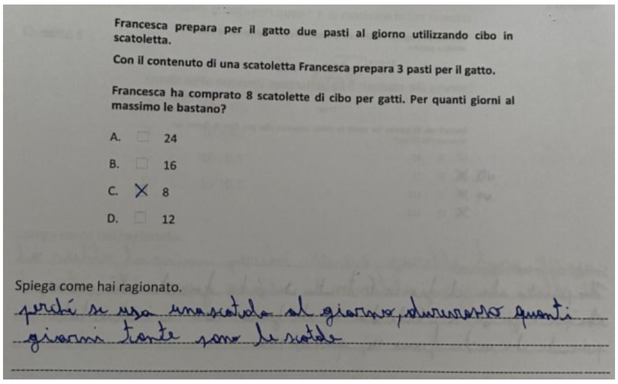
Francesca ha comprato 8 scatolette di cibo per gatti. Per quanti giorni al massimo le bastano?

A. 24
B. 16
C. 8
D. 12

Spiega come hai ragionato.

Ho diviso 8 per 2 e poi, però, dato che avrei dovuto dividerlo per 3 ma non è divisibile per questo numero, ho raddoppiato il risultato di 8:2 ed è venuto 8.

Figura 16b



Francesca prepara per il gatto due pasti al giorno utilizzando cibo in scatoletta.

Con il contenuto di una scatoletta Francesca prepara 3 pasti per il gatto.

Francesca ha comprato 8 scatolette di cibo per gatti. Per quanti giorni al massimo le bastano?

A. 24
B. 16
C. 8
D. 12

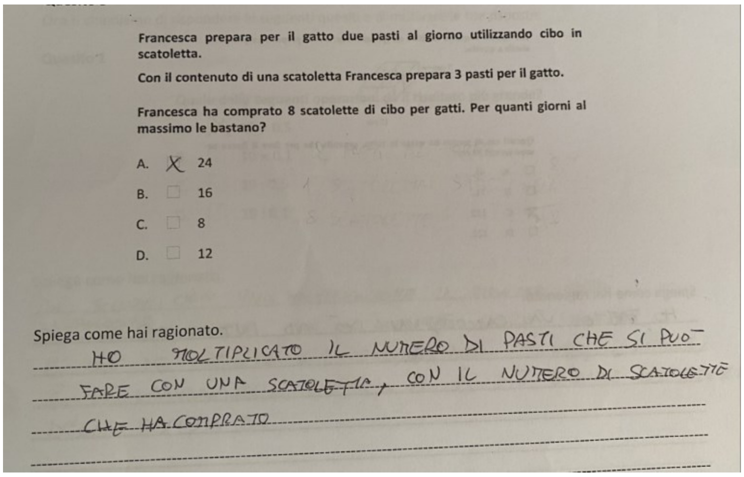
Spiega come hai ragionato.

perché si usa una scatola al giorno, dureranno quanti giorni tante sono le scatole

Figura 16c

Figura 16. Esempi di protocolli - opzione C.

Sono interessanti le risposte date da coloro che hanno scelto l'opzione di risposta A (19 studenti: nove studenti di grado 6, sei studenti di grado 7 e quattro studenti di grado 8).
Due studenti, uno di grado 6 e uno di grado 7, rispondono rispettivamente "ho seguito il mio istinto" (sic!) e "l'ho fatto a sentimento perché non me lo ricordavo". Questo potrebbe indicare che hanno scelto l'opzione di risposta che immediatamente poteva essere intuita come una relazione moltiplicativa tra gli unici dati numerici presenti nel testo.
Quattro studenti (tre di grado 6 e uno di grado 7) hanno invece esplicitamente motivato di aver determinato il numero totale di pasti, mostrando quindi di non aver letto o correttamente interpretato la domanda del quesito (Figura 17). Questo comportamento potrebbe essere collegato ad un comportamento prototipico dovuto alla prassi didattica di presentare problemi in cui il contesto e la domanda sono scollegate. In questo caso, potrebbe esserci stata una frattura tra il contesto e la domanda (Zan, 2012) per cui gli studenti hanno risposto ad una domanda diversa da quella presentata nel quesito; ad esempio: "Quanti pasti al massimo può dare al gatto?".


Francesca prepara per il gatto due pasti al giorno utilizzando cibo in scatoletta.

Con il contenuto di una scatoletta Francesca prepara 3 pasti per il gatto.

Francesca ha comprato 8 scatolette di cibo per gatti. Per quanti giorni al massimo le bastano?

A. X 24
B. ☐ 16
C. ☐ 8
D. ☐ 12

Spiega come hai ragionato.
HO MOLTIPLICATO IL NUMERO DI PASTI CHE SI PUO' FARE CON UNA SCATOLETTA, CON IL NUMERO DI SCATOLETTE CHE HA COMPRATO

Figura 17. Esempio di protocollo - opzione A

Il restante numero di studenti (quasi equidistribuiti rispetto ai tre gradi: cinque di grado 6, cinque di grado 7 e tre di grado 8) non ha fornito alcuna motivazione al di fuori della moltiplicazione tra 3 e 8. Tutte queste risposte presentano frasi del tipo: "tre per otto ventiquattro" oppure "ho fatto 3x8=24" oppure "8x4=24" ("forzando" il risultato dell'operazione). Un esempio in Figura 18.
I due terzi degli studenti che hanno scelto 24 come risposta sembrano rinunciare a fornire una giustificazione che vada oltre l'operazione eseguita, a differenza di quanto era emerso per le altre opzioni di risposta. Questo potrebbe rafforzare l'ipotesi che tali studenti abbiano individuato la risposta rinunciando a considerare il senso del contesto e della domanda e la richiesta di

giustificazione, individuando semplicemente la tipologia di problema (moltiplicativo) e i dati numerici presentati.

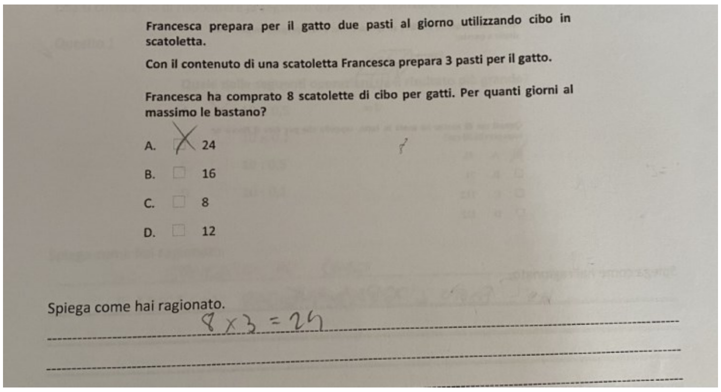


Figura 18. Esempio di protocollo - opzione A

Per tutti questi casi, una possibile interpretazione potrebbe essere quella per cui gli studenti hanno riconosciuto un problema moltiplicativo e hanno selezionato gli unici due dati nel testo espressi attraverso un registro numerico. Il fatto di utilizzare solo i dati numerici presenti nel testo e di non effettuare un controllo semantico del risultato, anche in termini di coerenza con lo stimolo, sono inquadrabili con alcuni degli aspetti del contratto didattico nel senso di Brousseau (1986), sia in termini di *Effetto dell'età del capitano* (Baruk, 1985), sia in termini di *Clausola di delega formale* (D'Amore, 1999).

Questa raccolta di esempi di risposte, con le tipologie di processi che fanno intravedere, mostrano come le traiettorie dei singoli studenti possano essere molto diverse e sembrare, in certi casi, addirittura casuali. D'altra parte, il fenomeno che emerge dal dato sistemico è molto evidente, in particolare nel mettere in luce, nel caso dei distrattori a pancia, come certe tipologie di "errore", che potrebbero essere considerate come tipici comportamenti di studenti particolarmente deboli, sono invece presenti in misura molto consistente anche in studenti collocati in decili "medi". Tutto questo conferma l'assunto generale alla base di queste ricerche: l'integrazione tra il dato quantitativo sistemico, e l'osservazione locale qualitativa, se supportata dal ricorso a costrutti teorici validati della didattica della matematica, permette di gettare luce su fenomeni complessi e fornisce indicazioni spendibili anche agli insegnanti. Nella sezione finale di questo seminario presenteremo esempi di come tutto questo possa integrarsi in percorsi di sviluppo professionale degli insegnanti.

L'analisi statistica effettuata attraverso il modello di Rasch e, in particolare, l'analisi dei distractor plot può essere particolarmente utile per evidenziare come rispondono gli studenti in funzione della loro abilità misurata sull'intera prova. Nello specifico, l'esempio presentato in questa sezione permette di focalizzare l’attenzione su specifici distrattori (quindi su specifiche difficoltà) facendo emergere su quali livelli di abilità questi distrattori risultino più attrattivi.

4.2 Un secondo esempio: andamento “a pancia” e differenze di genere

L’analisi statistica attraverso il modello di Rasch permette pertanto di avere informazioni su come rispondono gli studenti in funzione del loro livello di abilità, ma non solo: dall'analisi dei risultati delle prove INVALSI (e, in generale, delle valutazioni su larga scala) si possono ricavare importanti informazioni anche su come rispondono sottogruppi della popolazione. Questi sottogruppi possono essere individuati considerando altre caratteristiche, quali ad esempio il genere o lo status di cittadinanza. Anche in questo caso, il dato complessivo quantitativo è per certi versi "cieco", e mette in luce eventuali differenze tra sottogruppi. Un'analisi a livello dei singoli item e ricerche specifiche in contesto possono contribuire a interpretare i macro-fenomeni emergenti. Questo tipo di analisi è stato sviluppato per diverse tipologie di sottogruppi; presentiamo qui alcuni risultati relativi alle differenze di genere.

Il tema delle differenze di genere in matematica emerge chiaramente quando si prendono in considerazione i risultati delle valutazioni standardizzate, siano esse internazionali come le prove OCSE-PISA e IEA-TIMSS, o nazionali come le prove INVALSI. La forte disparità tra i risultati in matematica di maschi e femmine (a favore dei maschi), è affrontata nella letteratura internazionale ed i risultati delle valutazioni standardizzate confermano questa tendenza in tutti i gradi scolastici (Mullis, Martin, Foy & Hooper, 2016; OECD, 2016a). L'Italia, tra i paesi partecipanti alle rilevazioni internazionali, presenta uno dei divari di genere più marcati e il netto divario in matematica tra maschi e femmine in Italia è confermato anche nei vari gradi scolastici dai risultati delle indagini nazionali di valutazione standardizzata INVALSI (Mullis et al., 2016; OECD, 2016b; Giberti, 2019).
Sono molte le cause che vengono prese in considerazione da diverse prospettive teoriche, ideologiche (e anche politiche) per spiegare tale divario nei risultati. Una prima analisi interpretativa parte dal confronto tra paesi in cui questo gap si presenta in maniera diversa (Leder & Forgasz, 2008). In primo luogo, i dati delle prove internazionali mostrano che il divario tra i risultati dei maschi e delle femmine non è uniforme in tutti i sistemi scolastici analizzati: nella maggior parte dei Paesi, il divario è a favore dei maschi, ma ci sono eccezioni in cui è a favore delle femmine (OECD, 2016a). Questo fatto supporta la teoria secondo cui le cause di natura biologica e fisiologica, considerate in alcuni studi (Baron-Cohen & Wheelwright, 2004), non possano essere considerate un fattore predominante nell’emergere del divario. Se le cause fossero prevalentemente di natura biologica, le differenze dovrebbero essere più o meno le stesse in tutti i Paesi (Hill, Corbett, & Rose, 2010). La non

omogeneità nella distribuzione del gender gap supporta anche i numerosi studi che sostengono il forte impatto dei fattori sociali e culturali (Guiso, Monte, Sapienza & Zingales, 2008; Giberti & Spagnolo, 2021). Ad esempio, diversi studi hanno dimostrato che il divario di genere in matematica è strettamente legato all'emancipazione delle donne nella società e che nelle società in cui si raggiunge l'uguaglianza di genere, questo divario tende a scomparire (González de San Román & De La Rica, 2012).

Il divario delle performance è anche fortemente influenzato dalle convinzioni e dagli atteggiamenti di insegnanti, genitori e studenti nei confronti della matematica, che sono spesso dettati anche da stereotipi di genere (Jacobs & Bleeker, 2004; Tomasetto, 2013) e da fattori metacognitivi strettamente legati alla matematica (Herbert & Stipek, 2005; Hill et al., 2010; OECD, 2015). In attività di problem solving, un livello più basso di autoefficacia e di fiducia in sé stesse in matematica (OECD, 2015) può portare le femmine ad avere più paura di sbagliare, preferendo così strategie di risoluzione già note e adottate in classe piuttosto che sperimentarne di nuove (Bell & Norwood, 2007). Infine, fattori strettamente legati al contesto della classe, come il curriculum, le pratiche didattiche e valutative, possono avere un impatto diverso su maschi e femmine (Leder, 1992; Leder & Forgasz, 2008; Ferrara et al., 2021).

Anche fattori micro-sociali che nascono nel contesto della classe possono essere alla base delle differenze di genere in matematica: un maggiore legame delle ragazze con le pratiche didattiche e con l'insegnante può, ad esempio, portare le ragazze a essere più influenzate da misconcezioni e da effetti legati al contratto didattico (Ferretti & Giberti, 2021; Giberti, Zivelonghi & Bolondi, 2016; Giberti, 2019).

Pertanto, le valutazioni standardizzate risultano fondamentali nel confrontare le performance di maschi e femmine sull'intera prova, al fine di osservare l'emergenza e l'entità del gender gap in matematica. Un passo ulteriore però può essere fatto portando l'analisi del gender gap a livello dei singoli quesiti (Bolondi, Cascella & Giberti, 2017; Giberti, 2019). Nelle prove INVALSI, il gap tra maschi e femmine non è uniformemente distribuito su tutti i quesiti, e per alcuni quesiti si ha anche un risultato migliore per le femmine. Anche questa evidenza può essere studiata con due approcci complementari. Da un lato, si possono analizzare qualitativamente i quesiti per cui il gap è particolarmente rilevante, cercando di individuarne caratteristiche comuni (di formulazione, contesto, lunghezza, contenuti, ...). Dall'altro, questi quesiti possono essere analizzati dal punto di vista psicometrico, per vedere se è possibile individuare ulteriori elementi di caratterizzazione. Per una individuazione dei quesiti INVALSI che presentano un gap particolarmente rilevante, rimandiamo a Bolondi, Cascella e Giberti (2017). Per una ricerca nella seconda direzione, si veda il lavoro di

Cascella, Giberti e Bolondi (2020). Il gap tra maschi e femmine non è uniformemente distribuito tra tutti i quesiti e, per i quesiti in cui risulta particolarmente marcato, è possibile effettuare una analisi del gender gap in funzione dell'abilità degli studenti, sempre facendo uso dei distractor plots. In particolare, è possibile produrre nel medesimo distractor plot due curve empiriche distinte, una relativa ai maschi e una relativa alle femmine, per ogni opzione di risposta e osservare il funzionamento dell'item in base al genere dei rispondenti.

Un secondo esempio delle potenzialità didattiche di questo tipo di analisi è il quesito D21 (Figura 19) che appartiene alla prova INVALSI del grado 10 somministrata nell'anno scolastico 2011/2012.

D21. **L'espressione $a^{37} + a^{38}$ è uguale a**

A. ☐ $2a^{75}$

B. ☐ a^{75}

C. ☐ $a^{37}(a+1)$

D. ☐ $a^{37 \cdot 38}$

Figura 19. Quesito 21, Prova INVALSI di matematica Grado 10, 2012 www.gestinv.it

A differenza del quesito analizzato nella sezione precedente, in questo caso ci troviamo davanti ad un esercizio proposto in un contesto tipicamente matematico, dall'aspetto squisitamente scolastico. Per rispondere correttamente al quesito, lo studente deve utilizzare la proprietà distributiva della moltiplicazione sull'addizione, individuando l'equivalenza:

$$a^{37} + a^{38} = a^{37}(1 + a)$$

e successivamente la proprietà commutativa dell'addizione. La risposta corretta è l'opzione C, e solo 34,8% del campione nazionale l'ha scelta (Figura 20).

Come mostrato in Figura 20, si tratta di un quesito che può essere considerato difficile (delta=0,76), l'andamento della risposta corretta empirica rispetto alla curva teorica risulta essere più che accettabile (weighted = 0,93) e il grado di discriminazione è molto buono (0,5).

```
item:33 (M21)
Cases for this item  41812   Discrimination  0.50
Item Threshold(s):     0.76   Weighted MNSQ   0.93
Item Delta(s):      0.76
------------------------------------------------------------------------------
 Label    Score     Count   % of tot  Pt Bis     t  (p)    PV1Avg:1 PV1 SD:1
------------------------------------------------------------------------------
   1       0.00      8125      19.43   -0.20   -42.60(.000) -0.39     0.80
   2       0.00     11019      26.35   -0.18   -38.08(.000) -0.29     0.84
   3       1.00     14460      34.58    0.50   118.46(.000)  0.64     1.00
   4       0.00      6800      16.26   -0.15   -30.87(.000) -0.32     0.84
   7       0.00       111       0.27   -0.03    -6.83(.000) -0.71     1.19
   9       0.00      1297       3.10   -0.12   -24.06(.000) -0.67     1.05
==============================================================================
```
Figura 20. Analisi IRT, Quesito 21, Prova INVALSI di matematica Grado 10, 2012 www.gestinv.it

Tra le opzioni di risposta non corrette, l'opzione di risposta B è la più scelta (26,3%) ma risultano attrattive anche le opzioni A e D, scelte rispettivamente dal 19,1% e dal 16,1% degli studenti. La percentuale di risposte mancanti risulta inferiore al 3%, quasi tutti gli studenti forniscono una risposta e, pertanto, si può supporre che abbiano un certo grado di confidenza nell'indicare una risposta.

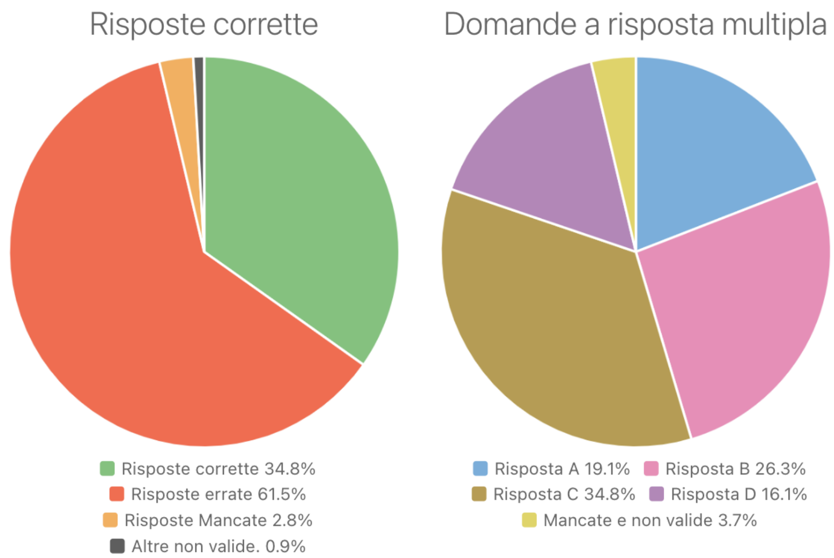


Figura 21. Percentuali nazionali, Quesito 21, Prova INVALSI di matematica Grado 10, 2012 www.gestinv.it

In termini di differenze di genere il quesito mostra delle particolarità già dall'analisi delle percentuali di scelta di ogni opzione di risposta (Tabella 3). Si osserva infatti una forte differenza tra maschi e femmine nella scelta della risposta corretta: il 38% dei maschi ha risposto correttamente a fronte del 31% delle femmine. L'opzione B è risultata la più attrattiva sia per i maschi che per le femmine, mentre per l'opzione A non si nota una differenza significativa. Gran parte del divario osservato nella scelta della risposta corretta è dovuto all'opzione D, scelta dal 14% dei maschi contro il 19% delle femmine.

Tabella 3. Percentuali di risposta, Quesito 21 Prova INVALSI di matematica Grado 10, 2012

	Percentuali di risposte	
	Femmine	Maschi
A	20%	19%
B	26%	27%
C	31%	38%
D	19%	14%
M	3%	3%

L'analisi del distractor plot con le curve empiriche rappresentate separatamente per maschi e femmine (Figura 22), può fornire maggiori informazioni in merito al divario relativo alle scelte delle diverse opzioni.

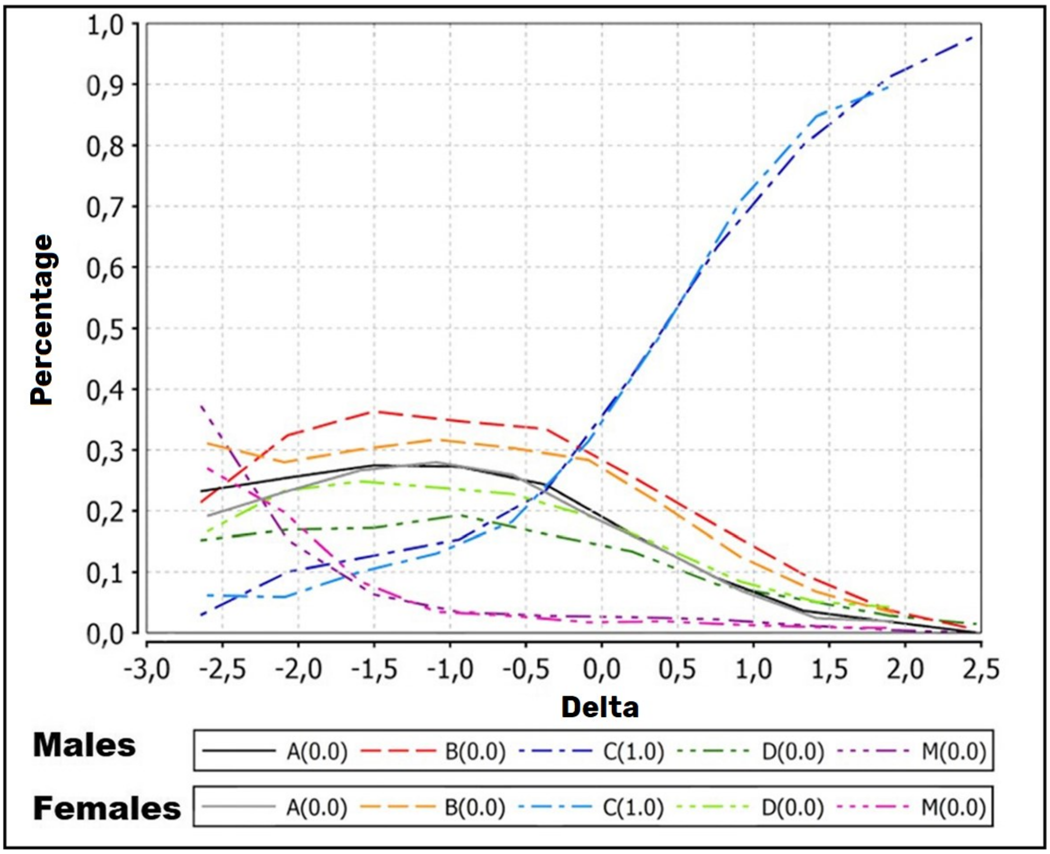


Figura 22. Distractor plot del Quesito 21, Prova INVALSI di matematica, Grado 10, 2012, percentuali di probabilità, curve empiriche rappresentate separatamente per maschi e femmine, nostra elaborazione dei dati INVALSI; elaborazione degli autori.

Nel distractor plot si può osservare che il divario di genere emerso in termini di probabilità di risposta corretta non risulta essere un divario costante su tutti i livelli di abilità. Se si confrontano infatti la curva blu (risposta corretta dei maschi) e la curva azzurra (risposta corretta delle femmine) si osserva una sovrapposizione delle due curve per livelli di abilità medi, un divario a favore dei maschi nei livelli di abilità bassi e un lieve divario a favore delle femmine per i livelli di abilità alti; il gap marcato osservato in termini di percentuali non è però tanto dovuto a questi lievi scostamenti tra le due curve quanto al fatto che i maschi raggiungono livelli di abilità superiori rispetto alle femmine: il decile più alto dei maschi raggiunge livelli medi di abilità appena inferiori a 2,5 e la probabilità di risposta corretta per tale decile è pari a quasi il 100%; il decile più alto delle femmine ottiene invece una abilità media inferiore a 2 e una probabilità di risposta corretta inferiore al 90%. Peraltro, le femmine più abili (nel complesso della prova) ottengono risultati migliori, su questa domanda, rispetto ai maschi di pari abilità. L'analisi delle curve che non c'è una differenza significativa nell'andamento dell'opzione A, mentre l'opzione B è preferita dai maschi specialmente per livelli di abilità medio bassi e l'opzione D è preferita dalle femmine, soprattutto per i livelli bassi e medi di abilità. Inoltre, per queste due opzioni possiamo osservare un andamento simile a quello descritto nella precedente sezione: si osserva infatti un andamento "a pancia" in particolare per quanto riguarda il distrattore D per le femmine e il distrattore B per i maschi.

Una volta individuato il divario di genere in relazione a uno specifico quesito e osservato le caratteristiche di tale divario in relazione alla scelta della risposta corretta, dei distrattori e dei relativi andamenti in funzione del livello di abilità, è interessante indagare quali possano essere i fattori che portano all'emergere di tale divario. Anche in questo caso un approccio *mixed method* si rivela il più adeguato. L'analisi qualitativa del quesito ha quindi previsto la somministrazione a 18 studenti di 6 diverse classi di Istituto tecnico e Liceo di una prova contenente, tra altri quesiti, il quesito D21 della prova INVALSI 2011, grado 10. Tra questi 18 studenti, solo 8 hanno fornito una risposta corretta, gli altri 10 hanno scelto le opzioni non corrette che risultavano più attrattive anche nella rilevazione nazionale: la B e la D.
Tutti gli studenti sono stati quindi intervistati individualmente ed è stato chiesto loro di spiegare le motivazioni alla base della risposta data al quesito; a coloro che avevano risposto correttamente sono stati successivamente mostrati i risultati a livello internazionale ed è stato chiesto di ipotizzare quali potessero, a loro avviso, essere le motivazioni alla base della bassa percentuale di risposte corrette rilevate dalla valutazione nazionale. I risultati dell'analisi qualitativa non possono ovviamente essere generalizzati e portare a considerazioni in termini di differenze di genere, ma hanno permesso di

individuare alcuni possibili fattori sottostanti la scelta del distrattore B e del distrattore D. Quest'ultimo era risultato essere particolarmente interessante nell'analisi quantitativa, e di fatto è stato anche il distrattore scelto da tutte le femmine intervistate che hanno risposto in modo errato.
Le due opzioni B (a^{75}) e D ($a^{37 \cdot 37}$) consistono in una potenza con la stessa base delle due di cui si richiede la somma e, come esponente, un'operazione tra i due esponenti. Le interviste hanno messo in luce come gli studenti che hanno scelto queste due opzioni cercano di risolvere il compito individuando un qualche tipo di regola (simile alle proprietà delle potenze studiate): questo processo è esplicitato chiaramente dalla studentessa ST_F_01 (queste sono le sigle utilizzate nella sperimentazione) che nell'intervista esprime la necessità di operare solo con gli esponenti e poi sceglie D, affermando "deve essere una proprietà!".
Si riconosce in questo comportamento un chiaro esempio di esigenza di giustificazione formale (D'Amore & Sandri, 1998). Infatti, questo comportamento è riconducibile agli atteggiamenti relativi al contratto didattico (Brousseau, 1988): secondo cui molti studenti (di fronte al testo di un problema) attivano una sorta di *lettura selettiva*, basata sull'identificazione dei dati numerici e su alcune regole che suggeriscono l'operazione giusta per "combinare" i numeri nel testo. Questi atteggiamenti sono spesso causati da norme implicite generate dal contratto didattico (D'Amore, Fandiño-Pinilla, Marazzani & Sarrazy, 2010), sono pertanto legate alle prassi scolastiche e non stupisce il fatto che gli studenti che incorrono maggiormente in questo tipo di errore siano gli studenti con livelli di abilità media, che sono influenzati da quanto proposto dall'insegnante in classe ma che non riescono a superare gli ostacoli legati a un'eccessiva aderenza alle pratiche didattiche.
La necessità di trovare norme e regole per risolvere il compito è espressa anche da altri studenti. Per esempio, la studentessa (ST_F_02) conosce le proprietà delle potenze e le cita affermando che "non si possono applicare le proprietà delle potenze perché non abbiamo la moltiplicazione; quindi non possiamo sommare le potenze, quindi non possiamo avere 75" (con riferimento all'opzione B), ma nonostante ciò sceglie poi l'opzione D.
Le interviste confermano dunque che gli studenti che affrontano questo compito non sono portati a ragionare sul significato dell'operazione, ad esempio lavorando sugli ordini di grandezza, ma si mettono immediatamente alla ricerca di una regola da applicare. Gli studenti spesso applicano regole e proprietà quando operano con le potenze, e poi presumono di doverlo fare sempre, anche quando capiscono che le proprietà non possono essere applicate in quel caso particolare.

Questo comportamento, riconducibile ad effetti del contratto didattico, ha fatto sì che gli studenti escludessero le opzioni A e C perché si aspettavano come risposta una potenza con la stessa base delle due sommate, come dichiarato esplicitamente dallo studente ST_M_01:

ST_M_01: cerco di fare lo stesso esponente... no beh la stessa base. Quindi questo no di sicuro questo no di sicuro [indicando la A e la C]

L'impatto delle pratiche didattiche nella non scelta della risposta corretta C è confermato anche dallo studente ST_M_07 che dichiara, riferendosi alla risposta corretta:

ST_M_07: Uno è abituato a vedere e a usare sempre le proprietà e quindi non la vede. Non viene in mente di raccogliere.

Gli studenti spesso applicano regole e proprietà quando operano con le potenze, e poi presumono di doverlo fare sempre, anche quando capiscono che le proprietà non possono essere applicate in quel caso particolare. Questo è il caso della studentessa ST_F_05, che risponde D e durante l'intervista afferma:

ST_F_05: Perché teoricamente ci sono le teorie tipo per le potenze solamente che non c'era un "per" in mezzo qua?.... o la B…
R: Perché la B come viene fuori?
ST_F_05: Ah, o sommando i due esponenti dei numeri. Solamente che se non mi ricordo male per sommare gli esponenti e dovrei avere un per qua in mezzo... Quindi forse è la D… o la B o la D…

Inoltre, notiamo che il fatto che la scelta dell'opzione D aumenti notevolmente l'ordine di grandezza non ostacola la scelta degli studenti. di grandezza non ostacola di fatto la scelta degli studenti. Le cause di questo comportamento possono essere ricondotte a una clausola del contratto didattico, quella della delega formale come definita da D'Amore (2008):

"The student reads the text, then decides on the operation to carry out and the numbers with which he/she has to work. At that point, indeed, the clause of the formal proxy is triggered. It is no longer for the student to reason and check; he/she no longer considers what follows as his/her personal responsibility […]. The involvement of the student is finished; now it's up to the algorithm or, better still, to the machine, to work for him. The student's next task will be that of transcribing the result, whatever it is, and it does not matter what it means within the problematic context he/she started with." (p. 16).

L'analisi qualitativa permette quindi di interpretare i motivi alla base della scelta del distrattore B e in particolare del distrattore D, in cui l'operazione tra gli esponenti è esplicitata, in termini di contratto didattico ed eccessiva aderenza alle pratiche didattiche.
Se si leggono pertanto le evidenze quantitative legate al gender gap alla luce di questa interpretazione possiamo affermare che l'importante divario tra le risposte corrette fornite da maschi e femmine e la maggiore attrattività del distrattore D per le femmine può essere in buona parte ricondotto a un maggiore impatto del contratto didattico per le studentesse con livelli di abilità medio bassi e medi. Questa interpretazione del divario di genere su questo specifico quesito risulta essere coerente con la letteratura nazionale e internazionale sulle differenze di genere, secondo cui sono particolarmente importanti nello spiegare tali differenze fattori di natura sociale, culturale ma anche micro-sociale, legati alle pratiche scolastiche, al contesto classe e a fattori di natura metacognitiva.
Altre evidenze relative alle differenze di genere in matematica, mostrate anche a livello locale, sono emerse all'interno del progetto GegaMATH, promosso dalla Libera Università di Bolzano. Nel corso del progetto è stata approfondita l'analisi a livello dei singoli quesiti in un contesto particolare come quello della provincia autonoma di Bolzano, in cui le differenze risultano essere piuttosto marcate e si è avviata una ricerca finalizzata anche ad analizzare le convinzioni dei docenti di matematica relativamente alle differenze di genere (Giberti, 2019).

5. L'effetto "Età della Terra"

5.1 Inquadramento del fenomeno

Il punto di partenza della ricerca[3] che ha condotto alla definizione dell'effetto è un fenomeno didattico che è stato osservato in due quesiti delle prove INVALSI di matematica, effettuate in Italia da tutti gli studenti di grado 10 (secondo anno della scuola secondaria di II grado). Il primo quesito, somministrato in modo censuario a circa 600.000 studenti italiani di grado 10 di tutti i percorsi scolastici (Licei, Istituti Tecnici e Istituti Professionali) è riportato in Figura 23.

D5. L'età della Terra è valutata intorno ai $4{,}5 \times 10^9$ anni. L'Homo Erectus è comparso circa 10^6 anni fa. Qual è la stima che più si avvicina all'età che la Terra aveva quando è comparso l'Homo Erectus?

- ☐ A. $4{,}5 \times 10^9$ anni
- ☐ B. $3{,}5 \times 10^9$ anni
- ☐ C. $4{,}5 \times 10^6$ anni
- ☐ D. $4{,}5 \times 10^3$ anni

Figura 23. Quesito 5, Prova INVALSI di matematica Grado 10, 2011, www.gestinv.it

Nel grafico seguente (Figura 24) sono riportati i risultati in riferimento a un campione statistico rappresentativo a livello nazionale e composto da 43.458 studenti (le caratteristiche tecniche del campione e dell'analisi dei dati della prova a cui appartiene il quesito si trovano nel Rapporto Tecnico (INVALSI, 2011).

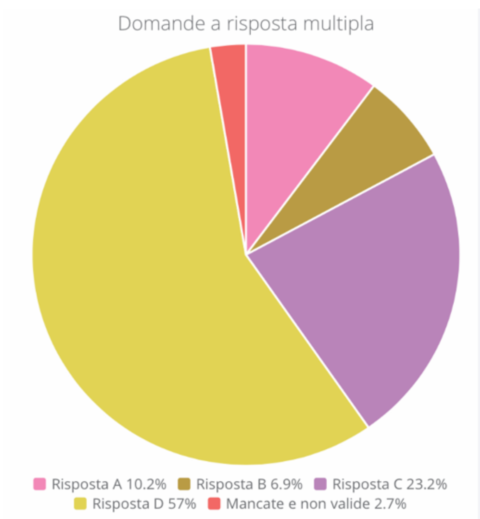


Figura 24. Percentuali nazionali, Quesito 5, Prova INVALSI di matematica Grado 10, 2011 www.gestinv.it

[3] La ricerca presentata è parte di un più ampio progetto, oggetto della tesi di dottorato di Federica Ferretti reperibile al seguente link: http://amsdottorato.unibo.it/7213/#

Come si vede dal grafico (Figura 24), le risposte corrette (opzione A) sono state il 10,2% e l'opzione errata più scelta è l'opzione D, in cui l'esponente della potenza è la differenza degli esponenti delle potenze nel testo. Come mostrato dall'analisi IRT (Figura 25), il quesito è risultato buono dal punto di vista misuratorio (Discrimination Index = 0,32, Weighted MSNQ = 0,97) e all'interno della prova stessa e presenta un indice di difficoltà molto elevato (Delta(s) = 2,56) (per i dettagli dell'analisi statistica si veda il Rapporto Tecnico, INVALSI, 2011).

Item:7 (MD5)
Cases for this item 43458 Discrimination 0.32
Item Threshold(s): 2.55 Weighted MNSQ 0.97
Item Delta(s): 2.56

Label	Score	Count	% of tot	Pt Bis	t (p)	PV1Avg:1	PV1 SD:1
1	1.00	4438	10.21	0.32	70.12(.000)	0.90	1.20
2	0.00	2992	6.88	-0.11	-24.05(.000)	-0.39	0.92
3	0.00	10084	23.20	-0.15	-31.01(.000)	-0.25	0.89
4	0.00	24831	57.14	0.02	4.97(.000)	0.02	0.88
7	0.00	41	0.09	-0.03	-5.55(.000)	-0.90	1.51
9	0.00	1072	2.47	-0.11	-22.17(.000)	-0.67	1.02

Figura 25. Analisi IRT, Quesito 5, Prova INVALSI di matematica Grado 10, 2011 www.gestinv.it

Come si può notare dal distractor plot (Figura 26), mentre il distrattore C è molto scelto dagli studenti situati nella parte bassa della scala di abilità, il distrattore D, in cui l'esponente della potenza è ottenuto per sottrazione degli esponenti delle potenze presenti nella consegna, è preferito a tutti i livelli di abilità. In particolare, la scelta di tale distrattore è massima per gli allievi situati nella parte medio-alta della scala e l'andamento è "a pancia" (analogo a quelli descritti nel Capitolo 4). La linea che rappresenta la risposta corretta (distrattore A), è bene interpolata dalla curva continua prevista dal modello e questo è indice di buona misura dal punto di vista psicometrico.

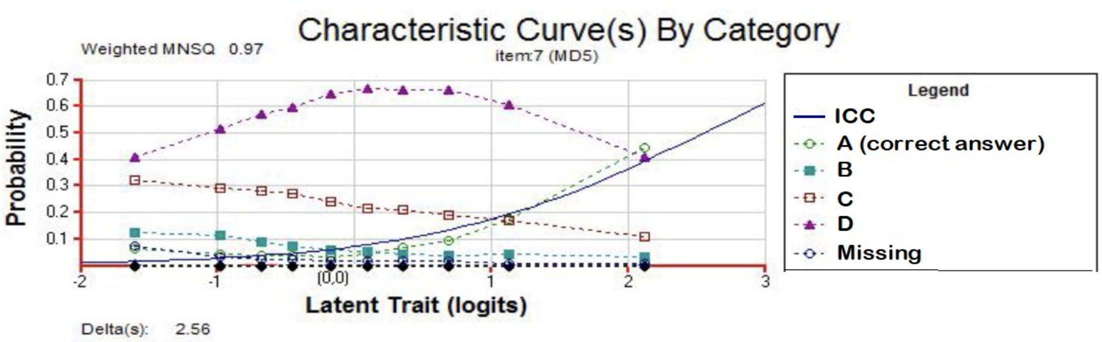
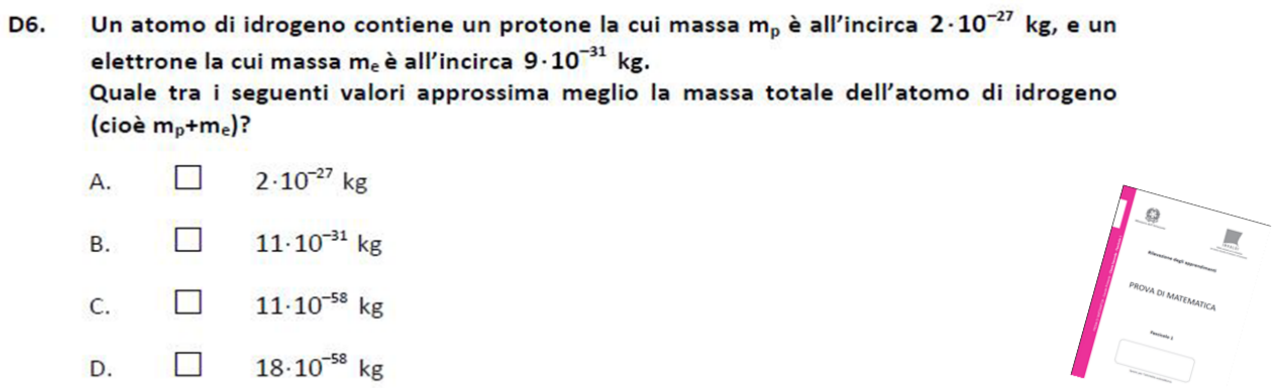


Figura 26. Curva caratteristica, Quesito 5, Prova INVALSI di matematica Grado 10, 2011, rielaborato da www.gestinv.it

Una situazione equivalente, dal punto di vista matematico, si è presentata nella prova INVALSI somministrata nell'anno scolastico 2012/2013, sempre al grado 10. Il quesito, che abbiamo denominato "Massa del Protone" (Figura 27) è stato somministrato a circa 560.000 studenti delle classi seconde del secondo ciclo di ogni percorso scolastico e il campione era costituito da 38.533 studenti.



D6. **Un atomo di idrogeno contiene un protone la cui massa m_p è all'incirca $2 \cdot 10^{-27}$ kg, e un elettrone la cui massa m_e è all'incirca $9 \cdot 10^{-31}$ kg.**
Quale tra i seguenti valori approssima meglio la massa totale dell'atomo di idrogeno (cioè m_p+m_e)?

A. ☐ $2 \cdot 10^{-27}$ kg

B. ☐ $11 \cdot 10^{-31}$ kg

C. ☐ $11 \cdot 10^{-58}$ kg

D. ☐ $18 \cdot 10^{-58}$ kg

Figura 27. Quesito 6, Prova INVALSI di matematica Grado 10, 2013, www.gestinv.it

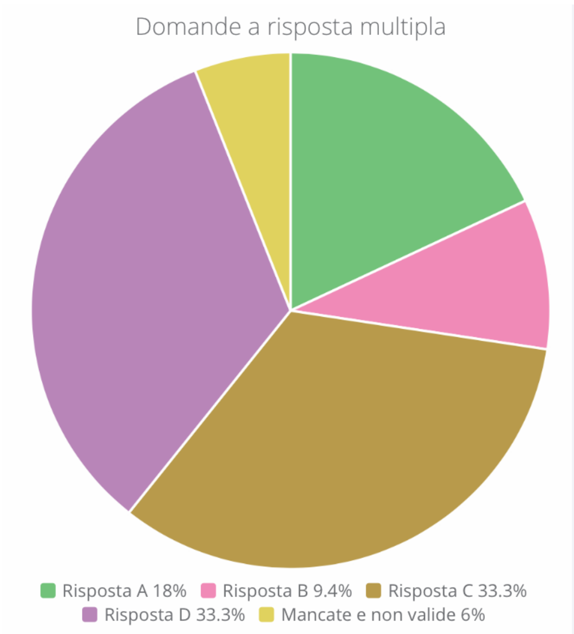


Figura 28. Percentuali nazionali, Quesito 6, Prova INVALSI di matematica Grado 10, 2013, www.gestinv.it

Come mostrano i risultati (Figura 8), le risposte corrette sono state il 18%. Analizzando le percentuali di scelta delle opzioni sbagliate, si può vedere come le opzioni C e D, in cui l'esponente della potenza è ottenuto per somma degli esponenti delle potenze presenti nel testo, sono i più scelti a livello nazionale (più del 30% di scelte ciascuno).

Come si può vedere dal grafico della curva caratteristica (Figura 9), i distrattori C (linea Item 7.3) e D (linea Item 7.4) in cui l'esponente della potenza è ottenuto per somma degli esponenti delle potenze presenti nel testo, oltre a essere i più scelti a livello nazionale, sono i preferiti dagli studenti a quasi tutti i livelli di abilità.

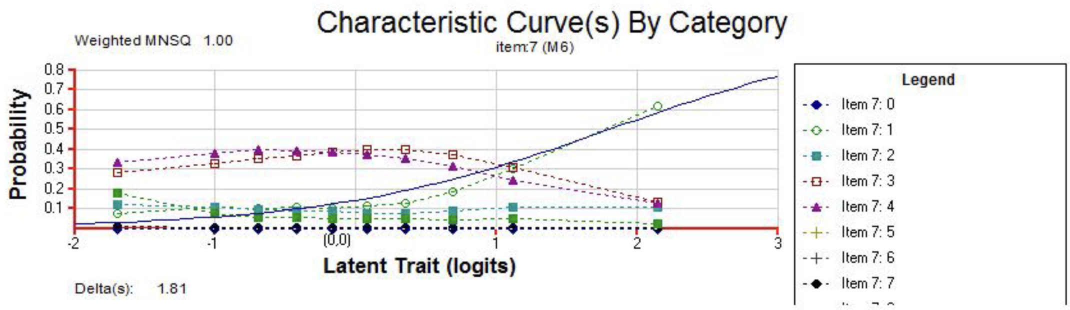


Figura 29. Curva caratteristica, Quesito 6, Prova INVALSI di matematica Grado 10, 2013, www.gestinv.it

Le prime interpretazioni (cfr. ad es. Impedovo, Orlandoni & Paola, 2011) del fenomeno evidenziato dal quesito "Età della Terra" hanno collegato il comportamento degli allievi genericamente a effetti di contratto didattico nel senso di Brousseau (Brousseau, 1997) e alla mancanza di controllo critico sui contenuti. In particolare, il fatto che la maggior parte degli allievi nel quesito "Età della Terra", il cui testo richiama intuitivamente una situazione sottrattiva, scelgano l'opzione D in cui l'esponente della potenza è la differenza degli esponenti delle potenze presenti nella consegna potrebbe apparire come un evidente esempio di presenza della clausola del contratto didattico denominata *delega formale* (D'Amore, 2007); così come nel quesito "Massa del protone" (situazione additiva) la maggior parte degli studenti scelgono le opzioni C e D (in cui compare la somma degli esponenti delle potenze).

Osserviamo però un altro fatto, manifestato in entrambi i casi: la risposta corretta è uno dei dati esplicitamente presenti nel testo del quesito.

In due quesiti somministrati agli studenti del grado 8 (classe terza della scuola secondaria di primo grado) si trovano situazioni similari (quesito E21, Grado 8 a.s. 2011/12 e quesito D23, Grado 8 a.s. 2009/10).

Comune a questi quattro esempi, differenti per molti aspetti di formulazione, contesto, contenuto, abilità richiesta, è la difficoltà che si manifesta quando la risposta coincide con uno dei dati di partenza, difficoltà che coinvolge inaspettatamente anche allievi di livello medio e alto di abilità in riferimento alle prove.

Da queste osservazioni, è nata la ricerca che ha portato alla definizione di un nuovo effetto di contratto didattico denominato effetto "Età della Terra".

Il macro-fenomeno osservato si presta ad essere descritto da differenti punti di vista, ognuno dei quali implica l'uso di una adeguata lente teorica. Di fatto, si è in presenza di un comportamento degli allievi (rilevato dall'analisi quantitativa nazionale e, come vedremo, confermato dall'indagine quantitativa della sperimentazione effettuata) che va interpretato anche in funzione del contesto in cui si manifesta. La ricerca parte dall'osservazione dell'esistenza di un *mismatch* tra il sapere atteso e il sapere appreso dagli studenti, l'uno misurato dalle valutazioni standardizzate nazionali e l'altro indicato dai curricoli sui quali sono costruite le valutazioni stesse. A partire da queste evidenze empiriche si è cercato di misurare l'effettiva distanza di questi due "saperi" e, successivamente, di indagarne le cause.
L'ipotesi fondamentale è che si è di fronte a un comportamento dell'allievo che non è spiegabile se non presupponendo, a monte del problema e in qualche modo indipendenti da esso, che esistano dei "principi regolativi" specifici che condizionano la sua azione, accettati esplicitamente o implicitamente, dinamicamente negoziati o profondamente interiorizzati e stabilizzati. Per essere più espliciti, con principio regolativo si intende un criterio organizzativo o normativo dell'azione dell'allievo che si innesta di fronte a una consegna di matematica. Le situazioni d'aula sono permeate da "negoziazioni" spesso implicite che influiscono in modo determinante sui contenuti e sui meta-contenuti in gioco. Questi criteri discendono dalle interazioni di classe tra insegnante, discente e sapere in gioco, e sono spesso decisi dal contratto didattico nel senso di Brousseau (Brousseau, 1986; EMS-EC, 2012). Queste norme possono riguardare più strettamente meta-contenuti, generando convinzioni o regole che convalidano in generale il discorso matematico determinando così l'instaurarsi di *norme socio-matematiche* (Yackel & Cobb, 1996; EMS-EC, 2013) o riguardare più direttamente i contenuti coinvolti, come nel caso di molte *misconcezioni* (vedi ad esempio Sbaragli, 2005; Stavy & Tirosh, 1996). Queste forme di organizzazione dei modelli sono dei principi regolativi (Garuti & Boero, 1994) che possono essere influenzati anche dalla prassi matematica della classe (è il caso del costume didattico di Balacheff, 1988).
Questa necessariamente breve rassegna di contributi mostra come negli ultimi decenni siano stati proposti diversi costrutti teorici che descrivono, di volta in volta, le diverse forme di questi principi regolativi. Il quadro che è stato proposto nella tesi di dottorato in cui si è studiato il fenomeno in oggetto (Ferretti, 2015) ha cercato di esplicitarne le relazioni in modo tale da poter arrivare alle possibili interpretazioni di un fenomeno che si può presentare e di fatto si presenta, come mostrano i risultati della sperimentazione, in contesti e situazioni molto diverse, dove un singolo costrutto non risulta esaustivo. L'interazione o quanto meno il confronto tra teorie e costrutti diversi non è sempre agevole (come argomenta Radford, 2005), ma in questo caso è indispensabile vista l'evidente varietà di manifestazioni di uno stesso fenomeno, quasi avatar diversi di uno stesso fatto.

In particolare, all'interno della ricerca sono stati indagati, e definiti, l'effetto e il principio regolativo che lo determina.

In dettaglio, l'effetto che si osserva può essere descritto in questo modo:

In una situazione didattica, di fronte a una consegna gli allievi tendono a non accettare una risposta che non sia identificabile chiaramente con un risultato distinto dai dati di partenza.

Ipotesi di ricerca è che questo effetto dipenda da un principio regolativo, ovviamente agente in maniera non consapevole (che a una prima analisi sembra discendere da pratiche d'aula) del tipo:

Il risultato di un problema o di una operazione non può essere lo stesso dato da cui si è partiti.

5.2 Breve inquadramento teorico

L'interpretazione "ortodossa" del costrutto di contratto didattico (Sarrazy, 1995) concentra l'analisi sulla relazione allievo-docente, e a partire da questa relazione evidenzia il ruolo della noosfera. Il contratto è in questo caso l'insieme delle convinzioni, delle pratiche messe in campo, dei comportamenti praticati da uno degli attori della situazione didattica (in questo caso l'allievo) in risposta alle aspettative reciproche che si instaurano nella relazione insegnante-allievo-sapere. Come messo in luce in Ferretti (2015), interpretazioni più ampie del costrutto lo estendono anche ai comportamenti generati dalle aspettative reciproche tra allievo e società. L'effetto descritto in questa ricerca può evidentemente originarsi in una situazione di interazione tra allievo e docente, soprattutto con allievi piccoli, abituati al fatto che l'insegnante propone di eseguire una operazione e *si aspetta un risultato*, quasi sempre diverso dai dati di partenza. Il principio regolativo che si è enunciato può facilmente iniziare a formarsi in situazioni di contratto didattico. Utilizzando come schema organizzativo il triangolo della didattica della matematica (Chevallard, 1992) si può asserire che quando viene proposta una consegna matematica all'allievo, questa viene vista da quest'ultimo posizionata in qualche modo lungo il lato insegnante-sapere. Se il processo per la soluzione della consegna (con tutti i suoi aspetti prasseologici[4], cfr. Chevallard, 1999) si svolge lungo il lato insegnante-allievo, i principi regolativi sono descritti efficacemente dalle categorie del contratto didattico. In questo caso, i principi regolativi sono negoziati, e la loro natura è essenzialmente dinamica (Brousseau, 1997; Sarrazy, 1995). Questa identificazione tra insegnante e sapere è particolarmente stretta per gli allievi piccoli, e di conseguenza per essi è naturale che i principi regolativi siano di tipo prevalentemente contrattuale e legati alla figura dell'insegnante. Usando l'immagine del triangolo, per essi l'angolo avente per vertice l'allievo è ancora molto stretto. D'altra

[4] In italiano, il termine francese *praxéologie* viene tradotto dai diversi autori con "prasseologia", "prassilogia" o "praxeologia". Noi abbiamo optato per la forma scelta da Nicola Abbagnano quando ha introdotto il termine nel dibattito culturale italiano.

parte, ben presto l'istituzione entro cui si esplica la dinamica didattica acquisisce importanza. Diversi studi hanno evocato la presenza di situazioni di contratto in contesto di istituzione; per fare solo due esempi, Elia, Van Den Heuvel-Panhuizen e Kolovou, 2009 e De Vleeschouwer e Gueudet, 2011. Si può affermare che anche l'*effetto Età del capitano* sia rilevabile al di fuori della relazione diretta docente-allievo, e arrivi a manifestarsi indipendentemente da essa (pur essendo, ovviamente, da essa originato).

Si assiste quindi a una sorta di accettazione e condivisione di questi principi regolativi, che all'origine sono continuamente negoziati ma poi diventano via via sempre più stabili. Per spiegare questi meccanismi a livello della prima istituzione, un microsistema in cui la presenza dell'insegnante è ancora esplicita e forte, vale a dire la classe, Balacheff (1988) ha proposto l'idea di *costume didattico*. Quando il fenomeno osservato si manifesta a livello di comportamento della classe il costrutto adatto sembra essere quindi piuttosto quello del costume didattico, che appare quasi come un primo sedimento delle situazioni di contratto vissute dall'allievo nel contesto della classe. Dagli insiemi di pratiche didattiche vissute in aula con l'insegnante resta un'impronta che viene in qualche modo interiorizzata come principio regolativo, una sorta di meta-oggetto.

Peraltro, il macro-fenomeno osservato si pone a un livello molto più ampio che non quello della classe, è stato rilevato in valutazioni standardizzate di dimensione nazionale, su campioni di diverse decine di migliaia di studenti. Siamo quindi in una situazione più ampia, che non può esser vista come una somma di tante situazioni-classe. Siamo di fronte all'emergere di un comportamento che deve rispondere a principi regolativi dell'azione degli studenti in cui l'insegnante (o l'insieme degli insegnanti della storia personale dello studente) non ha più un ruolo diretto. Elementi di questa storia si sono sedimentati nello studente, generando convinzioni e atteggiamenti, e si sono cristallizzati in principi regolativi che riguardano la matematica (scolastica) in quanto tale. Per tornare all'immagine del triangolo citata precedentemente, l'azione dell'allievo per rispondere alla domanda si sviluppa in questo caso lungo il lato allievo-sapere. Questo switching dell'effetto da un lato all'altro del triangolo passa necessariamente attraverso un processo di accettazione di norme e regole di comportamento, di cui il costume di Balacheff è un primo passo.

Anche Garuti e Boero (1994), studiando situazioni di modellizzazione matematica, evocano l'azione di "principi", intesi come criteri organizzativi delle operazioni di formulazione di una ipotesi, che esistono prima dell'analisi dello specifico problema esaminato

Un contributo alla comprensione di questo nuovo avatar del fenomeno può venire da teorie dei sistemi sociali, seguendo un approccio analogo a quello di Sierpinska, Bobos e Knipping (2008) e di Sensevy (2010), che utilizzano la strumentazione (in termini di norms e rules) sviluppata, tra gli altri, anche da Ostrom (2005) (per approfondimenti si rimanda a Ferretti, 2015). Appare infatti necessario

considerare in maniera più ampia le caratteristiche della noosfera e in particolare il ruolo delle regole e delle norme nell'istituzione in cui avviene l'interazione tra allievo, docente e matematica (anche in termini di norme socio-matematiche di Yackel & Cobb, 1996).
In questa ricerca, il principio regolativo enunciato appare effettivamente come un criterio organizzativo dell'azione di risoluzione del problema, esistente prima dell'analisi del problema stesso. L'ipotesi è quindi che questo principio regolativo del comportamento dell'allievo alle prese con la matematica sia alla base dell'ultimo avatar del fenomeno, manifestatosi in contesti sempre più allargati (prima il rapporto diretto studente-insegnante, poi l'ambiente di classe, infine il sistema scolastico nel suo insieme). Si presenta quindi come un qualcosa caratteristico del sistema, e non solo come la somma di tutte le specifiche situazioni classe. Obiettivo della ricerca è quindi indagare l'esistenza e le manifestazioni del fenomeno. In particolare, in primo luogo si è indagato se il fenomeno si manifesta in ambiti matematici diversi, se interessa tutti i livelli scolastici, se è presente in studenti con abilità matematiche diverse, se dipende dal controllo semantico che l'allievo ha sui contenuti della domanda e dalle diverse abilità matematiche e, infine, se è possibile quantificarlo. In secondo luogo, si è indagato se il comportamento degli allievi in cui si osserva l'effetto dipende da un principio regolativo vicino a quello enunciato.
Per rispondere alle domande di ricerca è stata condotta una sperimentazione, costituita da fasi quantitative e qualitative (Johnson & Onwuegbuzie, 2004), che ha coinvolto studenti del primo e del secondo ciclo d'istruzione[5]. La sperimentazione si è svolta in due fasi centrali: si sono prima somministrati dei questionari e successivamente, sono stati intervistati alcuni degli studenti che hanno risposto ai questionari, chiedendo l'esplicitazione della strategia messa in atto mediante brevi interviste mirate condotte a piccoli gruppi. Nella seconda fase sono stati intervistati studenti frequentanti la classe seconda della scuola secondaria di secondo grado che non avevano partecipato alla prima fase della sperimentazione per rilevare altri dati che permettessero di approfondire ulteriormente l'analisi del fenomeno.

5.3 La sperimentazione

La ricerca quantitativa è stata condotta in tre Istituti Comprensivi, Licei Scientifici e Istituti di Istruzione Superiore distribuiti nella regione Emilia-Romagna. In totale sono stati somministrati 124

[5] Si coglie l'occasione per ringraziare le scuole e i docenti delle classi che hanno partecipato alla ricerca: *Istituto Comprensivo "R. Gasparini" di Novi di Modena (MO), Istituto Comprensivo "C. Bassi" di Castel Bolognese (RA), Istituto Comprensivo "IC4" di Imola (BO), Liceo Statale "B. Rambaldi – L. Valeriani -Alessandro Da Imola" di Imola (BO), Istituto di Istruzione Superiore "E. Majorana" di San Lazzaro di Savena (BO), Istituto di Istruzione Superiore "E. Mattei" di San Lazzaro di Savena (BO), Istituto Statale di Istruzione Superiore "I.S.I.S. Archimede" di San Giovanni in Persiceto (BO), Liceo Scientifico Sportivo "Toniolo" di Bolzano, Liceo Scientifico Statale "A. Righi" di Cesena (FC).*

questionari che hanno svolto la funzione di pre-test e 612 questionari definitivi sui quali si è basata l'analisi quantitativa della ricerca. Successivamente sono state effettuate interviste a piccoli gruppi a 12 studenti del grado 10 tra gli studenti delle scuole secondarie superiori che hanno partecipato alla rilevazione. Per un migliore ancoraggio dei livelli di apprendimento in matematica di ognuna delle classi selezionate, sono stati confrontati i risultati ottenuti dalle classi nelle valutazioni standardizzate dell'INVALSI, sia rispetto alle medie nazionali e macro-regionali che rispetto alle 200 classi italiane più vicine, come indicatori socioeconomici di riferimento (metodologia *AnchorSA*, Ferretti, 2015). Nei questionari sono state proposte non solo domande indicative per la ricerca ma anche domande "schermo". Per poter disporre di termini di confronto per i risultati ottenuti sono stati proposti come domande schermo quesiti (in alcuni casi, leggermente modificati) presenti nei test INVALSI di anni precedenti di cui si posseggono risultati su scala nazionale; questo ha permesso di ancorare ulteriormente il livello degli studenti che hanno fatto parte della sperimentazione in riferimento all'intera popolazione italiana. Sono stati quindi somministrato i questionari in gradi scolastici di cui si hanno i risultati campionari nazionali, gradi 5, 6, 8, 10 (e all'inizio del grado 11). In questo modo si è effettuata la sperimentazione su una popolazione distribuita in tutti i segmenti della scuola italiana, frequentante classi per le quali erano disponibili i risultati delle valutazioni nazionali. Questo ha permesso di indagare l'ipotesi di ricerca per la quale l'effetto si manifesta indipendentemente dal grado scolastico. In particolare, si è focalizzata l'attenzione sui livelli di "passaggio" tra la scuola primaria e la scuola secondaria di primo grado e tra la scuola secondaria di primo grado e la scuola di secondaria di secondo grado in modo tale da cercare conferme dell'ipotesi per la quale l'effetto persiste indipendentemente dalla noosfera. Per cercare di validare l'altra ipotesi di ricerca, per la quale l'effetto sia presente indipendentemente dal contenuto matematico, nei vari questionari sono presenti quesiti che spaziano ambiti diversi della matematica, dall'algebra alla geometria alla probabilità. Sono stati predisposti tre questionari pre-test per i 5 gradi scolastici indagati; i risultati ottenuti nei pre-test e le difficoltà riscontrate dagli studenti (rilevate anche attraverso colloqui ex-post con insegnanti e studenti) hanno permesso di arrivare alla stesura dei questionari definitivi sui quali è stata basata la sperimentazione. Dopo la correzione dei questionari e la codifica dei risultati sono stati individuati 12 studenti che avevano svolto i questionari dei gradi 10 e 11 con i quali sono state condotte delle interviste semi-strutturate, task-based, a piccoli gruppi. Successivamente sono state effettuate interviste individuali, incentrate sulle domande delle valutazioni standardizzate da cui è nata la ricerca, a studenti del grado 10 che non hanno partecipato alla sperimentazione. Per quanto riguarda le interviste, come opzione metodologica è stata scelta l'intervista strutturata task-based (Goldin, 2000), conducendola sia in piccoli gruppi di studenti che individualmente.

Per la descrizione dei questionari, della struttura delle interviste e di tutti i risultati, si rimanda a Ferretti (2015). Per brevità espositiva si riportano di seguito solo alcuni dei risultati ottenuti.

5.4 Alcuni dei principali risultati

5.4.1 Il quesito "zero virgola"

Tutti i questionari si aprono con la stessa domanda filtro, che valuta la capacità di eseguire moltiplicazioni fra numeri razionali espressi in forma decimale, con la presenza di virgole e di zeri finali.

Esegui le seguenti moltiplicazioni:

2,5 x 32 = ...

1,9 x 4,1 = ...

2,5 x 320 = ...

19 x 0,41 = ...

25 x 0,32 = ...

1,9 x 41 = ...

2,5 x 3,2 = ...

1,9 x 410 = ...

Figura 30. Quesito "schermo" presente nei questionari di tutti i gradi

Questo quesito aveva lo scopo di filtrare gli studenti rispetto alla loro abilità nell'eseguire le moltiplicazioni con questa tipologia di numeri e controllare le interferenze, dovute a scarsa abilità nel calcolo, con l'effetto che si voleva osservare. In particolare, due degli item richiedono di gestire nella stessa operazione la presenza simultanea di un numero con la virgola e un numero terminante con 0, e quindi una delle funzioni di questo filtro è stata quella di eliminare interferenze nei risultati, dovute a scarsa padronanza del meccanismo di moltiplicazione con 0 finali.
In seguito, verranno indicati con "tipologia A" gli allievi che hanno risposto correttamente ad almeno 7 di questi 8 item, e con "tipologia A+" gli allievi che hanno risposto correttamente a tutti e 8 gli item. Il numero e la percentuale di studenti di ciascuna tipologia, nei diversi gradi, sono riportati nella seguente tabella (Tabella 4).

Tabella 4. Percentuali di risposte corrette al quesito “zero-virgola” degli studenti delle tipologie A e A+

	Gr 5	Gr 6	Gr 8	Gr 10	Gr 11
tip. A+	34 (28%)	42 (34%)	34 (30%)	42 (44%)	54 (44%)
tip. A	61 (50%)	66 (53%)	56 (50%)	57 (60%)	73 (60%)

La domanda filtro ha permesso quindi di mettere a fuoco, all″interno della popolazione, il comportamento degli strati di allievi con maggiore padronanza delle procedure di calcolo con zeri finali e virgole. La tipologia A raccoglie oltre la metà degli studenti, in tutti i livelli testati, mentre la tipologia A+ raccoglie all″incirca un terzo degli studenti della popolazione, in misura leggermente diversa a seconda dei livelli testati. Tutti i questionari contengono poi un quesito in cui questa situazione permette di osservare l’eventuale presenza dell’effetto “Età della Terra”.

Il quesito è presente nel questionario proposto ai gradi 5 e 6 con 4 item (di cui uno presentante l’effetto “Età della Terra” oggetto di studio, l’item c), e nei questionari somministrati ai gradi 8, 10 e 11 con 5 item (di cui due presentanti l’effetto “Età della Terra”, l’item *c* e, in aggiunta rispetto ai questionari dei gradi inferiori, l’item *e*).

Sapendo che 34 x 33 = 1122, determina il risultato delle seguenti operazioni (possibilmente senza fare i calcoli):

a) 34 x 3,3 = ...

b) 34 x 0,33 = ...

c) 3,4 x 330 = ...

d) 3,4 x 33 = ...

e) 0,34 x 3300 = ...

Figura 31. Quesito “zero-virgola” presente nei questionari dei gradi 8, 10 e 11

La seguente tabella (Tabella 5) e il grafico sottostante (Figura 12) riportano le percentuali di risposte corrette in ciascuno degli item della domanda sotto osservazione, distinte per i vari gradi testati.

Tabella 5. Percentuali di risposte corrette al quesito “zero-virgola” della popolazione della sperimentazione

	item a	item b	item c	item d	item e
Gr 05 - 06	87,04%	81,38%	48,99%	74,09%	
Gr 08	88,50%	75,22%	48,67%	76,99%	44,25%
Gr 10-11	88,07%	84,86%	73,85%	83,49%	72,48%

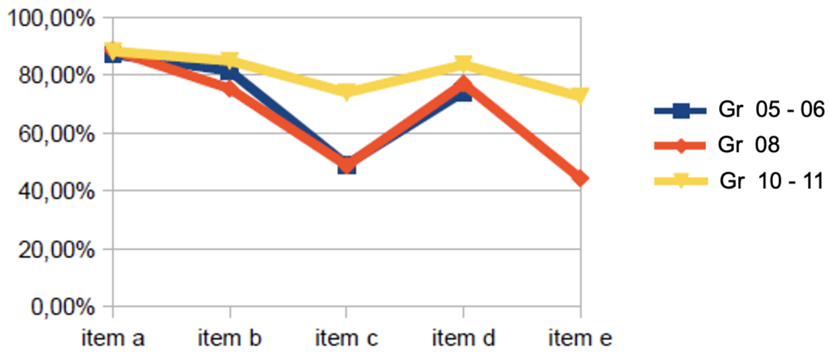


Figura 32. Rappresentazione grafica delle risposte corrette al quesito "zero-virgola" della popolazione della sperimentazione

È interessante notare che non c'è sostanziale differenza di comportamento tra gli allievi dei gradi 5 e 6 e quelli del grado 8 (nonostante intercorrano tra di essi due o tre anni di scolarità).

In tutti i livelli considerati la percentuale di risposte corrette negli item sotto osservazione è sensibilmente inferiore rispetto alle percentuali di risposte corrette negli altri item.

Gli studenti di grado 10 e 11 mostrano una flessione di risposte corrette più contenuta, negli item *c* ed *e*, ma pur sempre ben rilevabile. Si deve comunque tener conto che la somministrazione, nei gradi 10 e 11, è avvenuta in scuole di indirizzo scientifico, frequentate da studenti che, nei livelli precedenti, ottengono generalmente buoni risultati in matematica.

Si può notare a margine che i dati raccolti permettono anche di mostrare che non c'è sostanziale differenza tra i risultati degli studenti di grado 5 e quelli di grado 6, e neppure tra quelli di grado 10 e quelli di grado 11, nonostante queste coppie di gradi siano caratterizzate, nel sistema scolastico italiano, dal cambiamento dell'insegnante di matematica. In altre parole, gli studenti di grado 5 e quelli di grado 10 sono al termine un percorso con un insegnante, mentre quelli di grado 6 e di grado 11 sono all"inizio di un percorso con un nuovo insegnante; questo è un fattore importante per le interpretazioni in un'ottica interazionista.

Vengono di seguito riportate (Tabella 6) le percentuali di risposta agli item considerati degli studenti che hanno risposto correttamente ad almeno 7 degli 8 item della domanda filtro sopracitata (studenti di tipologia A).

Tabella 6. Percentuali di risposte corrette al quesito "zero-virgola" degli studenti di tipologia A

	item a	item b	item c	item d	item e
Gr 05 - 06	94,40%	90,40%	60,00%	85,60%	
Gr 08	96,43%	89,29%	62,50%	94,64%	66,07%
Gr 10-11	93,85%	93,08%	82,31%	93,08%	81,54%

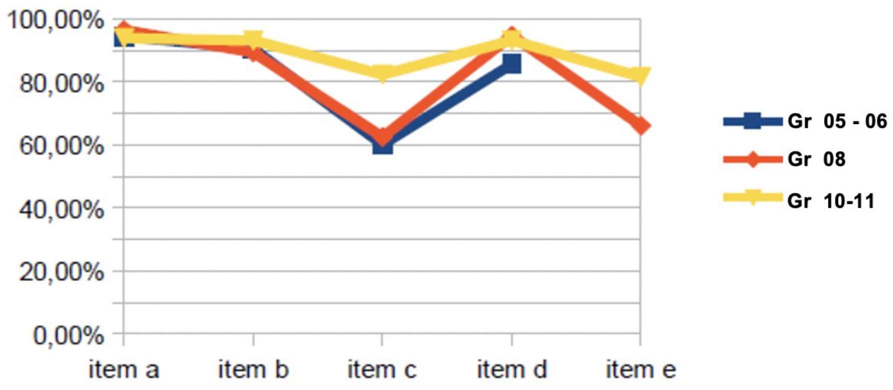


Figura 33. Rappresentazione grafica delle risposte corrette al quesito “zero-virgola” degli studenti di tipologia A

Anche in questo caso l’item *c* e l’item *e* (per i gradi 8, 10 e 11) riportano una percentuale di risposte corrette sensibilmente inferiore alle altre in tutti i livelli scolastici. Andando ad esaminare la distribuzione delle risposte sbagliate ai diversi item fornite da questi allievi di tipologia A, si rileva che oltre il 30% degli allievi di tipologia A sbaglia almeno uno degli item *c* ed *e* e il 22% sbaglia solo uno di questi due. In altre parole, tra gli allievi di Tipologia A che sbagliano almeno un item di questa domanda, oltre il 90% sbaglia proprio uno degli item *c* o *e* o entrambi. Viceversa, è irrilevante la percentuale di allievi che sbaglia sugli altri item e risponde correttamente agli item *c* ed *e*.
Per quanto riguarda gli allievi di tipologia A+ (gli allievi che hanno risposto correttamente a tutti gli 8 item della domanda filtro), le seguenti tabelle e i grafici (Tabella 7 e Figura 34) riportano, distinte per i diversi gradi, le percentuali di risposte corrette fornite agli item del quesito in osservazione.

Tabella 7. Percentuali di risposte corrette al quesito “zero-virgola” degli studenti di tipologia A

	item a	item b	item c	item d	item e
Gr 05 - 06	92,31%	88,81%	57,34%	86,01%	
Gr 08	95,38%	83,08%	60,00%	87,69%	63,08%
Gr 10-11	91,61%	90,32%	81,29%	88,39%	78,06%

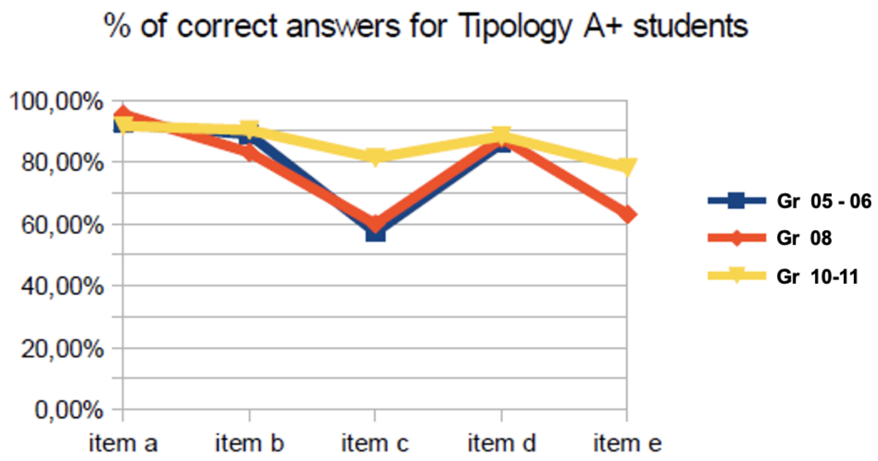


Figura 34. Rappresentazione grafica delle risposte corrette al quesito “zero-virgola” degli studenti di tipologia A+

Anche per gli studenti di tipologia A+ l’item *c* e l’item *e* (per i gradi 8 e 10-11) riportano una percentuale di risposte corrette sensibilmente inferiore alle altre in tutti i gradi scolastici. Andando ad esaminare la distribuzione delle risposte sbagliate ai diversi item fornite da questi studenti di tipologia A+, si rileva che oltre il 30% sbaglia almeno uno degli item *c* ed *e* e quasi il 20% sbaglia solo uno o due di questi item. Viceversa, è irrilevante la percentuale di studenti che sbaglia altri item e risponde correttamente agli item *c* ed *e*.

I seguenti grafici (Figura 35) mostrano comparativamente, per ogni grado, il confronto tra la percentuale di risposte corrette di tutta la popolazione e degli studenti di tipologia A+.

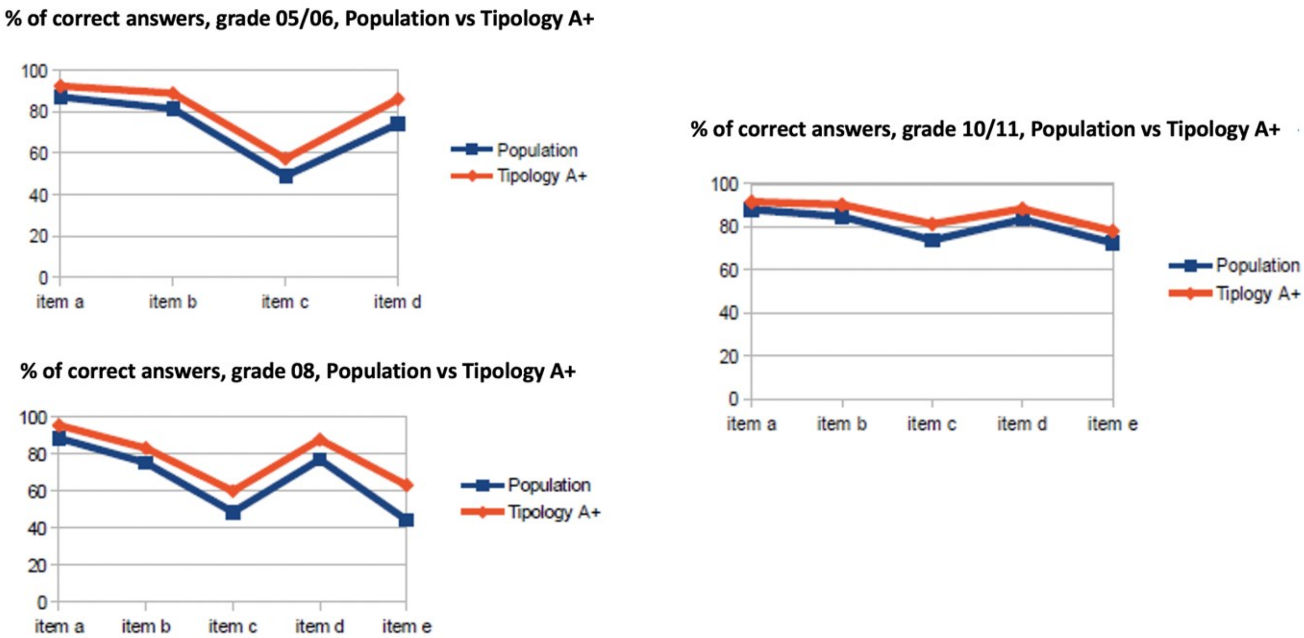
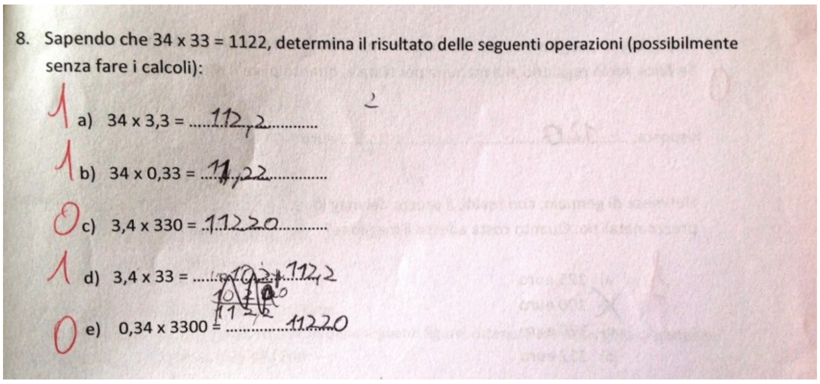


Figura 35. Confronto fra le percentuali di risposte corrette dell’intera popolazione e degli studenti di tipologia A+

Analizzando i protocolli degli studenti si hanno conferme delle analisi quantitative appena descritte. Ad esempio, le due immagini seguenti riportano gli atteggiamenti tipici degli studenti di fronte a questa domanda: gli item *c* ed *e* (in cui la risposta coincide con il dato fornito nello stimolo) vengono o sbagliati (Figura 36) o lasciati in bianco (Figura 37).


8. Sapendo che 34 x 33 = 1122, determina il risultato delle seguenti operazioni (possibilmente senza fare i calcoli):

a) 34 x 3,3 = 112,2

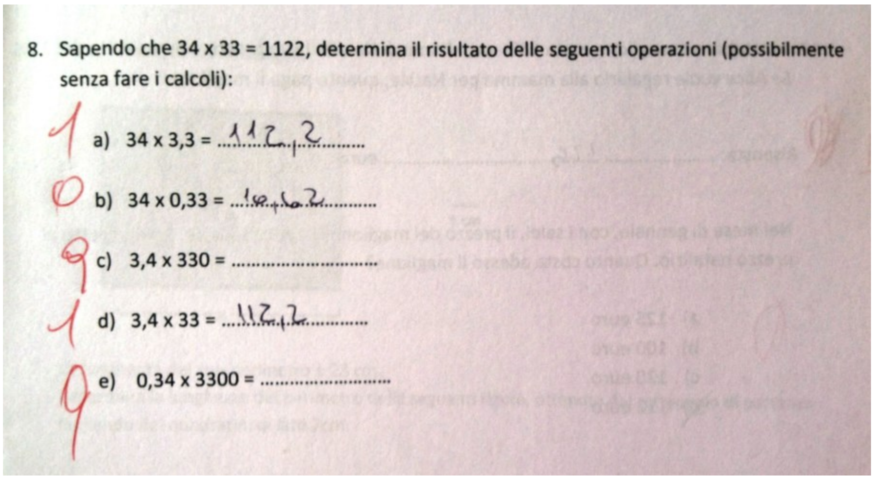
b) 34 x 0,33 = 11,22

c) 3,4 x 330 = 11220

d) 3,4 x 33 = 112,2

e) 0,34 x 3300 = 11220

Figura 36. Protocollo relativo al quesito "Zero-virgola" del questionario di grado 8, studente n° 1001


8. Sapendo che 34 x 33 = 1122, determina il risultato delle seguenti operazioni (possibilmente senza fare i calcoli):

a) 34 x 3,3 = 112,2

b) 34 x 0,33 = [illegible]

c) 3,4 x 330 = ...

d) 3,4 x 33 = 112,2

e) 0,34 x 3300 = ...

Figura 37. Protocollo relativo al quesito "Zero-virgola" del questionario dei gradi 10 e 11, studente n° 229

Le due immagini seguenti invece, (Figura 38 e Figura 39) mostrano un altro atteggiamento tipico riscontrato analizzando i protocolli: gli studenti ricorrono al calcolo in colonna solo per svolgere le operazioni degli item *c* ed *e*.

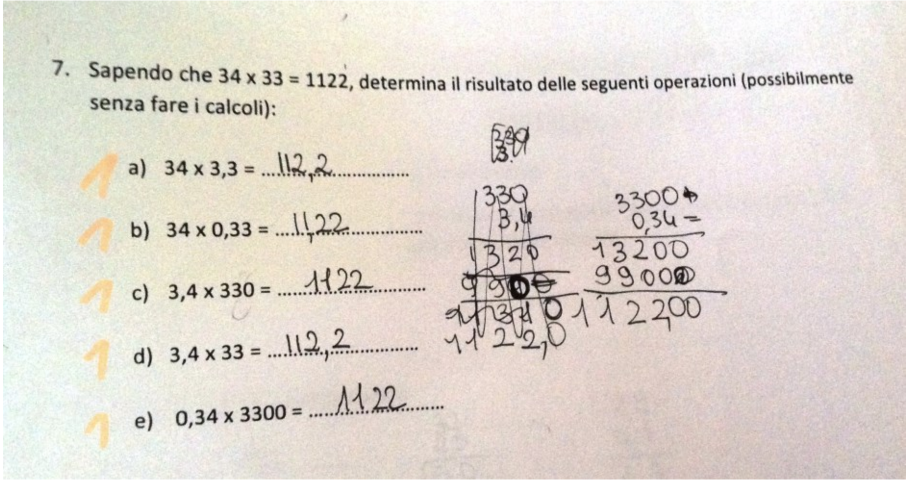


Figura 38. Protocollo relativo al quesito “Zero-virgola” del questionario dei gradi 10 e 11, studente n° 218

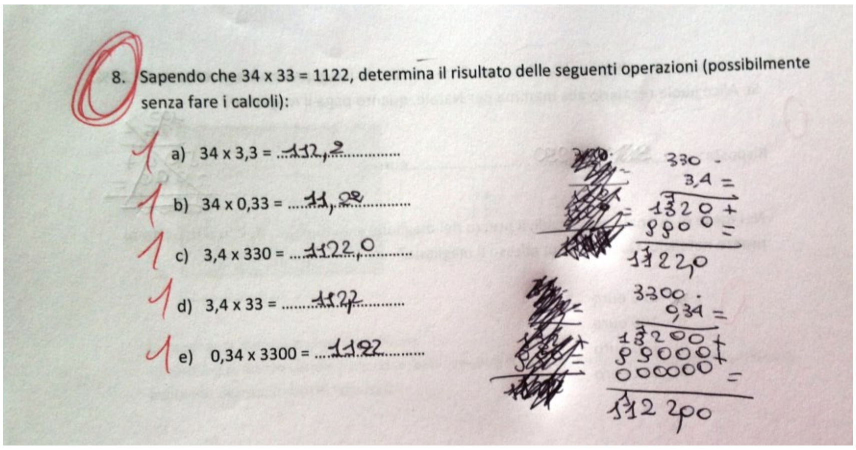


Figura 39. Protocollo relativo al quesito “Zero-virgola” del questionario di grado 8, studente n° 1061

5.4.2 Il quesito “Pianeta Marte”

Nei due questionari, quello rivolto agli studenti di grado 8 e quello rivolto agli studenti dei gradi 10-11, era presente un quesito, inerente alle proprietà delle potenze che richiamava la situazione di quesiti “Età della Terra” e “Massa del protone” delle prove INVALSI (Figura 3 e Figura 7). Il contesto del quesito del questionario è il pianeta Marte; per verificare quanto il controllo del contesto influisca sull’effetto indagato, il quesito è stato somministrato in due versioni.

La prima versione, “Pianeta Marte” è la seguente (Figura 40):

Il numero di atomi che compongono il pianeta Marte è stimato in circa 10^{54}. La sonda Voyager ha prelevato campioni di rocce composte da un numero di atomi stimato in 10^{28}, che ha portato sulla terra. Qual è la stima che più si avvicina al numero di atomi rimasti su Marte dopo che la sonda Voyager ha portato via i campioni?

a) 10^{54}

b) 10^{33}

c) 10^{28}

d) 10^{26}

Figura 40. Quesito "Pianeta Marte" presente nei questionari dei gradi 8, 10 e 11, versione A

Il contesto del quesito fa riferimento al numero di atomi del pianeta Marte, numero il cui significato (come grandezza) sfugge probabilmente a molti studenti di questi livelli scolastici.
La seconda variante, fa riferimento invece alla massa del pianeta. Si era infatti assunto che parlare di massa potesse portare a un maggiore controllo del significato rispetto al parlare di numero degli atomi. Il quesito è il seguente (Figura 41):

5. La massa del pianeta Marte è stimata in circa $6{,}4 \times 10^{26}$ g. La sonda Voyager ha prelevato campioni di rocce con una massa di circa 10^{5}g, che ha portato sulla terra. Quant'è, all'incirca, la massa di Marte dopo che la sonda Voyager ha portato via i campioni?

a) $6{,}4 \times 10^{26}$ g

b) 5.4×10^{26} g

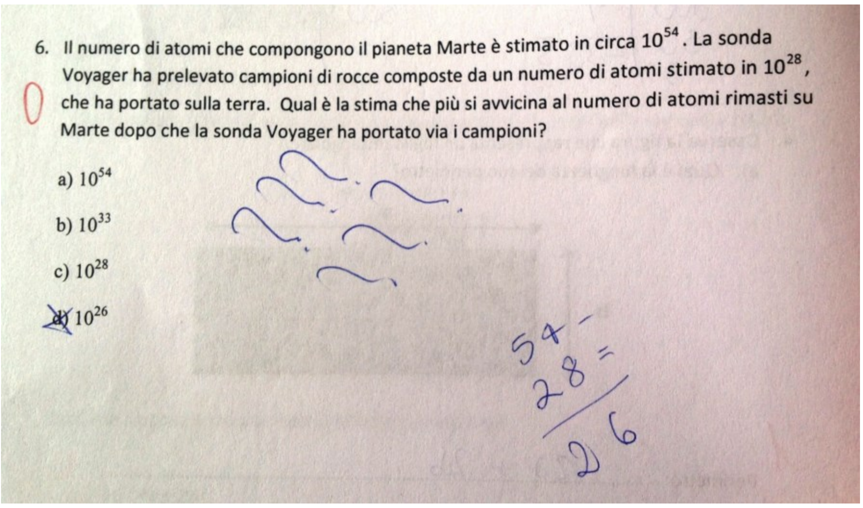
c) $6{,}4 \times 10^{5}$ g

d) $6{,}4 \times 10^{21}$ g

Figura 41. Quesito "Pianeta Marte" presente nei questionari dei gradi 8, 10 e 11, versione B

L'ipotesi era che risultasse per lo studente più naturale accettare il fatto che la massa di un pianeta non cambia se una sonda porta via un centinaio di kg di rocce, e quindi che non ci fosse difficoltà nell'accettare che la risposta fosse uno dei dati contenuti nell'enunciato del problema.
Complessivamente, al grado 11 le percentuali di risposte corrette sono del 40,5%, al grado 10 del 15,8% e al grado 8 del 8%. Nel grado 11, tra gli studenti che avevano la versione A, hanno risposto correttamente il 42,9% degli studenti mentre tra quelli che avevano la versione B hanno risposto correttamente il 38,5%. La situazione si rovescia al grado 10 dove solo il 13,3% degli studenti che avevano la versione A ha risposto correttamente e un po' di più, il 18%, la percentuale di risposte corrette degli studenti che avevano la versione B. I risultati ottenuti mostrano che, in questo caso,

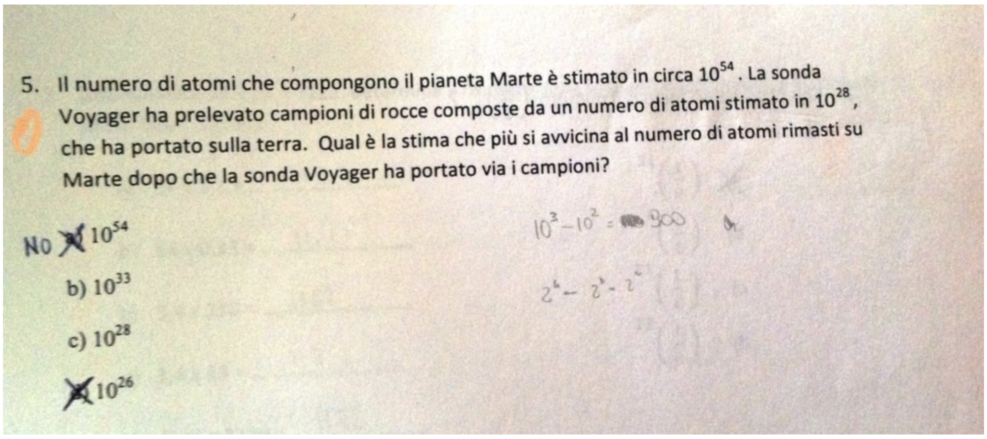
l’effetto indagato è indipendente dal controllo sul significato del contesto della domanda. A conferma del fenomeno rilevato nelle domande analizzate in precedenza, il distrattore d, nel quale l’esponente della potenza è ottenuto per sottrazione degli esponenti delle potenze presenti nel testo del quesito, è quello più scelto a tutti i gradi scolastici (è stato scelto dal 61% degli studenti di grado 8 e dal 49% degli studenti dei gradi 10 e 11).
Vengono di seguito riportati alcuni protocolli degli studenti (Figura 42 e Figura 43).


6. Il numero di atomi che compongono il pianeta Marte è stimato in circa 10^{54}. La sonda Voyager ha prelevato campioni di rocce composte da un numero di atomi stimato in 10^{28}, che ha portato sulla terra. Qual è la stima che più si avvicina al numero di atomi rimasti su Marte dopo che la sonda Voyager ha portato via i campioni?

a) 10^{54}
b) 10^{33}
c) 10^{28}
d) 10^{26}

54 - 28 = 26

Figura 42. Quesito “Pianeta Marte” del questionario di Grado 8, versione A, studente n°1024


5. Il numero di atomi che compongono il pianeta Marte è stimato in circa 10^{54}. La sonda Voyager ha prelevato campioni di rocce composte da un numero di atomi stimato in 10^{28}, che ha portato sulla terra. Qual è la stima che più si avvicina al numero di atomi rimasti su Marte dopo che la sonda Voyager ha portato via i campioni?

No a) 10^{54}
b) 10^{33}
c) 10^{28}
d) 10^{26}

$10^3 - 10^2 = 900$

Figura 43. Quesito “Pianeta Marte” del questionario dei Gradi 10 e 11, versione A, studente n°219

5.4.3 Qualche feedback dalle interviste di gruppo e individuali

I risultati rilevati nei protocolli delle interviste e le osservazioni e argomentazioni emerse durante le discussioni con gli intervistati hanno confermato l’ipotesi di partenza. La maggior parte degli studenti, 9 su 12, ha confermato (chi a livello più implicito, chi a livello più esplicito), in riferimento ad almeno una delle domande volte ad indagare l’effetto (nei protocolli della sperimentazione o nei protocolli delle interviste), che il fatto che il risultato fosse un dato esplicitamente presente nel testo dello stimolo ha costituito un elemento di disturbo. Un’intervista di gruppo particolarmente interessante è la seguente, effettuata a un gruppo di studenti del Liceo Statale “B. Rambaldi – L. Valeriani - Alessandro Da Imola” di Imola (BO). Il protocollo riportato fa riferimento alla fine dell’intervista; gli studenti stavano commentando i loro protocolli della sperimentazione. In dettaglio, stavano discutendo i risultati del quesito “Pianeta Marte”; entrambi gli studenti intervistati avevano svolto la versione A del questionario per il grado 10/11: lo studente S_2 ha risposto correttamente, lo studente S_1 ha prima messo la risposta giusta (A), poi si è corretto e ha scelto la risposta D (vedi protocollo n°219). Nel testo seguente R_1 e R_2 sono i due ricercatori che hanno condotto l’intervista ed S_1 lo studente e si fa riferimento al confronto tra le due versioni del quesito sul Pianeta Marte.

15:05 21 R_1: Dimmi se vedi delle differenze, qui come risponderesti.

15:28 22 S_1: Non mi ricordo mica sinceramente nemmeno quella che ho fatto io

23 R_1: Te la faccio vedere (*mostrando la domanda “gli atomi del Pianeta Marte” nel testo del questionario per i livelli 10/11 versione A, protocollo n° 219 di S_1*)

24 R_2: Eccola, leggila per bene

25 R_1: Ci spieghi come hai ragionato, quali sono le operazioni che hai fatto .. come mai hai fatto una scelta e poi l’hai corretta?

26 S_1: Innanzitutto questa domanda qui è stata una domanda un po’ particolare perché non ho capito se era una domanda trabocchetto o no perché quando ho letto che la sonda aveva prelevato una quantità di atomi pari a 10^{28}, ho messo inizialmente la risposta A perché ho pensato che comunque il numero generale di atomi del pianeta non cambiava! Poi invece ho capito che era una domanda… mmm… prettamente matematica perciò ho applicato le proprietà delle potenze perciò ho fatto 10^{54} meno 10^{28} e 54 meno 28 uguale 10^{26}

27 R_2: Quindi all’inizio hai scritto questo (*indicando la risposta A del protocollo n°219*) poi invece hai pensato che è una domanda di matematica e hai cambiato idea?

28 S_1: Si perché all’inizio proprio pensavo.. ho fatto nel senso.. ne ho fatto una questione fisica, nel senso se uno prende 10^{28} di atomi prende un tipo di atomo non è che li toglie tutti dal pianeta, anzi è come se non ne togliesse.

Come si evince dalle sue parole, lo studente S_1 leggendo il quesito ha subito interpretato il testo da un punto di vista fisico (lo studente frequentava la classe terza di un liceo scientifico e dalla classe prima ha seguito fisica come materia scolastica). Dal punto di vista fisico lo studente si è accorto che prelevando la quantità di atomi indicata nello stimolo, non sarebbe cambiato l'ordine di grandezza della quantità totale di atomi del Pianeta Marte. Però poi, quando ha osservato che stava risolvendo un quesito di matematica, ha cambiato strategia risolutiva, ha applicato le "proprietà delle potenze" e ha "fatto calcoli" che lo hanno portato a sbagliare. Alcune possibili cause del suo atteggiamento sono certamente riconducibili a due note clausole del contratto didattico, clausola di *delega formale* e clausola di *esigenza di giustificazione formale* (D'Amore, 2007). Indubbiamente, comunque si rileva il seguente fatto: quando lo studente si comporta in riferimento "al fare matematica", cambia idea e non sceglie più la risposta corretta contenente un dato esplicitamente presente nel testo.
Per quanto riguarda le interviste individuali (per i testi completi e le analisi puntuali delle interviste si veda Ferretti, 2015), le risposte fornite dagli studenti coinvolti si possono classificare principalmente con i costrutti di contratto didattico o del costume didattico (abitudini assunte nelle pratiche didattiche durante le lezioni di matematica oppure consuetudini apprese nello svolgere le operazioni e risolvere problemi matematici) e di norma socio-matematica (riferimento esplicito ai calcoli e a cosa è/non è ammesso mentre si fa matematica). Alcuni spiegano il proprio comportamento come se stessero enunciando una legge generale; in situazioni del genere è come se il contratto didattico avesse costruito un criterio di comportamento oggettivo, come se si fosse istituzionalizzato, confermando l'ipotesi interpretativa da cui si era partiti. C'è una costruzione di tipo "meta" e si è di fronte a un livello di generalità di secondo livello.
In altre interviste si può invece riscontrare la presenza del contratto didattico a livello implicito, con un riferimento generale a quello che si aspetta l'insegnante (o chi sta somministrando la prova); in linea con la definizione del costrutto teorico, infatti, spesso le "regole" che si instaurano tra allievo, insegnante e matematica spesso sono implicite. Altri studenti, infine, si riferiscono esplicitamente ai calcoli e a che cosa è lecito in matematica; si tratta di una norma socio-matematica unilaterale e distorta in quanto è una concezione degli studenti nei confronti della matematica e decide cosa non è lecito fare in matematica. Molto probabilmente questa norma si è instaurata dopo numerose esperienze vissute dagli studenti durante le quali non si sono mai trovati di fronte a problemi ed esercizi di matematica nei quali il risultato fosse un dato esplicitamente presente nello stimolo; questa regola si è quindi instaurata molto probabilmente in seguito a situazioni di contratto didattico.
Nelle parole degli studenti, in generale, si riscontra numerose volte la proiezione nel futuro di una regolarità vista nel passato; una trasposizione di un criterio matematico che ha funzionato nel passato (fenomeno che viene talvolta chiamato principio di continuità). Alcuni studenti fanno riferimento

esplicito a un principio regolativo come quello esplicitato e richiamano esplicitamente un comportamento legato al fatto che c'è un principio, altri devono invece essere sollecitati dalla ricercatrice.

Anche se manifestato in modo implicito, nella maggior parte delle interviste effettuate (22 su 24) si è quindi riscontrata la presenza di un principio regolativo che ha influenzato l'atteggiamento e il comportamento degli studenti di fronte ai task proposti. In generale, l'analisi dei protocolli ha mostrato come effettivamente fosse necessario ricorrere a diversi costrutti teorici per inquadrare i diversi livelli nei quali le parole degli studenti posizionavano questo principio.

5.5 Effetto "Età della Terra" - Conclusioni

Questa ricerca è stata per certi versi esemplare per il nostro approccio, da diversi punti di vista. In primo luogo, ha portato a confrontare diverse teorie con uno stesso fenomeno, aiutando a esplorare le sfaccettature del fenomeno stesso. Questo ha portato poi alla progettazione di altre situazioni, che hanno messo in luce altre caratteristiche del fenomeno. Si è così instaurata quella dinamica tra teorie e osservazione che ha caratterizzato molte delle ricerche del gruppo.

Inoltre, nella ricerca si sono succedute fasi sperimentali diverse, con un processo di definizione delle stesse determinato dai risultati ottenuti nelle fasi precedenti.

Partendo quindi da un macro-fenomeno emergente dai risultati di un quesito specifico, si è arrivati a vedere come questo stesso fenomeno si potesse verificare in gradi scolari diversi, su contenuti matematici vari, sia in situazioni contestualizzate che in domande senza contesto,

In particolare, la sperimentazione sul quesito "la massa del pianeta Marte" ha mostrato con chiarezza come i comportamenti di alcuni studenti fossero determinati esplicitamente dal fatto che "il quesito era di matematica". Infine, la ricerca ha mostrato come i problemi legati al contratto didattico siano presenti anche in livelli scolastici avanzati.

6. Prove INVALSI e formazione degli insegnanti

Quanto messo in luce dai macro-fenomeni emersi in sede di valutazione standardizzata, se opportunamente e consapevolmente integrato in ambito di formazione insegnanti, può contribuire a migliorare conoscenze e abilità degli insegnanti (Campbell & Levin, 2009; Di Martino & Baccaglini-Frank, 2017; Ferretti, Lemmo & Martignone, 2018a; Martignone, 2017). In particolare, l'analisi dei risultati di valutazioni su larga scala, interpretati alla luce dei singoli contesti, può essere utile agli insegnanti per riflettere sui processi di insegnamento-apprendimento e anche sui possibili punti di forza e sulle criticità di specifiche scelte didattiche. Per questo, già da molti anni, diversi membri del gruppo di ricerca hanno utilizzato le prove e i dati elaborati da INVALSI in programmi per la formazione insegnanti di matematica in formazione e in servizio (Bolondi, Ferretti & Spagnuolo, 2016; Bolondi et al., 2016; Ferretti, Lemmo & Martignone, 2018b; Lemmo et al., 2015; Martignone, 2016).

Con il passare del tempo e con l'aumento dell'esperienza sia in campo di formazione, sia in ambito di ricerca sul legame tra la didattica della matematica e le valutazioni su larga scala, le nostre esperienze nel campo della formazione insegnanti che utilizzavano le prove INVALSI sono confluite in ricerche sullo sviluppo professionale degli insegnanti e dei futuri insegnanti di matematica. In dettaglio, le attività che abbiamo progettato e studiato si basano su un utilizzo consapevole delle prove INVALSI di matematica per contribuire allo sviluppo delle conoscenze specialistiche degli insegnanti e dei futuri insegnanti. Nella sezione successiva sarà illustrato un modello di formazione insegnanti elaborato negli ultimi anni che collega l'analisi delle prove INVALSI di matematica e dei risultati di queste con lo sviluppo professionale degli insegnanti in formazione e in servizio innescando un ciclo virtuoso tra valutazione standardizzata, formazione insegnanti e insegnamento della matematica. Durante il seminario saranno discussi degli esempi di implementazione del modello con insegnanti in formazione e in servizio. Il seminario si concluderà con riflessioni in riferimento ad alcuni risultati di un progetto di ricerca condotto dal gruppo di ricerca dell'*Osservatorio SIRD – INVALSI e didattiche disciplinari* (Arzarello & Ferretti, 2021). Il progetto di ricerca è attualmente volto a raccogliere e analizzare informazioni sulle conoscenze e i *beliefs* degli insegnanti nei confronti dell'INVALSI: sia dell'INVALSI come Istituto, con le sue finalità e modalità di lavoro, sia nei confronti delle prove INVALSI di matematica e del loro impatto sulle pratiche didattiche.

6.1 Le ricerche sulla *Teacher Education* e lo strumento *Gestinv*

I nostri studi sulla *Teacher Education* sono incentrati da un lato sull'elaborazione di un approccio metodologico per progettare e sviluppare attività di formazione per insegnanti che utilizzino dati e risultati forniti dalle prove INVALSI, dall'altro sull'introduzione di risorse progettate per una collaborazione efficace tra insegnanti e tra insegnanti e ricercatori/formatori. Per quanto riguarda il primo aspetto, proponiamo un modello per la formazione degli insegnanti in cui le attività hanno le caratteristiche dei lavori di gruppo che si sviluppano in una *Community of inquiry* (Jaworski, 2006). Le nostre ricerche hanno voluto, in particolare, concettualizzare e delineare le conoscenze professionali specifiche che gli insegnanti potrebbero raggiungere come risultato del cambiamento (Guskey, 2002) innescato nel loro processo di formazione dall'incontro critico con le prove INVALSI, analizzate dalla ricerca didattica e confrontate con l'esperienza sul campo. Ci riferiamo a un'ampia gamma di conoscenze che includono matematica, epistemologia, pedagogia, didattica disciplinare, psicologia, ecc. I nostri studi si sono quindi avvalsi del modello MTSK - *Mathematics Teacher's Specialised Knowledge* (Carrillo-Yañez et al., 2018) per individuare e analizzare le conoscenze specialistiche di matematica degli insegnanti. All'interno di una *Community of inquiry* lo sviluppo di conoscenze specialistiche può richiedere strumenti e, in senso più ampio, risorse che mediano l'attività, contribuiscono allo scambio interpersonale tra i suoi attori e fanno emergere oggetti culturali e concettuali. Come sostenuto da Adler (2000), la formazione degli insegnanti di matematica dovrebbe incentrarsi sulle risorse, intese come strumenti che consentono agli insegnanti di ampliare il loro impatto nell'attività matematica in classe. In questa prospettiva, i nostri studi si sono rivolti verso la progettazione e utilizzo uno strumento/risorsa contenente tutti i quesiti delle valutazioni standardizzate INVALSI (Bolondi, Ferretti & Gambini, 2017): il database delle prove INVALSI *Gestinv* (www.gestinv.it).

Gestinv è un database strutturato, introdotto nel 2014, contenente tutti i dati rilasciati delle valutazioni standardizzate INVALSI, dalle rilevazioni del 2008 in poi e aggiornato costantemente, per tutti gli ambiti oggetto delle indagini INVALSI. Gestinv è corredato da un'interfaccia che permette di accedere a un motore di ricerca disegnato user-friendly. Nel caso specifico riguardante le prove INVALSI di matematica, il database contiene ad oggi 2340 item delle prove INVALSI di matematica. L'impatto del database è stato valutato sia quantitativamente che qualitativamente, attraverso indicatori standard come il numero di utenti registrati (oltre 35.000), il numero di accessi (in media 200 ogni giorno), il tempo trascorso sul sito e altri parametri. Questi dati, insieme alle sue informazioni strutturate, promuovono Gestinv come uno strumento da implementare nella progettazione di modelli di formazione degli insegnanti (Ferretti, Gambini & Santi, 2020).

Ci sono molti modi in cui si può utilizzare il database; al suo interno si possono infatti effettuare diverse forme di ricerca. In Gestinv ci sono i quesiti relativi alle prove INVALSI di tutte le materie oggetto di rilevazione, italiano, matematica e inglese. Noi presenteremo la sezione riguardante le prove INVALSI di matematica, l'unica oggetto del nostro modello di formazione.

Entrando nella sezione di Matematica è possibile effettuare ricerche in base a:

- le Indicazioni curricolari Ministeriali (le Indicazioni Nazionali per il Liceo; le Linee Guida per gli Istituti Tecnici e Professionali, gli Assi Culturali, le Indicazioni Nazionali per la scuola dell'infanzia e il primo ciclo d'istruzione);
- le parole chiave (ci sono circa 200 parole chiave che identificano uno o più contenuti in riferimento a ogni quesito);
- il testo dei quesiti (ricerche Full Text): il database permette di trovare il testo completo di un item digitando nel record di ricerca una o più parole contenenti in esso;
- i processi cognitivi individuati dal quadro teorico INVALSI;
- le percentuali di risposte corrette/errate e mancanti di ciascun item;
- le tipologie di quesiti (a scelta multipla, cloze, domande aperte, ecc.);
- la ricerca guidata: è possibile effettuare una ricerca incrociata con i connettori logici e/o coinvolgendo tutti i parametri sopra menzionati.

Nelle attività di formazione, Gestinv può svolgere un ruolo importante nel fornire agli insegnanti e ai ricercatori uno strumento interattivo per accedere a una vasta gamma di informazioni riguardanti le prove e i risultati delle prove INVALSI sul campione nazionale (Bolondi, Ferretti & Gambini, 2017; Ferretti, Gambini & Santi, 2020, Ferretti & Santo, in press). In particolare, come abbiamo visto, i risultati delle prove INVALSI evidenziano e quantificano macro-fenomeni rilevanti, che possono essere interpretati secondo i metodi e i risultati della ricerca in didattica della matematica. La struttura articolata e la ricchezza di informazioni fornite da Gestinv possono aiutare a collegare la pratica svolta in classe con la conoscenza specialistica degli insegnanti di matematica, attraverso la riflessione sui risultati delle prove. In linea con il quadro di riferimento dell'ICMI Study 25 Discussion document[6], Gestinv, se utilizzato in maniera consapevole, può diventare un *tool for teacher collaboration*. Riteniamo infatti che Gestinv sia una risorsa che intreccia strumenti "concettuali" e "pratici" (Grossman, Smagorinsky, & Valencia, 1999). Gestinv è uno strumento che può essere implementato, con un'utilità locale e immediata, come fonte di materiale didattico per l'insegnamento e l'apprendimento della matematica e per questo possiamo considerarlo uno strumento "pratico".

D'altra parte, il quadro teorico (INVALSI, 2018) che informa sia la costruzione che la selezione dei quesiti e la complessa struttura per la classificazione e ricerca delle informazioni sui quesiti nel database di Gestinv possono innescare l'uso di conoscenze di contenuto matematico, di strumenti teorici dell'educazione matematica, di idee sull'insegnamento e sull'apprendimento che complessivamente costituiscono strumenti concettuali.

[6] http://icmistudy25.ie.ulisboa.pt/wp-content/uploads/2019/03/190218-ICMI-25_To-Distribute_190304_edit.pdf

6.2 Un modello di formazione per insegnanti e futuri insegnanti di matematica che utilizza le prove INVALSI

Il modello che proponiamo per lo sviluppo professionale degli insegnanti in formazione e in servizio utilizza il database Gestinv all'interno di una *Community of inquiry* (Jaworski, 2006). I nostri studi riguardano l'influenza che l'uso consapevole di Gestinv può avere sulle conoscenze specialistiche degli insegnanti di matematica riflettendo criticamente sulla complessità delle valutazioni standardizzate con il supporto dei risultati della ricerca in educazione matematica. Come abbiamo già scritto, il quadro teorico con cui inquadriamo le conoscenze specialistiche degli insegnanti e dei futuri insegnanti di matematica è il modello *Mathematics Teacher Specialized Knowledge* di Carrillo-Yañez e colleghi (2018). Entriamo ora nel dettaglio della descrizione delle diverse componenti del nostro modello.

6.2.1 La Community of inquiry

L'interazione sociale all'interno di una *Community of practice* (Wenger, 1998) è alla base della costruzione della soggettività dell'insegnante. Le prospettive socioculturali nell'educazione matematica (Radford, 2008; Sfard, 2008) hanno mostrato il ruolo delle pratiche socio-comunicative in un contesto storico-culturale sia sui processi di apprendimento sia sulla costruzione dell'identità. Nel nostro modello, in un'ottica di formazione come duplice processo oggettivazione-soggettivazione, abbiamo esteso e adattato questi risultati della ricerca allo sviluppo professionale degli insegnanti di matematica. Inoltre, nella prospettiva socioculturale che stiamo sostenendo, la conoscenza matematica e la conoscenza pedagogica non sono entità fisse a priori che devono essere date per scontate; sono continuamente riflesse e rifratte nell'attività sociale e comunicativa. Lo sviluppo professionale degli insegnanti non può prescindere da questa caratteristica del pensare e del conoscere. Jaworski (2006, 2014) aggiunge un'importante caratteristica che contraddistingue una *Community of practice* (Wenger, 1998), cioè l'Inquiry – l'indagine, che si svolge in termini di pensiero critico, di investigare, dubitare, portare nuovi punti di vista, ecc.

"La trasformazione di una *Community of practice* in una *Community of inquiry* richiede che i partecipanti esaminino criticamente le loro pratiche mentre si impegnano in esse, mettano in discussione ciò che fanno mentre lo fanno ed esplorino nuovi aspetti della pratica. Queste forme di impegno basate sull'inquiry sono state definite "critical alignment" (Jaworski 2006). Il *critical alignment* è una necessità per sviluppare un modo di essere indagatore all'interno di una *Community of inquiry"* (Jaworski, 2014, p.77).[7]

[7] Trad. it a cura degli autori.

Pertanto, all'interno di questa concezione di *Community of practice*, possiamo pensare all'inquiry come "modo di essere", in cui gli insegnanti assumono il ruolo dell'indagine come centrale nel modo in cui pensano, agiscono e si sviluppano nella pratica e incoraggiano i loro studenti a fare lo stesso (Jaworski, 2014). L'appartenenza a una *Community of inquiry* si traduce quindi in un atteggiamento speciale, in un modo di essere e di diventare che definisce il modo in cui gli insegnanti agiscono, sentono, pensano, imparano e insegnano. Un approccio all'*inquiry* consente all'insegnante di sintonizzarsi con una situazione intrinsecamente imprevedibile, incontrollabile, fluida e flessibile, ossia l'aula di matematica, che richiede una costante interpretazione e rilettura per progettare e realizzare attività, prendere decisioni e gestire l'interazione sociale. La soggettività che l'insegnante realizza nel suo sviluppo professionale in una comunità di indagine non può essere separata dalla sua conoscenza matematica e dalla sua conoscenza per l'insegnamento. Nella sezione che segue, esamineremo l'altra componente del nostro modello, ossia la conoscenza per l'insegnamento che un insegnante acquisisce in una *Community of inquiry*. Per analizzare queste conoscenze useremo le lenti interpretative fornite dal modello MTSK.

6.2.2 Le conoscenze specialistiche degli insegnanti di matematica: il modello MTSK

L'importanza della conoscenza necessaria per l'insegnamento riguardante una specifica disciplina, è stata oggetto di molte ricerche: già a metà degli anni '80, Shulman (1986) propose riflessioni sulle aree di conoscenza che gli insegnanti dovrebbero possedere, in termini di *Pedagogical Content Knowledge* – conoscenza pedagogica del contenuto. La sua innovazione era la delineazione di questa nuova conoscenza del contenuto, specifica per l'insegnamento. In questa linea di ricerca, negli ultimi anni numerosi lavori hanno affrontato diversi aspetti riguardanti la conoscenza per e nell'insegnamento. Per indagare le conoscenze degli insegnanti, alcune ricerche non sono partite dai contenuti elencati nei programmi scolastici, ma si sono concentrate su approcci empirici al fine di individuare e comprendere le conoscenze matematiche per nell'insegnamento. Un esempio di ricerche di questo tipo sono quelle che hanno portato alla costruzione del modello MTSK (Carrillo-Yañez et al., 2018). Il modello MTSK è stato proposto dopo più di vent'anni di studi sulla formazione degli insegnanti dal gruppo di ricerca in didattica della matematica dell'Università di Huelva. Nei loro studi è centrale la riflessione sulle conoscenze usate dagli insegnanti per svolgere il loro lavoro, conoscenze specialistiche la cui specificità è appunto legata all'insegnamento della matematica. Il modello MTSK coordina due ampie aree di conoscenza, la Conoscenza Matematica (MK) e la Conoscenza Pedagogica di Contenuto (PCK) mettendo al centro i *beliefs* dell'insegnante (Figura 44).

La MK è la conoscenza posseduta da un insegnante di matematica in termini di una disciplina scientifica all'interno di un contesto educativo e PCK è la conoscenza relativa al contenuto matematico in termini di processi di insegnamento-apprendimento. I *beliefs* sulla matematica e sul suo apprendimento e insegnamento si trovano al centro del modello proprio per sottolineare la reciprocità tra *beliefs* e domini di conoscenza (Carrillo-Yañez et al., 2018). Nel modello, MK e PCK sono divisi in tre sottodomini. Il MK è composto da:

- Conoscenza degli argomenti – KoT (ad esempio la conoscenza di definizioni, proprietà, rappresentazioni, procedure…);
- Conoscenza della struttura della matematica – KSM (in questa conoscenza può rientrare la visione della matematica elementare da un punto di vista superiore, ma anche il saper passare e il collegare attività in domini diversi della matematica, ad esempio, tra algebra e aritmetica);
- Conoscenza delle pratiche in matematica – KPM (ed esempio il saper dimostrare, giustificare, definire, fare deduzioni e induzioni, dare esempi e comprendere il ruolo dei controesempi.).

I tre sottodomini del PCK sono:

- Conoscenza dell'insegnamento della matematica – KMT (ad esempio la conoscenza delle teorie sull'insegnamento della matematica, delle risorse e materiali e tecnologie, delle strategie per introdurre, rappresentare contenuti e concetti…);
- Conoscenza delle caratteristiche dell'apprendimento della matematica – KFLM (ad esempio sapere come potrebbero agire gli studenti, i loro errori e difficoltà in specifici argomenti e, in generale, come gli studenti interagiscono con la matematica tenendo conto anche degli aspetti affettivi);
- Conoscenza degli standard di apprendimento della matematica – KMLS (la conoscenza dei curriculum nazionali, ma anche di documenti internazionali e non solo per il proprio livello scolare ma per tutti).

I *beliefs* sulla matematica e sul suo apprendimento e insegnamento si trovano al centro del modello proprio per sottolineare la reciprocità tra *beliefs* e domini di conoscenza (Carrillo-Yañez et al., 2018)

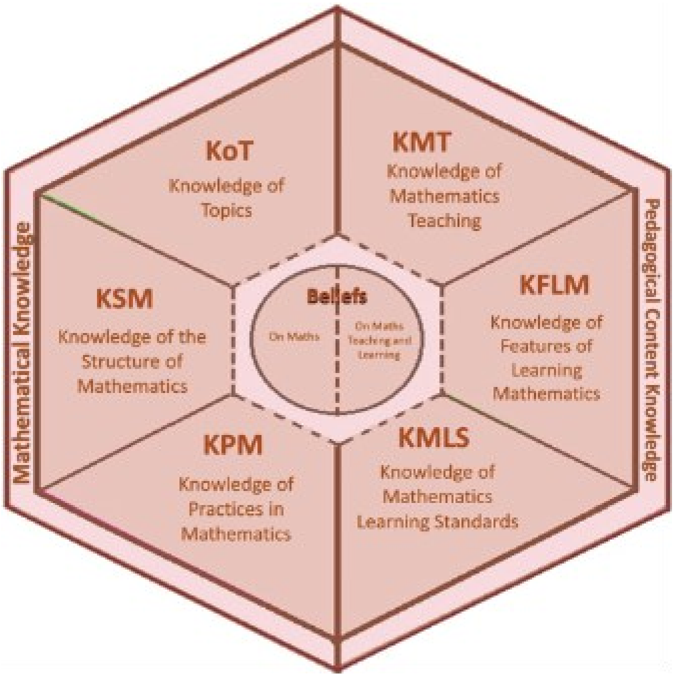


Figura 44. Il modello MTSK (Carrillo-Yañez et al., 2018, p. 241)

Questa distinzione tra i diversi sottodomini delle conoscenze specialistiche dell'insegnante e il ruolo centrale dei *beliefs* ci sono stati utili nelle nostre analisi per inquadrare i diversi aspetti delle conoscenze specialistiche degli insegnanti sviluppate nei programmi di formazione progettati secondo il nostro modello il cui nucleo è innescato da domande, dubbi, discussioni, confronti, indagini, ecc. riguardanti sia la MK sia la PCK all'interno di una *Community of inquiry*. Le caratteristiche della pratica "of inquiry", in cui gli insegnanti si impegnano e si confrontano, sono infatti legate al cambiamento e alla costruzione del sistema di conoscenze e *beliefs* che è descritto dal modello MTSK.

6.2.3 Il nostro modello per la formazione iniziale e in servizio

Nel nostro approccio per la formazione degli insegnanti, le attività svolte all'interno della *Community of inquiry* sfruttano dati e risultati forniti da INVALSI. Come abbiamo visto nelle sezioni precedenti, i risultati delle prove INVALSI evidenziano e quantificano macro-fenomeni rilevanti da un punto di vista didattico, ma che possono essere anche osservati dagli insegnanti, e collegati alla propria esperienza di classe. Il database Gestinv può fornire agli insegnanti e ai formatori uno strumento interattivo e condiviso per accedere a un'ampia gamma di informazioni e feedback. In tutto questo, il ruolo del formatore può diventare però cruciale, in quanto permette l'aggancio dell'esperienza del singolo insegnante alle evidenze sperimentali fornite dalle prove, sfruttando anche le lenti teoriche fornite dalla ricerca. In questi processi Gestinv funge da mediatore e materializza le attività della *Community of inquiry*. Questa si crea innescando il "questioning" - porre e cercare di rispondere alle domande (Jaworski, 2006) - il dubbio, il problem solving, la discussione, l'esplorazione e il confronto, promuovendo così l'apprendimento di conoscenze matematiche e pedagogiche dei contenuti. I percorsi sono concepiti con l'obiettivo di costruire comunità professionali di apprendimento in cui gruppi di insegnanti si impegnano a indagare sui processi di insegnamento e apprendimento della matematica (DuFour, 2004). Secondo il nostro modello di sviluppo professionale degli insegnanti, questi ultimi si sono impegnati in attività che hanno richiesto un uso interattivo di Gestinv, sfruttando la sua ricchezza di informazioni e la struttura degli strumenti di ricerca che abbiamo descritto nella sezione precedente. L'attività di inquiry è quindi svolta mediante l'interazione con i dati raccolti in Gestinv e potrebbe innescare un cambiamento delle conoscenze specialistiche e dei *beliefs* degli insegnanti per quanto riguarda sia aspetti relativi alla matematica sia al suo insegnamento e apprendimento. Durante le attività di formazione gli insegnanti possono mettersi in discussione e confrontarsi, discutendo sui contenuti matematici e sul loro insegnamento in verticale nei diversi gradi scolastici, sulle modalità e sui risultati della valutazione, sulle strategie e sulle metodologie di insegnamento, sui processi cognitivi e su errori e difficoltà degli studenti. La seguente immagine rappresenta le interazioni tra le diverse componenti del nostro modello di formazione (Figura 45).

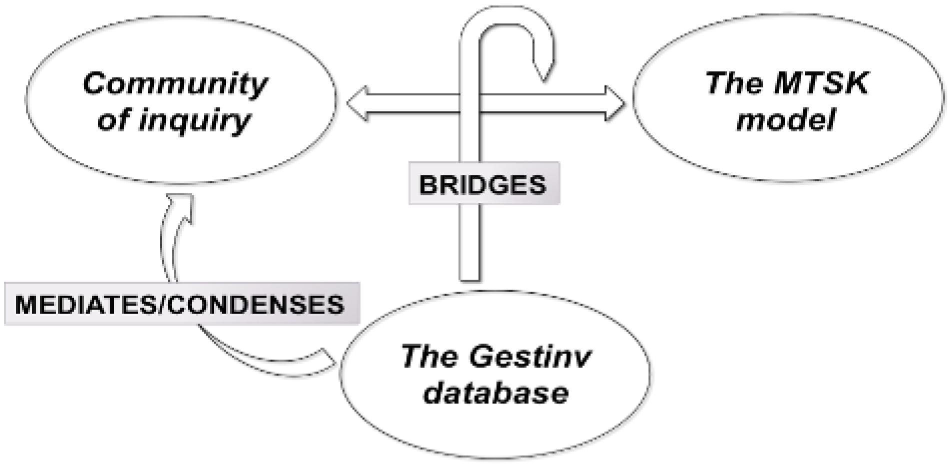


Figura 45. Componenti del nostro modello di formazione

Nello sviluppo del nostro modello adottiamo un approccio evolutivo che stimola gli insegnanti e i docenti a una più profonda consapevolezza delle proprie azioni, delle proprie motivazioni e dei propri obiettivi (Jaworski & Goodchild, 2006). Il nostro modello può chiamare in causa anche l'*Interpretative Knowledge* (Mellone et al., 2020), utile per la comprensione dei macro-fenomeni emersi in sede di valutazione standardizzata e nelle attività che caratterizzano la *Community of Inquiry*. Le pratiche sono state progettate al fine di migliorare e completare la consapevolezza degli insegnanti su entrambi i sottodomini del modello MTSK.

Arriviamo ora a descrivere in dettaglio la metodologia generale del nostro modello di formazione. Questa è costituita dalle seguenti fasi:

- *Introduzione dell'attività*. I ricercatori/formatori affrontano il contenuto matematico selezionato per l'attività da un punto di vista epistemologico e didattico. Successivamente vengono presentate alcune delle funzionalità di Gestinv che gli insegnanti useranno nella loro indagine e analizzano degli esempi di quesito. I ricercatori/formatori discutono con gli insegnanti i principali costrutti di didattica della matematica che forniscono possibili chiavi di lettura dei macro-fenomeni che emergono.
- *Analisi di un esempio*. I ricercatori discutono con l'intero gruppo di insegnanti un macro-fenomeno didattico utilizzando Gestinv stimolando riflessioni su uno o più sottodomini della conoscenza specialistica descritta dal modello MTSK.
- *Attività di gruppo*. I ricercatori assegnano un compito agli insegnanti che lavoreranno suddivisi in sottogruppi di massimo 4-5 persone. Il compito coinvolge un contenuto matematico in riferimento alle Indicazioni ministeriali del grado scolastico degli insegnanti coinvolti ed è inerente a una o più difficoltà di apprendimento individuate in letteratura. L'attività di gruppo mira alla costruzione di un prodotto multimediale, un artefatto, la progettazione di un'attività per gli studenti, ecc. Partendo da questi elaborati si discuteranno le conoscenze e i *beliefs* degli insegnanti coinvolti.
- *Discussione generale*. I sottogruppi presentano le loro produzioni al grande gruppo e al ricercatore. Ogni presentazione viene discussa all'interno della *Community of inquiry* al fine di evidenziare *beliefs*, affrontare dubbi, difficoltà e contenuti poco chiari riguardo sia la MK sia la PCK e delineare i sottodomini del MTSK che sono emersi durante tutto il percorso. La discussione finale, basata sulle presentazioni condivise, si svolge con le stesse caratteristiche di una *Community of inquiry*.

Il modello descritto è stato implementato con più di 4.000 insegnanti in servizio e in formazione negli ultimi anni; l'implementazione del modello ha fornito ulteriori materiali su cui sono state sviluppate ricerche (come, ad esempio, Ferretti, Martignone & Santi, 2022; Santi, Ferretti & Martignone, in press), e di cui una è oggetto della sezione successiva.

6.3 Esempi di ricerche sulla formazione insegnanti

Per descrivere e interpretare alcune variabili nei processi di formazione degli insegnanti basati sul nostro modello e poter analizzare le loro relazioni reciproche e le loro evoluzioni, abbiamo utilizzato il modello della *Trasposizione Meta-didattica* (Aldon et al., 2013; Arzarello et al., 2014; Robutti, 2020). Questo modello, che si basa sulla *Teoria Antropologica della Didattica* (TAD) elaborata da Chevallard (1999), prende in considerazione le pratiche e i processi che si sviluppano durante programmi di formazione di insegnanti. Lo sviluppo della matematica, l’insegnamento e l’apprendimento della matematica e la formazione degli insegnanti di matematica sono caratterizzati dalla dialettica tra la relazione personale e istituzionale con la conoscenza matematica (Chevallard, 1992). Con il modello della Trasposizione Meta-didattica possiamo descrivere e analizzare le prasseologie degli insegnanti di matematica e dei ricercatori-formatori sviluppate durante programmi per lo sviluppo professionale degli insegnanti o i progetti di ricerca-azione. In linea con la TDA, il modello della Trasposizione Meta-didattica si concentra sulla nozione di prasseologia per modellare le attività umane sviluppate all’interno di contesti istituzionali. In una prasseologia possiamo distinguere: il livello della prassi o del saper fare, che comprende il compito, o una famiglia di compiti, e le tecniche utilizzate per affrontarlo; e il livello del *logos* o della conoscenza, che comprende i "discorsi" sviluppati per giustificare o inquadrare le tecniche legate al compito (Chevallard, 1985; García, Gascón, Ruiz Higueras & Bosch, 2006). Nel modello della Trasposizione Meta-didattica le prasseologie sono sia didattiche, sia meta-didattiche: le prasseologie didattiche si riferiscono ai compiti relativi alla conoscenza da insegnare e alla tecnica che viene riconosciuta e giustificata all'interno di una specifica istituzione; le prasseologie meta-didattiche si riferiscono a compiti, tecniche e discorsi sulle tecniche legati alle riflessioni degli insegnanti e dei ricercatori sui contenuti da insegnare e sulle corrispondenti prasseologie didattiche. Nel nostro modello di formazione le prasseologie meta-didattiche riguardano le pratiche e le riflessioni teoriche sviluppate all’interno di una *Community of inquiry* (Jaworski, 2006). La comunicazione e la condivisione di riflessioni può essere favorita da strumenti, ideali o materiali, implementati nel processo di intermediazione: i *boundary objects* (Bowker, & Star, 1999). I *boundary objects* sono strumenti significativi che mettono in contatto insegnanti e formatori-ricercatori, anche se con sfumature e usi diversi e che caratterizzano le rispettive prasseologie. I boundary objects possono essere artefatti materiali, tecnologie digitali, procedure matematiche ecc. (Robutti et al., 2019). Rasmussen, Zandieh & Wawro (2009) caratterizzano i *boundary object* all’interno dell’educazione matematica e li definiscono come simboli matematici, tecnologie, documenti, software o altri elementi che consentono alle persone di collegare diverse comunità e di lavorare insieme. I *boundary objects* fungono da tali quando compiono un disvelamento, cioè quando fanno scorgere nuove idee e nuovi collegamenti. All’interno del nostro modello, se usato consapevolmente, Gestinv funge da *boundary object* nel contesto dei programmi di formazione degli insegnanti. Da un lato, Gestinv è prodotto e strumento di ricerca nell'ambito dei processi di valutazione su larga scala INVALSI; dall'altro, è uno strumento largamente utilizzato dagli insegnanti di matematica in Italia, dalla scuola primaria alla scuola secondaria di secondo grado. Nei nostri percorsi di formazione, Gestinv riveste il ruolo di risorsa di confine, tra il formatore-ricercatore e gli insegnanti (Wake, Foster & Swan, 2013). Gestinv diventa uno strumento di confine tra le due comunità, ne soddisfa le relative necessità informative e diventa mediatore di prasseologie, che nel nostro modello sono caratterizzate in termini di pratiche riflessive che crescono all'interno di una *Community of Inquiry*.

6.3.1 Gestinv come *boundary object*

In questa sezione mostreremo, attraverso un esempio, il ruolo di Gestinv come *boundary object* in un percorso di formazione per insegnanti in formazione (studenti di Scienze della formazione primaria della Libera Università di Bolzano e dell'Università di Verona) sviluppato a partire dal nostro modello (Ferretti, Gambini & Santi, 2020; Santi, Ferretti & Martignone, in press). Descriveremo brevemente come i dati strutturati forniti dal database Gestinv informino i contenuti delle attività e analizzeremo alcune delle risposte ottenute in un questionario somministrato al termine del percorso di formazione. Come vedremo, nelle risposte degli insegnanti in formazione abbiamo rilevato un incremento in termini del sottodominio della PCK definito nel modello MTSK come KMT - Conoscenza dell'insegnamento della matematica.

Uno dei quesiti INVALSI analizzato nel percorso di formazione è il seguente (Figura 46), somministrato a livello nazionale nella prova INVALSI di matematica di grado 5 del 2018.

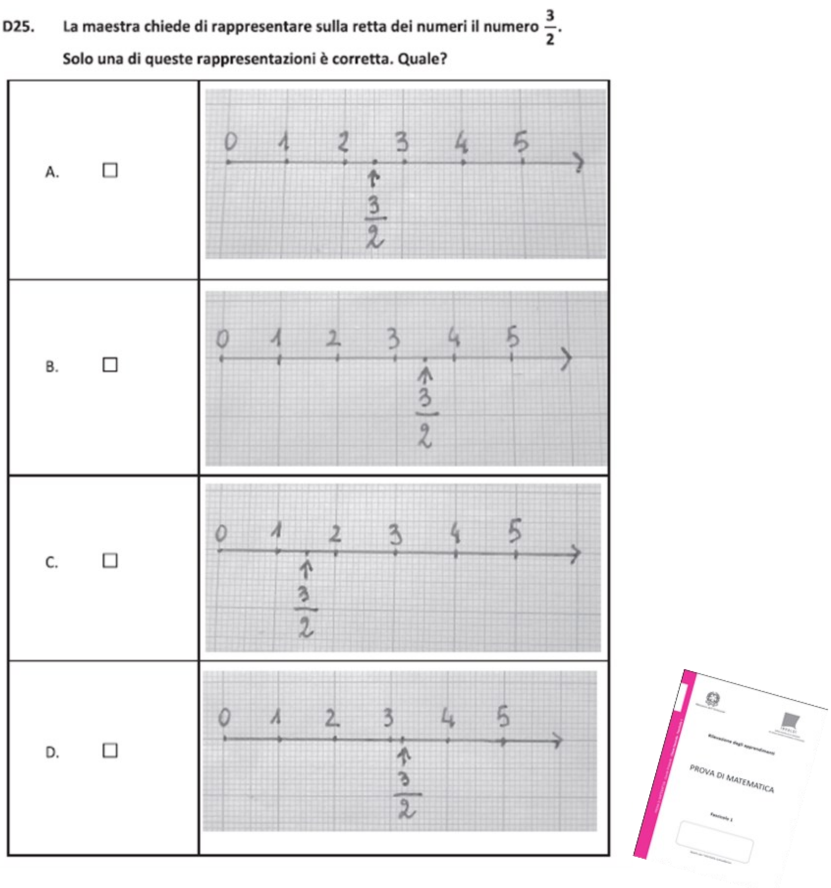


Figura 46. Quesito 25, Prova INVALSI di matematica Grado 5, 2018, www.gestinv.it

Il quesito è stato restituito dal database mediante ricerche, effettuate da diversi futuri insegnanti, focalizzate sull'individuazione di difficoltà diffuse sulle frazioni a livello di scuola primaria. Insieme ai ricercatori-formatori, si è deciso di analizzare il macro-fenomeno emerso in relazione al quesito in oggetto. Per i ricercatori il macro-fenomeno è interessante in quanto mette in luce alcuni dei possibili errori legati alle rappresentazioni dei numeri razionali. Il quesito richiede di riconoscere, tra le varie opzioni di risposta, quella che indica la posizione corretta della frazione 3/2 sulla linea dei numeri. Nell'immagine sottostante sono riportati i risultati ottenuti in riferimento a un campione statistico nazionale di quasi 500.000 studenti.

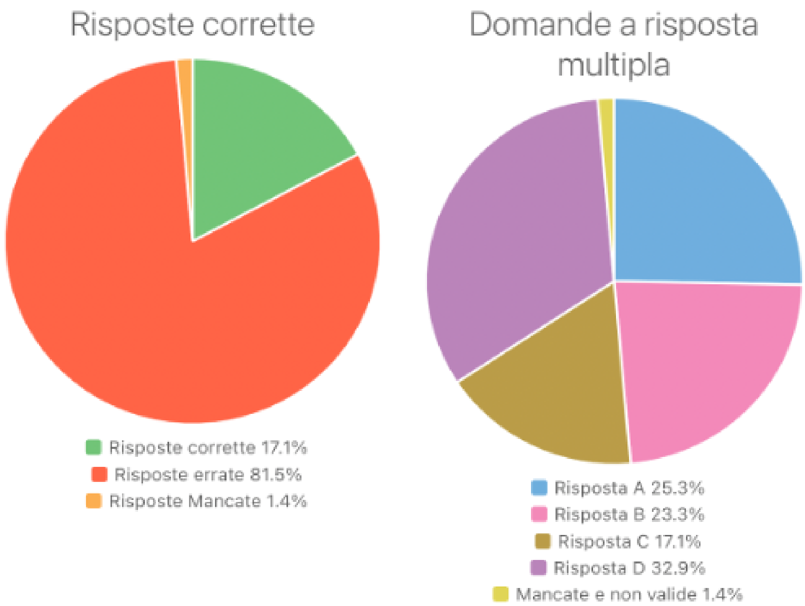


Figura 47. Percentuali nazionali, Quesito 25, Prova INVALSI di matematica Grado 5, 2018 www.gestinv.it

Vediamo (Figura 47) che circa il 17% degli studenti del campione nazionale ha fornito la risposta corretta. L'opzione D è stata scelta da più del 30% degli studenti e le opzioni A e B da circa il 25% e 23% rispettivamente. L'opzione A mostra il numero posizionato a metà tra 2 e 3 e potrebbe identificare gli alunni che interpretano 3/2 come un valore che sta a metà tra le cifre 3 e 2 che compongono la frazione. L'opzione B mostra il numero posizionato tra le cifre 3 e 4 e potrebbe identificare gli alunni che attribuiscono alla frazione 3/2 il significato di 3 e mezzo (3,5). L'opzione D mostra il numero posizionato due "tacche" dopo il 3 e potrebbe identificare gli alunni che interpretano la rappresentazione 3/2 come 3,2. Le tre opzioni sbagliate rappresentano quindi possibili interpretazioni della rappresentazione del numero razionale in forma di frazione e sono in linea con quanto emerso dalla letteratura sul tema che mette in luce le difficoltà nella gestione delle diverse rappresentazioni dei numeri razionali, del loro posizionamento sulla linea dei numeri e nel dare senso alle loro rappresentazioni, soprattutto quando sono sotto forma di frazione (Ni & Zhou, 2005; Pitta-

Pantazi, 2014; Saxe et al., 2007). Nel caso del quesito in esame siamo in presenza di una linea dei numeri graduata: le linee dei numeri graduate sono rappresentazioni ibride, costituite da una linea e da una scala, e la gestione di entrambi i significati veicolati dall'unica rappresentazione è di per sé un obiettivo di apprendimento importante (Skoumpourdi, 2010). In diverse situazioni, come in quella del quesito, lavorare sulla retta dei numeri permette di rendere particolarmente visibili alcune difficoltà relative alla gestione delle rappresentazioni numeriche.
Durante il percorso di formazione, i ricercatori-formatori hanno affrontato i contenuti matematici in riferimento al quesito individuato tramite Gestinv da un punto di vista epistemologico e didattico. Sono stati inoltre discussi con gli insegnanti in formazione i risultati della ricerca in educazione matematica. I futuri insegnanti hanno lavorato in gruppo concentrando le loro analisi e riflessioni sui macro-fenomeni relativi alle frazioni messi in luce dal quesito INVALSI presentato e da altri quesiti. Al termine di tutto il percorso, è stato somministrato un questionario volto a indagare le conoscenze specialistiche degli insegnanti in formazione per l'insegnamento della matematica inquadrate nel modello MTSK. In quasi tutte le risposte al questionario analizzate possiamo identificare delle frasi che mostrano una consapevolezza di un aumento della loro conoscenza dell'insegnamento della matematica e in cui dichiarano il ruolo cruciale di Gestinv nelle attività svolte. Riportiamo qui di seguito alcuni estratti delle risposte al questionario.

PS_BX_33: Per sviluppare meglio questi concetti abbiamo toccato con mano esempi tratti dalle prove INVALSI che riportavano argomenti come "frazioni di aree" o "linee di numeri". Le attività ci hanno permesso sia di analizzare i punti di forza e di debolezza della formulazione di quesiti valutativi e di poter entrare nel merito dei possibili errori commessi dagli studenti.

PS_VR_48: Lo studio dei diversi tipi di rappresentazione dei numeri razionali ha evidenziato l'importanza di gestire e apprendere le trasformazioni semiotiche all'interno di uno stesso registro. Poiché i bambini potrebbero non capire che stiamo parlando di rappresentazioni diverse dello stesso oggetto, è importante renderli consapevoli e proporre loro situazioni diversi, lavorando sul concetto attraverso le varie rappresentazioni. Quello che mi è rimasto di più è questo.

PS_BR_32: Il database è spettacolare. Il fatto di poter fare ricerca per parole chiave è la cosa che credo sia più utile in assoluto. Le nostre attività sono partite da errori commessi non di un singolo bambino ma da tantissimi bambini; per noi che ancora non abbiamo le nostre classi questo può aiutarci perché è probabile che anche i nostri futuri bambini incontrino quelle difficoltà.

Altri estratti, fanno riferimento ad altre conoscenze inquadrate nel MTSK. Per approfondimenti su questa ricerca e su altre che vertono sull'impatto dell'implementazione del nostro modello di formazione si rimanda a Santi, Ferretti e Martignone (2023) e Ferretti, Martignone e Santi (2022).

6.4 Analisi delle convinzioni e dei bisogni formativi degli insegnanti italiani in relazione alle prove INVALSI

In merito alla nostra esperienza nel campo del legame tra le valutazioni standardizzate e la formazione insegnanti, alcuni di noi sono stati coinvolti in un progetto di ricerca interdisciplinare ancora in corso di sviluppo condotto dal *Gruppo INVALSI-Didattiche Disciplinari, Osservatorio della S.I.R.D. - Società Italiana di Ricerca Didattica*, formato da esperti di didattica di matematica e informatica, pedagogisti sperimentali e insegnanti in servizio[8]. In questa sezione presenteremo i primi sviluppi del progetto.

Gli Osservatori della SIRD (https://www.sird.it/osservatorio-didattica-e-saperi/) nascono da un'esigenza di coesione emersa dall'esperienza condotta in questi anni dai diversi atenei italiani nel campo della formazione degli insegnanti che ha messo in evidenza con sempre maggiore forza, come la professionalità dei docenti si fondi sull'integrazione coerente, intelligente dei saperi pedagogico-didattici, con quelli dei diversi settori disciplinari.

Come abbiamo già discusso in precedenza, diverse ricerche mostrano come un utilizzo informato e consapevole delle valutazioni standardizzate italiane INVALSI in ambito di formazione insegnanti possa avere un impatto nella formazione degli insegnanti in servizio e dei futuri insegnanti di matematica (Di Martino & Baccaglini-Frank, 2017; Ferretti, Gambini & Santi, 2020; Ferretti, Martignone & Santi, 2022). Il progetto di ricerca condotto dall'Osservatorio SIRD Didattiche disciplinari e prove INVALSI si inserisce in questa prospettiva. La prima fase del progetto è stata incentrata sul rilevamento e l'analisi delle conoscenze e dei *beliefs* degli insegnanti sulle le prove INVALSI di Matematica e i loro possibili legami con i processi di insegnamento e apprendimento della Matematica (Ferretti, Funghi & Martignone, 2020). I principali risultati di questa prima fase sono stati oggetto di numerosi lavori a diffusione nazionale e internazionale (si veda, ad esempio, Arzarello & Ferretti, 2021; Faggiano, Ferretti & Arzarello, 2022; Faggiano, Monaco, Rizzo & Vaccaro, in press; Truffelli & Asquini, 2022; Truffelli & Vannini, 2021; Spagnolo, Vaccaro,

[8] Il gruppo di ricerca, coordinato da Ferdinando Arzarello (Università degli Studi di Torino) e Ira Vannini (Alma Mater Studiorum Università di Bologna), è composto da: Giorgio Asquini (Sapienza, Università di Roma), Barbara Balconi (Università di Milano Bicocca), Giorgio Bolondi (Libera Università di Bolzano), Eleonora Faggiano (Università di Bari Aldo Moro), Federica Ferretti (Università di Ferrara), Elisa Guasconi (Alma Mater Studiorum Università di Bologna), Violetta Lonati (Università degli Studi di Milano), Daniela Maccario (Università di Torino), Annarita Monaco (Sapienza, Università di Roma), Francesca Martignone (Università del Piemonte Orientale), Ottavio Rizzo (Università degli studi di Milano), Camilla Spagnolo (Libera Università di Bolzano), Elisa Truffelli (Alma Mater Studiorum Università di Bologna), Valentina Vaccaro (INVALSI e Università di Oviedo).

Faggiano, in press, Vaccaro, Faggiano & Ferretti, 2021); noi ci focalizzeremo su alcuni risultati ottenuti in riferimento alle conoscenze e *beliefs* degli insegnanti in riferimento a diversi aspetti collegati all'analisi di alcuni quesiti delle prove INVALSI di matematica.

6.4.1 Il progetto contenitore e il questionario

Nella prima fase di ricerca, il lavoro interdisciplinare è consistito nella definizione di un disegno di ricerca osservativa su un campione di convenienza di insegnanti di scuola primaria (circa 500), incentrato sulla rilevazione di opinioni, atteggiamenti e dichiarazioni di comportamento nei confronti delle rilevazioni INVALSI, sia riguardo le loro finalità valutative in termini istituzionali, sia in particolare riguardo le prove INVALSI di Matematica e le loro ricadute sulla didattica. L'obiettivo è quello di comprendere quali sono gli "strumenti" che gli insegnanti hanno a disposizione e, soprattutto, utilizzano per leggere e interpretare le rilevazioni INVALSI e di quali "strumenti" dispongono per individuare possibili ricadute delle prove sulla didattica della Matematica. Per raggiungere gli obiettivi prefissati sono state definite specifiche ipotesi correlazionali ed è stato progettato e somministrato un questionario volto ad indagare le convinzioni degli insegnanti per quanto riguarda le conoscenze e le abilità indagate dalle prove INVALSI, la loro vicinanza alle pratiche didattiche in Matematica e il ruolo che esse assumono all'interno del contesto scolastico. La progettazione del questionario si è fondata sulle variabili e sulle ipotesi di ricerca relative alle relazioni tra le variabili rappresentate in Figura 48.

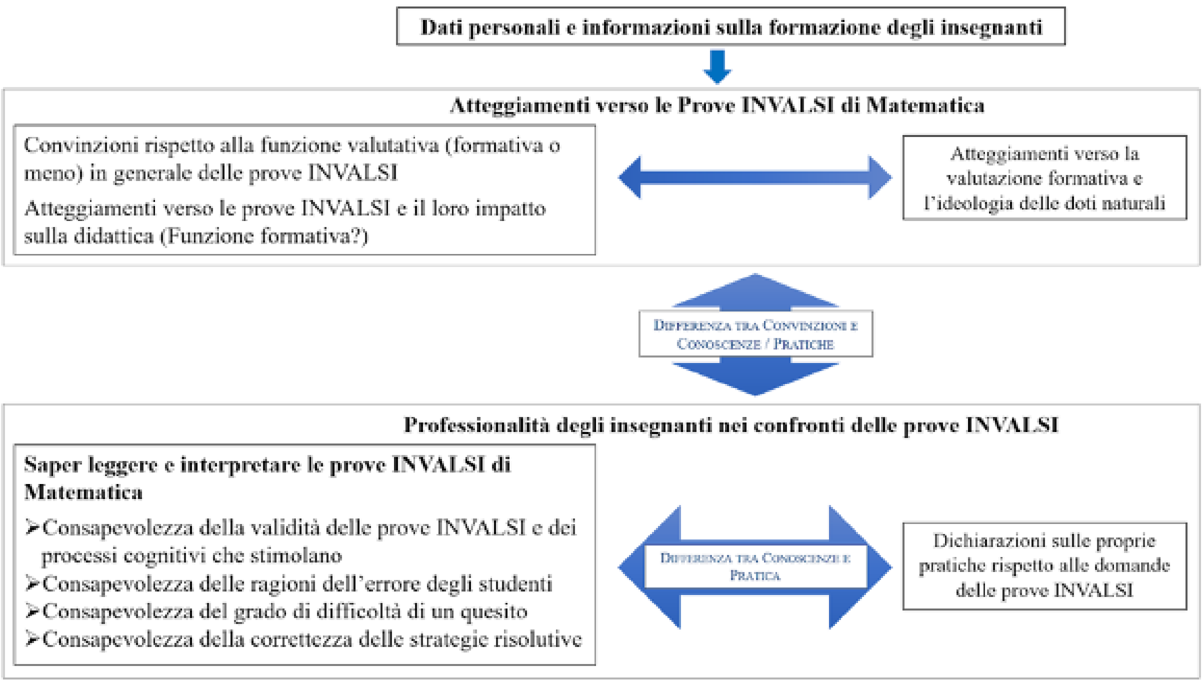


Figura 48. Quadro delle variabili del questionario

Il questionario è costituito da due ambiti di variabili, uno specificamente di Didattica della Matematica e uno relativo agli aspetti di Didattica generale. Questi due ambiti sono declinati in 51 domande (per la maggior parte a risposta chiusa e con scale di ranking), suddivisi in tre sezioni principali: una prima sezione di Didattica della Matematica, una seconda sezione relativa alle convinzioni e agli atteggiamenti degli insegnanti verso le prove INVALSI e verso l'uso di una valutazione con funzione diagnostico-formativa, una terza sezione che raccoglie dati personali e informazioni sui percorsi formativi e professionali.

In dettaglio, le variabili di Didattica della Matematica sono finalizzate a rilevare quanto i contenuti e le abilità matematiche rilevate con le prove INVALSI di Matematica siano: a) più o meno vicini alle pratiche didattiche personali quotidiane degli insegnanti; b) vissute come coerenti/incoerenti con le Indicazioni nazionali e siano riconosciute o meno in modo coerente con le intenzioni di INVALSI; c) ritenute utili per influenzare/innovare la pratica didattica personale. Nella prima sezione sono quindi proposti sette item delle prove INVALSI di matematica di grado 5 e grado 6, e per ciascun item vengono proposte domande volte ad indagare la *Pedagogical Content Knowledge* in relazione alle capacità di lettura e interpretazione dei testi e dei dati degli item INVALSI, della loro coerenza con le Indicazioni ministeriali e delle interpretazioni dei macro-fenomeni emersi in sede nazionale.
La seconda sezione, di Didattica generale, è volta ad indagare le opinioni, convinzioni e atteggiamenti degli insegnanti sulle e verso le prove INVALSI e gli atteggiamenti verso un uso diagnostico-formativo della valutazione in classe e verso l'ideologia delle doti naturali (Ciani & Vannini, 2017).
Il questionario è stato inizialmente somministrato in modalità telematica a 105 insegnanti. Questa prima fase di *Try Out* ha permesso di pre-testare il questionario e, in base alle analisi e all'interpretazione dei risultati ottenuti sono state apportate modifiche sia dal punto di vista contenutistico in un'ottica di maggiore aderenza agli obiettivi di ricerca sia per una maggiore validità e affidabilità da un punto di vista misuratorio. La nuova versione del questionario, oggetto del Main Study, è stata somministrata in modalità telematica a 427 docenti, con 421 casi validi. In entrambe le fasi, i partecipanti sono docenti provenienti da diverse regioni distribuite sul territorio nazionale e non costituiscono un campione rappresentativo da un punto di vista statistico. I dati raccolti sono stati codificati e analizzati utilizzando il software statistico *IBM® SPSS® Statistics 27*.

6.4.2 Il triplice conflitto metacognitivo

Focus della presentazione dei primi risultati ottenuti nel progetto sarà quello che Arzarello e Ferretti hanno denotato come "*three-fold meta-didactical conflict*" (Arzarello & Ferretti, 2021; Ferretti, Vannini & Arzarello, under review), costrutto emerso dall'analisi quantitativa delle risposte fornite nel questionario nella sezione di Didattica della matematica. La contestualizzazione di questo *conflitto meta-didattico* nasce dall'analogia tra il piccolo contesto di una classe di matematica e il contesto scolastico nazionale generale, che la ricerca sta prendendo in considerazione. Nell'analisi delle interazioni in classe è comune osservare quello che Anna Sfard chiama un "discorso incommensurabile" (Sfard, 2008) tra l'insegnante e gli studenti: spesso accade che – nei processi comunicativi nelle pratiche didattiche – gli studenti usino le stesse parole degli insegnanti ma con un significato diverso, e senza essere consapevoli della differenza. Si genera un conflitto che, se non viene superato, può avere severe conseguenze sul successo dei processi di insegnamento-apprendimento in classe.

Qualcosa di simile è successo quando abbiamo analizzato le risposte al questionario: anche in questo contesto abbiamo trovato linguaggi incommensurabili, che identificano il cosiddetto "three-fold meta-didactical conflict". Come suggerisce il nome, questo conflitto ha tre componenti ed è meta-didattico poiché riguarda i discorsi sui processi didattici come la valutazione, le competenze e gli errori degli studenti, ecc. e non i concetti matematici pensati in sé.

Di seguito, illustriamo dettagliatamente le tre singole componenti.

La prima componente: la percezione della difficoltà dei quesiti

La prima componente del conflitto riguarda il fatto che molti insegnanti percepiscono le difficoltà dei quesiti delle prove INVALSI di matematica in modo molto diverso da quanto emerge dai dati raccolti da INVALSI a livello nazionale. La domanda del questionario inerente questa prima componente si riferisce al quesito 10 della prova INVALSI di matematica a.s. 2008/09 di grado 5 (Figura 49).

10. A quale numero corrispondono "12 decine, 7 decimi e 2 millesimi"?

☐ A. 12,702.

☐ B. 120,702.

☐ C. 12,72.

☐ D. 120,72.

Figura 49. Quesito 10, Prova INVALSI di matematica Grado 5 - 2009, www.gestinv.it

In termini della teoria della gestione delle rappresentazioni di Duval (2006), la risoluzione del quesito richiede una *conversione*, una trasformazione semiotica da una rappresentazione in un registro (in questo caso il registro verbale) in un'altra rappresentazione in un registro differente (il registro numerico posizionale). Come mostrato dalla letteratura, le gestioni delle trasformazioni di conversione sono particolarmente difficili da realizzare perché non esistono regole sintattiche che leghino le due rappresentazioni in due diversi registri. In questo caso, una delle principali difficoltà è riconducibile alla gestione di conversione di “12 decine”. Le difficoltà incontrate dagli studenti sono confermate dalle percentuali nazionali rappresentate in Figura 50.

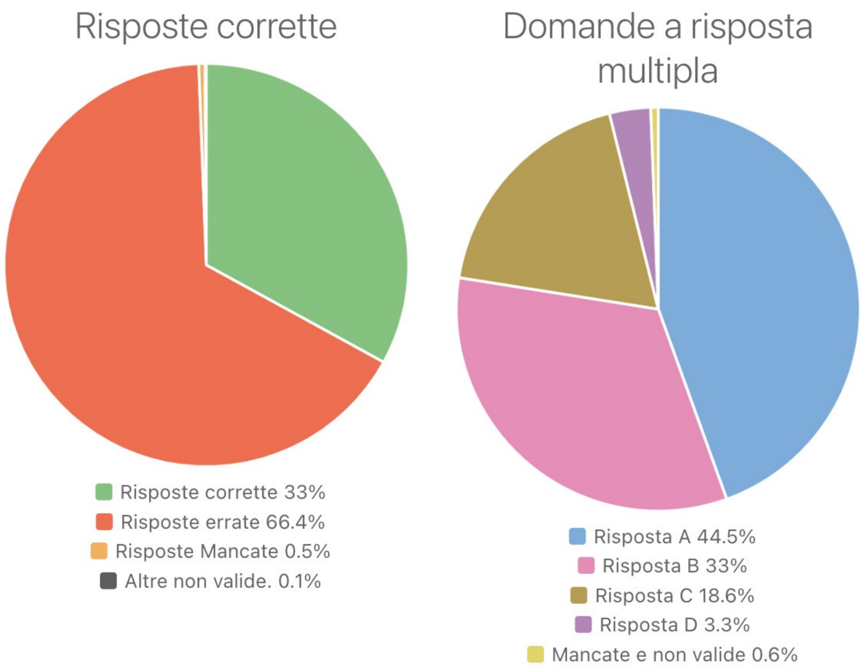


Figura 50. Percentuali nazionali, Quesito 10, Prova INVALSI di matematica Grado 5, 2009 www.gestinv.it

Come si può vedere in Figura 50 la percentuale di risposte corrette (opzione B) del campione nazionale analizzato da INVALSI è del 33%. Il 44% degli studenti del campione, più di quelli che hanno risposto correttamente, ha scelto l'opzione A in cui sono state gestite correttamente le conversioni di tutti gli altri elementi, tranne “12 decine”.

Nel questionario non sono state fornite le percentuali di risposte ottenute a livello nazionale ed è stato chiesto agli insegnanti di stimare su una scala da 1 a 10 il grado di difficoltà del quesito, in termini di “quanto considerino difficile questo quesito al termine della classe quinta della scuola primaria”.

Nella tabella seguente (Tabella 8) sono riportati i risultati con riferimento alle 516 risposte valide raccolte (1 = molto facile, 10 = molto difficile).

Tabella 8. Risposte ottenute in riferimento alla domanda del questionario sulla difficoltà del quesito INVALSI

valore	# risposte
1	75
2	143
3	86
4	39
5	67
6	29
7	37
8	30
9	7
10	3

Come si può osservare in Tabella 8, la percezione della difficoltà del quesito da parte degli insegnanti è molto distante dalla difficoltà effettiva misurata sul campo dall'INVALSI: 410 insegnanti su 516, quasi l'80%, ha scelto una risposta compresa tra da 1 a 5. Quasi la metà degli insegnanti, il 42%, sceglie i primi due valori, e il valore mediano della distribuzione di frequenze è 3. Il quesito viene quindi ritenuto molto facile.

La seconda componente: la consapevolezza delle cause degli errori degli studenti

La seconda componente riguarda le interpretazioni da parte degli insegnanti delle possibili cause degli errori degli studenti rilevati in sede di valutazione standardizzata INVALSI. In questa direzione, sono rilevanti i risultati in riferimento a quattro domande del questionario che fanno riferimento ai seguenti quesiti delle prove INVALSI di matematica: D3 e D25 grado 6 - a.s. 2011/12, D31, grado 5 – a.s. 2015/16, D17, grado 5 – a.s. 2008/09.

In riferimento a ciascuno dei quesiti è stata riportata la percentuale delle risposte corrette a livello nazionale ed è stato richiesto agli insegnanti di indicare quale, tra alcune motivazioni proposte, sia la principale causa del dato rilevato. I quesiti, tutti a risposta chiusa, sono stati individuati dai ricercatori in quanto in ciascuno di essi la causa della scelta di almeno una delle opzioni sbagliate è chiaramente inquadrabile con alcuni dei costrutti più noti e condivisi della letteratura in didattica della matematica a livello internazionale (come, ad esempio, il costrutto di Contratto Didattico di Brousseau, 1986, o la teoria delle *intuitive rules* di Stavy & Tirosh, 1996). Tra le motivazioni proposte ce n'è almeno una rappresentativa di inquadramento della causa degli errori in linea con la letteratura. Come messo in luce in Vaccaro, Faggiano e Ferretti (2021), il 35% degli insegnanti non fornisce la risposta in cui le possibili cause sono quelle individuate attraverso la letteratura di ricerca in nessuna delle quattro domande e solo l'1% risponde in modo coerente con la letteratura a tutte e quattro le domande.

La terza componente: la percezione dello scopo dei quesiti

La terza componente è una conseguenza delle due precedenti e riguarda il modo contradditorio con cui gli insegnanti interpretano lo scopo dei quesiti delle prove INVALSI di matematica.
In seguito a ogni quesito della prima sezione del questionario sono presenti due domande – le stesse per tutti i quesiti. Queste due domande sono volte a indagare quanto gli insegnanti ritengano ciascun quesito delle prove INVALSI di matematica adatto a valutare l'apprendimento degli studenti e quanto utilizzino questi quesiti nelle loro pratiche didattiche. I risultati ottenuti in riferimento ai quesiti INVALSI focus della prima sezione del questionario rivelano forti contraddizioni tra la risposta a queste due domande e la consapevolezza delle cause degli errori/la percezione della difficoltà dei quesiti. Ad esempio, per quanto riguarda il primo quesito indagato (il quesito inerente la gestione delle rappresentazioni semiotiche di numeri razionali), le cui basse percentuali di risposte corrette non corrispondono alla difficoltà ipotizzata dagli insegnanti, è stato considerato uno dei quesiti più adatti alla valutazione (445 insegnanti su 516, l'86%) e come tipologia, il più utilizzato nelle pratiche didattiche (452 insegnanti su 516, l'88%). Situazioni analoghe si sono verificate in riferimento agli altri quesiti presenti nella prima sezione.
Naturalmente, è possibile che le tre componenti siano solo l'epifenomeno di un conflitto più profondo, di cui al momento non si è ancora compreso la natura; finora è stato discusso solo un triplice conflitto, con le sue tre componenti profondamente intrecciate.
L'ipotesi su questo conflitto sarà confermata o smentita attraverso la somministrazione del questionario a una platea più ampia di soggetti e un'analisi approfondita dei relativi dati. In caso di conferma, sarà possibile affinare la stessa analisi del conflitto, decidendone la triplice o diversa natura, e chiarendone anche la struttura e la natura profonda, ad esempio rispetto alle conoscenze e alle convinzioni degli insegnanti. Sulla base dell'analogia con i conflitti in classe, per i quali una strategia di successo per il loro superamento è generata da una chiara comprensione della loro natura, sarà proprio a partire da un quadro pulito della struttura e della dinamica del conflitto meta-didattico che sarà possibile progettare linee guida adeguate per superarlo e ottenere un reale miglioramento delle pratiche relative all'uso delle prove INVALSI nella scuola.

7. Conclusioni

L'insieme delle ricerche presentate mostra, a nostro avviso, che il patrimonio di informazioni restituito sulla scuola italiana dalle prove INVALSI di matematica può costituire una risorsa preziosa per i ricercatori, una volta che questi materiali e questi risultati siano interpretati criticamente nel contesto specifico e con adeguati strumenti teorici. Può inoltre diventare uno strumento importante per la formazione degli insegnanti.

Lo sviluppo delle ricerche condotte in questi anni ha portato a sviluppare disegni di ricerca *mixed-methods* variamente articolati, con piani sperimentali aventi come elemento comune la presenza di analisi qualitativa di macro-fenomeni emergenti dai risultati delle prove INVALSI. I quadri teorici di riferimento hanno richiesto, di volta in volta, interazioni di vario tipo tra teorie. I risultati hanno poi portato a quantificare, in certe situazioni e nello specifico contesto della scuola italiana, quanto le teorie descrivono. In altri casi, i fenomeni studiati hanno permesso di mettere a confronto su casi concreti teorie diverse.

Tutto questo insieme di ricerche e riflessioni, costantemente ancorate al contesto, è poi confluito in un modello di formazione insegnanti e nella costruzione di specifici strumenti per la formazione e la riflessione sulle proprie pratiche didattiche.

Nel Seminario, necessariamente, abbiamo potuto presentare solo alcune di queste nostre ricerche. È rimasto ad esempio fuori tutto il tema, importante dal punto di vista teorico e cruciale per come tocca la sensibilità e l'esperienza degli insegnanti, di come le prove INVALSI abbiamo stimolato la riflessione sulla formulazione dei quesiti di matematica e in particolare sul ruolo della lingua, sia nel testo dei quesiti che nelle risposte degli allievi. Un campo di ricerca molto attivo anche a livello internazionale, e che nel caso italiano ha permesso anche di "mettere a fuoco" molte specificità dei materiali e delle pratiche nostrane. Questo filone di ricerca ha anche portato ad attivare collaborazioni molto proficue e reciprocamente arricchenti con linguisti e ricercatori in didattica dell'Italiano (tra cui ricordiamo in particolare Matteo Viale dell'Università di Bologna).

I risultati ottenuti hanno indicato anche molte piste di ricerca che stiamo esplorando: dall'utilizzo diretto in classe in funzione formativa delle Prove INVALSI, all'analisi di questioni specifiche come il ruolo del contesto in un *task* matematico (Giberti, Canalini, Cascella & Bolondi, 2021); dal cambiamento che il passaggio delle prove INVALSI al *Computer-Based-Testing* sta generando (Ferretti, 2020c; Lemmo, 2019, 2021), alla comprensione delle informazioni fornite dall'introduzione dei *livelli* (con la loro relazione con le nuove modalità di valutazione nella scuola primaria); da una riflessione sul quadro di riferimento INVALSI come strumento per la valutazione scritta di classe (Lemmo, 2023) al contributo che analisi basate su tecniche di *eye-tracking* possono dare all'esplorazione delle dinamiche di *problem-solving* (Gambini et al., 2022), o all'approfondimento di come tutto questo sta cambiando l'*identity* degli insegnanti italiani (Ferretti, Del Zozzo & Santi, 2020; Ferretti, Santi, Del Zozzo, Garzetti & Bolondi, 2021).

Crediamo anche, con un pizzico di orgoglio, che il nostro lavoro abbia contributo in questi anni a innescare una dinamica in cui molti insegnanti si sono impegnati in attività di riflessione e di ricerca attraverso il confronto con le prove, autonomamente, nei propri istituti, in collaborazione con gruppi di ricercatori universitari. Ne è testimonianza la presenza sempre più ampia in Convegni come *Incontri con la Matematica* di interventi di insegnanti in cui le prove Invalsi sono utilizzate come un elemento che arricchisce il lavoro in classe. Va ricordato anche soprattutto il ciclo di Seminari *I dati Invalsi*, organizzato dall'Istituto, in cui ogni anno vengono presentati decine di studi, esperienze, raccolte di dati e materiali, e in cui la presenza di insegnanti-ricercatori è sempre più massiccia.

Speriamo comunque di aver così contribuito, nel nostro piccolo, a innescare un circolo virtuoso di miglioramento, in cui ricerca scientifica, riflessione pratica, esperienza degli insegnanti interagiscono con le azioni istituzionali (il Sistema Nazionale di Valutazione). Collaborare all'azione di miglioramento delle scuole e degli insegnanti italiani era e resta la nostra principale ambizione.

Bibliografia

Adler, J. (2000). Conceptualising resources as a theme for teacher education. *Journal of Mathematics Teacher Education*, *3*(3), 205-224.

Aldon, G., Arzarello, F., Cusi, A., Garuti, R., Martignone, F., Robutti, O., Sabena, C. & Soury-Lavergne, S. (2013). The meta-didactical transposition: a model for analysing teachers education programs. In Lindmeier, A. M. & Heinze, A. (Eds.). *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, 97-124). Kiel, Germany: PME.

Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics. *For the learning of Mathematics*, 14(3), 24-35.

Arzarello, F., Bazzini, L., & Chiappini, G. (2001). A model for analysing algebraic processes of thinking. In R. Sutherland, T. Rojano, A. Bell, & R. Lins (Eds.), *Perspectives on school algebra*, 61-82. Dordrecht, The Nederlands: Kluwer.

Arzarello, F., & Ferretti, F. (2021). The connection between the mathematics INVALSI test and the teaching practices: an explorative study. *I dati INVALSI: uno strumento per la ricerca.* FrancoAngeli.

Arzarello, F., Robutti, O., Sabena, C., Cusi, A., Garuti, R., Malara, N., Martignone, F. (2014). Meta-didactical transposition: A theoretical model for teacher education programmes. In A. Clark-Wilson, O. Robutti, & N.Sinclair (Eds), *The Mathematics Teacher in the Digital Era*, 347-372. Germany, Dordrecht: Springer.

Balacheff N. (1988). Le contrat et la coutume: deux registres des interactions didactiques. In: Laborde C. (ed.) (1988). *Actes du premier colloque franco–allemand de didactique des mathématiques et de l'informatique*. Grenoble: La Pensée Sauvage. 15–26.

Banchelli, S., Garuti, R., & Nolli, N. (2021). Apprendimento a spirale in Matematica: alcuni spunti dalle prove INVALSI. In A cura di Patrizia Falzetti, *I dati INVALSI come strumento per migliorare la didattica della Matematica: IV Seminario" I dati INVALSI: uno strumento per la ricerca e per la didattica"*, 27-36, Franco Angeli.

Barana, A., & Marchisio, M. (2020). Le prove INVALSI per lo sviluppo di competenze matematiche e di problem solving. In P. Falzetti (a cura di) *Il dato INVALSI nella didattica delle discipline- II Seminario "I dati INVALSI: uno strumento per la ricerca"* (pp. 29-50). Milano: Franco Angeli.

Barana, A., & Marchisio, M. (2021). Dalle formule ai modelli. Un percorso interattivo con le domande INVALSI. In P. Falzetti (a cura di) *I dati INVALSI per indagare e migliorare l'insegnamento della matematica. III Seminario "I dati INVALSI: uno strumento per la ricerca"* (pp.9-26). Milano: Franco Angeli.

Barbaranelli, C., & Natali, E. (2005). I Test Psicologici. *Teorie e Modelli Psicometrici.* Roma: Carocci.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, 34, 163-175.

Baruk S. (1985). *L'âge du capitain.* Paris: Seuil.

Bell, K. N., & Norwood, K. (2007). Gender equity intersects with mathematics and technology: Problem solving education for changing times. In D. Sadker & E. S. Silber (Eds.), *Gender in the classroom* (pp. 225–258). LEA.

Bolondi, G., Branchetti, L., Ferretti, F., Lemmo, A., Maffia, A., Martignone, F., Matteucci, M. L., Mignani, S. & Santi, G. R. P. (2016). Un approccio longitudinale per l'analisi delle prove INVALSI di matematica: Cosa si può dire degli studenti in difficoltà. In *Concorso di idee per la ricerca: progetto Sistema informativo integrato I-3-FSE-2009-1: PON FSE Competenze per lo sviluppo: Convenzione MIUR 24/04/2009* (pp. 81-102). Cleup.

Bolondi, G., & Cascella, C. (2020). A mixed approach to interpret large-scale assessment psychometric results of the learning of mathematics. *La Matematica e la sua didattica,* vol. 28 (2).

Bolondi, G., Cascella, C., & Giberti, C. (2017). Highlights on gender gap from Italian standardized assessment in mathematics. In J. Novotnà & H. Moravà (Eds.), *Diversity in Mathematics Education. Proceedings of the International Symposium Elementary Maths Teaching*, SEMT 17 (pp. 81-90). Universita Karlova Press.

Bolondi, G., & Ferretti, F. (2019). L'algebra tra virtuosismi didattici e perdita di senso. Riflessioni su due domande INVALSI. *Nuova Secondaria, XXXVII*(3), 72-77.

Bolondi, G., & Ferretti, F. (2021). Quantifying Solid Findings in Mathematics Education: Loss of Meaning for Algebraic Symbols. *International Journal of Innovation in Science and Mathematics Education*, *29*(1).

Bolondi, G., Ferretti, F., & Gambini, A. (2017). Il database Gestinv delle prove standardizzate INVALSI: Uno strumento per la ricerca. In P. Falzetti (Ed.), *I dati INVALSI: Uno strumento per la ricerca* (pp. 33–42). Milano: FrancoAngeli.

Bolondi, G., Ferretti, F., & Giberti, C. (2018). Didactic contract as a key to interpreting gender differences in maths. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, (18), 415-435.

Bolondi, G., Ferretti, F., & Maffia, A. (2020). Monomials and polynomials: the long march towards a definition. *Teaching Mathematics and its Applications*, 39(1), 1-12.

Bolondi, G., Ferretti, F., & Santi, G. (2019). National standardized tests database implemented as a research methodology in mathematics education. The case of algebraic powers. In: U. T. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education*. Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.

Bolondi, G., Ferretti, F., & Spagnolo, C. (2021). Argomentare in Matematica: analisi di protocolli di studenti su catene di quesiti INVALSI proposti in diversi gradi scolastici. In *I dati INVALSI: uno strumento per lo sviluppo delle competenze trasversali: III Seminario" I dati INVALSI: uno strumento per la ricerca",* a cura di Patrizia Falzetti, 50-69, Franco Angeli.

Bolondi G., Ferretti F., & Spagnuolo A. (2016). Le prove INVALSI con Geogebra: trasformare la valutazione standardizzata in valutazione formativa, in Robutti O. (a cura di), *La formazione docenti con Geogebra. Atti del IV GeoGebra Italian Day 2014*, Ledizioni, Milano: 99-108.

Bosch, M., Dreyfus, T., Primi, C., & Shiel, G. (2017). *Solid findings in mathematics education: What are they and what are they good for?* In T. Dooley & G. Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education - CERME10,* (Vol.1, pp.53-56). Dublin, Ireland: DCU Institute of Education and ERME.

Bowker, G.C. & Star, S.L. (1999). *Sorting things out: classification and its consequences*. MIT Press, Cambridge, MA.

Brousseau, G. (1986). Fondements et méthodes de la didactique des mathématiques. *Recherches en Didactique des Mathématiques,* 7, 2, 33-115.

Brousseau, G. (1988). Le contrat didactique: le milieu [The didactic contract: The milieu]. *Recherches en Didactique des Mathématiques*, 9(3), 309–336.

Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Dordrecht: Kluwer Academic Publishers.

Camara, W. (2013). Defining and measuring college and career readiness: A validation framework. *Educational Measurement: Issues and Practice*, *32*(4), 16-27.

Campbell, C., & Levin, B. (2009). Using data to support educational improvement. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, *21*(1), 47-65.

Capraro, M. M., & Joffrion, H. (2006). Algebraic Equations: Can Middle-School Students Meaningfully translate from Words to Mathematical Symbols? *Reading Psychology, 27*, 147-164.

Carnoy, M. (2015). *International Test Score Comparisons and Educational Policy.* A Review of the Critiques. Boulder, CO: National Education Pol- icy Center. Retrieved from http://nepc.colorado.edu/files/pb_carnoy_international_test_scores_0.pdf

Carotenuto, G., Di Martino, P., & Lemmi, M. (2021). Students' suspension of sense making in problem solving. *ZDM–Mathematics Education, 53*, 817-830.

Carrillo-Yañez, J., Climent, N., Montes, M., Contreras, L. C., Flores-Medrano, E., Escudero-Ávila, D., Vasco, D., Nielka, R., Pablo, F., Álvaro, A. G., Ribeiro, M., Muñoz-Catalán, M. C. (2018). The mathematics teacher's specialised knowledge (MTSK) model. *Research in Mathematics Education*, *20*(3), 236-253.

Cascella, C., Giberti, C., & Bolondi, G. (2020). An analysis of Differential Item Functioning on INVALSI tests, designed to explore gender gap in mathematical tasks. *Studies in Educational Evaluation*, vol. 64, I

Chevallard, Y. (1985). La transposition didactique du savoir savant au savoir enseigné. Grenoble: La Pensée Sauvage.

Chevallard, Y. (1992). A theoretical approach to curricula. *Journal fuer Mathematik-didaktik*, 13(2-3), 215-230.

Chevallard, Y. (1999). El análisis de las prácticas docentes en la teoría antropológica de lo didáctico. *Recherches en didactique des mathématiques*, 19(2), 221-266.

Ciani, A., Vannini, I. (2017). Equità e didattica. Validazione di scale sulle convinzioni di insegnamento democratico, *Cadmo*, 25(2), pp. 5–32.

Cochran-Smith, M. (2001). Learning to Teach Against the (New) Grain. *Journal of Teacher Education, 52*(1), 3–4.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). London, UK: Sage.

D'Amore B. (1999). *Elementi di didattica della matematica*. Bologna: Pitagora.

D'Amore, B. (2007). Epistemologia, didattica della matematica e pratiche d'insegnamento. *La matematica e la sua didattica*, *21*(3), 347-369.

D'Amore, B. (2008). Epistemology, didactics of mathematics and teaching practices. *Mediterranean Journal of Research in Mathematics Education*, 7(1), 1-22.

D'Amore, B., Fandiño-Pinilla, M. I., Marazzani, I., & Sarrazy, B. (2010). *Didattica della matematica alcuni effetti del "contratto"* [Mathematics education some effects of the "contract"]. Bologna, Italy: Archetipolibri.

D'Amore, B., & Sandri, P. (1998). Risposte degli allievi a problemi di tipo scolastico standard con un dato mancante. *La Matematica e la sua didattica, (1)*, 4-18.

De Lange, J. (2007). Large-scale assessment and mathematics education. *Second handbook of research on mathematics teaching and learning*, *2*, 1111-1144.

De Vleeschouwer, M., & Gueudet, G. (2011). Secondary-tertiary transition and evolutions of didactic contract: The example of duality in linear algebra. *In Seventh Congress of the European Society of Research on Mathematics Education* (pp. 2113-2122). University of Rzesvow, Poland.

Di Martino, P., & Baccaglini-Frank, A. (2017). Beyond performance results: Analyzing the informational and developmental potentials of standardized mathematics tests. *For the Learning of Mathematics*, *37*(3), 39-44.

Di Martino, P., & Gregorio, F. (2019). The mathematical crisis in secondary–tertiary transition. *International Journal of Science and Mathematics Education*, *17*(4), 825-843.

Di Martino, P., Gregorio, F., & Iannone, P. (2022). The transition from school to university in mathematics education research: new trends and ideas from a systematic literature review. *Educational Studies in Mathematics*, 1-28.

Di Martino, P., & Signorini, G. (2019). Beyond the standardized assessment of mathematical problem solving competencies: From products to processes. *Mathematical Problem Solving: Current Themes, Trends, and Research*, 209-229.

Doig, B. (2006). Large-scale mathematics assessment: looking globally to act locally. *Assessment in Education: Principles, Policy & Practice, 13*(3), 265–288.

Dreyfus, T. (2017). What are solid findings in mathematics education? In T. Dooley & G. Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10)*, (Vol.1, pp.57-62). Dublin, Ireland: DCU Institute of Education and ERME.

DuFour, R. (2004). What is a" professional learning community"?. *Educational leadership*, 61(8), 6-11.

Duval, R. (1993). Registres de représentations sémiotique et fonctionnement cognitif de la pensée. *Annales de Didactique et de Sciences Cognitives, 5*(1), 37–65.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational studies in mathematics*, *61*(1), 103-131.

Duval, R. (2017). *Understanding the mathematical way of thinking: The registers of semiotic representations*. Springer International Publishing.

Elia, I., Van Den Heuvel-Panhuizen, M., & Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM*, 41, 605-618.

EMS (2011). "Solid Findings" in Mathematics Education. *Newsletter of the European Mathematical Society*, *81*, 46-49.

EMS-EC (Education Committee of the EMS) (2012). What are the Reciprocal Expectations between Teacher and Students? Solid Findings in Mathematics Education on Didactical Contract. *Newsletter of the European Mathematical Society*, 84, 53-55.

EMS-EC (Education Committee of the EMS) (2013). Sociomathematical Norms: In Search of the Normative Aspects of Mathematical Discussions. *Newsletter of the European Mathematical Society*, 88, 59-61.

Eurydice, I. U. I. (2016). La valutazione delle scuole in Europa: politiche e approcci in alcuni Paesi europei. *Firenze, Eurydice Italia*, 83.

Faggiano, E., Ferretti, F., & Arzarello, F. (2022). How do primary teachers interpret and use standardized assessment: the case of the crochet placemats. In Hodgen, J., Geraniou, E., Bolondi, G. & Ferretti, F. (Eds.), *Proceedings of the Twelfth Congress of European Research in Mathematics Education (CERME12*). Bozen-Bolzano, Italy: ERME/Free University of Bozen-Bolzano.

Faggiano, E., Monaco, A., Rizzo, O. G. & Vaccaro, V. (in press). An exploratory study on the connection between INVALSI assessment and mathematics teaching and learning processes at the Primary School level. *I dati INVALSI: uno strumento per la ricerca*.

Ferrara, F., Ferrari, G., Robutti, O., Contini, D., & Di Tommaso, M. L. (2021). When gender matters: A study of gender differences in mathematics. In M. Inprasitha, N. Changsri, N. Boonsena (Eds.), *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 255-263). Khon Kaen, Thailand: IGPME

Ferrara, F., Ferrari, G., & Savioli, K. (2021). Sviluppare competenze sul senso del grafico. In *I dati INVALSI per indagare e migliorare l'insegnamento della matematica: III Seminario "I dati INVALSI: uno strumento per la ricerca e per la didattica"*, a cura di Patrizia Falzetti, 44-62, Franco Angeli.

Ferrara, F., Gilardi, M., & Savioli, K. (2021). "Continua oltre la figura": come un quadrilatero diventa un triangolo. In *I dati INVALSI come strumento per migliorare la didattica della Matematica: IV Seminario" I dati INVALSI: uno strumento per la ricerca e per la didattica"*, a cura di Patrizia Falzetti, 85-99, Franco Angeli.

Ferrara, F., & Pozio, S. (2021). Risposte errate e livelli di abilità degli studenti del grado 8 nelle prove nazionali di matematica. In B. D'Amore (ed.), *La didattica della matematica: riflessioni teoriche e proposte concrete. Incontri con la matematica n.35* (pp. 131-132). Bologna: Pitagora.

Ferretti, F. (2015). *L'effetto "Età della Terra". Contratto didattico e principi regolativi dell'azione degli studenti in matematica* (Doctoral thesis), Alma Mater Studiorum Università di Bologna. Retrieved from http://amsdottorato.unibo.it/7213/4/Ferretti_Federica_Tesi.pdf https://doi.org/10.6092/unibo/amsdottorato/7213.

Ferretti F (2020a). Valutazioni standardizzate di matematica: evidenze di difficoltà dalla scuola primaria alla scuola secondaria di secondo grado [Standardized assessment in Mathematics: highlights of difficulties from primary school to upper secondary school]. *Didattica della matematica*, 7, 36- 49.

Ferretti, F. (2020b). The Manipulation of Algebraic Expressions: Deepening of a Widespread Difficulties and New Characterizations. *International Electronic Journal of Mathematics Education, 15*(1), em0548. https://doi.org/10.29333/iejme/5884

Ferretti F (2020c). Le prove INVALSI CBT: riflessioni sulle prove della scuola secondaria di primo grado. *L'insegnamento della matematica e delle scienze integrate*, 43, 34-54.

Ferretti, F., Bolondi, G. (2019). This cannot be the result! The didactic phenomenon "the Age of the Earth". *International Journal of Mathematical Education in Science and Technology. 52*(2), 194-207.

Ferretti, F., Del Zozzo, A., & Santi, G. (2020). La didattica della matematica a distanza ai tempi del Covid-19 e la sua interazione con l'identità docente. *Annali online della Didattica e della Formazione Docente*, *12*(20), 84-108.

Ferretti, F., Funghi, S. & Martignone, F. (2020). How standardised tests impact on teacher practices: an exploratory study of teachers' beliefs. In C. Andrà. D. Brunetto & F. Martignone (Eds.), *Theorizing and Measuring Affect in Mathematics Teaching and Learning*, 139-146.

Ferretti, F. & Gambini, A. (2017). A vertical analysis of difficulties in mathematics by secondary school to level; some evidence stems from standardized assessment. In T. Dooley & G. Gueudet (Eds.), *Proc. of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 3492- 3499). Dublin, Ireland: DCU Institute of

Education and ERME.

Ferretti, F., Gambini, A., & Santi, G. (2020). The Gestinv Database: A Tool for Enhancing Teachers Professional Development within a Community of Inquiry. In H. Borko and D. Potari (Eds.), *Proceedings of the Twenty-fifth ICMI Study School Teachers of mathematics working and learning in collaborative groups* (pp.621–628). Portugal: University of Lisbon.

Ferretti, F., & Giberti, C. (2021). The properties of powers: Didactic contract and Gender Gap. *International Journal of Science and Mathematics Education*, *19*(8), 1717-1735.

Ferretti, F., Giberti, C., & Lemmo, A. (2018). The Didactic Contract to interpret some statistical evidence in mathematics standardized assessment tests. *EURASIA Journal of Mathematics, Science and Technology Education, 14*(7), 2895-2906.

Ferretti, F., Giberti, C., & Lemmo, A. (2020). I costrutti di didattica della matematica come chiavi di lettura di alcune evidenze statistiche nelle prove invalsi. In Falzetti P. (Ed.). *"I dati INVALSI: uno strumento per la ricerca*", November 17, 2017, Firenze. Milano: FrancoAngeli.

Ferretti, F., Lemmo, A., & Martignone, F. (2018a). Attained curriculum and external assessment in Italy: how to reflect on them. *Proceedings of the Twenty-fourth ICMI Study School Mathematics Curriculum Reforms: Challenges, Changes and Opportunities*, 381-388.

Ferretti, F., Lemmo, A. & Martignone, F. (2018b). La probabilità nelle prove INVALSI: analisi in verticale. *Induzioni*, *55*, 27–47.

Ferretti, F., Martignone, F., & Santi, G. (2022). Analysis of standardized tests and pre-service teacher education: reflections on developed teachers' specialized knowledge. In Hodgen, J., Geraniou, E., Bolondi,G. & Ferretti, F. (Eds.), *Proceedings of the Twelfth Congress of European Research in Mathematics Education (CERME12*). Bozen-Bolzano, Italy: ERME/Free University of Bozen-Bolzano.

Ferretti, F., Michael-Chrysanthou, P., & Vannini, I. (2018). *Formative assessment for mathematics teaching and learning: Teacher professional development research by videoanalysis methodologies*. FrancoAngeli.

Ferretti, F., & Santi, G: (in press). Pre-service teachers' subjectification: a dialogue between LSA and mathematics education. *For the learning of Mathematics*.

Ferretti, F., Santi, G., & Bolondi, G. (2022). Interpreting difficulties in the learning of algebraic inequalities, as an emerging macro-phenomenon in Large Scale Assessment. *Research in Mathematics Education*, 1-23.

Ferretti, F., Santi, G., Del Zozzo, A., Garzetti, M., & Bolondi, G. (2021). Assessment Practices and Beliefs: Teachers' Perspectives on Assessment during Long Distance Learning. *Education Sciences*, vol. 11(6).

Ferretti, F., Vannini, I., & Arzarello, A. (under review). La formazione degli insegnanti di matematica e le valutazioni standardizzate: primi risultati di un progetto di ricerca nazionale interdisciplinare.

Ferri, F. & Martignone, F. (2018). Analisi di quesiti Invalsi per la scuola primaria. In "Verso le prove Invalsi", *Supplemento alla La vita Scolastica 6*, (pp.1-24). GiuntiScuola.

Gambini, A. (2018). *Working Paper - I test di ingresso universitari di Matematica: confronto con il framework delle prove di matematica INVALSI.*

Gambini, A., Banchelli, S., Nolli N. (2020). Analisi verticale del concetto di pendenza: dalla scuola secondaria di primo grado all'università. In P. Falzetti (a cura di), in *Il dato nella didattica delle discipline* (pp. 184-200). Milano: FrancoAngeli Editore.

Gambini, A., Desimoni, M., & Ferretti, F. (2022). Predictive tools for university performance: an explorative study. *International Journal of Mathematical Education in Science and Technology,* 1-27.

Gambini, A., Spagnolo, C., Capone, R., Ferretti, F., & Casalvieri, C. (2022). Which strategies do students adopt when reading a text containing a mathematical problem? Initial results of an exploratory study with eye-tracking. *Quaderni di Ricerca in Didattica,* pp. 1-13.

García, F.J., Gascón, J., Ruiz Higueras, L., & Bosch, M. (2006). Mathematical modelling as a tool for the connection of school mathematics. *ZDM, 38*(3), 226-246.

Garuti, R., & Boero, P. (1994). Mathematical modelling of the elongation of a spring: given a double length spring. In *Proc. of PME-XVIII*, vol. 2, 384–391. Lisbon, Portugal: IGPME.

Garuti, R., Lasorsa, C., & Pozio, S. (2017). The Italian national education assessment system: building mathematics items. In Dooley, T., & Gueudet, G. (Eds.), *Proc. of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 3545–3552). Dublin, Ireland: DCU Institute of Education and ERME.

Garuti, R., & Martignone, F. (2015). The SNV (INVALSI) experience. Teaching and learning mathematics: resources and obstacles. *Proc. of CIEAEM 67, Quaderni di ricerca didattica, 25*(2), 9598.

Garuti, R. & Martignone, F. (2019). Assessment and argumentation: an analysis of mathematics standardized items. In: U. T. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education*, (pp. 4075-4082). Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.

Gerofsky, S. (1996). A linguistic and narrative view of word problems in mathematics education. *For The Learning of Mathematics, 16* (2), 36-45.

Giberti, C. (2018). Differenze di genere e misconcezioni nell'operare con le percentuali: Evidenze dalle prove INVALSI [Gender differences and misconceptions in working with percentages: Evidence from INVALSI tests]. *CADMO, 2018*(2), 97–114.

Giberti, C. (2019). Gegamath, Gender Gap in Mathematics achievement in Südtirol. Bolzano: UniBZ.

Giberti, C., Canalini, R., Cascella, C., & Bolondi, G. (2021). L'influenza del contesto sulla risoluzione dei problemi. *L'insegnamento della Matematica e delle scienze integrate*, *44A*(1), 37-70.

Giberti, C., & Spagnolo, C. (2021). Educazione matematica e differenze di genere: capire il presente per cambiare il futuro. In *Disparità di genere: processi identitari, dinamiche interpersonali e cornici socioculturali*, a cura di Roberta Di Pasquale. Lubrina Editore, Bergamo, 53-81.

Giberti, C., Zivelonghi, A., & Bolondi, G. (2016). The role of metaphors in interpreting students' difficulties in operating with percentages: A mixed method study based on large scale assessment. In C. Csikos, A. Rausch & J. Szitanyi (Eds.), *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education* (p. 275). Szeged, Hungary: IGPME.

González de San Román, A., & De La Rica, S. (2012). Gender gaps in PISA test scores: The impact of social norms and the mother's transmission of role attitudes. *IZA Discussion Paper No. 6338.*

Graziani, I., & Babini, S. (2018). Analisi di errori su area e perimetro in alcuni quesiti Invalsi. *EDiMaST: Experiences of Teaching with Mathematics, Sciences and Technology*, 4, 609-622.

Grossman, P. L., Smagorinsky, P., & Valencia, S. (1999). Appropriating tools for teaching English: A theoretical framework for research on learning to teach. *American journal of Education*, *108*(1), 1-29.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science, 320*(5880), 1164-1165.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and teaching*, *8*(3), 381-391.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (Vol. 2). Sage.

Hart, L.C., Smith, S.Z., Swars, S. L., & Smith, M.E. (2009). An Examination of Research Methods in Mathematics Education (1995-2005). *Journal of Mixed Methods Research, 3*(1), 26-41.

Herbert, J., & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of applied developmental Psychology, 26*(3), 276-295.

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. Washington, D.C: American Association of University Women.

Impedovo, M., Orlandoni, A., & Paola, D. (2011). *Quaderni SNV N. 1-MAT. Guida sintetica alla lettura della prova di Matematica*. Frascati: Invalsi.

INVALSI (2011). Rilevazioni nazionali degli apprendimenti 2011-2012. Rapporto tecnico. Retrieved from https://www.invalsi.it/snv2012/documenti/Rapporti/Rapporto_tecnico_SNV2012.pdf

INVALSI (2017). Rilevazioni nazionali degli apprendimenti 2016-2017. Rapporto tecnico. Retrieved from https://www.invalsi.it/invalsi/doc_eventi/2017/Rapporto_tecnico_SNV_2017.pdf

INVALSI (2018). Quadro di riferimento delle prove di INVALSI di matematica. Retrieved from https://invalsi-areaprove.cineca.it/docs/file/QdR_MATEMATICA.pdf

Jacobs, J. E., & Bleeker, M. M. (2004). Girls' and boys' developing interests in math and science: Do parents matter?. *New directions for child and adolescent development*, (106), 5-21.

Jaworski, B. (2006). Theory and practice in mathematics teaching development: Critical inquiry as a mode of learning in teaching. *Journal of mathematics teacher education*, *9*(2), 187-211.

Jaworski, B. (2014). Unifying complexity in mathematics teaching-learning development: A theory-practice dialectic. *Transforming mathematics instruction: Multiple approaches and practices*, 439-457.

Jaworski, B., & Goodchild, S. (2006). Inquiry community in an activity theory frame. In *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 353-360). Prague, Czech Republic: IGPME.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher*, *33*(7), 14-26.

Kieran, C. (1992). The learning and teaching of school algebra. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp.390-419). New York: MacMillan.

Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: a project of the National Council of Teachers of Mathematics*, (Vol. 2, pp. 707-762). Charlotte, NC: Information Age Pub.

Leder, G. C. (1992). Mathematics and gender: Changing perspectives. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 597–622). Macmillan.

Leder, G., & Forgasz, H. (2008). Mathematics education: New perspectives on gender. *ZDM – The International Journal on Mathematics Education*, 40(4), 513–518.

Lemmo, A. (2023) Le prove INVALSI come strumento di riflessione per una valutazione multidimensionale in matematica. *Nuova Secondaria*, n. 5, gennaio 2023 - anno XL -, 88-93.

Lemmo, A. (2021). A Tool for Comparing Mathematics Tasks from Paper-Based and Digital Environments. *International Journal of Science and Mathematics Education, 19(8)*, 1655 - 1675.

Lemmo, A. (2019) Il passaggio alla somministrazione informatizzata delle prove INVALSI di matematica: l'analisi di alcuni quesiti. *L'insegnamento della matematica e delle scienze integrate* 42, pp.323.

Lemmo, A., Branchetti, L., Ferretti, F., Maffia, A., & Martignone, F. (2015). Students' difficulties dealing with number line: a qualitative analysis of a question from national standardized assessment", Quaderni di Ricerca in Didattica (Mathematics). In *CIEAEM: 20-24 Luglio* (Vol. 25, No. Supplement 2, pp. 143-150). GRIM.

Looney J.W. (2011). Integrating Formative and Summative Assessment: Progress Toward a Seamless System?, *OECD Education Working Papers, 58,* OECD Publishing.

Martignone, F. (2016). Un'attività di formazione per insegnanti di scuola secondaria di primo grado: analisi di prove Invalsi di matematica. *Form@ re*, *16*(1).

Martignone, F. (2017). Analysis of mathematics standardized tests: examples of tasks for teachers. In Dooley, T., & Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 3344–3351). Dublin, Ireland: DCU Institute of Education and ERME.

Meinck, S., Neuschmidt, O., Taneva, M. (2017). Workshop Theme: "Use of Educational Large-Scale Assessment Data for Research on Mathematics Didactics". In: Kaiser, G. (eds) *Proceedings of the 13th International Congress on Mathematical Education.* ICME-13 Monographs. Springer, Cham.

MIUR - Ufficio Statistica e Studi (2017). Gli immatricolati nell'a.a. 2016/2017 il passaggio dalla scuola all'università dei diplomati nel 2016. http://ustat.miur.it/media/1116/notiziario-statistico-2017-1.pdf.

Mellone, M., Ribeiro, M., Jakobsen, A., Carotenuto, G., Romano, P., & Pacelli, T. (2020). Mathematics teachers' interpretative knowledge of students' errors and non-standard reasoning. *Research in Mathematics Education, 22*(2), 154-167.

Morris, A. (2011). *Student Standardised Testing: Current Practices in OECD Countries and a Literature Review*. OECD Education Working Papers, No. 65, OECD Publishing.

Mullis, I. V., Martin, M. O., Foy, P., & Hooper, M. (2016). TIMSS 2015 international results in science. C*hestnut Hill: TIMSS & PIRLS International Study Centre, Boston College*.

Ni, Y. & Zhou, Y-D. (2005). Teaching and learning fraction and rational numbers: the origins and implications of whole number bias. *Educational Psychologist*, 40, 27–52.

Niss, M. (2012). Models and Modelling in Mathematics Education. *Newsletter of the European Mathematical Society, 86*, 49-52.

OECD (2015). *The ABC of Gender Equality in Education: aptitude, behaviour, confidence.* OECD Publishing.

OECD (2016a). PISA 2015 – *Results (volume I): Excellence and equity in education.* OECD Publishing.

OECD (2016b). ITALY – Country note – *Results from PISA 2015*. OECD Publishing.

Ostrom, E. (2005). Doing institutional analysis digging deeper than markets and hierarchies. *Handbook of new institutional economics*, 819-848.

Pitta-Pantazi D. (2014). Number Teaching and Learning. In: Lerman S. (Ed.) *Encyclopedia of Mathematics Education*. Springer, Dordrecht.

Pozio, S., & Bolondi, G. (2019). Difficulties in formulating a geometric situation algebraically; hints from a large-scale assessment. In: M. Graven, H. Venkat, A. Essien & P. Vale (Eds.) *PME 43. Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp 225-232). Pretoria, South Africa: IGPME.

Primi, C. (2017). Solid findings in mathematics education: A psychometric approach. In T. Dooley & G. Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education-CERME10*, (Vol.1, pp.63-67). Dublin, Ireland: DCU Institute of Education and ERME.

Radford, L. (2003). Gestures, speech, and the sprouting of signs: A semiotic-cultural approach to students' types of generalization. *Mathematical thinking and learning*, *5*(1), 37-70.

Radford, L. (2005). Why do gestures matter? Gestures as semiotic means of objectification. *In Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 143-145).

Radford, L. (2006). Elements of a cultural theory of objectification. In L. Radford, & B. D'Amore (Eds.), *Semiotics, culture and mathematical thinking [special issue]. Revista Latinoamericana de Investigación en Matemática Educativa* (RELIME), 103–129.

Radford, L. (2008). Connecting theories in mathematics education: Challenges and possibilities. *ZDM*, 40(2), 317-327.

Radford, L. (2021). *The theory of objectification: A Vygotskian perspective on knowing and becoming in mathematics teaching and learning.* Brill.

Rasch, G. (1960). *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: University of Chicago Press.

Rasmussen, C., Zandieh, M., & Wawro, M. (2009). How do you know which way the arrows go? The emergence and brokering of a classroom mathematics practice. In Roth WM (Ed.) *Mathematical representations at the interface of the body and culture*, (pp. 171–218). Charlotte: Information Age Publishing.

Robutti, O. (2020). Meta-didactical Transposition. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (pp. 611–619). Cham: Springer.

Robutti, O., Aldon, G., Cusi, A., Olsher, S., Panero, M., Cooper, J., Carante, P., & Prodromou, T. (2019). Boundary objects in mathematics education and their role across communities of teachers and researchers in interaction.

In G. Lloyd & O. Chapman (Eds.), *International handbook of mathematics teacher education* (Vol. 3, pp. 211–240). Rotterdam: SensePublishers.

Santi, G., Ferretti, F., & Martignone, F. (2023). Mathematics Teachers Specialised Knowledge and Gestinv Database. *VI Seminario "I dati INVALSI: uno strumento per la ricerca"*.

Sarrazy, B. (1995). Le contrat didactique. Revue française de pédagogie, 112(85), 118.

Saxe, G.B., Shaughnessy M.M., Shannon A., Langer-Osuna J., Chinn R., Gearhart M. (2007), Learning about fractions as points on a number line, in P. Elliot, W.G. Martin, M.E. Strutchens (Eds.), *NCTM Sixty-ninth Yearbook*. Reston: NCTM.

Sbaragli, S. (2005). Misconcezioni inevitabili e misconcezioni evitabili. *La matematica e la sua didattica*, 1, 57-71.

Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(2), 107-131.

Sensevy, G. (2010). Outline of a joint action theory in didactics. *In Proceedings of CERME 6* (Vol. 6, pp. 1645-54).

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coins. *Educational Studies in Mathematics.* 22, 1–36.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing.* Cambridge university press.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, 15(2), 4-14.

Sierpinska, A., Bobos, G., & Knipping, C. (2008). Sources of students' frustration in pre-university level, prerequisite mathematics courses. *Instructional Science*, 36, 289-320.

Skoumpourdi, C. (2010). The number line: An auxiliary means or an obstacle. *International Journal for Mathematics Teaching and Learning*, 270.

Spagnolo, C., Capone, R., Ferretti, F., & Gambini, A. (2021). Quali strategie adottano gli studenti per leggere il testo di un problema? Primi risultati di uno studio di eye-tracking con quesiti INVALSI e OCSE-PISA. In P. Falzetti (a cura di), *I dati INVALSI come strumento per migliorare e valutare le competenze trasversali. IV Seminario "I dati INVALSI: uno strumento per la ricerca"* (pp. 68-99). Milano: Franco Angeli.

Spagnolo, C., Capone, R., & Gambini, A. (2021). Where do students focus their attention on solving mathematical tasks? An eye tracker explorative study. Inprasitha, M., Changsri, N., & Boonsena, N. (Eds). (2021). *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol.4). p. 89-96. Khon Kaen, Thailand: IGPME.

Spagnolo, C., Vaccaro, V., Faggiano, E. (in press). An exploratory study on the connection between teachers' training and meta-didactical conflict. *VII Seminario I dati INVALSI: uno strumento per la ricerca*.

Stavy, R. & Tirosh, D. (1996). Intuitive rules in science and mathematics: the case of 'More of A More of B'. *International Journal for Science Education, 18*(6), 553-667.

Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., Silver, E., Ufer, S., Vos, P. (2016). *Assessment in mathematics education: Large-scale assessment and classroom assessment*. Springer Nature.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. New York: Sage Publishing.

Tomasetto, C. (2013). Matematica per i maschi, italiano per le femmine: Stereotipi di genere e atteggiamenti verso le materie scolastiche tra genitori e figli. *IN-MIND ITALIA*, 5, 19-24.

Truffelli, E., Asquini G., (2022). Concezioni della valutazione e dell'apprendimento in insegnanti di matematica nella scuola primaria. *CADMO*, 1, 51-66.

Truffelli, E., Vannini, I. (2021). Convinzioni e atteggiamenti degli insegnanti di scuola primaria italiani e orientamento all'uso formativo delle prove INVALSI di matematica. in L. Lucisano (ed), *Ricerca e Didattica per promuovere intelligenza comprensione e partecipazione*. Atti del X convegno della SIRD, aprile 2021, (pp.376-395). Lecce, Rovato: Pensa Multimedia, Lecce.

Vaccaro, V., Faggiano, E. & Ferretti, F. (2021). Consapevolezza degli insegnanti delle ragioni degli errori degli studenti in matematica: uno studio esplorativo nella scuola primaria., in L. Lucisano (ed), *Ricerca e Didattica per promuovere intelligenza comprensione e partecipazione. Atti del X convegno della SIRD,* aprile 2021, (pp. 411-430). Lecce, Rovato: Pensa Multimedia.

Wake, G., Foster, C., & Swan, M. (2013). A theoretical lens on lesson study: professional learning across boundaries. In A. Lindmeier, & A. Heinze (Eds.), *Proceedings of the 37th conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 369–376).

Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems thinker*, 9(5), 2-3.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for research in mathematics education, 27*(4), 458-477.

Zan, R. (2012). La dimensione narrativa di un problema: il modello C&D per l'analisi e la (ri) formulazione del testo. *L'insegnamento della matematica e delle scienze integrate, 35*(2), 107-126.

Zan, R. (2016). *I problemi di matematica: difficoltà di comprensione e formulazione del testo*. Roma: Carocci.