

18
December 2018

Gaetano Domenici

Editoriale / *Editorial*

«Comportamento insegnante» e sviluppo del pensiero scientifico 11
(The Attitude that it Teaches and the Development of Scientific Thought)

STUDI E CONTRIBUTI DI RICERCA

STUDIES AND RESEARCH CONTRIBUTIONS

Paola Ricchiardi - Federica Emanuel

Soft Skill Assessment in Higher Education 21
(Valutare le soft skill in Università)

Gamal Cerda Etchepare - Carlos Pérez Wilson

Karina Pabón Ponce - Verónica León Ron

Análisis de los esquemas de razonamiento formal 55
en estudiantes de Educación Secundaria Chilenos
mediante la validación del Test of Logical Thinking (TOLT)

*(Formal Reasoning Schemes Analysis in Chilean Secondary Education
Students through the Validation of the Test of Logical Thinking - TOLT)*

*(Analisi degli schemi di ragionamento formale degli studenti
della Scuola Secondaria cilena attraverso la validazione
del Test del Pensiero Logico - TOLT)*

- Laura Occhini*
Orientamento universitario in entrata: misurare l'efficacia 75
(*University Incoming Orientation: Measure Forcefulness*)
- Giulia Bartolini - Giorgio Bolondi - Alice Lemmo*
Valutare l'apprendimento strategico: uno studio empirico 99
per l'elaborazione di uno strumento
(*Evaluating Strategic Learning: An Empirical Study for the Elaboration of an Instrument*)
- Kenneth T. Wang - Tatiana M. Permyakova*
Marina S. Sheveleva - Emily E. Camp
Perfectionism as a Predictor of Anxiety in Foreign Language 127
Classrooms among Russian College Students
(*Il perfezionismo come predittore di ansia nei corsi di lingua straniera per studenti universitari russi*)
- Li-Ming Chen - Li-Chun Wang - Yu-Hsien Sung*
Teachers' Recognition of School Bullying According 147
to Background Variables and Type of Bullying
(*Riconoscimento da parte degli insegnanti del bullismo scolastico in relazione a variabili di sfondo e tipo di bullismo*)
- Laura Girelli - Fabio Alivernini - Sergio Salvatore*
Mauro Cozzolino - Maurizio Sibilio - Fabio Lucidi
Affrontare i primi esami: motivazione, supporto all'autonomia 165
e percezione di controllo predicono il rendimento degli studenti
universitari del primo anno
(*Coping with the First Exams: Motivation, Autonomy Support and Perceived Control Predict the Performance of First-year University Students*)
- Nicoletta Balzaretto - Ira Vannini*
Promuovere la qualità della didattica universitaria. 187
La Formative Educational Evaluation in uno studio pilota
dell'Ateneo bolognese
(*Promoting Quality Teaching in Higher Education. A Formative Educational Evaluation Approach in a Pilot Study at Bologna University*)
- Emanuela Botta*
Costruzione di una banca di item per la stima dell'abilità 215
in matematica con prove adattative multilivello
(*Development of an Item Bank for Mathematics Skill Estimation with Multistage Adaptive Tests*)
-

<i>Rosa Cera - Carlo Cristini - Alessandro Antonietti</i> Conceptions of Learning, Well-being, and Creativity in Older Adults	241
<i>(Concezioni dell'apprendimento, benessere e creatività negli anziani)</i>	
<i>Marta Pellegrini - Giuliano Vivanet - Roberto Trincherro</i> Gli indici di effect size nella ricerca educativa. Analisi comparativa e significatività pratica	275
<i>(Indexes of Effect Sizes in Educational Research. Comparative Analysis and Practical Significance)</i>	
<i>Antonio Calvani - Roberto Trincherro - Giuliano Vivanet</i> Nuovi orizzonti della ricerca scientifica in educazione. Raccordare ricerca e decisione didattica: il Manifesto S.Ap.I.E.	311
<i>(New Horizons for Scientific Research in Education. Linking Research and Educational Decision: The Manifesto S.Ap.I.E.)</i>	
<i>Giusi Castellana</i> Validazione e standardizzazione del questionario «Dimmi come leggi». Il questionario per misurare le strategie di lettura nella scuola secondaria di primo grado	341
<i>(Validation and Standardization of the Questionnaire «Tell Me How You Read». The Questionnaire on Reading Strategies in the Lower Secondary School)</i>	
<i>Laura Menichetti</i> Valutare la capacità di riassumere. Il Summarizing Test, uno strumento per la scuola primaria	369
<i>(Evaluating Summarizing Skills. The Summarizing Test, a Tool for Primary School)</i>	

NOTE DI RICERCA

RESEARCH NOTES

<i>Elsa M. Bruni</i> La valutazione vista da lontano: lo sguardo della pedagogia generale (II)	399
<i>(Evaluation Viewed from a Distance: The Vision of General Pedagogy - II)</i>	
<i>Giorgio Bolondi - Federica Ferretti - Chiara Giberti</i> Didactic Contract as a Key to Interpreting Gender Differences in Maths	415
<i>(Il contratto didattico come una chiave di lettura per interpretare le differenze di genere in matematica)</i>	

<i>Elisa Cavicchiolo - Fabio Alivernini</i> The Effect of Classroom Composition and Size on Learning Outcomes for Italian and Immigrant Students in High School <i>(L'impatto della composizione e della dimensione della classe sugli apprendimenti degli studenti italiani e immigrati nella scuola secondaria di secondo grado)</i>	437
<i>Marta Pellegrini - Lucia Donata Nepi - Andrea Peru</i> Effects of Logical Verbal Training on Abstract Reasoning: Evidence from a Pilot Study <i>(Effetti di un training logico verbale sulle capacità di ragionamento astratto: risultanze da uno studio pilota)</i>	449
<i>Massimiliano Smeriglio</i> Porta Futuro Lazio: l'innovazione possibile nel servizio pubblico per lo sviluppo dell'occupabilità in ottica lifelong learning <i>(Porta Futuro Lazio: A Possible Public Service Innovation for Employability's Development in a Lifelong Learning View)</i>	459
<i>Giorgio Asquini</i> Osservare la didattica in aula. Un'esperienza nella scuola secondaria di I grado <i>(Classroom Observation. A Study in Lower Secondary School)</i>	481
COMMENTI, RIFLESSIONI, PRESENTAZIONI, RESOCONTI, DIBATTITI, INTERVISTE COMMENTS, REFLECTIONS, PRESENTATIONS, REPORTS, DEBATES, INTERVIEWS	
<i>Antonio Calvani</i> Per un nuovo dibattito in campo educativo <i>(For a New Debate in the Educational Field)</i>	497
<i>Journal of Educational, Cultural and Psychological Studies</i> Notiziario / News	503
Author Guidelines	505

Valutare la capacità di riassumere

Il Summarizing Test, uno strumento per la scuola primaria *

Laura Menichetti

Università degli Studi di Firenze, Department of Education and Psychology (Italy)

DOI: <http://dx.doi.org/10.7358/ecps-2018-018-meni>

laura.menichetti@unifi.it

EVALUATING SUMMARIZING SKILLS. THE SUMMARIZING TEST, A TOOL FOR PRIMARY SCHOOL

ABSTRACT

The synthesis of the texts is recognised as one of the main processes that contribute to the Reading Comprehension, however in Italian scientific literature (with reference to the context in which the test is to be applied) there is no widespread evidence of suitable tools to evaluate the achievement and possession of this specific competence, nor to evolve and improve it in children. The Summarizing Test, developed within a program aiming at evaluate the effectiveness of the Reciprocal Teaching at primary school, tries to fill this gap. The test validation, realized with data collected from fourth grades in primary schools geographically distributed throughout the Country, gives a glimpse of a promising tool, that could integrate the other reading comprehension tools already available. This paper describes the results of the item analysis carried out in accordance with the Classical Test Theory, and of the split half to allow the usage of two equivalent versions for the pre-test and post-test in the experimental programs.

* Il Summarizing Test è stato ideato dal prof. A. Calvani per un intervento sperimentale all'interno dei programmi S.Ap.I.E. Desidero ringraziare sentitamente, oltre al professore per aver consentito l'uso dello strumento e per i suggerimenti ricevuti, la prof.ssa L. Cajola che coordina il programma di ricerca RC-RT, la prof.ssa R. Cardarello con la quale sono in corso ulteriori approfondimenti, nonché le dott.sse A. Rizzo e M. Traversetti (Università Roma Tre) e i dott. C. Bertolini, A. Pintus, A. Vezzani (Università di Modena e Reggio Emilia) per i preziosi contatti con le scuole.

Keywords: Item analysis; Primary school; Reading comprehension; Summarizing; Test validation.

1. COMPRENDERE E RIASSUMERE UN TESTO

1.1. *Leggere e comprendere un testo*

La competenza alfabetica funzionale, considerata dal Consiglio dell'Unione Europea (2018) tra le competenze chiave per il lifelong learning, «presuppone la conoscenza della lettura e della scrittura e una buona comprensione delle informazioni scritte». In contrasto con tale raccomandazione, le rilevazioni internazionali riferiscono che in Italia un quindicenne su quattro raggiunge solo il livello 2 della scala OCSE, sui 7 previsti, collocandosi così significativamente sotto la media dei Paesi sviluppati, e anche il trend evolutivo non mostra segni di miglioramento (OCSE, 2015).

Nel novero delle indagini internazionali, la reading literacy già nel 2000 ha costituito il focus della prima rilevazione del Programme for International Student Assessment (PISA) rivolto ai quindicenni: «Reading literacy is understanding, using, and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society» (OCSE, 1999). Negli anni successivi la definizione di reading literacy ha incluso l'engagement in reading, sottolineando la motivazione e la consapevolezza dei soggetti che, in quanto competenti, considerano la lettura rilevante e coinvolgente (OCSE, 2009), mentre il Programma nelle revisioni più recenti, oltre a introdurre forme di digital reading, ha preso in considerazione anche l'interpretazione e la valutazione dei contenuti (OCSE, 2018).

Agli studenti, in effetti, è richiesto di mostrare un certo livello di competenza nei diversi aspetti che concorrono alla comprensione del testo – individuare le informazioni rilevanti, comprendere il significato generale, ravvisare l'interpretazione dell'autore, riflettere sul contenuto e valutarlo, riflettere sulla forma e valutarla (Langer, 1995) – perché tali competenze sono state identificate come rilevanti e strettamente funzionali per conseguire apprendimenti (Dole *et al.*, 1991; Pressley & Woloshyn, 1995).

La comprensione del testo è un'attività complessa, all'interno della quale entrano in gioco diversi processi cognitivi e linguistici, studiati fin dai precursori della psicologia cognitiva (Bartlett, 1932; Piaget, 1936) e in particolare oggetto di analisi più approfondita dalla metà degli anni Settanta-

ta (Bransford & Johnson, 1973; Schank & Abelson, 1977; Kintsch & Van Dijk, 1978).

Al di là della diversità dei modelli elaborati, la ricerca scientifica concorda sul ravvisare nella comprensione del testo alcuni momenti caratteristici che si dispongono secondo un ordine crescente di complessità, partendo da operazioni di maggiore aderenza alla superficie del testo, per completarsi secondo meccanismi di graduale progressiva astrazione.

La lettura e la comprensione presuppongono le capacità di interazione dinamica e di integrazione tra le informazioni dedotte via via dal testo e gli schemi e i copioni cognitivi presenti nella mente del lettore (processi bottom-up), ma implicano anche, viceversa, l'applicazione al contenuto del testo degli schemi mentali già posseduti e delle situazioni già conosciute per formulare ipotesi e andare alla ricerca di conferme (processi top-down).

La fase iniziale è rivolta alla decodifica e alla identificazione delle unità lessicali e i primi processi restano ancorati alla superficie del testo, alla ricerca di informazioni esplicite, eventualmente da rintracciare facendo appello alla conoscenza di piccole variazioni sinonimiche (Van den Broek *et al.*, 2005). Nelle fasi successive la mente cerca di dar luogo a rappresentazioni proposizionali dotate di senso sulla base dei legami che si possono costituire tra frasi adiacenti, processo che porta alla formulazione di microstrutture (Van Dijk & Kintsch, 1983; Kintsch & Van Dijk, 1978; Levorato 1988; 2000). Avanzando con la lettura, questi segmenti sono via via riesaminati nello sforzo di selezionare le informazioni principali, immagazzinarle in memoria, raccordarle con quadri più complessi già presenti nella mente del lettore (schemi o copioni), al fine di generare macrostrutture che la mente riconosce come similari a sezioni di realtà o a situazioni già conosciute (Johnson-Laird, [1983] 1988; Garnham & Oakhill, 1996). Solo le ultime fasi consentono la messa in relazione con altre fonti, la valutazione dell'attendibilità dell'informazione, il riconoscimento del registro comunicativo, fino a cogliere – con un massimo del livello di astrazione – il messaggio che l'autore vuol trasmettere.

Riuscire a selezionare le informazioni più importanti in un segmento di testo dipende dalla capacità di cogliere i nessi tra frasi adiacenti e soprattutto di inferirli dove esistano lacune o elementi di incoerenza, eventualmente includendo fatti ante/post/paralleli all'azione descritta. I processi cognitivi che caratterizzano la comprensione passano così attraverso ipotesi o inferenze cognitive che vengono via via avanzate e confrontate con le rappresentazioni situazionali presenti in memoria, alla ricerca di una composizione coerente e di un affinamento del senso. Questo processo si avvale del sostegno fornito dalla flessibilità propria del linguaggio; la mente ha infatti bisogno di fissare e trattenere il significato con piccole frasi sintetiche che

hanno un'importante funzione ergonomica, perché evitano che la memoria di lavoro sia travolta da un sovraccarico informativo e fanno da sostegno nella progressiva riformulazione complessiva del senso che si viene costruendo e che è sottoposto a continue verifiche di coerenza formale e semantica.

1.2. *Riassumere un testo*

All'interno dell'attività di comprensione di un testo, è da sempre riconosciuta di grande importanza la capacità di *riassumere*, cioè di formulare linguisticamente un secondo testo, più breve, che non si limiti ad escludere alcune frasi, ma che ristruttururi il contenuto e che incorpori il senso complessivo del testo originario, attraverso il modello della trasformazione delle microstrutture in macrostrutture o secondo altri modelli (Brown & Day, 1983; Brown, Day, & Jones, 1983; Winograd, 1984; Benvenuto, 1987; Anderson & Hidi, 1988; Lumbelli, 2009).

Nell'ambito delle scienze cognitive è proprio sul riassumere che si è giocata la sfida dell'intelligenza artificiale e anzi si è generata la sua crisi negli anni Ottanta, fino a quando le tecnologie non sono state in grado di cominciare a gestire il semantic web, con le sue più recenti evoluzioni del Natural Language Programming e soprattutto del Natural Language Understanding, che costituiscono una delle maggiori sfide dell'informatica contemporanea (Schank, 1992; Allen, 1995; Berners Lee *et al.*, 2001; Green, 2012). Nel tentativo di implementare programmi informatici capaci di riassumere il testo si è arrivati anche a scoprire la rilevanza dell'enciclopedia di riferimento del lettore, con i significativi concetti, appunto, di script o copione (Schank & Abelson, 1977; McKoon & Ratcliff, 1992; Schank, 1992).

Sul piano dell'intervento didattico, oggi i modelli più rilevanti ed efficaci di educazione alla lettura come il Reciprocal Teaching (Palincsar & Brown, 1984; Brown & Palincsar, 1987) hanno incorporato il summarizing come una delle quattro strategie di base – questioning, clarifying, summarizing, predicting – e, in genere, tutti i metodi che si occupano di sviluppare abilità di studio considerano la presa di appunti e il riassumere come attività cognitive fondamentali. Anzi, il summarizing può entrare a far parte anche delle strategie metacognitive perché, come indicato da Palincsar e Brown (1984) nell'ambito delle pratiche suggerite per l'educazione alla lettura, chiedere agli studenti di riassumere un testo significa non solo indurli a focalizzare l'attenzione sui contenuti principali tralasciando i dettagli, ma anche sensibilizzarli ad interrogarsi e ad autovalutarsi circa la comprensione del testo.

1.3. *Le prove di comprensione*

In Italia sono già diffusi diversi strumenti standardizzati per valutare la capacità dei bambini della scuola primaria di comprendere i testi letti. Se ne sono esaminati alcuni per decidere se utilizzarli nell'ambito del progetto sperimentale di Reciprocal Teaching (in corso di attuazione), che è rivolto a soggetti di 8-10 anni.

Tra le più note si possono indicare le Prove INVALSI (<http://www.invalsi.it>), che fanno parte delle rilevazioni obbligatorie al termine della quinta primaria: esse, avvalendosi di brani di media lunghezza, prevedono quesiti a scelta multipla sui significati letterali e figurati di espressioni polisemiche, sulla ricerca di informazioni esplicite nel testo, sulla coerenza interna, sull'efficacia espressiva e sulle scelte stilistiche, ma soprattutto richiedono numerose inferenze dirette (partendo dal testo) e inverse (fornendo l'informazione e chiedendo da quale frase del testo è inferita).

In ambito scolastico sono molto diffuse anche risorse per la valutazione delle abilità di lettura realizzate dal prof. C. Cornoldi e dal Gruppo MT. Adeguate a bambini e ragazzi dal primo anno della scuola primaria fino al biennio della scuola secondaria di secondo grado, sono prove centrate prima di tutto sul ritrovamento di informazioni all'interno del brano e quindi orientate alla comprensione generale, più che su forme di parafrasi e di sintesi. Sono pensate come strumento di programmazione per l'intera classe, per l'identificazione precoce di alunni con basso livello di competenza e a rischio di disturbi specifici dell'apprendimento e nelle versioni più recenti sono affiancate da indicazioni per il potenziamento (Cornoldi, Colpo, & Carretti, 2017). Ulteriori materiali riabilitativi progettati dagli stessi ricercatori sono disponibili a scopo diagnostico, e per il trattamento individuale specialistico, come supporto al recupero.

L'Assessment di Lettura e Comprensione per l'Età evolutiva (prove ALCE) è invece un test che bilancia comprensione locale, comprensione globale e anche espressione del bambino, attraverso domande orali a risposta aperta relativamente a brani narrativi e descrittivi letti o ascoltati; è però meno adatto ad un'applicazione su larga scala per valutazione di interventi didattici, ed è stato pensato, invece, per l'utilizzo da parte di psicologi, logopedisti, neuropsichiatri, nell'ambito di protocolli diagnostici e durante il trattamento di disturbi specifici dell'apprendimento o per funzionamento intellettivo limite o in caso di ritardi mentali lievi (Bonifacci *et al.*, 2014).

È utile ricordare, per l'attenzione che è necessario prestare al curriculum verticale, che anche per la scuola secondaria di primo grado esistono prove standardizzate, con domande a scelta multipla o aperte, orientate a valutare la capacità di interpretare diverse tipologie di testo scritto e insieme

a fornire spunti operativi e riflessioni utili per i processi di insegnamento (in Giovannini & Ghetti, 2016, ad esempio, si sviluppano prove rivolte al primo anno della scuola secondaria di primo grado).

Nessuno dei test fin qui esaminati, però, appare specificatamente disegnato per mettere a fuoco la capacità di riassumere e per essere usato come strumento di valutazione a supporto di interventi didattici orientati a questo scopo. È possibile semmai ipotizzare l'uso combinato del Summarizing Test con queste prove validate, per misurare la correlazione tra le diverse dimensioni caratteristiche della comprensione del testo.

Nella letteratura internazionale, viceversa, sono documentati studi sperimentali, condotti a vari livelli scolastici e universitari, relativi alla competenza nel sintetizzare testi, prevalentemente realizzati con strumenti di tipo qualitativo (Friend, 2001; Pečjak & Pirc, 2018).

Nel caso del programma italiano è stato preferito però uno strumento a risposte chiuse, più rapidamente fruibile da parte di ricercatori e scuole. La scelta che ha orientato ad un questionario a struttura chiusa, rispetto ad altre soluzioni notoriamente più seguite nella letteratura scientifica con prova di sintesi di natura aperta, è legata alla necessità di disporre di uno strumento che possa dare con immediatezza risultati riferiti ad applicazioni su larga scala. Su campioni più limitati, o per alcuni specifici aspetti, è in corso anche un'attività di correlazione con strumenti diversi.

2. LE CARATTERISTICHE DEL SUMMARIZING TEST

Il Summarizing Test è uno strumento che ha come scopo la valutazione della capacità del bambino di riconoscere tra varie formulazioni quelle che offrono una migliore qualità di sintesi di un testo narrativo o di altra tipologia funzionale (Werlich, 1975): la prova è stata progettata ad hoc, per verificare l'efficacia di un programma formativo basato sul Reciprocal Teaching (in corso di attuazione), per essere utilizzata in classi di quarta primaria.

Nella sua globalità la prova è costituita da otto testi, ciascuno di circa 700-1000 caratteri (spazi inclusi), privi di particolari difficoltà per quanto riguarda la terminologia usata e le situazioni rappresentate. Per non sovraccaricare e stancare i bambini si stima che sia adeguato proporre non oltre quattro brani per volta, disponendo complessivamente il test in due sessioni in giorni diversi (consecutive, oppure distanziate dagli interventi didattici se in modalità pre/post).

Per ogni brano il bambino deve leggere il testo da solo, in silenzio, e poi rispondere ad una serie di domande scritte.

La struttura è quella di un questionario con risposte a scelta multipla, in cui l'alunno deve selezionare tra le proposte formulate quelle che appaiono le migliori: ogni quesito presenta infatti sei alternative di risposta, tutte con un certo grado di accettabilità, e chiede al bambino di segnare le tre che più delle altre risultano confacenti a rappresentare in maniera sintetica il brano letto, quelle che vanno oltre il semplice riconoscimento letterale e la corrispondenza dei termini e incorporano nel massimo grado possibile il senso complessivo del testo, evitando di includere, o di mettere a fuoco oltre misura, dettagli solo apparentemente rilevanti.

Per ogni brano il questionario presenta le stesse tre domande, che intercettano gli indicatori ritenuti fondamentali, corrispondenti alla concettualizzazione del costrutto di summarizing adottata:

- individuare le informazioni più importanti contenute nel brano proposto, quelle che possono essere identificate da chi sa cogliere il senso complessivo del testo, eliminando dettagli, particolarità, specificazioni, pervenendo a macroproposizioni in grado di coprire tutti gli elementi di rilievo;
- trovare i titoli più adeguati, cioè quelli in grado di rappresentare la situazione chiave, l'intero brano e non solo una sua parte, di mantenere il focus sull'azione principale;
- trovare le parole chiave, cioè delle triplete di parole che nel loro insieme possono servire a rappresentare l'intero testo.

Al bambino si chiede quindi di trovare gli elementi rilevanti per produrre la sua sintesi¹.

Prima della somministrazione ai bambini, il test è stato sottoposto ad una validazione da parte di adulti esperti² per rilevare

- l'adeguatezza del contenuto; il test deve essere rappresentativo della capacità di riassumere e in maniera economica cogliere gli aspetti rilevanti del processo;
- la comprensibilità; il test deve prevedere contenuti, linguaggio, registro, grafica, e ogni altra componente sintonizzati per un target di bambini di quarta primaria³;

¹ Il programma di Reciprocal Teaching, inizialmente accompagnato da un processo di thinking aloud del docente, segue un percorso analogo. In aggiunta, il programma prevede che i bambini, lavorando a coppie, talvolta si esercitino anche a modificare o integrare sintesi avviate da altri, fino ad ottenere elaborati che corrispondono alla loro migliore sintesi.

² Tre docenti universitari dell'area 11 (Scienze storiche, filosofiche, pedagogiche e psicologiche) e tre docenti di scuola primaria.

³ Nei brani si rappresentano contesti con cui bambini di scuola primaria hanno familiarità e nelle classi in cui il test è stato validato è stato verificato che il vocabolario usato non costituisce un ostacolo alla interpretazione del testo. Sono state effettuate prove preliminari

- la sensibilità; il test deve riuscire a discriminare le differenze tra le performance ottenute dai vari allievi;
- l'affidabilità; il test deve essere ripetibile in vista di applicazioni successive a distanza di tempo, con una ragionevole garanzia di continuare ad ottenere gli stessi risultati a fronte di livelli di competenza invariati.

La prova è stata quindi progressivamente affinata, e considerata accettabile solo quando i valutatori esperti hanno espresso il loro giudizio positivo e hanno dato risposte allineate. Tra le soluzioni scelte dai valutatori si è accettata una discordanza non superiore ad una sola risposta, e con il vincolo che la risposta discordante non fosse sempre la stessa.

L'artificio di presentare più brani, di contenuti e tipologie diverse, ha lo scopo di consentire un range di punteggio sufficientemente ampio e aumentare la discriminatività del test.

Le consegne per chi somministra la prova agli alunni prevedono che prima dell'avvio della prova egli/ella segua una procedura standard: (i) si accerta, eventualmente anche con esempi, che tutti i bambini sappiano gestire un questionario con risposte a scelta multipla; (ii) illustra la struttura del test indicando il numero dei brani, delle risposte possibili, di quelle attese e segnala in particolare che potrebbero esserci diverse risposte plausibili, tra cui il bambino deve selezionare le tre migliori; (iii) fa compilare la prima pagina del fascicolo con i dati anagrafici, utili particolarmente in applicazioni pre/post, avendo cura eventualmente di prendere provvedimenti per la tutela della privacy; (iv) raccomanda la lettura silenziosa, comunica l'impossibilità di porre domande di qualsiasi tipo durante lo svolgimento, e fa iniziare tutti contemporaneamente.

La prova non ha una durata predefinita, anche se il tempo necessario è stato stimato in circa 30 minuti per 4 brani. Quando i 4/5 della classe hanno terminato, il somministratore sollecita gli alunni rimasti a concludere e dopo 3-5 minuti ritira i fascicoli. La modalità di raccolta dei dati è ovviamente un fattore molto delicato di cui preoccuparsi e consegne chiare e dettagliate servono per evitare, o almeno contenere, errori casuali e sistematici.

Nella fase pilota, servita per l'acquisizione dei dati per l'item analysis (try-out) e svoltasi nell'anno scolastico 2017/18, le prove sono state somministrate in forma cartacea e il data entry è stato realizzato secondo un tracciato comune. È allo studio la possibilità di gestire tutto in formato elettronico.

A supporto dell'intervento di ricerca è necessario disporre di due metà *equivalenti* del test, da applicare nelle classi sperimentali all'inizio del pro-

per verificare che i bambini conoscessero sia la terminologia, sia i copioni situazionali di riferimento.

gramma e alla fine degli interventi didattici e nelle classi di controllo negli stessi periodi, raccogliendo informazioni sulle strategie didattiche adottate dai vari docenti. Per quanto riguarda il Summarizing Test è necessario quindi procedere allo split half.

In *Figura 1* è riportato l'intero test. Si indicano direttamente le due versioni consolidate – ST_A e ST_B – ottenute dopo l'item analysis trattata nei paragrafi successivi, che comunque ha indotto modifiche contenute.

I fascicoli forniti agli studenti prevedono font e spaziature diversi da quelli che compaiono in questo contributo, specificatamente pensati per facilitare la lettura. In *Figura 2* sono riportate le griglie di correzione.

ST_A - Primo testo

Luisa ha invitato tutta la classe in un locale per il suo compleanno.

Ha fatto preparare una torta gigantesca ricoperta di cioccolato, panna montata e fiorellini di zucchero oltre a pasticcini di ogni tipo. Tutta la stanza è decorata di palloncini, filamenti appesi, lampioni di carta colorata e scritte di “Buon Compleanno” attaccate ai muri.

Ha chiamato persino un mago per far divertire con alcuni trucchi e due amici più grandi che hanno suonato e cantato. Davvero bravi!

I suoi genitori devono aver speso un sacco di soldi.

1) scegli le TRE informazioni più importanti:

- Ai muri erano attaccate le scritte di “Buon compleanno”
- Nella torta c'erano anche fiorellini di zucchero
- Molti addobbi, cibi ed animatori
- Luisa ha festeggiato il compleanno
- I genitori hanno speso molto per la festa
- Un mago ha presentato alcuni trucchi

2) scegli i TRE titoli migliori:

- Gli animatori alla festa di Luisa
- Un compleanno in grande
- Luisa e i suoi amici
- Le spese dei genitori di Luisa
- Il compleanno di Luisa
- Una festa costosa

3) scegli i TRE gruppi di parole che riassumono meglio il significato dell'intero brano:

- Luisa, mago, trucchi
- Torta, cioccolato, panna
- Compleanno, addobbi, animatori
- Palloncini, mago, cantanti
- Compleanno, festa, divertimento
- Luisa, compleanno, grandi spese

SEGUE 

Figura 1. – Summarizing test nella sua versione finale, dopo l'item analysis.

ST_A - Secondo testo

Alla curva il ciclista si imbatte in una piccola buca. La bicicletta sbanda. Alla fine il ciclista cade per terra e rotola per alcuni metri.

Fa per rialzarsi ma un dolore molto forte alla gamba lo costringe a rimanere fermo. Una signora anziana, uscita a fare la spesa, lo vede e si ferma a soccorrerlo. Un altro passante chiama l'autoambulanza. Passano alcuni minuti mentre il ferito si lamenta dal dolore.

Si sente in lontananza la sirena dell'autoambulanza che a mano a mano si avvicina. Ecco, ci siamo. Scendono i barellieri dall'autoambulanza.

Il ciclista viene con cura portato sulla barella e poi sull'autoambulanza e via, a tutta velocità verso l'ospedale dove, per prima cosa, il medico di guardia richiede una radiografia. Il risultato arriva rapidamente: c'è una frattura alla gamba destra.

1) scegli le TRE informazioni più importanti:

- Il ciclista ha avuto un incidente
- Si sente la sirena dell'autoambulanza
- Una signora uscita a fare la spesa si ferma
- Arrivano i soccorsi
- I barellieri poggiano il ciclista sulla barella
- Il ciclista viene trasportato all'ospedale

2) scegli i TRE titoli migliori:

- Una brutta caduta
- Un intervento di soccorso
- La radiografia
- La sirena dell'autoambulanza
- La velocità dei barellieri
- L'incidente

3) scegli i TRE gruppi di parole che riassumono meglio il significato dell'intero brano:

- Incidente, soccorso, ospedale
- Ciclista, donna anziana, barellieri
- Sirena, dolore, ospedale
- Incidente, assistenza, diagnosi
- Medico di guardia, radiografia, frattura
- Caduta, ospedale, gamba rotta

ST_A - Terzo testo

Il pescatore prese la canna da pesca e si diresse verso il lago. Era una bella giornata di primavera, e questo lo metteva di buon umore.

Uscì speranzoso di fare una buona pesca. Gli uccellini svolazzavano e i prati cominciavano a rivestirsi di fiori. Tutto intorno la natura si stava risvegliando. Mentre camminava il pescatore fischiava allegro con il suo cane che lo accompagnava, anch'esso felice, scodinzolando. Arrivato sul lago cominciò a gettare la canna. Il tempo passava ma non prendeva nulla.

Ad un certo punto la lenza si incagliò sul fondo e si spezzò. Che sfortuna!

Poi il pescatore riuscì a prendere un pesciolino ma così piccolo che gli faceva pena, così lo ributtò nel lago. Di nuovo un altro pesciolino troppo piccolo, ed ancora un altro... e sempre il pescatore lo ributtò nel lago.

Ad un tratto dal lago sente venire una vocina. "Sei stato bravo, hai compiuto una buona azione. Ne sarai ricompensato".

Mentre il pescatore se ne tornava a casa tutto triste per non aver preso nulla, sul sentiero vide qualcosa che luccicava: era una moneta d'oro!

1) scegli le TRE informazioni più importanti:

- I prati si rivestono di fiori
- La lenza si spezza
- Il pescatore fischiava
- Il pescatore ributta nel lago i pesciolini piccoli

- Il pescatore va a pesca
- Il pescatore trova una moneta d'oro

2) scegli i TRE titoli migliori:

- La primavera
- Una buona azione ricompensata
- Il buon pescatore
- Una moneta d'oro
- Una giornata magica
- La storia dei piccoli pesciolini

3) scegli i TRE gruppi di parole che riassumono meglio il significato dell'intero brano:

- Campi fioriti, lenza, pesciolini
- Buona azione, magia, premio
- Primavera, uccellini, pesciolini
- Pesca, lago, lenza rotta
- Pesca, incantesimo, moneta d'oro
- Pesciolini magici, bontà, ricompensa

ST_A - Quarto testo

La neve si manteneva fino all'estate. Mio padre mi accompagnava a scuola con la slitta, nel paese più a valle. Ogni giorno, ci dovevamo alzare alle 6 per essere a scuola in tempo.

Che spettacolo esaltante! Quegli sterminati campi silenziosi e innevati davano un senso di pace. Gli stambecchi si avvicinavano spesso alla casa in cerca di cibo. La casa era grande e d'inverno ospitavamo anche comitive di sciatori.

Io e i miei fratelli più piccoli passavamo gran parte dei pomeriggi intorno casa a giocare liberamente, a fare pupazzi con la neve, ad andare sugli slittini. A volte ci si trovava con gli altri ragazzi delle case più vicine a fare gare di sci. Quando sono diventato grande e sono andato a vivere in città quelle immagini e quella sensazione di pace non mi hanno mai abbandonato.

La vita in città si svolgeva in un piccolo appartamento, in un condominio rumoroso; appena sceso nella strada rimanevi avvolto in un grande ingorgo di auto, stridio di freni e clacson in continuazione.

1) scegli le TRE informazioni più importanti:

- Si stava a lungo a giocare con la neve
- A volte si facevano gare di sci
- Ospitavamo comitive di sciatori
- I campi innevati davano un senso di pace
- Mi ricordo spesso di quelle immagini
- Nella strada rimanevi avvolto in un ingorgo di auto

2) scegli i TRE titoli migliori:

- Campi innevati
- Il ricordo dell'infanzia
- Pupazzi di neve e slittini
- Due fase diverse della mia vita
- Un'infanzia con la natura
- La vita assordante della città

3) scegli i TRE gruppi di parole che riassumono meglio il significato dell'intero brano:

- Ricordo, infanzia, libertà
- Neve, pupazzi, stambecchi
- Giochi, gare di sci, infanzia
- Quiete, città, confronto
- Vita in città, condominio, clacson
- Vita dell'infanzia, vita da adulto, differenze

ST_B - Primo testo

I genitori hanno preparato una festa per la fine della scuola elementare.

Hanno fatto arrivare nell'aula pizzette, gelati, bibite, focacce, dolci a volontà. Tutta la stanza è addobbata con palloncini, filamenti appesi, lampioni di carta colorata, tanti disegni appesi alle pareti ed altri oggetti prodotti dalla classe (sculture in pongo, burattini di stoffa, plastici in legno).

Le maestre hanno fatto un regalino ad ogni bambino. Le maestre hanno anche cantato canzoni divertenti e fatto girotondi coi bambini. Il babbo di Luigi ha suonato la chitarra.

Davvero una festa ben organizzata! I bambini avranno così un buon ricordo della loro scuola.

1) scegli le TRE informazioni più importanti:

- I genitori hanno preparato la festa di fine scuola
- Le canzoni cantate dalle maestre erano divertenti
- Nell'aula c'erano anche lampioni di carta colorata
- Ci sono stati musica, giochi e consegna di regali
- Un babbo ha suonato la chitarra
- La festa si è svolta in un clima di grande allegria

2) scegli i TRE titoli migliori:

- La festa di fine scuola
- I canti e balli di maestre e genitori
- L'addio alla scuola elementare
- Gli oggetti prodotti dalla classe in bella mostra
- Maestre e genitori suonano e cantano
- Una festa piena di allegria

3) scegli i TRE gruppi di parole che riassumono meglio il significato dell'intero brano:

- Pizzette, lampioni colorati, burattini
- Festa della scuola, divertimento, ricordo
- Fine della scuola, festa, regali
- Cibo, disegni, sculture
- Canti, balli, musica
- Scuola, festa, affetto

ST_B - Secondo testo

Le previsioni del tempo non sono buone, c'è allerta in mare, ma Mario decide lo stesso di uscire in mare per fare surf.

Ad un certo punto il vento aumenta di intensità; le onde diventano alte anche più di 2 metri e Mario è spinto sempre più al largo. Per fortuna ha con sé il cellulare e può telefonare per chiamare i soccorsi.

Passano minuti interminabili, infine in lontananza vede la motovedetta della guardia costiera che arriva, sempre più vicina. Ecco ci siamo; la piccola motovedetta raggiunge Mario. Mario si accosta. Viene fatto salire a bordo e riportato a riva.

Se la caverà con un grosso spavento ed una bella multa per non aver rispettato l'avviso meteo.

1) scegli le TRE informazioni più importanti:

- Mario esce in mare a fare surf
- Mario non riesce più a controllare il surf
- Le onde diventano alte anche più di 2 metri
- Mario vede la guardia costiera che si avvicina
- Mario si accosta alla motovedetta
- Mario viene salvato dalla guardia costiera

2) scegli i TRE titoli migliori:

- La motovedetta
- Una brutta avventura

- Una multa meritata
- Il surf
- Mare in tempesta
- L'imprudenza

3) scegli i TRE gruppi di parole che riassumono meglio il significato dell'intero brano:

- Imprudenza, salvataggio, multa
- Surf, allerta meteo, motovedetta
- Rischio, aiuto, salvezza
- Vento, motovedetta, guardia costiera
- Onde, cellulare, ritorno
- Pericolo, soccorso, punizione

ST_B - Terzo testo

C'era una volta un contadino che possedeva un piccolo pezzetto di terra.

In una bella giornata di primavera, mentre gli uccellini svolazzavano felici e i prati si riempivano di fiori, vi scorse una rondine con l'ala spezzata; la sollevò con delicatezza e la portò a casa sua.

La curò a lungo finché la rondine non poté volare.

L'anno dopo, la rondine ritornò e portò con sé un semino che il contadino, incuriosito, piantò tra le zolle. Da quel semino nacque un cocomero che in breve diventò gigantesco.

Il contadino invitò poi gli amici per una festa con cibo, carne e vino in grande quantità. Alla fine il contadino prese il cocomero e lo spezzò. Che sorpresa! Era pieno di monete d'oro. Il contadino divise le monete fra i suoi ospiti.

Finalmente poté vivere senza tanti pensieri.

1) scegli le TRE informazioni più importanti:

- Gli uccellini svolazzavano felici
- Il contadino curò la rondine
- Alla festa con gli amici c'era cibo e vino in grande quantità
- Il cocomero era gigantesco
- Il cocomero era pieno di monete d'oro
- Il contadino può vivere senza pensieri

2) scegli i TRE titoli migliori:

- La primavera
- Una buona azione ricompensata
- Il buon contadino
- La rondine guarita
- La rondine e le monete d'oro
- Il contadino e i suoi amici

3) scegli i TRE gruppi di parole che riassumono meglio il significato dell'intero brano:

- Primavera, rondine ferita, rondine guarita
- Primavera, cocomero, monete d'oro
- Bontà, incantesimo, benessere
- Buona azione, seme magico, ricompensa
- Cicogna, seme, cocomero
- Contadino, bontà, premio

ST_B - Quarto testo

Le abitazioni erano isolate a poche centinaia di metri dal mare.

I nostri genitori dovevano badare alle loro cose, così lasciavano noi ragazzi liberi di muoverci come volevamo; l'unica regola era di tornare a casa all'imbrunire. Gran parte del tempo eravamo in acqua. Ci si tuffava e si usciva dall'acqua di continuo. Si prendevano molluschi e piccoli pesci che mangiavamo sulla spiaggia dopo averli cotti con piccoli legni raccolti qua e là. A volte facevamo delle battaglie tra di noi senza che intervenissero i genitori, lotta e pugni a volontà, poi amici come prima.

Se ci facevamo male, ci disinfettavamo con l'acqua di mare senza tornare a casa. Da grande sono andato a vivere in città, in un appartamento in una strada molto trafficata. Rumori e clacson di continuo. Poi tutta una corsa, da casa al lavoro! Le immagini di quelle spiagge e di quelle giornate in piena libertà mi ritornano spesso in mente.

- 1) scegli le TRE informazioni più importanti:
- Si trascorreva la giornata in piena libertà
 - Si raccoglievano legnetti per cucinare i pesci
 - Ci si tuffava e si usciva dall'acqua di continuo
 - Abitavamo vicino al mare
 - Lotta e pugni a volontà
 - Mi ricordo spesso quelle giornate
- 2) scegli i TRE titoli migliori:
- Una bella spiaggia
 - Il ricordo dell'infanzia
 - Le battaglie tra ragazzi
 - Due mondi del tutto diversi
 - Un'infanzia libera
 - Il caos della vita in città
- 3) scegli i TRE gruppi di parole che riassumono meglio il significato dell'intero brano:
- Mare, pesca, rumori
 - Natura, città, confronto
 - Ricordo, infanzia, libertà
 - Spiaggia, tuffi, traffico
 - Giochi, battaglie, infanzia
 - Infanzia, vita da grande, memoria

Figura 1. – Summarizing test nella sua versione finale, dopo l'item analysis.

ST_A			
Primo testo	C-D-E	B-E-F	C-E-F
Secondo testo	A-D-F	A-B-F	A-D-F
Terzo testo	D-E-F	B-C-E	B-E-F
Quarto testo	A-D-E	B-D-E	A-D-F

ST_B			
Primo testo	A-D-F	A-C-F	B-C-F
Secondo testo	A-B-F	B-C-F	A-C-F
Terzo testo	B-E-F	B-C-E	C-D-F
Quarto testo	A-D-F	B-D-E	B-C-F

Figura 2. – Griglie di correzione.

3. ITEM ANALYSIS DEL SUMMARIZING TEST E SPLIT HALF

3.1. *Dati raccolti e loro preparazione per l'analisi*

I dati utilizzati per la validazione del test sono stati raccolti nel secondo quadrimestre dell'anno scolastico 2017/18 in 11 classi quarte della scuola primaria di Istituti Comprensivi distribuiti sul territorio nazionale (8 al nord e 3 al centro-sud, 4 in capoluoghi di provincia e 7 in centri minori), con i quali le Università facenti parte del progetto avevano già relazioni, ma in cui non era stato precedentemente condotto nessun programma specifico per la comprensione del testo.

Per mettersi in una situazione ottimale dal punto di vista del carico di lavoro, e del tutto analoga a quella che sarà adottata nel programma sperimentale, si è deciso di proporre quattro brani e relativi quesiti nel corso di una giornata e altri quattro in una giornata successiva. Le rilevazioni sono avvenute nell'arco della stessa settimana in modo da non far intercorrere interventi didattici significativi tra le due prove.

Per escludere di rilevare distorsioni dovute alla ripetizione di test identici nella struttura e simili come difficoltà a breve distanza di tempo, in oltre un terzo dei casi è stata invertita la sequenza di applicazione delle due metà (ST_A-ST_B oppure ST_B-ST_A).

In *Figura 3* sono riportati i campi acquisiti⁴. Non tutti i dati codificati nel tracciato record sono indispensabili per l'item analysis, ma questa prima fase di acquisizione è servita anche per testare parte delle procedure organizzative che saranno usate nella fase sperimentale; inoltre, i dati raccolti consentono la possibilità di successive letture differenziate per geografia, per scuola, per genere, o altro.

⁴ Il tracciato è fornito alle scuole e ai gruppi di ricerca dei vari Centri universitari con una legenda esplicativa. Raccogliere le risposte ST_A T1 D2, ad esempio, significa mostrare se il bambino ha scelto oppure no ciascuno dei distrattori a b c e f nella seconda domanda, quella relativa ai titoli ritenuti più appropriati da attribuire al primo testo nella prova ST_A , quello che parla del compleanno di Luisa. Per quanto riguarda le consegne più importanti adottate, ho previsto che per IC e Scuola si usino i codici meccanografici del Ministero, nel campo ID per lo studente si debba scegliere un identificativo che tuteli la privacy ma che possa essere riapplicato a distanza di tempo per le valutazioni pre e post, per i bisogni educativi speciali si distinguano almeno le tre macrotipologie segnalate, per la sequenza di applicazione delle due metà del test si indichi se prima ST_A e poi ST_B oppure viceversa, etc. Poi per ogni testo e per ogni distrattore è richiesta l'immissione esplicita del risultato, in modo tra l'altro da distinguere le risposte non date dalle assenze e dalle mancate tabulazioni.

CITTA'		IC		SCUOLA		CLA SSE	SEZ	PERSONA DI CONTATTO										
								COGNOME	NOME									
DATA STA		MINUTI		DATA STB		MINUTI		SEQUENZA		STUDENTE								
gg-mmm-aa		STA		gg-mmm-aa		STB		AB/BA		ID	GENERE M/F	BES A/B/C						
STA T1 D1			STA T1 D2			STA T1 D3												
a	b	c	d	e	f	a	b	c	d	e	f	a	b	c	d	e	f	... RIPETUTO PER TUTTI I TESTI

Figura 3. – Tracciato dei campi acquisiti.

L'item analysis consiste nell'applicazione di un insieme di tecniche statistiche che consentono di ricavare informazioni circa l'affidabilità di un test nel suo complesso e circa il funzionamento dei singoli item, in modo che non vi siano quesiti troppo facili o troppo difficili e che sia possibile distinguere in maniera più certa possibile coloro che superano la prova dagli altri. In fase di costruzione di un questionario è possibile stabilire in via teorica le difficoltà dei vari item, la progressione secondo cui disporli, i punteggi medi attesi, ma occorre sempre effettuare una prova pilota (try-out) su un campione sufficientemente ampio di soggetti con caratteristiche simili a quelle della popolazione su cui lo strumento dovrà operare, in modo da verificare il corretto funzionamento e riservarsi di applicare piccole eventuali correzioni agli item che dovessero presentare qualche criticità.

Per l'item analysis del Summarizing Test si adotta qui la Teoria Classica dei Test.

Rispetto ai dati raccolti sono stati esclusi per la validazione del test i record relativi alle prove di bambini assenti a una delle due applicazioni, perché non avrebbero contribuito correttamente allo split half, quelli dei bambini che hanno dato più di tre risposte a una o più domande perché non ci sarebbero stati elementi per decidere quali risposte includere e quali escludere, quelli di soggetti con bisogni educativi speciali perché il test non è al momento realizzato per supportare questa casistica⁵. I record da esaminare risultano quindi 174.

La particolarità di aver previsto tre risposte appropriate per ciascuna domanda (e non una sola), induce a considerare – oltre alla prova nel suo complesso – gli indici sia per singolo distrattore sia per intero quesito.

Per quanto riguarda lo split half merita ricordare che la numerosità dei brani proposti consente di migliorare l'attendibilità del test nel suo

⁵ Per i bambini con bisogni educativi speciali (BES) sono in fase di completamento testi adattati mediante semplificazioni e riduzioni, mentre laddove necessario sarà previsto l'uso di misure compensative.

complesso, ma tutti i brani sono orientati ad intercettare un unico costrutto, sono sempre seguiti dalle stesse tipologie di risposta (informazioni, titoli, parole chiave), e sono stati ideati con un livello di difficoltà molto simile tra loro, perciò in via teorica si dovrebbe poter procedere ad un'attribuzione pressoché casuale all'una o all'altra metà test per sorteggio. Nei paragrafi che seguono si provvede a verificare la bontà dello split.

3.2. Risultati conseguiti

Rispetto ad un punteggio massimo teorico di 72 raggiungibile da ciascun bambino sul test completo, i risultati hanno una mediana di 49, che si ritiene adeguata: si colloca oltre la metà del punteggio per non demotivare gli studenti, ma consente ancora molto spazio di crescita per poter applicare un programma di miglioramento (Fig. 4). Analoghe considerazioni possono essere fatte singolarmente per ST_A e ST_B .

	ST totale	ST_A	ST_B
minimo conseguito	25	12	11
massimo conseguito	66	36	33
massimo teorico	72	36	36
media	49,10	24,47	24,63
deviazione standard	7,73	4,50	4,28
mediana	49	25	25
1° quartile	tra 66 e 55	tra 36 e 28	tra 33 e 28
2° quartile	tra 55 e 49	tra 28 e 25	tra 28 e 25
3° quartile	tra 49 e 44	tra 25 e 21	tra 25 e 22
4° quartile max	tra 44 e 25	tra 21 e 12	tra 22 e 11

Figura 4. – Punteggi conseguiti dagli alunni.

3.3. Split half

In Figura 5 si dà una rappresentazione grafica della distribuzione dei punteggi tra ST_A e ST_B .

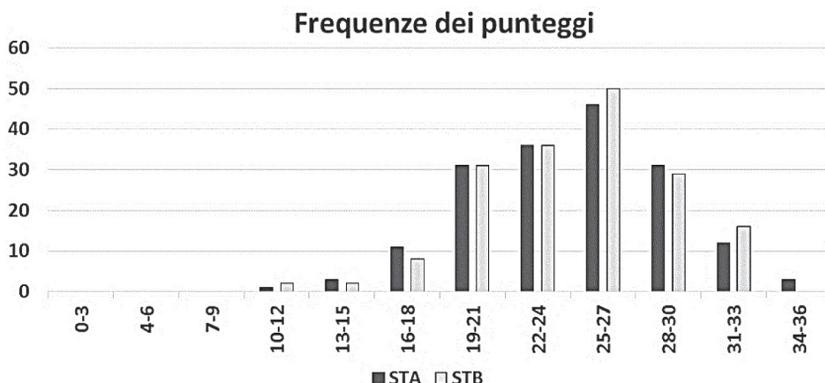


Figura 5. – Distribuzione dei punteggi conseguiti dagli alunni in ST_A e in ST_B .

Visivamente si può apprezzare dalla *Figura 5* che le due distribuzioni sono entrambe monomodali, quasi sovrapponibili (curtosi rispettivamente $-0,95$ e $-0,50$; asimmetria rispettivamente $0,82$ e $0,92$). Dalla *Figura 4* si rileva che media, mediana e deviazione standard sono pressoché le stesse, quindi le due metà ST_A e ST_B possono essere considerate un adeguato split half del Summarizing Test.

In circa $1/3$ dei risultati raccolti (63 vs 111) il test ST_B è stato somministrato prima del test ST_A e questo di per sé annulla l'effetto di addestramento al test che potrebbe influire positivamente sulla seconda applicazione.

3.4. Indice di distrattività

Per avviare l'item analysis, la prima verifica riguarda l'uso dei diversi distrattori: l'indice di distrattività corrisponde alla percentuale degli alunni che sceglie un certo distrattore (*Fig. 6*)⁶.

Se i distrattori fossero tutti plausibili in uguale misura, la presenza di sei alternative per ogni quesito porterebbe il valore teorico di utilizzo di ciascun distrattore a circa il 16%. In *Figura 6* sono segnalati in grigio i distrattori usati meno del 4%, cioè meno del 25% del valore teorico.

Si rileva che nel Summarizing Test tutti i distrattori sono utilizzati e soltanto sette tra loro (il 5%) sono utilizzati in maniera ridotta.

⁶ La somma delle percentuali a b c d e f per ciascuna domanda relativa ad uno stesso brano fa 100 (es. nella prima riga $13 + 7 + 21 + 25 + 17 + 17 = 100$).

	D1: Informazioni importanti						D2: Titolo						D3: Parole chiave					
	a	b	c	d	e	f	a	b	c	d	e	f	a	b	c	d	e	f
STA Test01	13%	7%	21%	25%	17%	17%	5%	29%	14%	2%	31%	19%	11%	8%	27%	12%	25%	18%
STA Test02	30%	4%	19%	16%	5%	26%	33%	26%	6%	3%	1%	31%	29%	15%	11%	13%	12%	21%
STA Test03	7%	16%	2%	30%	15%	30%	5%	31%	26%	14%	19%	5%	13%	26%	9%	17%	20%	15%
STA Test04	22%	12%	14%	23%	16%	14%	18%	26%	12%	17%	19%	9%	25%	12%	24%	10%	9%	20%
STB Test01	31%	7%	5%	24%	8%	25%	32%	3%	27%	2%	3%	31%	7%	29%	24%	7%	13%	20%
STB Test02	25%	11%	23%	10%	5%	26%	4%	27%	17%	11%	25%	16%	27%	17%	18%	7%	9%	22%
STB Test03	6%	29%	11%	9%	26%	19%	6%	29%	21%	15%	24%	5%	20%	19%	9%	28%	2%	21%
STB Test04	29%	8%	18%	13%	9%	22%	14%	27%	10%	20%	22%	7%	11%	15%	27%	10%	19%	17%

Figura 6. – Percentuali di scelta di ciascun distrattore.

3.5. *Indice di facilità*

L'indice di facilità è definito come $IF = X_{medio} / X_{max}$, dove X_{medio} è il punteggio medio ottenuto considerando tutti i rispondenti ad un certo item e X_{max} è il punteggio massimo che in quell'item può essere conseguito in via teorica. Questo indice serve a individuare item non efficaci perché troppo facili o troppo difficili.

Nel caso del Summarizing Test è stato calcolato l'IF per ogni domanda (es. Testo 1– informazioni importanti), ma anche per singolo distrattore (a, b, c, ...). Nel primo caso il punteggio massimo è 3: cioè 1 per ciascuna risposta corretta. Nel secondo caso il punteggio massimo è 1 che risulta attribuito ad ogni scelta appropriata effettuata dal bambino (per i distrattori corrispondenti a scelte corrette si attribuisce 1 quando il bambino li segna come rappresentativi del brano, per i distrattori corrispondenti a scelte errate si attribuisce 1 se il bambino correttamente non li ha segnati).

Il risultato è rappresentato in *Figura 7* per tutti gli otto brani presi in esame. In questa figura a livello di singolo distrattore gli IF sono disposti per ogni brano su due righe, per evidenziare le risposte corrette (di sopra) rispetto a quelle errate (di sotto).

Si assumono come soglie 0,1 (troppo difficile) e 0,9 (troppo facile) e in *Figura 7* risultano segnate in grigio le occasioni al di sotto e al di sopra delle soglie prefissate.

Come si nota, non ci sono item risultati troppo difficili.

A livello di singolo distrattore, quindi di una risposta rispetto alle tre richieste, qualcuna risulta abbastanza facile, ma per oltre il 90% non è questo il caso e infatti a livello di domanda o di intero brano (ultime quattro colonne sulla destra) non risultano item troppo facili.

3.6. *Indice di discriminazione*

L'indice di discriminazione di un item esprime la capacità di contribuzione di un item nel distinguere i soggetti che complessivamente rispondono meglio all'intero test da quelli che complessivamente rispondono peggio.

Per calcolarlo, infatti, si procede ordinando i risultati sulla base del punteggio finale conseguito nel test da ciascun alunno, in modo che nelle righe più in alto compaiano le risposte degli studenti che hanno conseguito i punteggi più alti e nelle righe più in basso quelle degli studenti che hanno complessivamente ottenuto i punteggi più bassi.

D1: Informazioni importanti						D2: Titolo						D3: Parole chiave						D1	D2	D3	Totale
a	b	c	d	e	f	a	b	c	d	e	f	a	b	c	d	e	f				
STA Testo 1																					
	0,36	0,30	0,43		0,28		0,32		0,17	0,13	0,04		0,32	0,26	0,00	0,17	0,19	0,36	0,18	0,20	0,25
0,53	0,32					0,26						0,32									
STA Testo 2																					
	0,09	0,26	0,28	0,15	0,32		0,04	0,28	0,32	0,17	0,09	0,21	0,09	0,40	0,17	0,34	0,43	0,23	0,18	0,30	0,24
STA Testo 3																					
	0,30	-0,04	0,09		0,13		0,17	0,30	0,23	0,28	0,21	0,53	0,34	0,19	0,47	0,34	0,12	0,25	0,38	0,25	
						0,32						0,21	0,34	0,19	0,26	0,40					
STA Testo 4																					
	-0,04		0,36	0,13	0,19		0,32	0,43	0,53	0,26	0,34	0,21	0,47	0,19	0,34	0,26	0,25	0,37	0,32	0,31	
						0,28						0,21	0,47	0,19	0,34	0,26					
STB Testo 1																					
	0,11	0,34	0,17	0,43	0,47		0,04	0,34	0,15	0,17	0,11	0,49	0,26	0,43	0,28	0,30	0,35	0,16	0,38	0,30	
						0,11						0,49	0,26	0,43	0,28	0,30					
STB Testo 2																					
	0,11	0,19	0,36	0,17	0,11		0,06	-0,09	0,21	-0,09	0,40	0,17	0,32	0,23	0,13	0,40	0,23	0,13	0,31	0,22	
						0,21						0,17	0,32	0,23	0,13	0,40					
STB Testo 3																					
	0,21	0,19	0,21	0,43	0,19		0,28	0,23	0,13	0,19	0,17	0,26	0,06	0,09	0,21	0,09	0,23	0,23	0,13	0,20	
						0,30						0,26	0,06	0,09	0,21	0,09					
STB Testo 4																					
	0,17	0,34	0,17	0,43	0,49		0,32	0,43	0,43	0,40	0,26	0,53	0,36	0,15	0,11	0,23	0,33	0,38	0,31	0,34	
						0,40						0,53	0,36	0,15	0,11	0,23					

Figura 8. – Indice di discriminazione per distrattore, per domanda, per brano.

Si segmenta poi l'elenco in 3 gruppi (Kelley, 1939; Feldt, 1961; Wiersma & Jurs, 1990), prendendo in considerazione il 27% dei punteggi più alti, il 46% dei punteggi intermedi, il 27% dei punteggi più bassi e si calcola così l'indice di discriminazione per singolo item:

$$ID = [\Sigma(\text{punteggi alti}) - \Sigma(\text{punteggi bassi})] / p_{\max} / n$$

dove p_{\max} è il massimo punteggio conseguibile e n è il numero dei soggetti che fanno parte dei gruppi superiore e inferiore. In *Figura 8* sono rappresentati i valori ottenuti.

Si definiscono come soglie $-0,1$ e $+0,1$: all'interno dell'intervallo così definito si assume che l'item non discrimini, mentre al di fuori di esso può discriminare in senso positivo (quando il risultato per il singolo item è concorde a quello dell'intero test), o in senso negativo (quando il risultato per il singolo item è in controtendenza rispetto al punteggio complessivo).

Nelle prime colonne si mostrano i risultati per singolo distrattore, nelle ultime a destra quelli per intero quesito o intero brano.

Come si nota, non ci sono item che invertono il senso della discriminazione (avrebbero segno negativo), mentre ci sono alcuni item con un basso potere discriminante.

3.7. Consistenza interna del test

La prova multi-item potenzialmente serve a dare maggiore robustezza ai risultati raccolti, anche laddove si debba misurare un certo costrutto monodimensionale; occorre controllare però che tutti gli item contribuiscano in maniera coerente.

Nella matrice di correlazione item-item non si notano particolari correlazioni tra gruppi di item, se non tra quelli che si riferiscono ad uno stesso brano, evidentemente collegati dal meccanismo di comprensione dello stesso.

Per valutare il questionario nel suo insieme, però, si può usare il coefficiente α di Cronbach (Cronbach, 1951), indice sintetico che verifica la consistenza interna del test, mettendo in relazione le varianze dei singoli item con quella dell'intero test:

$$\alpha = [1 - \Sigma(\text{varianze item}) / \text{varianza test}] N / (N - 1)$$

dove N = numero di item del test.

Sempre per la particolarità delle tre risposte positive per ciascun quesito, si considerano le risposte a ciascuno dei distrattori. I valori ottenuti sono riportati in *Figura 9*.

	ST _A	ST _B	ST _{totale}
alfa di Cronbach	0,83	0,80	0,89

Figura 9. – Alfa di Cronbach per determinare la consistenza interna.

Valori al di sopra di 0,8 esprimono una buona coerenza interna, che quindi qui rileviamo sia per il test nella sua totalità, sia nel caso di split ST_A e ST_B.

Il peso degli errori casuali tende a diminuire al crescere del numero degli item e questo giustifica un valore più alto di consistenza interna quando si considera il test completo, ma l' α si mantiene buono anche in caso di split half.

4. CONCLUSIONI

Tra i vari processi cognitivi e linguistici che caratterizzano la piena comprensione di un testo, la capacità di riassumere è riconosciuta come uno dei più ricchi di significato ma anche dei più complessi, perché richiede di saper iterare un lavoro di decostruzione e ricostruzione del testo attraverso micro e macro proposizioni che in parte possono poggiare sul linguaggio ma che devono fare appello ad un continuo procedere di interazioni tra il testo in esame e gli schemi mentali e i copioni del soggetto.

Saper valutare questa specifica competenza, e in particolare i delta incrementali che caratterizzano i percorsi di apprendimento, costituisce elemento imprescindibile per la scuola primaria. La sfida consiste nel riuscire a operazionalizzare sufficientemente il *riassumere*, in modo da rendere intersoggettivamente valido il test – che come prima fase è stato sottoposto ad un team di esperti – ma nello stesso tempo occorre strutturare la prova con dei gradi di libertà (analoghi ai diversi riassunti possibili), proponendo alternative di risposta che non siano del tutto giuste o sbagliate, ma con gradi diversi di rappresentatività del testo, tra cui dover scegliere le migliori interpretazioni: la prova deve risultare uno strumento a sostegno di un percorso di apprendimento e non lo scopo di esso.

A supporto di un programma sperimentale sul Reciprocal Teaching sviluppato all'interno dell'Associazione S.Ap.I.E., è stato quindi realizzato il Summarizing Test, i cui passi di validazione sono oggetto del presente contributo.

Il risultato dell'item analysis è stato utilizzato per migliorare gli indici di facilità e di discriminazione della prova. Il lavoro è servito anche per

mostrare l'equivalenza delle due metà della prova, ST_A e ST_B , che potranno essere applicate come pre-test e post-test per valutare i miglioramenti conseguenti a interventi didattici. Dal punto di vista metodologico, a livello complessivo del programma sperimentale, si suggerisce comunque di applicare un piano di somministrazione incrociato – per alcune classi ST_A - ST_B e per altre ST_B - ST_A – in modo da evitare che i dati raccolti siano influenzati dalla ripetizione del test.

Il Summarizing Test copre un ambito nuovo nel panorama italiano rispetto ai test già presenti in letteratura; esso potrà essere somministrato in contemporanea ad altri test di comprensione del testo, progettati con obiettivi affini, per verificare eventuali correlazioni. In ogni caso la validazione lascia intravedere la disponibilità di uno strumento adeguato per la rilevazione dei miglioramenti ottenuti a seguito di interventi formativi.

RIFERIMENTI BIBLIOGRAFICI

- Allen, J. (1995). *Natural language understanding*. London: Pearson.
- Anderson, V., & Hidi, S. (1988). Teaching students to summarize. *Educational leadership*, 46(4), 26-28.
- Bartlett, F. C. (1932). *Remembering: A Study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Benvenuto, G. (1987). *Insegnare a riassumere. Proposte per un itinerario didattico di prove di riassunto*. Torino: Loescher.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, 284, May, 34-43.
- Bonifacci, P., Tobia, V., Lami, L., & Snowling, M. J. (2014). *ALCE. Assessment di Lettura e Comprensione in Età evolutiva*. Boston, MA: Hogrefe & Huber.
- Bransford, J. D., & Johnson, M. K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing*. Orlando: Academic Press.
- Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, 22, 1-14.
- Brown, A. L., Day, J. D., & Jones, R. S. (1983). The development of plans for summarizing texts. *Child Development*, 54(4), 968-979.
- Brown, A. L., & Palincsar, A. S. (1987). Reciprocal teaching of comprehension strategies: A natural history of one program for enhancing learning. In J. D. Day & J. G. Borkowski (Eds.), *Intelligence and exceptional ability: New*

- directions for theory, assessment, and instructional practices* (pp. 81-132). Westport, CT: Ablex.
- Consiglio dell'Unione Europea (2018). *Raccomandazione del Consiglio del 22 maggio 2018 relativa alle competenze chiave per l'apprendimento permanente*.
- Cornoldi, C., Colpo, G., & Carretti, B. (2017). *Prove MT – Kit scuola*. Firenze: Giunti EDU.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research*, 61, 239-264.
- Feldt, L. S. (1961). The use of extreme groups to test for the presence of a relationship. *Psychometrika*, 26(3), 307-316.
- Friend, R. (2001). Effects of strategy instruction on summary writing of college students. *Contemporary Educational Psychology*, 26(1), 3-24.
- Garnham, A., & Oakhill J. (1996). The mental models theory of language comprehension. In B. K. Britton, & A. C. Graesser, *Models of understanding text* (pp. 313-339). Mahwah, NJ: Lawrence Erlbaum Associates.
- Giovannini, M. L., & Ghetti, M. (2016). *Prove standardizzate di comprensione dei testi per la scuola secondaria di I grado*. Milano: LED.
- Green, G. M. (2012). *Pragmatics and natural language understanding*. London: Routledge.
- Johnson-Laird, P. N. (1988). *Modelli mentali*. Bologna: il Mulino (ed. orig., *Mental models: Toward a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press, 1983).
- Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*, 30(1), 17-24.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model text comprehension and production. *Psychological Review*, 85, 363-394.
- Langer, J. (1995). *Envisioning literature*. Newark: International Reading Association.
- Levorato, M. C. (1988). *Racconti, storie e narrazioni. I processi di comprensione dei testi*. Bologna: il Mulino.
- Levorato, M. C. (2000). *Le emozioni nella lettura*. Bologna: il Mulino.
- Lumbelli, L. (2009). *La comprensione come problema. Il punto di vista cognitivo*. Bari: Laterza.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99, 440-466.

- OCSE (1999). *Measuring student knowledge and skills: A new framework for assessment*. <http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33693997.pdf> (accessed 26/10/2018).
- OCSE (2009). *PISA 2009 assessment framework: Key competencies in reading, mathematics and science*. <http://www.oecd.org/dataoecd/11/40/44455820.pdf> (accessed 26/10/2018).
- OCSE (2015). *Indagine OCSE PISA 2015: i risultati degli studenti italiani in scienze, matematica e lettura*. http://www.invalsi.it/invalsi/ri/pisa2015.php?page=pisa2015_it_07 (accesso 31/08/2018).
- OCSE (2018). *PISA 2018: Global competence*. <https://www.oecd.org/pisa/pisa-2018-global-competence.htm> (accessed 26/10/2018).
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal Teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117-175.
- Pečjak, S., & Pirc, T. (2018). Developing summarizing skills in 4th grade students: Intervention programme effects. *International Electronic Journal of Elementary Education, 10*(5), 571-581.
- Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant*. Neuchâtel: Delacha et Niestlé.
- Pressley, M., & Woloshyn, V. (1995). *Cognitive strategy instruction that really improves children's academic performance*. Cambridge, MA: Brookline Books.
- Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology, 3*(4), 552-631.
- Schank, R. C. (1992). *Il lettore che capisce. Il punto di vista dell'intelligenza artificiale*. Firenze: La Nuova Italia.
- Schank, R. C., & Abelson, R. (1977). *Scripts, plans goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Van den Broek, P., Kremer, K. E., Lynch, J. S., Butler, J., White, M. J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. In S. Paris & S. Stahl, *New directions in assessment of reading comprehension* (pp. 107-130). Mahwah, NJ: Lawrence Erlbaum Associates.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Werlich, E. (1975). *Typologie der Texte: Entwurf e. textlinguist. Modells z. Grundlegung e. Textgrammatik*. Heidelberg: Quelle Meyer.
- Wiersma, W., & Jurs, S. G. (1990). *Educational measurement and testing*. Boston, MA: Allyn & Bacon.
- Winograd, P. N. (1984). Strategic difficulties in Summarizing texts. *Reading Research Quarterly, 19*(4), 404-425.

RIASSUNTO

La rielaborazione sintetica dei testi è riconosciuta come uno dei processi principali ai fini della comprensione della lettura, tuttavia nella letteratura scientifica italiana (con riferimento al contesto in cui il test deve essere applicato) non sono diffusi strumenti atti a valutare l'esercizio di questa competenza, né a supportare il suo sviluppo e il suo miglioramento nei bambini. Il Summarizing Test, nato nell'ambito di un programma volto a sperimentare l'efficacia del Reciprocal Teaching nella scuola primaria, cerca di colmare questa lacuna. La validazione del test, compiuta con dati acquisiti in classi di quarta primaria distribuite geograficamente sul territorio nazionale, lascia intravedere la disponibilità di uno strumento promettente, da affiancare ad altri strumenti già diffusi sulla comprensione della lettura. L'articolo riporta i risultati dell'item analysis condotta secondo la Teoria Classica dei Test e lo split half per consentire di utilizzare due versioni equivalenti, applicabili in ingresso e in uscita di un percorso sperimentale.

Parole chiave: Comprensione del testo; Item analysis; Riassunto; Scuola primaria; Validazione di un test.

How to cite this Paper: Menichetti, L. (2018). Valutare la capacità di riassumere. Il Summarizing Test, uno strumento per la scuola primaria [Evaluating summarizing skills. The Summarizing Test, a tool for primary school]. *Journal of Educational, Cultural and Psychological Studies*, 18, 369-396. doi: <http://dx.doi.org/10.7358/ecps-2018-018-meni>