European Journal of Special Education Research
ISSN: 2501 - 2428
ISSN-L: 2501 - 2428
Available on-line at: www.oapub.org/edu

DOI: 10.46827/ejse.v9i3.5009 Volume 9 | Issue 3 | 2023

DYNAMIC ASSESSMENT IN INCLUSIVE ELEMENTARY EDUCATION: A SYSTEMATIC LITERATURE REVIEW OF THE USABILITY, METHODS, AND CHALLENGES IN THE PAST DECADE

Hang Le[1],
Juliene Madureira Ferreira[2i],
Elina Kuusisto[3]
[1]Faculty of Education and Culture,
Tampere University,
Finland
orcid.org/0009-0008-0870-8034
[2]Faculty of Education and Culture,
Tampere University,
Åkerlundinkatu 5, Office 509
Tampere, 33100,
Finland
orcid.org/0000-0002-8977-5982
[3]Faculty of Education and Culture,
Tampere University,
Finland
orcid.org/0000-0001-5003-547X

Abstract:
Dynamic Assessment (DA) provides alternative strategies for assessing children in inclusive education but hasn't been widely incorporated in teacher education as a tool for mediating learning in special education. This systematic literature review investigates empirical studies reported in the last decade to provide a comprehensive overview of the applications DA in inclusive elementary education, reviewing common methodologies and tools, challenges, and recommended solutions to develop DA. In the final analysis, 24 articles published between 2010- 2021 met the study's inclusion criteria. Results indicate that DA is mainly used to identify students' educational needs, predict students' future performance and responsiveness to intervention, and assess academic development, but rarely to guide intervention, inform instruction, or develop individual educational planning. DA is mainly used by trained researchers and their assistants, instead of teachers, despite its relevance to educational processes. Challenges employing DA and recommendations for developing its practice are critically discussed.

[i] Correspondence: email letruongthuhang93@gmail.com, juliene.madureiraferreira@tuni.fi, juliene.madureira.ferreira@gmail.com, elina.kuusisto@tuni.fi

Keywords: dynamic assessment, systematic literature review, elementary education, special education

1. Introduction

Dynamic Assessment (DA) is a method for interactive, non-standardized, and mediated assessment in education (Elliott, 2003; Elliott et al., 2018; Haywood & Tzuriel, 2002), particularly relevant in the context of inclusive education (Al-Hroub & Whitebread, 2019; Gellert & Elbro, 2018; Kirschenbaum, 1998; Regalla & Peker, 2017). DA methods positively impact testing performance over static testing conditions (Swanson & Lussier, 2001), offering the opportunity to understand students' thinking processes during the execution of a task (Feuerstein et al., 2003). In the context of special education, DA methods offer a means for measuring the transformations in the thinking process supported by alternative communicative strategies. The mediated nature of DA supports the teacher in exploring various ways to understand students' ability to grasp abstract concepts, solve problems, systematize strategies, and transcend knowledge during the evaluation process. Therefore, DA methods can reveal the potential and cognitive skills of children with disabilities, developmental disorders, or learning difficulties (Al-Hroub & Whitebread, 2019; Gellert & Elbro, 2018; Kirschenbaum, 1998; Regalla & Peker, 2017), becoming an important tool in the promotion of inclusive practices in elementary education.

The field of DA has developed in many divergent directions and varied applications. Different DA models have been employed in educational contexts, focusing on speech/language, neuropsychological, or psychological assessments (Haywood & Lidz, 2006). However, earlier studies have not systematically researched how DA has been used in inclusive contexts, what challenges are encountered, and what advancements in the field allow for overcoming such barriers. Thus, the present study investigates the application of DA in inclusive elementary education by looking at empirical studies conducted from 2010 to 2021. The study is grounded in cultural-historical theoretical frameworks. It aims to offer a comprehensive understanding of the application of DA in inclusive elementary education, providing teachers and educators with knowledge on how to improve assessment, teaching, and learning practices.

2. Literature review

2.1 Dynamic assessment as an interactive approach to assessment and teaching/learning

Dynamic Assessment (DA) was originally elaborated as an interactive approach to assessing students' learning and learning potential (Kozulin, 2015). It is structured by a procedure that consists of a pre-test/learning intervention/post-test dynamic (Haywood & Lidz, 2006), which allows the teacher to understand students' reasoning processes during the execution of a task. The term dynamic expresses that assessment happens

within an interaction, which involves exchanging information, intentions, behaviors, and actions between two or more people in a continuum flow (Feuerstein et al., 2003). Thus, assessment is not localized to the individual or reduced to recalling information about a specific situation or problem (Al-Hroub & Whitebread, 2019).

DA is theoretically based on Lev Vygotsky's concept of the Zone of Proximal Development (ZPD) (Vygotsky, 1978). The ZPD conceptualizes the difference between children's previously formed mental functions and potential ones (Kozulin, 2011), focusing on the fundamental role of joint activity of learners and teachers and the appropriation of mediation in developing higher mental functions. The ZPD explains how students can perform differently under the condition of assisted (scaffolded) learning (Kozulin, 2015; Kozulin & Garb, 2016), thus impacting how development, learning, and assessment-related activities are implemented in educational contexts (Kozulin, 2011; 2015), creating "*a comprehensive dynamic paradigm*" (Alony & Kozulin, 2007, p.324). Furthermore, the idea of ZPDs emphasizes the interactive nature of learning, bringing meaningful interpersonal communication between teachers and learners as well as learners with their peers to the front stage of the learning processes (Kozulin, 2015; Kozulin & Garb, 2016) and setting the premises for the DA.

DA first gained momentum in the late 1960's parallel to the rising criticism of standardized psychometric tests (Kozulin & Garb, 2016). The first complete versions of dynamic cognitive assessment were elaborated by Budoff and Friedman in 1964, followed by Feuerstein and Shalom in 1968 (Kozulin & Garb, 2016). Feuerstein's (Jensen & Feuerstein, 1987) version of DA was formulated grounded on the Theory of Structural Cognitive Modifiability, which proposed that "*the individual is an open system susceptible to influences that can produce permanent structural changes in cognitive functioning"* (p. 360) when the learning experience provides a combination of cognitive processes. Feuerstein and colleagues proposed that learning experiences must be mediated by other human beings following a set of specific criteria and procedures to achieve structural cognitive modifiability (Feuerstein et al., 1999). Feuerstein was also responsible for adapting these procedures used in clinical cases to the first models of DA applicable in educational contexts. Within this approach, the essential element determining the development of humans' higher levels of cognitive functioning depends on one's opportunity to benefit from mediated learning experiences (Jensen & Feuerstein, 1987; Feuerstein et al., 1999). Thus, DA should not merely address specific academic contents that require memory or automatic skills but should concentrate on tasks demanding cognitive strategies (Feuerstein et al., 1999).

At first, the tasks utilized in DAs were similar to those of standardized psychometric tests. DA was either presented as a supplemental method to be utilized in addition to the static test or as a substitute for standardized psychometric assessment (Samran & Mehdi, 2018). As the work with DA increased in the late 1980s, various new methods emerged using significantly different techniques (Poehner, 2018). The more progressive approaches to DA concentrated on analyzing the students' socio-emotional and cognitive developments, adjusting instructional needs, disapproving of the

prediction goal, or using DA as a method to categorize the child's needs (Haywood, 2012). At this point, it was possible to recognize a more prominent role of the evaluator (mediator), with less standardized intervention rules in the learning phase. In many situations, DA was addressed as a teaching/learning process more than an assessment method (Kozulin & Garb, 2016). Despite the variety of methods and strategies, the way DA is used tells how assessment is contextualized within the teaching and learning process.

2.2 Dynamic assessment in inclusive education

The definition of inclusive education varies among researchers. A narrow definition of inclusion means taking the perspective of some specific groups of students, while a broad perspective aims to provide non-discriminatory education for all students (Ainscow et al., 2006) regardless of their cultural, linguistic, worldview, or socioeconomic backgrounds, disabilities, learning difficulties, and diverse giftedness profiles. In this study, we corroborate with a broader perspective of inclusion, undertaking the complexities and tensions emerging from the demands of special needs education within an inclusive perspective. Therefore, within this perspective, educational support should be offered to all students who need it at any given moment during schooling through individualized or collective activities, content-oriented, activity-oriented, or pedagogical differentiation (Ferreira et al., 2022).

Nevertheless, it is important to recognize the significantly different developmental paths that children can have and act according to the educational needs circumscribed in such cases. Under this perspective, the tripod - access, participation, and achievement (UNESCO, 2005), is a key sign of inclusion in educational contexts (Ferreira, 2018). Access includes but is not restricted to actions that facilitate and support the child's physical presence in school; it focuses on identifying the obstacles that prevent children from attending school consistently. Participation is related to the quality of the learners' educational experiences, defined by how well the student is inserted into school activities and participates effectively in the learning situation. Thus, defining participation integrates learners' perspectives. The third element of the tripod is achievement, and it concerns identifying the learning outcomes across a curriculum, not targeting only the results from tests or exams but also the learning processes (UNESCO 2005). It is within the helm of improving tools and techniques to measure achievement that dynamic assessment becomes significant.

2.3 Assessing children with significantly different developmental paths

In the context of assessing children with significantly different developmental paths, such as children with intellectual disabilities and developmental disorders or children that have different cultural backgrounds and that have not familiarized themselves with the demands of the new socio-cultural environment, traditional assessment imposes great challenges. Leeber et al. (2011) call attention to the inadequacy of utilizing standardized or individual-performance tests to assess students with disabilities and, to a greater

extent, merely employing such measurements to identify and justify the need for educational support or special education. Leeber et al. (2013) argue that standardizing assessment processes for children who otherwise have individualized learning plans only reinforces labeling and children's limitations. Consequently, it embraces a threat of arousing low educational expectations or self-fulfilling prophecies, pointing out that one of the hindrances to inclusive education is generated by how these children are examined before and throughout schooling.

From an inclusive perspective, ideally, assessment should be linked with practices and policies established to enhance the learning of all learners regardless of their disabilities (Watkins, 2007). The assessing procedures should assist and promote effective participation, value students' achievements according to the plans stipulated for children individually, and target to celebrate diversity in learning. Thus, students should be involved in the decision-making of their assessment (Watkins, 2007). This approach to assessment explicitly aims to diminish isolation by avoiding labeling students and concentrating on teaching and learning practices that further inclusion in mainstream contexts.

DA has been used to understand how learners can demonstrate their knowledge and academic capabilities with the teacher's assistance (Caffrey et al., 2008; Grigorenko, 2009; Wagner & Compton, 2011). The significant difference between DA and traditional methods for student assessment can be summarized in its process and content. Traditional assessment methods usually rely on recollecting information and focus on the individual performance in applying the knowledge in a specific situation; the focus is on the product of learning (Cho et al., 2020). In contrast, DA privileges the learners' learning potential during the teacher-student interaction created in the evaluation. DA methods focus on assessing the cognitive skills to respond to intervention (Al-Hroub & Whitebread, 2019; Feuerstein et al., 2003), thus measuring the learning potential.

Considering the theoretical premises and the practices constituting inclusive education, pioneer researchers studying DA (e.g., Haywood, 2012; Haywood & Lidz, 2006; Jensen & Feuerstein, 1987) have recommended it to be particularly beneficial in four situations. First, when scores on static tests (traditional methods or standardized tools) are poor, particularly when they do not correspond to information from other sources of assessment of students' learning and development (e.g., the child can solve problems during unformal situations but cannot when tested). Second, learning seems hindered by a lack of motivational, emotional, and behavioral disorders, learning difficulties, intellectual disabilities, or developmental disorders. In such cases, identifying how the child thinks and what parameters they use to think about the given task informs the teacher about the precise element that needs intervention (Kaniel et al., 1999). Third, when students encounter problems in languages, such as developmental language disorders, restricted vocabulary, and differences between the learner's mother tongue and the language of instruction used in school. Lastly, DA is recommended when teacher and student do not share the same cultural background and thus do not share language nor a similar understanding of the cultural signs surrounding them. In these

circumstances, static testing is prone to produce poor academic results, which might not represent the child's cognitive capabilities or development but rather show the state of cultural deprivationii (Feuerstein & Feuerstein, 2001). Within these contexts, the role of DA is to identify hindrances to more successful learning and performance, discover strategies to surmount them and evaluate the effects of removing those hindrances on the effectiveness of future learning and performance (Haywood & Lidz, 2006). Therefore, DA is a useful method to assess students' thinking processes and an effective teaching tool to improve students' effectiveness at learning (Lawrence & Cahill, 2014).

3. Material and Methods

3.1 Research design and questions

The present study consists of a Systematic Literature Review (SLR) (Boland et al., 2017; Zawacki-Richter et al., 2020). SLRs can efficiently create a firm groundwork for advancing knowledge and facilitating theory development by effectively collecting, synthesizing, and integrating findings from previous research (Snyder, 2019; Tranfield et al., 2003). This SLR focuses on analyzing previous studies that show the application of DA and its implications for schooling in elementary education. Two research questions guided the study:

RQ1: What does empirical literature reveal about the applicability of DA in elementary inclusive education?

a) What methodologies are applied?
b) What purposes are DAs serving in schooling processes?

RQ2: What are the challenges for DA in elementary inclusive education and the solutions for developing its practice?

The study draws upon the historical-cultural approach to reflect upon the evolution of the application of DA in inclusive elementary education within the time frame of 2010-2021. The historical-cultural theory frames this work as it provides two fundamental ideas – learning and development are intrinsically related and dependent on the interactions one experiences in life, and the idea that interactive processes between people are considered embedded in and emerging from a socio matrix of semiotic nature (Ferreira, 2018). Within this perspective, the aim is to reflect upon the socio-economic, historical, and cultural elements that can explicate different applications of DA across the decade.

[ii] Cultural deprivation within the framework of the Theory of Mediated Learning Experience (Feuerstein et al., 1999) is the state in which a person has limited access or opportunity to participate and construct cultural references; they have not been inducted into their own culture due to the inadequate provision of mediated learning experiences. According to Feuerstein and colleagues (1999), cultural deprivation can be associated with negative predictions of learning development and academic achievement but is remediated by a mediated learning approach. Different from other sociological concepts of cultural deprivation, this one is not associated with low social-economic classes or inferior norms and values.

3.2 Search strategy and study selection

The protocol for this systematic review was prepared following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). Studies were retrieved from five databases (Education Resources Information Center [ERIC], ScienceDirect, PubMed, Web of Science, and PsycINFO), the direct archive of four publishing companies (SAGE Journals Online, Taylor & Francis Online, SpringerLink, and Wiley Online Library) and one journal (Journal of Cognitive Education and Psychology), which devotes particular attention to this approach to assessment. The search string was developed by extracting search terms employed in previous reviews carried out in the field of dynamic assessment (Swanson & Lussier, 2001; Rezaee & Ghanbarpour, 2016; Ghanbarpour, 2017) and through discussion between the authors. The search terms used were [Special education OR Inclusive education OR special class] AND [dynamic assessment OR mediated assessment OR ZPD-based assessment OR mediated learning experience OR scaffolded assessment OR classroom assessment] AND [historical-cultural theory OR activity theory OR "zone of proximal development"].

From the search, only studies reported in English and published in peer-reviewed academic forums were selected. Conference presentations, thesis, and dissertations are not included. From this selection, 24,951 articles were retrieved, of which 14,050 were duplicates, leaving 10,901 articles ready for the title and abstract screening process.

During title and abstract screening, the first author used Rayyan software (Ouzzani et al., 2016) to support the procedure of deciding whether the studies identified met any inclusion criteria for the review. The application of a list of criteria framed the screening process. The study focused on the use of DA in primary and elementary education, in whole classrooms, small groups, and individual assessments of students attending special educational support due to temporary or permanent disabilities and learning difficulties. Any standardized evaluation was admitted, and when it was unclear if the study included participants' evaluation, the first author contacted the paper's authors and clarified the information. Clinical studies and those targeting students from other levels of education were excluded. Also, the study acknowledged that there are many different scenarios and settings in which DA can be implemented. However, the focus remained on DA methods mediated directly by human activity in the educational context; thus, it eliminated all the reports addressing findings with Computerized Dynamic Assessment (C-DA).

After the screening procedure, 10,861 articles were eliminated, leaving 40 studies for the whole text analysis. The first author screened all articles for inclusion, and the second independently screened 30%. Inter-observer agreement for inclusion was 95% at the title and abstract stage and 98% at the full stage. Disagreements were resolved with discussion until consensus. Following PRISMA guidelines (Moher et al., 2009), the next step was to establish the eligibility criteria, identifying (1) a detailed description of the DA methodology and (2) high-quality standards. In this phase, 40 articles were analyzed in full, and from this pool, 24 met all the eligibility criteria (see Figure 1).

3.3 Quality appraisal

An adapted version of the Mixed Methods Appraisal Tool (MMAT) (Hong et al., 2018) was employed to assess the quality of articles retrieved for this study. The MMAT is a quality assessment tool developed to provide a critical appraisal of the methodological quality of empirical studies (Hong et al., 2018) designed as (a) mixed-methods studies, (b) quantitative descriptive, (c) randomized controlled trials, (d) nonrandomized and (e) qualitative. The MMAT (version 2018) encompasses 25 criteria (5 criteria for each type of study design) and two screening questions. In this study, the authors adapted the instrument (see Table 1 for the adapted MMAT), eliminating the two initial screening questions. In this case, the first step became the categorization of the study in each type of study design, and posteriorly, the study was rated following the criteria of the selected category. For instance, if the article was a mixed-methods study, the researcher only rated the study according to the five criteria elaborated for the mixed-methods category.

3.4 Data extraction, analysis, and synthesis of the analysis

The data was extracted, summarized, and organized in tables. Data relating to research questions on all selected studies were organized across the following categories: (a) year of publication and authorship, (b) location where the study was conducted, (c) research setting, (d) characteristics of participants, (e) DA procedure, (f) the utility of DA, (g) methodologies employed, (h) DA tools and techniques. The categories c, e, and h were analyzed using a predetermined coding scheme (see Table 2). For the remaining categories, the data extracted across the included studies were synthesized narratively to portray a general understanding of the content of the included studies. Narrative synthesis refers to presenting the results of included studies utilizing words merely concerning data tables (Boland et al., 2017). Due to the diverse study designs across the twenty-four studies selected in this systematic review, a narrative synthesis is considered an appropriate procedure to report the research findings (Schwarz et al., 2019).

After extracting all relevant data, the analysis was conducted to identify the significant contents answering RQ1 and RQ2. The data from the description of the research settings and DA methods were significant in answering RQ 1. The description of methodologies and the utility of DA (its implications) were significant to answering RQ 2. It is important to mention that the second author redid the data extraction after three days and cross-checked that both sets of extracted data were the same (Boland et al., 2017). The summaries of included studies are displayed in Table 3, organized chronologically by publishing years and alphabetically by the first author's last name.

4. Results and Discussion

4.1 Overall description of the studies

From the 24 included studies, eleven (46%) were conducted in the U.S., two in the Netherlands, and two in Denmark. Four of the studies (17%) did not mention the location. The remaining five studies were conducted in the UK, Jordan, Singapore, South Africa,

and Spain, covering distinct parts of the world with diverse cultures and educational practices. Participants in these studies were children from 6 to 16 years of age primarily attending primary education, presenting various special educational needs. DA is used for *assessing learning* of students with reading disabilities (20,21), math difficulties (9,10,13,15), language impairment (12), and with limited English proficiency (16,23). To *assess performance* in reading (1,5,6,11) and in math (19,24). To *assess high abilities* (2,4,17,22) and to *identify learning strategies* of students on autism spectrum disorder (3) and children belonging to ethnic minorities (7,15). One study targeted students with a range of barriers to learning but did not state the participants' special educational needs in detail (18). In this case, specific information was collected via email with the study's authors. The remaining study (8) included subgroups of children with reading disabilities, and one included teachers in their sample (7). Furthermore, seven studies (5,10,11,14,16,20,21) specifically *compared DA methods with standardized assessment* to measure students' learning.

4.2 DA methods applied in inclusive elementary education

DA is methodological flexible and resourcefulness in different conditions. This systematic literature review identified variances concerning the *types of assistance to students* through which the DA is methodologically implemented, types of *DA integration* to other methods for teaching-learning and student assessment, and concerning the *DA application*, which is defined by who is responsible for applying DA in empirical studies (e.g., teacher, researcher, or assistant) and how the application is organized (e.g., individually or in-group).

4.2.1 Diverse types of student assistance during DA

Concerning the *type of assistance*, this study reveals five categories of methods: (1) systematized and standardized prompts (e.g., graduated prompts), (2) instructional scaffolding, (3) feedback, (4) mediated learning, and (5) free-form cues, probes, or reminders. Combinations of two or more types of assistance appeared in eighteen out of twenty-four selected studies.

The most-reported method was systematized and standardized prompts (n = 18) (1,4,5,6,8,9,10,11,12,13,14,15,16,17,19,20,21,23). Standardized prompts are applied when the students hesitate on the test items or provide an incorrect answer right away. The evaluator supports the student by offering a series of prompts until the learner can master the test item. Different types of prompts are described in this study data set, including strategic, modeling, cognitive, metacognitive, increasingly explicit, indirect, verbal, and visual prompts. All prompts were provided in the form of questions and clues to evoke the students' constant awareness of their possible learning difficulties and measure their learning potential. While most studies employed a single structured protocol of prompts, multiple protocols allowed a more comprehensive approach in the intervention phase. They gave students multiple opportunities for knowledge construction (19). Consequently, it allows the evaluators to classify groups of students with different needs

for instruction. Nevertheless, this specific technique was found to be difficult for the evaluators to master, conduct, and score, especially in group interventions. To respond to each individual answer, the evaluator had to select the appropriate protocol, and, in some cases, the selected prompts were not useful for all students (e.g., Bosma et al., 2017).

The second most common method implemented involved instructional scaffolding. Instructional scaffolding is an interactive process through which teachers gradually provide increasingly explicit scaffolded instruction for students to improve learning and assist with the mastery of tasks (Vygotsky, 1978; Walqui, 2006). Twelve studies that used instructional scaffolding employed scripted mediating tools, creating a standardized approach (1,6,9,10,11,13,14,15,16,20,21,23). All these studies utilized 3 to 5 levels of standardized prompts to scaffold students during instruction. Other instructional methods to facilitate the instructional process included additional visual or pictorial representation (i.e., lines and pictures) (15,23), corrective feedback (21), non-contingent feedback (14), and cooperative learning (13,15).

Another type of assistance in DA methods is framed as giving feedback. Giving feedback included immediate and positive feedback (2), corrective feedback (20,21), positive and informative feedback (19), feedback on whether students succeeded or failed in their learning (3), instructional feedback (15), and performance-contingent feedback (9,13) versus noncontingent feedback (14). Different from the previous types of assistance in DA methods that mainly focused on students' ongoing thinking processes or their answers to testing, giving feedback entails a broader scope of elements. For example, evaluators can address students' behaviors during the task, focus on developing mindsets for testing, or support students in adopting strategies to cope with difficulties during testing situations. Feedback is an important part of the learning process (Hattie & Timperley, 2007; Svanes & Skagen, 2017); it allows students to modify their approach to learning and mindset, re-direct the behaviors related to learning, and adjust the overall experience of knowledge construction (Hattie & Clarke, 2019).

The fourth type of assistance the evaluators employed during the intervention or the test was mediated learning (2,3,4,7,12,17,18). In the mediated learning situation, the evaluators interpose between the environmental stimuli and the student and modify the stimuli for the student, acting as a mediator (Kozulin & Presseisen, 1995). Instead of explicit teaching, the mediators ask questions to enhance students' strategic thinking, logical thinking, and reflection to help them plan, analyze, deduce, and monitor their answers. To understand students' metacognition, mediators typically ask students to explain and discuss their strategies and methods to solve the problems (2,7,17,18). Students can also be asked to reflect on the requirements and difficulties of the tasks (12) to assess their metacognitive awareness.

The last type of assistance identified across the selected studies entailed using free-form cues, probes, and reminders (3,24). This method involves delivering fragments of information that direct the students' thinking process. It is a free-form and context-based approach that can be used and designed by the evaluator at any given moment of the intervention. The interaction between students and the evaluators is individualized and

customized to the needs of individual students. It is preferably used before the student provides the final answer to a test. According to the studies in this systematic literature review, the advantage of using cues/probes is allowing the evaluator to make decisions based on what is happening at the moment of the test. Albeit these strategies are reported to support the students to reach their potential in most cases, studies involving autistic students (3) point out the challenge for them to understand the meaning of the cues/probes used in mediation and make meaningful connections to solve the tasks. Thus, there is a possibility that the employment of DA fails to measure the learning potential of autistic children. The solutions recommended in the study were relatively vague. Thus, further attention needs to be given to developing specific mediation strategies that are appropriate for students with autism.

4.2.2. Different ways of integrating DA and other methods

DA was implemented both as integrated with other teachings (n=3) or assessment methods (n=16), as a framework for the research study (n=2), and as an independent method, i.e., standing alone as the main assessment process (n=3). As part of a multidimensional assessment (1,2,3,4,5,6,8,10,11,12,14,16,20,21,22,23). Out of ten studies comparing DA to other types of assessment, nine studies (5,6,10,11,14,16,20,21,23) focused on whether DA yielded higher predictive validity than static assessments. The remaining studies integrated with other assessment methods by employing IQ tests to classify participants into groups before utilizing DA to measure students' learning potential (3,4,8) or by integrating DA into part of a test battery to evaluate different aspects of students' learning potential (2,12,22).

Three studies (7,9,24) integrated DA with other teaching methods. Jeltova et al. (2011) conducted DA at the end of each instructional intervention in the regular classroom. In contrast, DA in Moscardini and Moscardini's case study (2020) was embedded during the collaborative teaching sessions between a support teacher and the class teacher within the real-life context of the classroom. In Orosco et al. (2011) study, DA was integrated with the strategies instruction and the framework for the intervention procedure investigated in the study. In this intervention, students took dynamic testing after the examiners pre-taught the concepts and instructed the strategies.

In only two studies (13,15), DA was a framework for a research study. In those studies, the structure and principles of DA were used to guide and frame a math comprehension strategy procedure. In the intervention, the researcher modified the main concepts and vocabulary students need to acquire to the individual student's understanding level of the specific domain and then gave strategy instruction with probes that tested students' problem-solving ability (13,15). No specific assessment process involves DA in these studies.

Three studies reported implementing DA solely (17,18,19). In Amod et al. (2017), DA was used as a pre and post-test of the Planning, Attention, Simultaneous and Successive Processing (PASS) model of cognitive processing to develop an intervention

program by employing the Cognitive Assessment System (CAS; Naglieri & Das, 1997), a test instrument guided by PASS theory.

4.2.3 Different ways to apply DA

In terms of the application of DA, two factors were considered. First, the professional responsible for the application of the method, and second, the organization of its application, e.g., individually or in-group. One-to-one DA was used in the majority (n = 17) (1,2,3,5,6,8,9,10,11,12,13,16,19,20,21,22,24) of the studies, and in most cases (ten out of seventeen), the members of the research team implemented the procedure. They were specifically trained in the chosen methods and instructed on the type of assistance necessary for their implementation. The teacher assisted in three of these studies (2,9,13,24). In Moscardini and Moscardini (2020)'s study, although the testee attended a regular classroom with her classmates, she was the only student attending the dynamic process with the support teacher. In Orosco et al. (2011) and Orosco (2014)'s studies, the trained teacher and researcher alternated sessions conducting the intervention. The remaining two studies applying DA individually did not mention who was responsible for the assessment (12,19).

Only two studies investigated group-administered DA (7,15), and those DA methods were performed by teachers and trained evaluators, respectively. Teachers were either divided into groups and responsible for assessment under different conditions (7), or all teachers participated in the intervention throughout the entire process (15). Furthermore, five studies did not state whether students were assessed individually or in a group in their reports. Trained research assistants and researchers conducted the DA assessment in two studies (14,23). In Al-Hroub and Whitebread's (2019) 's study, a subject teacher was in charge of DA assessment. Calero et al. (2011) and Vogelaar et al. (2016) did not mention the examiners' information.

4.3 Purposes of using DA in schooling processes

The results point out that DA contributes to four different aspects of schooling: First, identifying students' learning potential; second, planning and guiding interventions; third supporting Responsiveness-To-Intervention (RTI) decision-making processes; and forth, a framework for developing intervention procedures and guide pedagogical practices.

4.3.1 Identification of students' learning potential

DA is used to identify students' learning potential in four studies (3,4,17,22) addressing specific subjects such as mathematics and more general cognitive competencies. In these studies, participants who experienced dynamic training demonstrated larger improvements from the pre-test to the post-test than those with unguided practice only (17). The results from these studies indicate that by targeting what students can learn within the moment of mediation rather than what students have previously learned, the dynamic assessment identifies children's learning potential and provides a more

comprehensive picture of students' cognitive abilities than conventional static assessments.

In combination with other methods, DA was used because of its possibility to refine the results of standardized testing, supporting the elaboration of prediction models. For example, Al-Hroub (2011) applied DA to identify learning difficulties in gifted children sensitively, as DA offers multiple ways to address children's very specific difficulties. Gellert and Elbro (2018) used DA because of its accuracy, applying it to identify children at risk for reading disabilities. The detailed data that is produced from applying DA affords a deeper and better understanding of children's difficulties, and it can be used to elaborate prediction models for learning and future performance (5,8,10,14,16,20,23). The unique prediction value of the dynamic assessments varied in different studies ranging from 1 year to 3 years. DA was reported to explain unique variance in students' future performance and have adequate sensitivity and specificity above 80% (e.g., Petersen & Gillam, 2015).

Likewise, DA was employed to support the teachers in the teaching process by evaluating students' learning potential and guiding intervention (12), informing immediate and long-term instructional decisions (24), and identifying the differences in the need for instruction (19). DA supports teachers in further interventions mainly by enabling teachers to learn about the child's learning process and by allowing the identification of the child's specific needs. It provides a more focused and tailored educational service, which, in the long run, saves time and increases the chances of children's full participation and achievement in classroom activities.

When added to a Responsiveness-To-Intervention (RTI) decision-making process, DA was used to predict students' responsiveness to identify who would not be in response to Tier 2 and, therefore require Tier 3 (1,6,11). RTI is a multi-tier approach for early identifying and supporting students with learning disabilities (Fletcher et al., 2004; Fuchs & Vaughn, 2012). In these studies, participants had been previously identified as unresponsive to general instruction (Tier 1) and moved to more intensive and explicit instruction (Tier 2). After being assessed dynamically, students identified as non-responders to Tier 2 were moved to more intensive, individualized interventions (Tier 3) (Fuchs & Vaughn, 2012). When compared to other static and well-established predictors, DA was a significant predictor of students' growth, uniquely explaining the variance in students' Tier 2 responsiveness (6) or significantly decreased false positives when being added to the procedure (1).

Lastly, DA was also used to frame research studies (9,13,15), particularly to develop a math comprehension strategy called Dynamic Strategic Math (DSM), which scaffolds students in all education levels. This intervention procedure was reported to improve students' problem-solving ability and give teachers a better understanding of students' levels, their learning potential, and their cognitive processes of solving problems.

4.4 Challenges and recommended solutions for implementing DA in inclusive elementary education

Across all twenty-four included studies, only nine explicitly mentioned challenges for applying DA (see Table 4). Of the challenges, only three are directly related to practical implications.

Although DA methods are initially designed as a qualitative tool for specifying educational support, designing individual pedagogical planning, and intervening at an individual level (Feuerstein et al., 2003), the studies utilizing DA still prioritize metrics relevant to standardized (quantitative) methods and tools. For example, Cho and colleagues (2014) report challenges regarding the validity of DA procedures, pointing out that the subjective mediation that is part of DA as a relational process imposes questions on how to ensure the results represent what they are supposed to measure, e.g., learning potential. Fuchs and colleagues (2011) see challenges dealing with the floor effect of DA with few low-performing participants where measurements failed to provide evidence of *how much* (quantified number) poor readers learned the decoding skill. Furthermore, Jeltova and colleagues (2011) indicate the challenge to make DA methods as trustworthy as traditional static assessment methods due to the impossibility of replicating mediated interactions. Such challenges are deeply connected to the epistemological reasoning in framing the studies with DA. The recommendations to overcome these challenges include the triangulating analysis employing standardized and decoding tests to measure the same construct (11), and increasing standardized procedures during the application of the method, such as providing more straightforward probes earlier in a hierarchy of prompts (6).

Challenges regarding the application of DA, from a practical point of view, are noted in three studies. Petersen and Gillam (2015) found it challenging to deal with participants who were occasionally uncommunicative or unresponsive. To remedy the situation, evaluators offered moderately neutral prompts such as “You can guess—what do you think this word says?” or “Remember what I told you?” to guide the students. Elleman and colleagues (2011) pointed out the challenge of identifying whether and when the prompts were adequate (supporting the mediation) or hindering students' independent, reflective process. Similarly, Bosma and colleagues (2017) noticed difficulty in learning, performing, and scoring students' responses to the prompts and feedback. For both, the challenge lay in the lack of a standardized or systematic approach to applying the prompts and feedback. The recommended solutions pointed out in such studies entailed offering evaluators more training or applying DA supported by a third person who would be responsible for administrating and scoring prompts and responses. The challenges impacting the use of DA in research and practice are generated due to epistemological conflicts. There is an underlying assumption that DA should be framed by the same rules of standardized static testing (i.e., generalizable), even though this is a qualitative method designed to enhance individual assessment shifting the focus from the product to the learning process. This result may indicate contradictions between research designs and understanding of the purpose of DA in practice.

5. Recommendations

This study reviewed the literature on dynamic assessment in elementary inclusive education in the past decade, answering two main research questions: (1) What does empirical literature reveal about the applicability of DA in elementary inclusive education? (2) What are the challenges for DA in elementary inclusive education and the solutions for developing its practice? This systematic literature review aimed to shed light on the current knowledge on DA, supporting teachers and educators to improve assessment, teaching, and learning practices. Based on our findings, we make three remarks.

First, *DA is a powerful tool for assessing students' learning*. DA is used in various ways, carrying methodological and practical value in identifying and forecasting students' learning potential, guiding intervention, individual planning and instruction, and predicting students' responsiveness in an RTI decision-making process. Consistent with previous studies (Elliott, 2003), this literature review also shows that DA can support teachers in addressing learning needs and providing a tangible idea of their learning potential. The variety of types of assistance and implementation (as a sole method or in combination with other teaching) makes DA resourceful and flexible, which can support educators at all levels of schooling. All studies concluded that DA practices effectively achieved their pre-determined aims, indicating that regardless of what type of DA method is applied, it offers positive results.

Interestingly, despite DA being used mainly to investigate students learning potential and create individual pedagogical planning to enhance the schooling experience, teachers are not the main users of the method in empirical studies. This literature review points out that the majority of the users of DA methods are, in fact, researchers or professionally trained evaluators. Teachers may not often conduct research on their practices with DA, or researchers find it easier to carry out research that they are in control of the mediation process. Nevertheless, considering the strong evidence of the benefits of DA as an assessment tool, we argue in favor of expanding the use of DA in elementary education and encourage future studies to explore whether and how DA methods are being introduced in teacher education and training. Alternatively, future studies could investigate how teachers, educators, and school psychologists use the information obtained from the DA process in classroom contexts and the effectiveness of its application.

The variety of DA methods shows how flexible and resourceful this tool can be. However, it also opens to multiple interpretations of its use. Our second remark regards the *importance of keeping epistemological consistency*. Many of the studies in this literature review were concerned with testing DA methods as an alternative to standardized methods and evaluating their efficiency in predicting learning (e.g., Cho et al., 2014; Fuchs et al., 2011; Compton et al., 2010). In doing so, studies framed their investigation with the same parameters of static assessment methods, being concerned, for example, with the generalizability of the findings and operationalizing DA in a standardized way.

We argue that it is important to keep epistemological consistency. DA is, at its core, a qualitative approach to assessment grounded on an interactional process that is imperatively subjective. This means that even following a similar mediation structure, such as a script for a prompt, the interaction is unique and can result in different experiences, influencing the assessment and its end results. Therefore, considering the qualitative and relational nature of DA methods, we argue that future studies and applications of the methods should explore the advantages of the methods, focusing on the qualities of the method: singularities instead of generalizations. DA has shown the potential to provide information that can help teachers determine students' educational needs, focusing on analyzing learning processes rather than pointing out shortcomings of learning outcomes. Thus, it is a tool that overcomes the limitations of standardized assessment in informing "*how to understand the child's problems, how to work with a child, what the child may be able to do, or what the real needs are*" (Leeber et al., 2013, p.5). DA, as an assessment tool, is used for providing a type of knowledge that is essentially unique - an individual's reasoning during learning situations.

The third and last remark is that *applying DA demands training*. When analyzing the challenges in implementing DA reported by the selected studies, we see that the core issue is the relational feature of DA - the mediation in the interaction. In any context or type of DA method, the mediation between the evaluator/mediator and the student is key for the method to work (Feuerstein & Lewin-Benham, 2012). Thus, evaluators must trust their ability to establish a close relationship with the student. This process is not about identifying a list of pre-defined behaviors but about learning to know the student and using the mediation process to measure the learning potential (Feuerstein et al., 2003). The challenges discussed in the studies (Aljunied & Frederickson, 2011; Bosma et al., 2017; Elleman et al., 2011; Petersen & Gillam, 2015) often refer to this *knowing,* which is interpreted as understanding that the mediation is correct and effective.

Although the selected studies did not raise specific discussions on the need for training, we believe teacher education and training in DA are key elements to bridge implementation gaps. To increase the applications of DA methods as assessment tools in inclusive elementary education, this relational factor of DA must be addressed together with the operational training that DA demands. The training can be incorporated into the teacher education curriculum or as specialized pedagogies accessible in continuous education.

6. Conclusion

To conclude, this systematic literature review contributes to a further understanding of the complexity involving the elaboration and application of dynamic assessment in elementary school. Particularly, by targeting studies on inclusive education, the analysis and discussions in this study indicate a set of strategies that could be used as guidelines for teachers in developing assessments in inclusive classrooms. The present study stands by other academic works arguing in favor of providing individualized educational

support for all students in need, regardless of their gender, possible developmental specificities, or cultural or social background at any moment of their schooling path. This systematic literature review shows a wide variety of ways DA can provide such educational assistance.

6.1 Limitations

Regarding the limitations of this study, like in many systematic literature reviews, this one is limited by the margins of its scope. Particularly, this study excluded practices of DA in other levels of education than elementary school and excluded studies that were not directly related to educational practices. Also, despite researchers' efforts to synthesize all possible relevant findings from an extensive array of eligible literature, potential studies may be missed due to the search terms employed. Especially, there is the possibility that the search strings ["Special education" OR "Inclusive education" OR "special class"] are relatively insufficient to capture all relevant papers investigating DA in inclusive education. To fully grasp the development of DA, further studies should include those used in clinical settings and thus, reported as clinical cases instead of empirical research.

Conflict of Interest Statement

The authors declare no conflicts of interest.

About the Author(s)

Hang Le is a graduate of Tampere University with a Master of Arts degree in Teacher Education. Her academic interests are focused on the field of dynamic assessment and the advancement of inclusive practices in developing countries.

Juliene M. Ferreira is an assistant professor (tenure track) in Education - focus area of Educational Psychology. Her research investigates children's sociability in challenging interactions and contexts (e.g., special education), analyzing how intersubjectivity influence/impact learning and development. She is equally interested in improving inclusive education in teacher education and training.

Elina Kuusisto is a university lecturer (diversity and inclusive education) at the Faculty of Education and Culture, Tampere University. She holds the title of Docent at the University of Helsinki. Her academic writings deal with teachers' professional ethics and school pedagogy, with a special interest in purpose in life, moral sensitivities, and a growth mindset.

References

Ainscow, M., Booth, T., Dyson, A., Farrell, P., Frankham, J., Gallannaugh, F., Howes, A., & Smith, R. 2006. Improving schools, developing inclusion. Routledge. https://doi.org/10.4324/9780203967157

Al-Hroub, A. 2011. Developing assessment profiles for mathematically gifted children with learning difficulties at three schools in Cambridgeshire, England. Journal for the Education of the Gifted, 34(1), 7–44. https://doi.org/10.1177/016235321003400102

Al-Hroub, A., & Whitebread, D. 2019. Dynamic assessment for identification of twice-exceptional learners exhibiting mathematical giftedness and specific learning disabilities. Roeper Review, 41(2), 129–142. https://doi.org/10.1080/02783193.2019.1585396

Aljunied, M., & Frederickson, N. 2011. Does central coherence relate to the cognitive performance of children with autism in dynamic assessments? Autism, 17(2), 172–183. https://doi.org/10.1177/1362361311409960

Alony, S., & Kozulin, A. 2007. Dynamic assessment of receptive language in children with Down syndrome. International Journal of Speech-Language Pathology, 9(4), 323–331. https://doi.org/10.1080/14417040701291415

Amod, Z., Heafield, D., & Seabi, J. 2017. Assessing a REMEDIAL intervention programme in developing the planning skills of grade 4 and 5 Learners. International Journal of Disability, Development and Education, 65(4), 428–441. https://doi.org/10.1080/1034912x.2017.1406067

Boland, A., Cherry, M. G., & Dickson, R. (Eds.). 2017. Doing a systematic review: A student's guide (2nd ed.). Sage Publications, USA.

Bosma, T., Stevenson, C. E., & Resing, W. C. 2017. Differences in need for instruction: Dynamic testing in children with arithmetic difficulties. Journal of Education and Training Studies, 5(6), 132. https://doi.org/10.11114/jets.v5i6.2326

Caffrey, E., Fuchs, D., & Fuchs, L. S. 2008. The predictive validity of Dynamic Assessment. The Journal of Special Education, 41(4), 254–270. https://doi.org/10.1177/0022466907310366

Calero, M. D., Belen, G.-M. M., & Robles, M. A. 2011. Learning potential in high IQ children: The contribution of dynamic assessment to the identification of gifted children. Learning and Individual Differences, 21(2), 176–181. https://doi.org/10.1016/j.lindif.2010.11.025

Cho, E., Compton, D. L., Fuchs, D., Fuchs, L. S., & Bouton, B. 2014. Examining the predictive validity of a dynamic assessment of decoding to forecast response to tier 2 intervention. Journal of Learning Disabilities, 47(5), 409–423. https://doi.org/10.1177/0022219412466703

Cho, E., Fuchs, L. S., Seethaler, P. M., Fuchs, D., & Compton, D. L. 2020. Dynamic assessment for identifying spanish-speaking english learners' risk for mathematics disabilities: Does language of administration matter? Journal of Learning Disabilities, 53(5), 380–398. https://doi.org/10.1177/0022219419898887

Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Gilbert, J. K., Barquero, L. A., Cho, E., & Crouch, R. C. 2010. Selecting at-risk FIRST-GRADE readers for Early intervention: Eliminating false positives and exploring the promise of a two-stage

GATED screening process. Journal of Educational Psychology, 102(2), 327–340. https://doi.org/10.1037/a0018448

Elleman, A. M., Compton, D. L., Fuchs, D., Fuchs, L. S., & Bouton, B. 2011. Exploring dynamic assessment as a means of identifying children at risk of developing comprehension difficulties. Journal of Learning Disabilities, 44(4), 348–357. https://doi.org/10.1177/0022219411407865

Elliott, J. 2003. Dynamic assessment in educational settings: Realising potential. Educational Review, 55(1), 15–32. https://doi.org/10.1080/00131910303253

Elliott, J. G., Resing, W. C., & Beckmann, J. F. 2018. Dynamic assessment: A case of unfulfilled potential? Educational Review, 70(1), 7–17. https://doi.org/10.1080/00131911.2018.1396806

Ferreira, J. M. 2018. Inclusive Early Childhood Education and the role of peer interaction: Brazil and Finland in dialogue [Published Doctoral dissertation]. University of Tampere. University Press (Acta Electronica Universitatis Tamperensis 1947). https://trepo.tuni.fi/handle/10024/104470

Ferreira, J. M., Mäkinen, M., & Mäkihonko, M. 2022. Equality, equity, and the child's best interest as guiding principles in the implementation of the Finnish inclusive education system. In G. M. Kraemer, K. Wunder & L. B. Lopes, A educação das pessoas com deficiência: Desafios, Perspectivas e Possibilidades (pp. 134-159). Pimenta Cultural Publishing, Brazil.

Feuerstein, R., Klein, P. S., & Tannenbaum, A. J. 1999. Mediated Learning Experience: theoretical, psychological and learning implications. Freund Publishing House, Israel.

Feuerstein, R., & Feuerstein, R. S. 2001. Is dynamic assessment compatible with the Psychometric Model? Specific Learning Disabilities and Difficulties in Children and Adolescents, 218–246. https://doi.org/10.1017/cbo9780511526794.008

Feuerstein, R., Feuerstein, R., Falik, L. H., & Rand, Y. 2003. The dynamic assessment of cognitive modifiability: The learning propensity assessment device, theory, instruments, and techniques. International Center for the Enhancement of Learning Potential.

Feuerstein, R., & Lewin-Benham, A. 2012. What Learning Looks Like – Mediated Learning in Theory and Practice K-6. Teachers College Press, UK.

Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S. 2004. Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. Annals of Dyslexia, 54(2), 304–331. https://doi.org/10.1007/s11881-004-0015-y

Fuchs, D., Compton, D. L., Fuchs, L. S., Bouton, B., & Caffrey, E. 2011. The construct and predictive validity of a Dynamic assessment of young children learning to Read: Implications for RTI Frameworks. Journal of Learning Disabilities, 44(4), 339–347. https://doi.org/10.1177/0022219411407864

Fuchs, L. S., & Vaughn, S. 2012. Responsiveness-to-intervention. Journal of Learning Disabilities, 45(3), 195–203. https://doi.org/10.1177/0022219412442150

Gellert, A. S., & Elbro, C. 2017. Does a dynamic test of phonological awareness predict early reading difficulties? Journal of Learning Disabilities, 50(3), 227–237. https://doi.org/10.1177/0022219415609185

Gellert, A. S., & Elbro, C. 2018. Predicting reading disabilities using Dynamic assessment of decoding before and after the onset of reading instruction: A longitudinal study from kindergarten through Grade 2. Annals of Dyslexia, 68(2), 126–144. https://doi.org/10.1007/s11881-018-0159-9

Ghanbarpour, M. 2017. A qualitative meta-synthesis of research on Dynamic assessment of second/foreign language Learning: Implications for language teachers. Journal of Language Teaching and Research, 8(4), 731. https://doi.org/10.17507/jltr.0804.12

Grigorenko, E. 2009. Dynamic Assessment and Response to Intervention: Two Sides of One Coin. Journal of Learning Disabilities, 42(2), 111–132. https://doi.org/10.1177/0022219408326207

Hasson, N., & Dodd, B. 2014. Planning intervention using dynamic assessments: A case study. Child Language Teaching and Therapy, 30(3), 353–366. https://doi.org/10.1177/0265659014521642

Hattie, J., & Timperley, H. 2007. The power of feedback. Review of Educational Research, 77(1), 81-112. https://doi.org/10.3102/003465430298487

Hattie, J., & Clarke, S. 2019. Visible learning: Feedback. Routledge, USA. https://doi.org/10.4324/9780429485480

Haywood, H. C., & Tzuriel, D. 2002. Applications and challenges in dynamic assessment. Peabody Journal of Education, 77(2), 40–63. https://doi.org/10.1207/s15327930pje7702_5

Haywood, H., & Lidz, C. 2006. Dynamic Assessment in Practice: Clinical and Educational Applications. In Dynamic Assessment in Practice. Cambridge University Press, UK. https://doi.org/10.1017/CBO9780511607516

Haywood, H. C. 2012. Dynamic Assessment: A history of fundamental ideas. Journal of Cognitive Education and Psychology, 11(3), 217-229. http://dx.doi.org/10.1891/1945-8959.11.3.217

Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. 2018. Mixed methods appraisal tool (MMAT), version 2018. Registration of copyright, 1148552(10).

Jeltova, I., Birney, D., Fredine, N., Jarvin, L., Sternberg, R. J., & Grigorenko, E. L. 2011. Making instruction and assessment responsive to diverse students' progress: Group-administered dynamic assessment in teaching mathematics. Journal of Learning Disabilities, 44(4), 381–395. https://doi.org/10.1177/0022219411407868

Jensen, M., & Feuerstein, R. 1987. The Learning Potential Assessment Device: From philosophy to practice. In C.S. Lidz (Ed.), Dynamic assessment: An interactional approach to evaluating learning potential. Guilford Press, USA.

Kaniel, S., Tzuriel, D., Feuerstein, R., Ben-Shachar, N., & Eitan, T. 1999. Dynamic assessment: Learning and transfer abilities of Ethiopian Immigrants to Israel. In R.

Feuerstein, P. S. Klein & A. J. Tannenbaum, Mediated Learning Experience: theoretical, psychological and learning implications (pp. 179–209). Freund Publishing House.

Kirschenbaum, R. J. 1998. Dynamic assessment and its use with underserved gifted and talented populations. Gifted Child Quarterly, 42(3), 140–147. https://doi.org/10.1177/001698629804200302

Kong, J. E., & Orosco, M. J. 2016. Word-problem-solving strategy for minority students at risk for math difficulties. Learning Disability Quarterly, 39(3), 171–181. https://doi.org/10.1177/0731948715607347

Kozulin, A., & Presseisen, B. Z. 1995. Mediated learning experience and psychological tools: Vygotsky's and Feuerstein's perspectives in a study of student learning. Educational Psychologist, 30(2), 67–75. https://doi.org/10.1207/s15326985ep3002_3

Kozulin, A. 2011. Learning potential and cognitive modifiability. Assessment in Education : Principles, Policy & Practice, 18(2), 169–181. https://doi.org/10.1080/0969594X.2010.526586

Kozulin, A. 2015. Dynamic Assessment of Adult Learners' Logical Problem Solving: A Pilot Study with the Flags Test. Journal of Cognitive Education and Psychology, 14(2), 219–230. https://doi.org/10.1891/1945-8959.14.2.219

Kozulin, A., & Garb, E. 2016. Dynamic Assessment of EFL Text Comprehension. School Psychology International, 23(1), 112–127. https://doi.org/10.1177/0143034302023001733

Lawrence, N., & Cahill, S. 2014. The impact of dynamic assessment: An exploration of the views of children, parents and teachers. British Journal of Special Education, 41(2), 191–211. https://doi.org/10.1111/1467-8578.12060

Lee Swanson, H. 2011. Dynamic testing, working memory, and reading Comprehension growth in children with Reading Disabilities. Journal of Learning Disabilities, 44(4), 358–371. https://doi.org/10.1177/0022219411407866

Leeber, J., Candeias, A., & Grácio, L. 2011. With a different glance: Dynamic assessment of functioning of children oriented at development & inclusive learning. Antwerp: Garant Publishing.

Leeber, J., Partanen, P., Candeias, A., Grácio, M. L., Bohacs, K., Sonnesyn, G., Van de Veire, H., Van Trimpont, I., Orban, R., János, R., Demeter, K., Schraepen, B., & Dawson, L. 2013. The need for a more dynamic and ecological assessment of children experiencing barriers to learning to move towards inclusive education: A summary of results of the Daffodil Project. Transylvanian Journal of Psychology, Special Issue, 175-205.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. British Medical Journal, 339(7716), b2535. http://dx.doi.org/10.1136/bmj.b2535

Moscardini, L., & Moscardini, C. 2020. Dynamic assessment and teachers' knowledge of children's mathematical thinking: A case study in children's mathematics. Support for Learning, 35(4), 522–541. https://doi.org/10.1111/1467-9604.12331

Naglieri, J. A., & Das, J. P. 1997. Cognitive assessment system. Riverside Publishing Company.

Orosco, M. J., Swanson, H. L., O'Connor, R., & Lussier, C. 2011. The effects of dynamic strategic math on english language learners' word problem solving. The Journal of Special Education, 47(2), 96–107. https://doi.org/10.1177/0022466911416248

Orosco, M. J. 2014. Word problem strategy for latino english language learners at risk for math disabilities. Learning Disability Quarterly, 37(1), 45–53. https://doi.org/10.1177/0731948713504206

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. 2016. Rayyan—a web and mobile app for Systematic Reviews. Systematic Reviews, 5(1). https://doi.org/10.1186/s13643-016-0384-4

Petersen, D. B., & Gillam, R. B. 2015. Predicting reading ability for bilingual latino children using dynamic assessment. Journal of Learning Disabilities, 48(1), 3–21. https://doi.org/10.1177/0022219413486930

Poehner, M. E. 2018. Probing and provoking L2 Development: the object of mediation in dynamic assessment and mediated development. In J. P. Lantolf, M. E. Poehner, & M. Swain, The Routledge Handbook of Sociocultural Theory and Second Language Development (pp. 249-265). Routledge. https://doi.org/10.4324/9781315624747

Regalla, M., & Peker, H. 2017. Prompting all students to learn: Examining dynamic assessment of special needs and typical students in a prekindergarten inclusive French program. Foreign Language Annals, 50(2), 323–338. https://doi.org/10.1111/flan.12261

Rezaee, A. A., & Ghanbarpour, M. 2016. The Measurement Paradigm and Role of Mediators in Dynamic Assessment: A Qualitative Meta-Synthesis. International Journal of Language Studies, 10(4), 77–108.

Samran, M., & Mehdi, M. 2018. Dynamic Assessment in Vygotsky's Sociocultural Theory: Origins and Main Concepts. Journal of Language Teaching and Research, 9(3), 600-607. https://dx.doi.org/10.17507/jltr.0903.20

Schwarz, C. M., Hoffmann, M., Schwarz, P., Kamolz, L.-P., Brunner, G., & Sendlhofer, G. 2019. A systematic literature review and narrative synthesis on the risks of medical discharge letters for patients' safety. BMC Health Services Research, 19(1). https://doi.org/10.1186/s12913-019-3989-1

Seethaler, P. M., Fuchs, L. S., Fuchs, D., & Compton, D. L. 2012. Predicting first graders' development of Calculation VERSUS word-problem performance: The role of dynamic assessment. Journal of Educational Psychology, 104(1), 224–234. https://doi.org/10.1037/a0024968

Seethaler, P. M., Fuchs, L. S., Fuchs, D., & Compton, D. L. 2016. Does the value of Dynamic assessment in predicting end-of-first-grade mathematics performance differ as a function of English Language proficiency? The Elementary School Journal, 117(2), 171–191. https://doi.org/10.1086/688870

Snyder, H. 2019. Literature review as a research methodology: An overview and guidelines. Journal of business research, 104, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Svanes, I. K., & Skagen, K. 2017. Connecting feedback, classroom research and Didaktik perspectives. Journal of Curriculum Studies, 49(3), 334-351, https://doi.org/10.1080/00220272.2016.1140810

Swanson, H. L., & Lussier, C. M. 2001. A selective synthesis of the experimental literature on dynamic assessment. Review of Educational Research, 71(2), 321–363. https://doi.org/10.3102/00346543071002321

Tranfield, D., Denyer, D., & Smart, P. 2003. Towards a methodology for developing evidence-informed management knowledge by means of a systematic review. British Journal of Management, 14, 207–222. https://doi.org/10.1111/1467-8551.00375.

UNESCO. 2005. Guidelines for Inclusion. Ensuring Access to Education for All (Orientaciones para la inclusión. Asegurar el acceso a la Educación para Todos)

Vogelaar, B., Bakker, M., Elliott, J. G., & Resing, W. C. 2016. Dynamic testing and test anxiety amongst gifted and average-ability children. British Journal of Educational Psychology, 87(1), 75–89. https://doi.org/10.1111/bjep.12136

Vygotsky, L., 1978. Mind in Society. Harvard University Press.

Wagner, R., & Compton, D. 2011. Dynamic Assessment and Its Implications for RTI Models. Journal of Learning Disabilities, 44(4), 311–312. https://doi.org/10.1177/0022219411407859

Walqui, A. 2006. Scaffolding instruction for English language learners: A conceptual framework. International Journal of Bilingual Education and Bilingualism, 9(2), 159–180. https://doi.org/10.1080/13670050608668639

Watkins, A. 2007. Assessment in inclusive settings: Key issues for policy and practice. Odense. European Agency for Development in Special Needs Education.

Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., & Buntins, K. 2020. Systematic reviews in educational research: Methodology, perspectives and application. Springer. http://dx.doi.org/10.1007/978-3-658-27602-7

Appendix: Elements

a. Figures

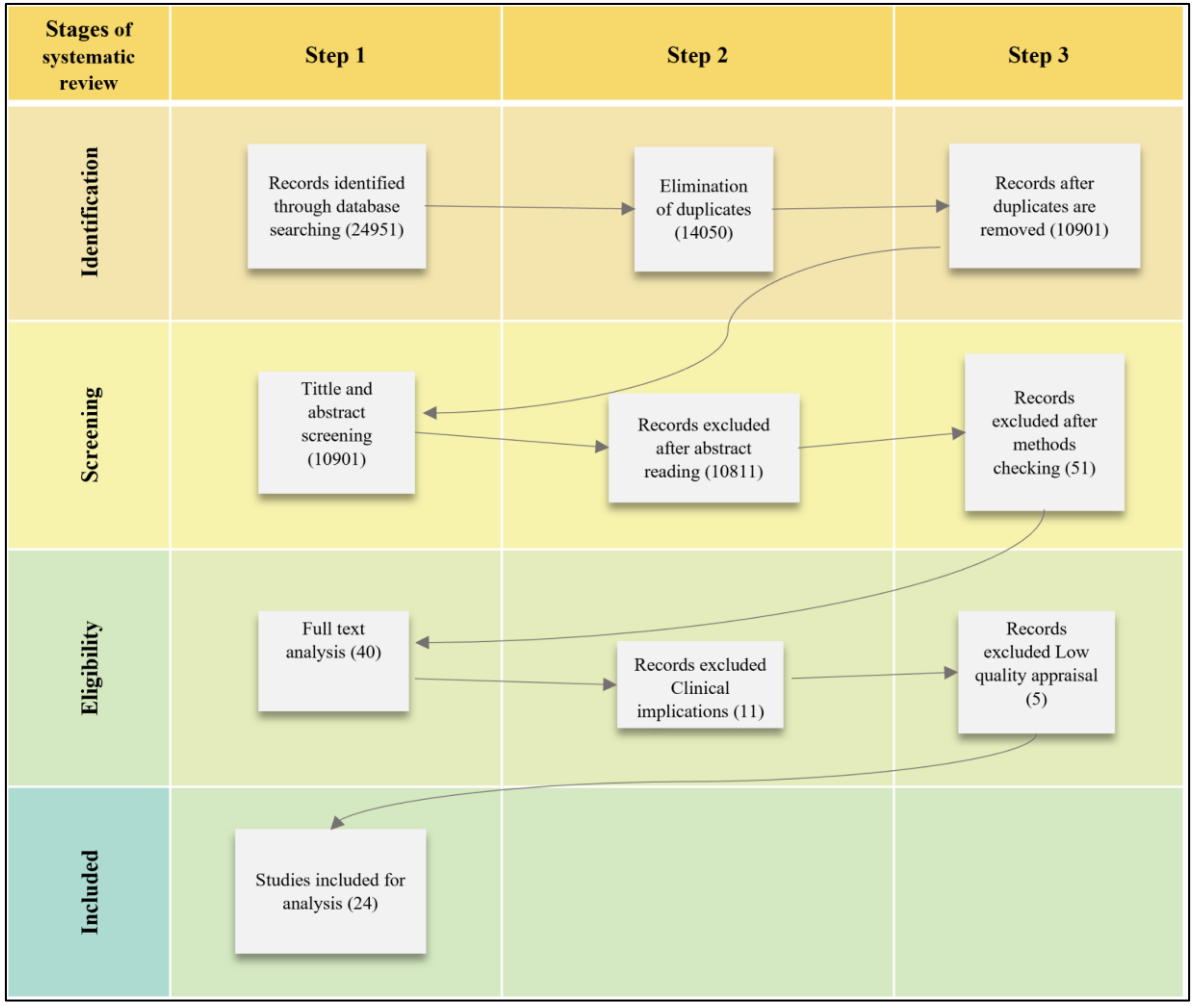

Figure 1: Flowchart for study inclusion

b. Tables

Table 1: Adapted Mixed Methods Appraisal Tool (MMAT, Hong et al., 2018)

Category of study designs	Methodological quality criteria	Responses		
		Yes	No	Cannot tell
1. Qualitative	1.1. Is the qualitative approach appropriate to answer the research question/ test the hypothesis/ achieve the aim?			
	1.2. Are the qualitative methods to collect data adequate to answer the research question/ hypothesis/ aim?			
	1.3. "Are the findings adequately derived from the data?" (Hong et al., 2018, p.2)			
	1.4. "Is the interpretation of results sufficiently substantiated by data?" (Hong et al., 2018, p.2)			
	1.5. Is the coherence between qualitative data sources, collection, analysis, and interpretation? (Hong et al., 2018, p.2)			
2. "Quantitative randomized controlled trials" (Hong et al., 2018, p.2)	2.1. Is randomization appropriately conducted?			
	2.2. "Are the groups comparable at baseline?" (Hong et al., 2018, p.2)			
	2.3. Are the outcome data completed?			
	2.4. "Are outcome assessors blinded to the intervention provided?" (Hong et al., 2018, p.2)			
	2.5. Did the participants adhere to the assigned intervention?			
3. "Quantitative nonrandomized" (Hong et al., 2018, p.2)	3.1. Do the participants represent the target population?			
	3.2. Are measurements suitable in terms of both the outcome and intervention/ DA process?			
	3.3. Are the outcome data completed?			
	3.4. "Are the confounders accounted for in the design and analysis?" (Hong et al., 2018, p.2)			
	3.5. During the research period, is the intervention/the DA process conducted according to the plan?			
4. Quantitative descriptive	4.1. Does the sampling strategy relate to answering the research question/ hypothesis/ aim?			
	4.2. Does the sample represent the target population?			
	4.3. Are the measurements suitable?			
	4.4. Is the risk of nonresponse bias poor?			
	4.5. Is the statistical analysis suitable to answer the research question/ test the hypothesis/ achieve the aim?			
5. Mixed methods	5.1. Is there an adequate explanation for utilizing a mixed-methods design to address the research question/ hypothesis/ aim?			
	5.2. Are the different components of the study effectively integrated to answer the research question/ test the hypothesis/ achieve the aim?			
	5.3. "Are the outputs of the integration of qualitative and quantitative components adequately interpreted?" (Hong et al., 2018, p.2)			

	5.4. "Are divergences and inconsistencies between quantitative and qualitative results adequately addressed"? (Hong et al., 2018, p.2)			
	5.5. "Do the different components of the study adhere to the quality criteria of each tradition of the methods involved?" (Hong et al., 2018, p.2)			

Table 2: Guiding coding scheme for data analysis of DA methods

Category	Codes	Code descriptions
DA Integration	1. DA is implemented solely	There is no specific teaching involved. DA is the **only assessment method** used.
	2. DA is integrated into a multidimensional assessment	DA is used together with other assessment methods, such as standardized tests.
	3. DA is integrated into a learning program	DA is used in the assessment process of a specific **Learning Program.**
	4. DA is a framework for the study	DA **guides interventions and assessment tools** throughout the study, but there isn't a specific assessment process involving DA
DA application	1. Individual sessions	DA is applied individually; the evaluator only assists one student at the time.
	2. Small group	DA is applied in groups of 3 to 6 students simultaneously. The evaluator assists all students at the same time.
	3. Classroom	DA is applied in the entire classroom. The evaluator assists all students in the class at the same time.
	4. Implemented by teachers	Teachers conduct the assistance in DA
	5. Implemented by trained evaluators	Trained evaluators conduct the assistance in DA
	6. Implemented by researchers	The researcher conducts the assistance in DA
DA types of assistance	1. Free-form cues, probes, or reminders	Context-based, **created by the evaluator at any given moment of the intervention**.
	2. Systematized and standardized prompts (e.g., graduated prompts)	Specific **standardized prompting protocol.** This means that the prompts used follow a pre-defined prompt.
	3. Instructional scaffolding	Intervention is **based on the content** of what is being evaluated. Scaffolding happens during the execution of the task or test.
	4. Mediated learning	Interaction is **based on the cognitive process (metacognition is involved)**. Mediation happens during the execution of the task or test.
	5. Feedback	**Insights are provided after the students have given their answers**. The feedback supports the regulation of future thinking processes.

Table 3: Summary of included studies and their respective methods

Study	Date/ Location	Participant	DA types of assistance	Purpose of DA	DA Application		DA integration
(1) Compton, Fuchs,. Fuchs, Bouton, Gilbert, Barquero, Cho, & Crouch	2010/U. S. A.	355 first graders, with low performing scores	- Instructional scaffolding - Systematized and standardized prompts	Predict second grade RD status	Implemented by trained evaluators	Individual sessions	DA is integrated into multidimensional assessment.
(2) Al-Hroub	2011/UK	5 gifted children with learning difficulties, 4-6 grades	- Feedback - Mediated learning	Assess learning potential	Implemented by teacher	Individual sessions	DA is integrated into multidimensional assessment.
(3) Aljunied & Frederickson	2011/Singapore	52 children with autism from both mainstream and special schools in Singapore, 8-12 years old	- Mediated learning - Free-form cues, probes, and reminders - Feedback	Assess gains in learning	Implemented by trained evaluators	Individual sessions	DA is integrated into multidimensional assessment.
(4) Calero, Belen, & Robles	2011/Spain	127 children, 64 high-IQ and 63 average-IQ, 7-11 years old	- Mediated learning - Systematized and standardized prompts	Identify gifted children, regardless of the IQ they show	not mentioned	not mentioned	DA is integrated into multidimensional assessment.
(5) Elleman, Compton, Fuchs, Fuchs, & Bouton	2011/U.S.A.	100 second graders including high, average, and low performers in reading, 12% received special education services	Systematized and standardized prompts	Predict students' reading disabilities	Implemented by trained evaluators	Individual sessions	DA is integrated into multidimensional assessment.
(6) Fuchs, Compton, Fuchs, Bouton, & Caffrey	2011/U.S.A.	318 first graders including high, average, and low performers in reading	- Instructional scaffolding - Systematized and standardized prompts	Predict responsiveness to reading instruction	Implemented by trained evaluators	Individual sessions	DA is integrated into multidimensional assessment.
(7) Jeltova, Birney, Fredine, Jarvin, Sternberg, & Grigorenko	2011/U.S.A.	-1,332 students from 4 ethnic groups. Some participants have learning disabilities or low SES.	Mediated learning	Evaluate a group of students in regular classrooms while teaching	Implemented by teacher	Classroom	DA is integrated into a learning program.

		-63 classroom teachers					
(8) Lee Swanson	2011/U.S.A.	78 students, including students with RD merely; students struggling with both reading and calculation; low verbal IQ readers, and skillful readers Mean age: 11.6 (SD = 2.28)	Systematized and standardized prompts	Predict later reading comprehension performance	Implemented by trained evaluators	Individual sessions	DA is integrated into multidimensional assessment.
(9) Orosco, Lee Swanson, O'Connor, & Lussier	2011/U.S.A.	6 Latino ELLs at risk for MD Grade: 2	- Instructional scaffolding - Systematized and standardized prompts - Feedback	DA as a framework for developing a word-problem-solving strategy	Implemented by trained evaluators	Individual sessions	DA is integrated into a learning program.
(10) Seethaler, S. Fuchs, Fuchs, & Compton	2012	184 high, average, and at-risk students in math, in which 9 students received special education services a learning, speech, or language disability. Grade: 1	- Instructional scaffolding - Systematized and standardized prompts	Predict 1st-grade mathematics development	Implemented by trained evaluators	Individual sessions	DA is integrated into multidimensional assessment.
(11) Cho, Compton, Fuchs, Fuchs, & Bouton	2014	134 low performing students in reading Grade: 1	- Instructional scaffolding - Systematized and standardized prompts	Identify tier 2 nonresponders	Implemented by trained evaluators	Individual sessions	DA is integrated into multidimensional assessment.
(12) Hasson & Dodd	2014	1 student with a specific language impairment Age: 9	- Mediated learning - Systematized and standardized prompts	Evaluate learning potential to guide intervention	-not mentioned	Individual sessions	DA is integrated into multidimensional assessment.
(13) Orosco	2014/U.S.A.	6 Latino ELLs at risk for MD Grade: 3	- Instructional scaffolding	DA as a framework for developing a word-problem-solving strategy	Implemented by trained evaluators	Individual sessions	DA is a framework for the study

			- Systematized and standardized prompts - Feedback				
(14) Petersen & Gillam	2015/U.S.A.	63 bilingual Latino kindergarten students "at risk for language impairment" (Petersen & Gillam, 2015, p.4)	- Feedback - Instructional Scaffolding - Systematized and standardized prompts	Predict reading ability in bilingual students	Implemented by trained evaluators	Not mentioned	DA is integrated into multidimensional assessment.
(15) Kong & Orosco	2016/U.S.A.	8 minority students at risk for MD Grade: 3	- Instructional scaffolding - Systematized and standardized prompts - Feedback	DA as a framework for developing a word-problem-solving strategy	Implemented by trained evaluators	Small group	DA is a framework for the study
(16) Seethaler, Fuchs, Fuchs, & Compton	2016/U.S.A.	129 LEP qualified to receive Second Language services, and 163 non-LEP children Grade: 1	- Instructional scaffolding - Systematized and standardized prompts	Predict individual differences in year-end performance on WP and calculation	Implemented by trained evaluators	Individual sessions	DA is integrated into a multidimensional assessment
(17) Vogelaar, Bakker, Elliott, & Resing	2016/The Netherlands	113 gifted and average-ability children Age: 7 and 8 years	- Systematized and standardized prompts - Mediated learning	Measure the cognitive abilities of intellectually gifted children	not mentioned	not mentioned	DA is implemented solely.
(18) Amod, Heafield, & Seabi	2017/South Africa	51 students presenting a variety of obstacles to learning Grade: 4-5	Mediated learning	Measure students' learning potential	Implemented by trained evaluators	Small group	DA is implemented solely.
(19) Bosma, Stevenson, & Resing	2017/The Netherlands	120 low-performing students Grade: 2	- Systematized and standardized prompts - Feedback	Identify differences in instructional needs	not mentioned	Individual sessions	DA is implemented solely.

(20) Gellert & Elbro	2017/Denmark	160 kindergarten students, including 76 children at risk for RD and 84 not-at-risk children	-Instructional Scaffolding - Systematized and standardized prompts - Feedback	Predict children's early reading development	Implemented by trained evaluators	Individual sessions	DA is integrated into a multidimensional assessment
(21) Gellert & Elbro	2018/Denmark	158 kindergarten students. The majority were students at risk for reading difficulties.	- Instructional Scaffolding - Systematized and standardized prompts - Feedback	Predict difficulties with reading accuracy	Implemented by trained evaluators	Individual sessions	DA is integrated into a multidimensional assessment
(22) Al-Hroub & Whitebread	2019/Jordan	30 students with high potential in mathematics and learning difficulties Grade: 5-6	Feedback	Identify the students' mathematical giftedness in twice-exceptional learners	Implemented by teachers	Individual sessions	DA is integrated into a multidimensional assessment.
(23) Cho, Fuchs, Seethaler, Fuchs, & Compton	2020/U.S.A.	368 students with LEP in which 15 students received special education services Grade: 1	- Instructional Scaffolding - Systematized and standardized prompts	Predict later-emerging mathematics disabilities as a supplement screener	Implemented by trained evaluators	Not mentioned	DA is integrated into a multidimensional assessment.
(24) Moscardini & Moscardini	2020/	1 low-performing student in math Age: 8	Free-form cues, probes, and reminders	Inform instructional decisions	Implemented by teachers	Individual sessions	DA is integrated into a learning program.

***Note:** *WP* stands for the word problem, *LEP* limited English proficiency, *ELLs* English language learners, *RD* reading disabilities, *MD* mathematics difficulties, *low SES* low-socioeconomic status, *RTI* responsiveness to intervention, *L1* the first language, which in this context is Spanish, *DA* dynamic assessment, *MLE* Mediated Learning Experience, *LD* learning difficulties.The studies are numbered to facilitate referencing in the results section.

Table 4: Description of practical challenges in the application of DA and possible solutions pointed out by empirical studies

Study	Challenges	Solutions
Problems relating to the using DA as an assessment method in research		
Cho et al., 2014	Challenge 1: The construct validity of DA Challenge 2: Whether the emphasis of instruction moderates DA's predictive validity.	Solution for challenge 1: 1. "Using a structural equation modeling approach to derive a latent construct of responsiveness by using multiple indicators of responsiveness" (Cho et al., 2014, p.421) 2. "Testing whether decoding and responsiveness represent a unitary or distinct construct using static and decoding measures of the same construct, decoding" (Cho et al., 2014, p.421).
Fuchs et al., 2011	The floor effect of DA with few low-performing participants	2 solutions: 1. To replace a less difficult task for the CVC(C)ing task. 2. To "keep the CVC(C) ing task but to give children more helpful guidance—more explicit information—earlier in the sequence of graduated prompts" (p.346).
Problems relating to applying DA		
Jeltova et al., 2011	Using a static test as a criterion measure make DA a traditional static form of assessment.	Not mentioned
Petersen & Gillam, 2015	During the post-test, in some cases, children were reluctant and did not want to conduct the test independently.	Assessor offered some neutral prompts, for instance, "You can guess—what do you think this word says?" or "Remember what I told you?"
Aljunied & Frederickson, 2011	"Providing prompts, cues, or reminders in mediation may impose greater demands on central coherence as an increasing difficulty is experienced. Weak central coherence may make the gains less effective. The use of DA may underrate the learning potential of children with autism" (p.182)	Developing specific mediation strategies that can minimize the impacts of weak central coherence.
Bosma et al., 2017	The multiple protocols are difficult for evaluators to learn, apply, and score.	Increase training, computerized scoring, and more evaluators
Elleman et al., 2011	There are uncertainties about the effectiveness of the prompts, and whether or not they are supporting or hindering reflective processes	Not mentioned