iejee

June 2023, Volume 15, Issue 5

www.iejee.com

# international electronic journal of elementary education

KURA group Education & Publishing

# INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION




ISSN: 1307-9298

www.iejee.com
iejee@iejee.com

KURA group
Education & Publishing

iejee

International Electronic Journal of Elementary Education
June 2023, Volume 15, Issue 5

# Table of Contents




ISSN: 1307-9298

iejee

International Electronic Journal of Elementary Education
June 2023, Volume 15, Issue 5, 371-382

# Teacher Professional Development for STEM Integration in Elementary/ Primary Schools: A Systematic Review

Tugba Boz*

Received : 27 October 2022
Revised : 5 April 2023
Accepted : 17 June 2023
DOI : 10.26822/iejee.2023.306

* **Correspondance Details:** Tugba Boz, Department of Educational Theory and Practice, University of Georgia, Athens GA, United States.
E-mail: tugba.boz@uga.edu
ORCID: https://orcid.org/0000-0002-8020-127X

## Abstract

There has been a recent increase in studies devoted to integrated science, technology, and engineering mathematics (STEM) education. However, as the field is emerging, there has not been a large amount of scholarship on how to best provide teacher professional development (PD) on integrated STEM education. On the other hand, well-grounded research is available in the field of effective teacher professional development. Therefore, in this study, a systematic review of empirical studies on training elementary/primary school teachers in integrated STEM education was undertaken. Using an adapted version of Lawless and Pellegrino's analytical framework (2007), the emerging commonalities among eleven studies in conjunction with the literature on effective teacher professional development were analyzed and discussed. This study aims at connecting current STEM-integrated PD activities with the literature on effective PD and laying the groundwork for future professional development programs and evaluations in integrating STEM subject areas in elementary or primary schools.

**Keywords:**

Teacher Professional Development, Integrated STEM, Review of Literature, Elementary Teachers

## Introduction

In recent years, there has been an increase in studies dedicated to exploring the conceptual, theoretical, and practical implications of integrated STEM education which refers to "the approach to teaching the STEM content of two or more STEM domains, bound by STEM practices within an authentic context for the purpose of connecting these subjects to enhance student learning" (Kelly & Knowles, 2016, p.3). The literature (e.g., English, 2016; Nadelson & Seifert, 2017; Pearson, 2017) argues that integrated STEM education increases students' interest in and engagement with STEM fields and better prepares students for the future STEM workforce in which disciplines are interrelated and integrated. This attention to integrating STEM subject areas has brought about the development and implementation of integrated STEM programs in real classrooms. Engineering design projects or the development of robotics projects in science and mathematics classrooms are a few examples. As there is more emphasis on initiating integrated STEM

education at early ages (e.g., Cunningham, 2018), programs have been designed and implemented to train elementary school teachers to implement integrated STEM education in their classrooms. However, since the field is emerging, there has been a limited amount of research on teacher professional development in integrated STEM education. On the other hand, the literature provides well-grounded and comprehensive answers on the essential characteristics of effective teacher professional development (PD) and their effects on teacher beliefs and classroom practices.

Therefore, this study begins with an in-depth literature review on the key characteristics of effective teacher professional development (PD). The review draws on highly influential and frequently cited studies, authored by the leading figures in the field systematic review is then presented to identify the current empirical studies on professional development activities in integrated STEM education. The discussion focuses on the connection between the current STEM-integrated professional development activities and the literature on effective PD characteristics. This inquiry is aimed to lay the groundwork for future professional development programs and evaluations in STEM integration since, as Desimone and Garet (2015) argued, specific inquiries in teacher PD research are necessary to "translate the broad features into specific, effective activities in varying contexts" (p. 260). Furthermore, this study aims to provide a framework that outlines the effective components of professional development in the integration of STEM subjects and presents guidelines for program/project developers and policymakers in their consideration of teacher professional development in STEM integration.

## Overview of Teacher Professional Development

### *Definition and History*

Professional development (PD) is an activity that aims to improve paid staff members' performance (Little, 1987). It is broadly defined as teachers' professional growth resulting from their formal or informal experiences, such as attending workshops, professional meetings, mentoring, reading teacher magazines, and watching documentaries (Villegas-Reimers, 2003; OECD, 2019). Some scholars prefer using the term professional learning instead of professional development to emphasize "the nature and ownership of teacher learning" and position teachers as "the key decision makers about what matters for their growth" (Smith, 2017, p.2). According to Borko (2004), teacher learning can occur in various contexts, such as brief hallway conversations or counseling a troubled child after school. his broad definition of PD recognizes that professional development can occur in multiple settings and ways."

It is worth noting that teacher professional development or learning has not always been defined and considered in such a broad way or as an ongoing process. In the first half of the 20th century, research on teacher professional development began with large survey studies (Cochran-Smith, 2009). Starting in the 1970s, a new approach started to govern teacher education. According to Cochran-Smith and Demers (2008), this new model was "more constructivist than transmission-oriented—the recognition that both prospective and experienced teachers (like all learners) brought prior knowledge and experience to all new learning situations, which are social and contextually specific" (p.1011). Therefore, in this new model, teacher education was not a one-time event but an ongoing and complex process which "was marked by an equally important set of factors" embedded in teachers' everyday practices and interactions (Avalos, 2011, p.17).

Today, it is widely accepted that teacher professional development plays a significant role in enhancing the quality of school, influencing teachers' beliefs and practices (Desimone, 2009, 2011; Garet et al., 2001; Villegas-Reimers, 2003), and improving student achievement (Desimone et al., 2005). Effectively designed and implemented PD programs can help teachers in their day-to-day work (Fullan & Miles, 1992). Therefore, PD is not only a government or state-mandated mass learning practice but is also highly desired by teachers.

### *Characteristics of Effective Professional Development*

The literature proposes well-grounded recommendations for designing and delivering effective teacher PD. One such recommendation is to focus on active learning in PD, as research indicates thar active learning is more likely to result in intended teacher knowledge and skills (Bates & Morgan, 2018; Garet et al., 2001; Penuel et al., 2007). In other words, PD should offer teachers many opportunities to observe the modeling of intended changes in instructional practices, incorporate lesson planning and enactments, and practice and reflect on them. (Birman et al., 2000; Darling-Hammond et al., 2017; Darling-Hammond & Richardson, 2009; Darling-Hammond & McLaughlin, 1995; Desimone & Pak, 2017; Desimone, 2009; Desimone et al., 2013; Knapp, 2003; Popova et al., 2018). Additionally, PD should encourage teachers to discuss their work and their students' work, in order to foster collaboration and shared learning (Desimone et al., 2002).

Second, PD should focus on student learning-oriented teaching practices and activities (Guskey, 2000; Darling-Hammond & Richardson, 2009; Desimone et al., 2013; Knapp, 2003): As widely discussed in the literature, the main objective behind PD is to improve student learning. Guskey (2000) emphasizes that all PD

efforts should primarily focus on learning and learners, and that even in planning PD, educators should first decide on the specific student learning outcomes they expect to achieve (Guskey, 2014). Knapp (2003) claimed that professional learning experiences should concentrate on classroom teaching and the evidence of student learning (e.g., Gupta & Lee, 2020). Additionally, analyzing student work can be incorporated into PD and used to develop a shared understanding among teachers of what constitutes good work, common misconceptions students may have, and which instructional strategies are effective (Darling-Hammond & Richardson, 2009).

Third, PD should have a specific subject focus (Desimone et al., 2013; Garet et al., 2001; Villagas-Reimeners, 2013): Popova et al. (2018) have indicated that having a specific subject focus is positively associated with student learning. Villagas-Reimeners (2013) has indicated that a specific subject matter is important to choose for PD because the content, design, and implementation of PD vary according to the subject matter. For example, Villagas-Reimeners found that concentrated time for PD is more effective for mathematics, while distributed time is more effective for science.

Fourth, PD should be a collaborative and collegial act (Darling-Hammond & Richardson, 2009; Desimone et al., 2013; Garet et al., 2001; Knapp, 2003): Collaborative and collegial learning environments create communities of practice and help improve/increase the change process beyond individual classrooms. However, for such collaborative interactions to emerge, professional development leaders should ensure a safe place to nurture trust and critical dialogue (Borko, 2004).

Fifth, teacher learning in PD should be a coherent part of and seamlessly linked to the curriculum, assessment, and standards in place (Darling-Hammond & Richardson, 2009; Desimone, 2009): Desimone and Garet (2015) emphasized the importance of linking professional development (PD) explicitly to curriculum and classroom realities, as this connection can improve the success of PD initiatives. Additionally, PD programs should be designed to ensure that there are no disparities between what teachers learn and what they can implement in their classrooms (Desimone, 2009). To achieve this goal, PD should be connected to other changes that are already in place within schools (e.g., Garet et al., 2001). By making these connections, teachers can more easily integrate new knowledge and skills into their classroom practices.

Finally, PD should be sustained over time through on-demand in-site expert support and follow-ups (Darling-Hammond et al., 2017): While there is no consensus on the duration of PD required to lead to positive outcomes, some scholars suggest that 80 hours (Darling-Hammond & Richardson, 2009), 20 hours (Desimone, 2009), or even 8 hours (Parsad, 2001) can yield improvements in teaching in real-world classrooms. It is worth noting that the effectiveness of PD is also influenced by the content, format, and quality of the activities offered to teachers (OECD, 2019). Therefore, PD programs should offer teachers a variety of activities that allow for active learning, such as observation of effective teaching practices, lesson planning, and opportunities for reflection and collaboration with peers.

As Darling-Hammond (2009) claimed, the academic community is on its way toward reaching a consensus on the content, design, and implementation of effective professional development. However, contextual and background factors should also be considered when designing and delivering PD programs (e.g., Altun et al., 2021). Research suggests that teacher motivation plays an important role in the success of PD initiatives (Guskey, 2002; Shin & Jun, 2019). The content of PD also influences the impact of PDs on teacher change. According to Desimone and Garet (2015), it is easier for teachers to change procedural classroom behavior than to acquire inquiry-oriented instructional techniques. Boyd et al. (2009) claimed that professional development that focuses on practical aspects of teaching is more likely to produce positive effects on students than professional development that primarily emphasizes teacher behaviors. It seems that the existing body of literature on teacher professional development is still growing. Still, available literature pinpoints the ways in which PD evaluations can be undertaken and PD effectiveness can be determined.

### *Evaluating the Effectiveness of Professional Development*

As argued by Knapp (2003), "the most immediate target of professional development is pro-learning" (p.112). Therefore, PD evaluations aim to explore the "changes in the thinking, knowledge, skills, and application that form practicing teachers' or administrators' repertoire" P.112). Historically, many early PD evaluations consisted of self-reported changes in the participating teachers' behavior and knowledge at the end of a workshop or series (Desimone, 2011). Some scholars (e.g., Guskey & Sparks, 1991; Smylie, 1989) subsequently made this criticism and called for more in-depth studies on the subject. Today, scholars often suggest that rigorous evaluations should be done during the process of planning and during the process of delivery through formative and summative assessments. For example, Guskey (2014) listed the following questions to be considered in planning a PD: "Is this experience or activity leading to the intended results? Is it better than what was done in the past? Is it better than another, competing activity? Is it worth the costs?" (p.1219). He

also added that "Answers to these [the following] questions require more than a statement of findings. They demand an appraisal of quality and judgments of value, based on the best evidence available" (p.1219).

Additionally, more rigorous evidence is expected from researchers or PD practitioners about evaluation. It is recommended that these evaluations should be rigorously handled, including the use and administration of many tools (Desimone, 2009, 2011; Guskey, 2014). Depending on the goal of PD, questionnaires, structured interviews, lesson plans, "oral or written personal reflections, or examinations of participants' journals and portfolios...direct observations, either by trained observers or using digital recorders" are recommended in the literature to be employed in evaluations (Guskey, 2014, p.1227). Furthermore, more agents are expected to be involved, such as teachers, administrators, parents, and/or students (Guskey, 2014). Furthermore, student learning is recommended as a priority in PD evaluations. However, student learning outcomes are not only confined to increases in test scores or cognitive outcomes, but affective and behavioral outcomes should also be considered (Guskey, 2002), such as "their perceptions of teachers, fellow students, and themselves; their sense of self-efficacy and their confidence in new learning situations can be especially informative" (Guskey, 2014, p.1228).

### *Purpose of the Study*

After elaborating on the elements of effective professional development and evaluation, the author now turns to what is considered effective in PD planning, implementation, and evaluation in STEM integration. For this particular study, a broad and widely-agreed upon the conceptualization of integrated STEM education was adopted: STEM integration is the intentional and explicit connection of STEM disciplines—science, mathematics, engineering, and technology (English, 2016; Kelly & Knowles, 2016; National Academy of Sciences, 2014; Pearson, 2017). Under this concept of integrated STEM education, the following research questions were posed:

1. What is the current state of teacher PD in integrated STEM education in elementary schools?
2. What connections are present between the current state of PD in integrated STEM education and the literature on effective teacher PD in terms of types, design and evaluation of PD effectiveness?
3. What implications/recommendations can be offered to improve the current state of PD in integrated STEM education for elementary teachers based on the recommendations offered by the literature on effective teacher PD?

## Methods

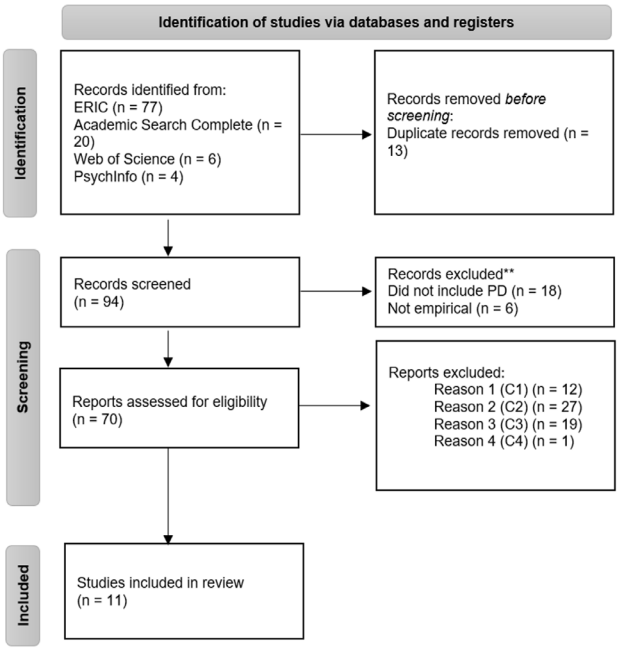
### *Screening and Eligibility Process of the Review*

The main databases in educational research (ERIC, Web of Science, PsychInfo, Academic Search Complete) were used to find the empirical articles published in peer-reviewed journals published between 1999 to 2022. Empirical was defined as having qualitative and quantitative data collection and analysis. Generalists: Elementary classroom teachers (K-5) and primary school teachers (K-6) were targeted. Using the Boolean operators, the following keywords and a combination of keywords were used: (STEM integration* OR interdisciplinary STEM) AND (professional development OR continuing training) AND (elementary teachers* OR primary teachers*). Based on this initial query, 94 articles were retrieved after removing duplicates. For this study, abstracts were examined based on the following criteria:

> *C1. Integration was aimed extensively and intentionally in PD, and PD was the reviewed study's focus.*
>
> *C2. More than one subject area was explicitly targeted and was used as integral to the activities rather than peripheral. For example, technology or engineering alone was not considered integrated.*
>
> *C3. Only in-service elementary or primary school teachers were targeted.*
>
> *C4. The PD aimed to change teacher knowledge, skills, or behaviors.*

At the end of this review process, 11 studies were included in this current review (see Table 1 below).

**Figure 1.**
*Identification of the Reviewed Studies (Page et al., 2021)*



*Analytical Framework and Analysis*

Lawless and Pellegrino (2007) suggested a schema to review technology TPD programs. In this schema, they recommended evaluating PD programs or research studies based on:

> *Three critical dimensions ... [which are] [a] type of professional development, which includes issues of delivery, duration, and content... [b] the unit (or units) of analysis that serves as the focus of any research/ evaluation of the outcomes and efficacy of that program ... [c] the nature of the research/evaluation study design and method, above and beyond the issues of a unit of analysis and measures. (p. 582)*

In this study, the schema proposed by Lawless and Pellegrino was partially adapted. The investigation focused on (a) the types of professional development, including duration, content, and delivery, (b) the unit of analysis or the outcomes evaluated in the professional development, such as student, program, or teacher outcomes, and (c) the designs and methods employed to research the effectiveness of professional development in integrated STEM education. In order to ensure the study's rigor and repeatability (Xiao & Watson, 2019), the author utilized the PRISMA framework for finding and selecting the reviewed studies and strictly followed the schema by Lawless and Pellegrino. Thematic analysis (Joffe, 2011) was employed to identify the main themes of the schema in the 11 reviewed studies, and the similar and different themes for each component of the schema were then reported. The following section discusses the similarities and the differences among the 11 studies in terms of their type, unit of analysis, and design of professional development (PD) programs. Then, the connections between these STEM-integrated PD programs and the literature on effective PD characteristics are explored. Finally, the implications are presented.

**Table 1**

*Lists of Empirical Studies on STEM integrated Professional Development in K-5.*

| Study | Duration & Delivery | Participants | Design & Methods |
|---|---|---|---|
| Baxter et al., 2014 | A two-day workshop in the summer, a six-day workshop during the school year, and two classroom observation sessions for ten contact days | 44 K-5 classroom teachers | Field notes of workshops, an online survey of teacher confidence in integrated mathematics and science, self-reported responses to changes in practice |
| Boice et al., 2021 | Five-week summer PD, ongoing support during the school year, | 17 STEM and art teachers from nine elementary schools | Surveys focus groups, written reflections |
| Cook et al., 2020 | Two-year PD program that included approximately 130 hours of PD during the school year. | 25 classroom teachers | Lesson plans, reflections |
| Guzey et al., 2016 | A three-week summer institute and year-long support from a coach through monthly meetings | 48 K-8 teachers from three large school districts | The researchers developed a STEM integration assessment tool |
| Havice et al., 2018 | Two-day PD | 33 K-5 teachers | Pre-Post Survey |
| McFadden and Roehrig, 2017 | Three-week intensive summer PD, developing STEM-integrated curricular units in teams, pilot teaching over the university-based summer camp and in their classrooms over the following academic year | Four elementary teachers in one group and three elementary teachers in the other group | Participant reflections, field notes, recordings of PD days, interviews, team conversations, curriculum development artifacts |
| Parker et al., 2015 | Two-week summer workshop, application during 4-week summer school in 22 schools | Six K-5 teachers from each of the 22 participant schools | Focus group interviews and teacher surveys over the academic year |
| Rich et al., 2017[a] | Year-long series of weekly PD followed by classroom practices (Year 1 of a longitudinal study) | 27 teachers of K-6 | Teacher self-efficacy and beliefs survey for integrating computing and engineering, Interviews and documented observations |
| Rich et al., 2018[a] | Weekly PD meetings for 45 minutes with a PD researcher on Friday afternoons over an academic year (Year 2 of a longitudinal study) | 17 elementary teachers | Interviews at the end of the lesson study/application, PD practitioners' fieldnotes |
| Sias et al., 2017 | One-week-long PD | 39 teachers of the third to fifth grades | all 39 lesson plans were coded/analyzed by the researchers developing rubric. |
| Wentworth and Monroe, 2011 | Series of PD over two years (Details are not specified) | 38 in-service elementary teachers | Ten lesson plans (along with enactments) produced by 38 teachers in groups of three to four; coded by two mathematics instructors based on the four criteria outlined at the beginning of the given PD |

Notes. [a] = As these two papers report on the same PD, they will be considered together in reporting their findings.

## Findings and Discussion

In this section, the similarities and the differences among the 11 studies regarding the type, unit of analysis, and the design features of their teacher professional development (PD) programs. For example, under the "type" of PD theme, the reviewed studies were categorized and presented based on the delivery medium employed, the duration of the PD offered, and the content (curriculum). The "unit of analysis" theme revealed the tools employed in the reviewed studies to evaluate the effectiveness of PD programs. The connections between these STEM-integrated PD programs and the literature on effective PD traits were presented for each theme (type, design, and unit of analysis).

### *Type of PD: Delivery, duration, and content*

#### *Reviewed Studies*

Regarding the delivery and duration, the majority of the reviewed PD studies were delivered during the summer months and consisted of workshops lasting from two days to five weeks. In addition, the majority of PD providers included year-long support for teachers to ensure ongoing engagement with the PD content and implementation in their classrooms. For example, Baxter et al. (2014) provided a two-day summer PD and workshops for six days during the school year. In addition, their participants engaged in two classroom observations for ten contact days over the year. Guzey et al. (2016) offered a three-week-long summer PD activity and continued to support teachers through monthly meetings throughout the year. Rich et al. (2017; 2018) did not hold a summer academy but provided teachers with weekly 45-minute workshops throughout the school year. Two reviewed programs (McFadden & Roehrig, 2017; Parker et al., 2015) employed one- to five-week-long workshops followed by classroom implementation during the academic year. Only one of the reviewed PD programs (Havice et al., 2018) offered a two-day workshop without any follow-up activity, and they discussed that their two-day-long PD was successful as there were successful changes in their participants' understanding of integrated STEM education.

Regarding the content and curriculum, all the reviewed PD programs aimed to address the existing grade-level curriculum and standards. This aspect of STEM-integrated PD literature agrees with the literature (e.g., Desimone, 2009), as PDs should be naturally linked to the curriculum and standards in place. In addition, as offered by the literature on effective PD, some PD programs included community-building activities such as field trips and family STEM activities.

Some studies emphasized particular STEM domains to be integrated. For example, McFadden & Roehrig (2017) focused on the integration of elementary engineering into content areas and standards in place (math, science, and technology). Two of the PD programs were geared toward integrating two domains, mathematics and science (Baxter et al., 2014), or technology and mathematics (Wentworth & Monroe, 2011). In the PD workshops by Rich et al. (2017), teachers were introduced to the integration of engineering and computing units, and in their 2018's study, their participant teachers worked on incorporating engineering or computing into their teaching through lesson study. Parker et al. (2015) focused on fully integrating science with reading, math, and engineering. All studies encouraged integration to align with the existing grade-level curriculum and standards.

The majority of the reviewed PD programs consisted of introducing example units and lessons to their participant teachers, such as Elementary is Engineering units, researchers-developed integrated STEM lesson plans (McFadden & Roehrig, 2017; Parker et al., 2015; Rich et al., 2017; Wentworth & Monroe, 2011). In Boice et al.'s study (2021), teachers felt more comfortable surrounded by support staff. When they worked with innovators outside the field of education and collaborated to develop integrated STEM lessons, they were inspired to prepare creative integrated STEM lesson plans. Three of the reviewed PD programs (Boice et al., 2021; Sias et al., 2017) ran special events such as field trips to local companies or engineering design facilities, engineering panels, STEM-related conferences, and family science activities.

Overall, the reviewed programs provided some details about the content of the PD activities. However, the majority of the reviewed PD programs did not provide sufficient information about the aspects that could make critical differences in the implementation and consequences of PD programs, such as PD instructors' background, available resources (material and staff capacity), presence/types of instruction and activities during PD.

#### *Connections with the literature on effective PD*

The reviewed studies align with the literature on effective PD that recommends PD to be run over the course of time with an on-site support (e.g., Darling-Hammond & Richardson, 2009; Penuel et al., 2007). However, providing encompassing PD programs is related to the availability of human, physical and financial resources, and as observed in Havece et al.'s study (2018), short workshops might still enhance teachers' understanding of STEM integration and support teachers to become comfortable with integrated STEM curriculum.

All the reviewed PD programs aimed to address the existing grade-level curriculum and standards. This

aspect of STEM-integrated PD literature agrees with the literature (e.g., Desimone, 2009), as PDs should be naturally linked to the curriculum and standards in place. In addition, as offered by the literature on effective PD, some PD programs included community-building activities such as field trips and family STEM activities.

The reviewed studies did not emphasize some design features of effective PD. For example, the following concerns were not located in the reviewed studies: how did the content present in PD? What types of inquiry prompts and tasks were presented? What was the level of modeling provided? What kind of opportunities were teachers granted for reflection on their work or their students' work? In particular, the reviewed studies do not explicitly discuss whether and how they employed active learning in their PD setting, an element frequently acknowledged as crucial for effective PD (Bergh et al., 2014).

*Unit of analysis*

*Reviewed Studies*

Researchers utilized a variety of tools to evaluate the effectiveness of their PD programs, including analysis of lesson plans created by participant teachers, pre-post surveys, field notes, interviews, and teacher debriefs. However, only one study focused on evaluating program outcomes (Parker et al., 2015), while the others evaluated teacher outcomes, such as changes in beliefs, understandings, self-reported knowledge, perceptions, confidence, self-efficacy, and self-reported levels of practice/implementation.

None of the reviewed studies collected student data to assess changes in knowledge, skills, or attitudes resulting from their teachers' participation in integrated STEM programs. Only Baxter et al. (2014) included student work in their PD to allow participant teachers to reflect on students' understanding of integrated science and mathematics concepts, but it was unclear whether participant teachers evaluated their students' work. Additionally, none of the studies included external evaluators to assess the effectiveness of their PD practices.

*Connections with the literature on effective PD*

The PD literature encourages the use of multiple tools and methods to evaluate changes in teachers' knowledge, skills, attitudes, classroom teaching, and student learning. The majority of the reviewed PD studies employed various data collection methods, such as field notes and teacher debriefs.

However, as recommended in the literature on effective PD, rigorous methods are needed to examine the direct influence of PD on real-life classroom instruction and learning (e.g., Guskey, 2014). None of the reviewed studies examined student learning, and it is unclear whether the integrated STEM activities continued to be taught by teachers after support from the research team was discontinued. Furthermore, the long-term effects of PD on teacher practice were not clearly addressed in the studies.

*Design*

*Reviewed Studies*

The majority of the studies discussed the importance of teacher collaboration in STEM-integrated PD. It was argued that having a supportive communal learning environment in the PD setting helped teachers develop confidence and comfort in implementing the targeted curriculum. Boice et al. (2021) argued that collaborative lesson planning was critical to ensure equal STEM integration and create a sense of community among teachers. Parker et al. (2015) emphasized that teachers benefitted from talking about their successes or challenges with the implementation of STEM integration with their colleagues who taught the same grade.

The assessment was considered to be a challenging aspect of integrated STEM education. Guzey et al. (2016) recommended to support teachers in assessing student learning in integrated STEM lessons. Cook et al. (2021) claimed that in implementing integrated STEM lessons, their participant teachers used assessments that depended on one STEM subject. They recommended performance-based assessment types (in PD and classroom) that "include student choice on the product they produce or the process through which they showcase their content understanding" (p. 206).

Studies focused on the elements of PD that help participants establish links between PD and their teaching and student learning. For example, noticing positive student reactions/responses to integrated STEM materials and lessons increased teacher willingness to dive deeper into STEM integration (Boice et al., 2021; Cook et al., 2020; Rich et al., 2017, 2018). Boice et al. (2021) proposed to present teachers with many hands-on activities to first-hand experience integrated STEM learning and even use these activities in their teaching of integrated STEM lessons. Upon their assessment of the lesson plans, Wentworth and Monroe also (2011) recommended that teachers can develop more complex integrated lesson plans if PD addresses how and where teachers can find and examine integrated STEM lesson examples.

The reviewed studies included PD features that supported their participants' learning and teaching of integrated STEM education. Modeling the STEM-integrated curriculum was discussed in the reviewed

articles as necessary for teachers to become familiar with the new content and pedagogy of integrated STEM education (Parker et al., 2015; Sias et al., 2016; Wentworth & Monroe, 2011). The authors recommended that the PD practitioners need to model the skills and practices that teachers are expected to demonstrate in PD, and teachers should be clear about what skills and knowledge teachers will gain at the end of PD. Sias et al. (2017) even claimed, "without experience or models of the implementation of the educational innovations, teachers are less likely to conceive of how they might integrate the innovations in their lessons" (p. 235).

Studies discussed the importance of examining and presenting reflexive connections across STEM subject areas in PD settings (e.g., Baxter et al., 2014; Cook et al., 2020). They also suggested that PD that focuses on the natural relationship between mathematics and science will enhance teacher confidence and practice. Guzey et al. (2016) indicated that science teachers need more subject-matter knowledge to teach mathematics effectively, and teachers might mainly focus on engineering and science integration. Therefore, it is necessary for PD providers to provide opportunities for teachers to have the chance to increase their understanding of all the subject areas that will be integrated. Wentworth and Monroe (2011) similarly emphasized that PD that integrates mathematics and technology should focus on the use of technology in an integral way that "allows candidates to think more deeply about the mathematics in ways they could not without the technology" (p. 271).

Some contextual factors were proposed to consider in organizing and implementing PDs. The provision of necessary physical resources (i.e., supplies, materials for engineering lessons, printouts, and teacher and student workbooks) enabled teachers to incorporate computing and engineering in their classes (e.g., Rich et al., 2017, 2018). Boice et al. (2021) indicated that "ongoing financial and material support allowed teachers to engage students in new and otherwise impossible ways" (p. 17).

Being provided with support staff (researchers, PD practitioners, or coaches) or having administration/peer support at school also assisted teachers in teaching integrated STEM (Boice et al., 2021; Rich et al., 2018). Time constraints in developing integrated lessons, unanticipated changes during trainings (participant turnovers, changing expectations), and lack of shared vision among the actors of PD implementation inhibited teachers' efforts to integrate STEM in implementation (Guzey et al., 2016; Rich et al., 2018).

Parker et al. (2015) discussed that their participant teachers expected coaches to have a strong content area and district-level contextual knowledge. In implementing an integrated STEM lesson, teachers surely needed to believe that integration of STEM lessons would help them reach their instructional goals and state standards and that the provided/presented materials would align well with their classroom realities.

Two studies concluded that teachers were not ready to think in nontraditional ways as integrated curriculum and teaching just started to be included in elementary schools (McFadden & Roehrig, 2017; Sias et al., 2017). In addition, working in large schools was correlated with increased engineering integration activity per teacher. Only one study (Baxter et al., 2014) included information about the nature of integration for a successful PD: opportunities of infusion and transfer in connecting mathematics and science.

*Connections with the literature on effective PD*

Some implications and recommendations offered in the reviewed studies agreed with the literature on effective PD. For example, modeling the STEM-integrated curriculum was discussed as crucial in the reviewed studies. Collaborative practices and peer/administration support were also discussed as critical enablers of STEM-integrated curriculum and lesson implementation in most of the reviewed studies. The PD literature also discusses the importance of collaborative and collegial learning environments as conducive to creating communities of practice and increasing the change process beyond individual classrooms (e.g., Bates & Morgan, 2018).

The literature on effective PD recommends that PD should have a specific subject focus (e.g., Desimone et al., 2013). Some studies focused on the integration of certain STEM subject areas, such as mathematics and science, and those studies suggested that it would be important to focus on the natural connections across STEM subject areas so that elementary/primary teachers increase their cross-curricular subject matter knowledge. Some studies focused on STEM without any concerns to increase teachers' knowledge of particular subjects (e.g., Boice et al., 2020). From the reviewed studies, it is not possible to conclude whether focusing on individual subject areas or STEM overall leads to different results in teacher knowledge and skills in integrated STEM education.

Cochran-Smith and Demers (2018) also underline the importance of "the recognition that both prospective and experienced teachers (like all learners) brought prior knowledge and experience to all new learning situations, which are social and contextually specific"

(p.1011). Therefore, context plays an important role in PD practitioners/researchers' successes and challenges in the field. However, only a few studies reviewed (i.e., Parker et al., 2015) included detailed information regarding teacher, school, and district backgrounds.

In alignment with the literature on effective PD, the reviewed studies showed that examining intended change in student work in PD and positive student reactions/responses to integrated STEM materials and lessons increased teacher willingness and motivation to engage with STEM integration in classrooms (e.g., Rich et al., 2017, 2018).

Some practices that were raised in the reviewed studies were interesting. The assessment was one area that was discussed as important to focus on in integrated STEM PD. The reviewed studies recommended that teachers are challenged with developing assessment practices that will equally assess student learning in more than one subject area in integrated STEM lessons, and they should be supported in terms of how to assess student learning in integrated STEM classes.

Human and material resources made the difference in the way teachers employed integrated STEM education in their practice. The provision of materials, supplies and support from the school community (i.e., administrators) enabled teachers to incorporate intended PD knowledge and skills in their practice. In particular, it requires more time for teachers to plan and develop integrated STEM lessons; therefore, time constraints can significantly influence teachers' efforts to teach integrated STEM lessons.

**Implications and Conclusion**

This systematic review identified several key implications for improving professional development (PD) practices and advancing research on integrated STEM education. First, it is important to recognize that some teachers may lack confidence or expertise in certain subject areas, which can hinder their ability to create balanced and effective integrated STEM lessons. To address this issue, PD providers can offer targeted workshops to enhance subject matter knowledge and encourage collaborative planning with content experts. Additionally, PD developers should model active learning strategies and provide teachers with a variety of examples of effective integrated STEM curriculum for primary/elementary classrooms, including lessons, videos, and student artifacts. It is also crucial to ensure ongoing support and highlight successful practices for teachers implementing integrated STEM activities.

Furthermore, PD programs should align with school or district-level knowledge and provide opportunities for teachers to receive continuous support from administrators and parents. Field trips to local businesses and community-wide seminars and workshops can also be organized to establish a shared vision and sense of community around integrated STEM education. PD providers should also emphasize assessment practices that equally assess student learning and understanding across each subject area integrated into STEM lessons.

Finally, it is important to support teachers in planning and implementing integrated STEM lessons. One way to achieve this is by involving support staff from integrated disciplines such as engineering, who can assist teachers in developing creative ideas for lesson planning and delivery. To further advance research on effective STEM-integrated PD, future studies should explicitly describe the PD process, including details about teacher, instructor, school, and community backgrounds, available resources, and types of instruction and tasks.

Despite the insights gained from this systematic review, there are several limitations to consider. The review was based on a relatively small sample of 11 studies; therefore, more research is needed to establish consensus on effective components of STEM-integrated PD. Nevertheless, this review provides a solid foundation for future PD development and evaluation in STEM integration and can inform the development of more comprehensive research designs to explore the important characteristics of effective teacher professional development in integrated STEM education.

**References**

References marked with an asterisk indicate studies included in the systematic review

Altun, S., Yabas, D., & Bal-Nayman, H. (2021). Teachers' experiences on instructional design based professional development: A narrative inquiry. *International Electronic Journal of Elementary Education, 14*(1), 35–50.

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education, 27*(1), 10–20.

Bates. C. C., & Morgan, D. N. (2018). Seven elements of effective professional development. *The Reading Teacher, 71*(5), 623–626. https://doi:10.1002/trtr.1674

*Baxter, J. A., Ruzicka, A., Beghetto, R. A., & Livelybrooks, D. (2014). Professional development strategically connecting mathematics and science: The impact on teachers' confidence and practice. *School Science and Mathematics, 114*(3), 102–113. https://doi:10.1111/ssm.12060

Van den Bergh, L., Ros, A., & Beijaard, D. (2014). Improving teacher feedback during active learning: Effects of a professional development program. *American Educational Research Journal, 51*(4), 772–809. https://doi.org/10.3102/0002831214531322

Birman, B., Desimone, L., Porter, A., & Garet, M. (2000). Designing professional development that works. *Educational Leadership,57*(8), 28–33.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3–15. https://doi:10.3102/0013189X033008003

*Boice, K. L., Jackson, J. J., Alemdar, M., Rao, A. E., Grossman, S., & Usselam, M. (2021). Supporting teachers on their STEAM journey: A collaborative STEAM teacher training program. *Education Sciences, 11*, 105. https://doi.org/10.3390/educsci11030105

Cochran-Smith, M., & Demers, K.E. (2008). How do we know what we know? Research and teacher education. In M. Cochran-Smith, S. Feiman-Nemser, & D. J. McIntyre (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (3rd ed., pp. 1009–1016). Routledge.

Cochran-Smith, M., & Zeichner, K. (2005). *Studying teacher education: The report of the aera panel on research and teacher education.* Lawrence Erlbaum Associates, Inc.

*Cook, K., Bush, S., Cox Jr, R., & Edelen, D. (2020). Development of elementary teachers' science, technology, engineering, arts, and mathematics planning practices. *School Science and Mathematics, 120*(4), 197–208. https://doi:10.1111/ssm.12400

Darling-Hammond, L. (2000). How teacher education matters. *Journal of Teacher Education, 51*(3), 166–173. https://doi:10.1177/002248710005100300 2

Darling-Hammond, L. (2006). Constructing 21st-century teacher education. *Journal of Teacher Education, 57*(3), 300–314. https://doi:10.1177/0022487105285962

Darling-Hammond, L., & Richardson, N. (2009). Research review/teacher learning: What matters? *Educational Leadership, 66*(5), 46–53.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development.* Learning Policy Institute.

Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181–199. https://doi:10.3102/0013189X08331140

Desimone, L. M., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society, & Education, 7*(3), 252–263. https://doi:10.25115/psye.v7i3.515

Desimone, L.M., & Pak, K. (2017). Instructional coaching as high-quality professional development. *Theory into Practice, 56*(1), 3–12.

Desimone, L., Porter, A.C., Garet, M.S., & Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*(2), 81–112. https://doi:10.3102/01623737024002081

Desimone, L., Smith, T.M., & Phillips, K. (2013). Linking student achievement growth to professional development participation and changes in instruction: A longitudinal study elementary students and teachers in Title I schools. *Teachers College Record, 115*(5), 1–46.

Desimone, L. M., Smith, T. M., Hayes, S., & Frisvold, D. (2005). Beyond accountability and average math scores: Relating multiple state education policy attributes to changes in student achievement in procedural knowledge, conceptual understanding and problem solving in mathematics. *Educational Measurement: Issues and Practice, 24*(A), 5–18.

English, L. (2016). STEM education K-12: Perspectives on integration. *International Journal of STEM Education, 3*(1), 1–8. https://doi:10.1186/s40594-016-0036-1

Fullan, M., & Milles, M. (1992). Getting reform right: What works and what doesn't. *Phi Delta Kappan, 73*, 744–752.

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? *American Educational Research Journal, 38*(4), 915–945. https://doi:10.3102/00028312038004915

Gupta, A., & Guang-Lea, Lee. (2020). The effects of a site-based teacher professional development program on student learning. *International Electronic Journal of Elementary Education, 12*(5), 417–428. https://doi.org/10.26822/iejee.2020562132

Guskey, T. R. (2000). *Evaluating professional development.* Thousand Oaks, CA: Corwin.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching, 8*(3), 381–391. https://doi:10.1080/135406002100000512

Guskey, T. R. (2014). Evaluating professional learning. In Billett, S., Harteis, C., & Gruber, H. (Eds.). *International handbook of research in professional and practice-based learning.* (pp. 1215–1235). Springer.

*Guzey, S. S., Moore, T. J., & Harwell, M. (2016). Building up STEM: An analysis of teacher-developed engineering design-based STEM integration curricular materials. *Journal of Pre-College Engineering Education Research (J-PEER), 6*(1), 11–29.

*Havice, W., Havice, P., Waugaman, C., & Walker, K. (2018). Evaluating the effectiveness of integrative STEM education: Teacher and administrator professional development. *Journal of Technology Education, 29*(2), 73–90.

Joffe, H. (2011). Thematic analysis. In D. Harper & A. R. Thompson (Eds.), *Qualitative methods in mental health and psychotherapy: A guide for students and practitioners* (pp. 209-223). Wiley.

Knapp, M.S. (2003). Professional development as a policy pathway. *Review of Research in Education, 27,* 109–157.

Kelley, T., & Knowles, J. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*(1), 1–11. https://doi:10.1186/s40594-016-0046-z

Lawless, K.A., & Pellegrino, J.W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of Educational Research, 77*(4), 575–614. https://doi:10.3102/0034654307309921

Little, J. W. (1987). Teachers as colleagues. In V. Richardson-Koehler (Ed.), *Educators' handbook: A research perspective* (pp. 491–518). Longman.

*McFadden, J., & Roehrig, G. (2017). Exploring teacher design team endeavors while creating an elementary-focused STEM-integrated curriculum. *International Journal of STEM Education, 4*(1), 1–22. https://doi:10.1186/s40594-017-0084-1

Nadelson, L. S., & Seifert, A. L. (2017). Integrated STEM defined: Contexts, challenges, and the future. *The Journal of Educational Research, 110*(3), 221–223. https://doi:10.1080/00220671.2017.1289775

OECD (2019). *TALIS 2018 Results (Volume I): Teachers and school leaders as lifelong learners.* OECD Publishing. https://doi:10.1787/1d0bc92a-en

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T.C., & Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*(71). https://doi:10.1136/bmj.n71

*Parker, C., Abel, Y., & Denisova, E. (2015). Urban elementary STEM initiative.

*School Science and Mathematics, 115*(6), 292–301. https://doi:10.1111/ssm.12133

Parsad, B., Lewis, L., Farris, E., & Greene, B. (2001). *Teacher preparation and professional*

*Development* (NCES 2001-088). U.S. Department of Education, National Center for Education Statistics.

Pearson, G. (2017). National academies piece on integrated STEM. *The Journal of Educational Research, 110*(3), 224–226. https://doi:10.1080/00220671.2017.1289781

Penuel, W. R., Fishman, B. J., & Yamaguchi, R. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal, 44*(4), 921–958. https://doi:10.3102/0002831207308221

Popova, A., Evans, D., Breeding, M. E., & Arancibia, V. (2018). *Teacher professional development around the World: The gap between evidence and practice* (WPS 8572). World Bank Group. http://documents.worldbank.org/curated/en/349051535637296801/Teacher-Professional-Development-around-the-World-The-Gap-between-Evidence-and-Practice

*Rich, P. J., Jones, B., Belikov, O., Yoshikawa, E., & Perkins, M. (2017). Computing and engineering in elementary school: The effect of year-long training on elementary teacher self-efficacy and beliefs about teaching computing and engineering. *International Journal of Computer Science Education in Schools, 1*(1), 1–20. https://files.eric.ed.gov/fulltext/ED572692.pdf

*Rich, P.J., Belikov, O., Yoshikawa, E., & Perkins, M. (2018). Enablers and inhibitors to integrating computing and engineering lessons in elementary education. *Journal ofTechnology and Teacher Education, 26*(3), 437–469. https://www.learntechlib.org/primary/p/181979/

*Sias, C. M., Nadelson, L. S., Juth, S. M., & Seifert, A. L. (2017). The best laid plans: Educational innovation in elementary teacher generated integrated STEM lesson plans. *The Journal of Educational Research, 110*(3), 227–238. https://doi:10.1080/00220671.2016.1253539

Smith, K. (2017). *Teachers as self-directed learners: Active positioning through professional learning.* Springer. https://doi:10.1007/978-981-10-3587-6

Villegas-Reimers, E. (2003). *Teacher professional development: An international review of the literature*. UNESCO International Institute for Educational Planning.

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research, 39*(1), 93–112. https://doi.org/10.1177/0739456X17723971

Shin, Y.-S., & Jun, J.S. (2019). The hierarchical effects of individual and organizational variables on elementary school teachers' lifelong learning competence. *International Electronic Journal of Elementary Education, 12*(2), 205–212. https://doi.org/10.26822/iejee.2019257668

*Wentworth, N., & Monroe, E. E. (2011). Inquiry-based lessons that integrate technology: Their development and evaluation in elementary mathematics teacher education. *Computers in the Schools, 28*(4), 263–277. https://doi:10.1080/07380569.2011.620938

iejee

International Electronic Journal of Elementary Education
June 2023, Volume 15, Issue 5, 383-398

# Pre-Service Primary School Teachers' Interdisciplinary Competence and their Interest, Self-Concept, and Sense of Belonging Regarding Natural and Social Sciences: Findings from a Longitudinal Study in Germany

Markus Sebastian Feser[a,*], Kerstin Michalik[b]

Received : 9 February 2023
Revised : 30 May 2023
Accepted : 18 June 2023
DOI : 10.26822/iejee.2023.307

[a,*] **Corresponding Author:** Markus Sebastian Feser, Faculty of Education, Universität Hamburg, Hamburg, Germany.
E-mail: markus.sebastian.feser@uni-hamburg.de
ORCID: https://orcid.org/0000-0001-8503-0951

[b] Kerstin Michalik, Faculty of Education, Universität Hamburg, Hamburg, Germany.
E-mail: kerstin.michalik@uni-hamburg.de
ORCID: https://orcid.org/0000-0003-0258-7688



## Abstract

In German primary schools, natural sciences and social studies are learned and taught in an integrative manner within a subject called Sachunterricht. To teach Sachunterricht in a high-quality manner, it is reasonable to assume that primary school teachers themselves require—among other things, such as knowledge about pedagogy, teaching Sachunterricht, and the various content areas of Sachunterricht—a distinct interest, academic self-concept, and sense of belonging regarding natural and social sciences. Furthermore, they should possess a solid interdisciplinary competence that enables them to teach natural and social sciences in an integrative way. In the present study, we conducted a longitudinal survey of pre-service primary school teachers from a German university over a period of 2 years to investigate the changes in their (self-evaluated) interdisciplinary competence; the changes in their interest, academic self-concept, and sense of belonging regarding natural and social sciences; and the correlations between these constructs. Our data analysis revealed a decrease over time in participants' sense of belonging to natural and social sciences, as well as their (self-evaluated) interdisciplinary competence, while their academic self-concept in natural and social sciences remained stable. Participants' interest in social sciences decreased, while their interest in natural sciences increased. Moreover, we found varying degrees of correlation between these constructs. In summary, the results of the present study provide important insights into the professional development of pre-service primary school teachers within university-based teacher education for teaching natural and social sciences in primary school. The implications of these findings are discussed in detail at the end of this paper.

**Keywords:**

Sachunterricht, Pre-Service Teacher Education, Natural Sciences, Social Sciences, Interest, Self-Concept, Sense of Belonging, Interdisciplinary Competence, Longitudinal Study

## Introduction

In Germany, children's formal education typically begins at the age of six, when they start primary school. In total, primary school in Germany comprises four to six school years, and it is the only school type in Germany in which nearly all children of the same age cohort learn together. In addition to German and mathematics, the core curriculum of primary school in Germany explicitly includes the subject Sachunterricht (Kultusministerkonferenz, 2015). The aim of Sachunterricht[a] is to enable children to acquire elementary knowledge about physical and social aspects of the world, develop competencies for navigating and participating in their everyday lives, and obtain the perquisites for later learning in natural and social sciences (Köhnlein, 2015). Therefore, Sachunterricht is a multidisciplinary subject in which natural and social sciences are learned and taught in an integrated manner[b] (Kahlert, 2014; Tänzer, 2014; Thiel, 2003).

Due to its core curriculum status, teaching Sachunterricht is one of the key responsibilities of primary school teachers in Germany. In order to teach Sachunterricht in a high-quality manner, primary school teachers require pedagogical professionalization, as well as professionalization in teaching Sachunterricht (Gesellschaft für Didaktik des Sachunterrichts, 2003; 2019). Most importantly, however, primary school teachers also require professionalization in the content areas of Sachunterricht, namely in natural sciences, such as biology, chemistry, and physics, as well as in social sciences, such as economics, political science, history, geography, and sociology (Gesellschaft für Didaktik des Sachunterrichts, 2003; 2019). This professionalization in the content areas of Sachunterricht involves not only the acquisition of knowledge and skills but also affective-motivational orientations related to natural and social sciences (Blömeke et al., 2015; Kleickmann, 2015; Lange, 2015).

In particular, these affective-motivational orientations include the constructs of interest, academic self-concept, and sense of belonging. On one hand, within educational research, interest is usually conceptualized as the "specific and distinguished relationship between a person and [...] concrete things, a topic, a subject-matter or an abstract idea" (Krapp & Prenzel, 2011, p. 31). On the other hand, academic self-concept refers to the cognitive representations of one's own abilities in academic performance situations (Gabriel-Busse et al., 2018; Dickhäuser et al., 2002; Shavelson et al., 1976), and sense of belonging to an academic domain is defined as individuals' feeling or sensation of connectedness and social affiliation with and within an academic domain (Feser, 2021; Pendergast et al., 2020). The domain-specific manifestations of these three constructs (i.e., teachers' interest, academic self-concept, and sense of belonging regarding natural sciences and, analogously, teachers' interest, academic self-concept, and sense of belonging regarding social sciences) are considered key features of teachers' domain-specific professional identity. Therefore, they are essential for teachers attempting to acquire knowledge in natural and social sciences themselves and provide high-quality teaching in natural and social sciences to their students (for details see Feser & Haak, 2022).

In addition to the above, to teach natural and social sciences within Sachunterricht in an integrative manner, primary school teachers themselves must be able to link their knowledge of multidisciplinary issues within the various content areas of Sachunterricht (e.g., issues related to sustainability, the digital world, or climate change; Künzli David et al., 2016; Lenoir & Hasni, 2016; Wilhelm & Brühwiler, 2016). Consequently, it is reasonable to claim that primary school teachers also require interdisciplinary competence (Hasni et al., 2015; Pharo, et al., 2012), meaning the knowledge and skills required to address multidisciplinary issues in an interdisciplinary way, as well as the motivational and volitional readiness to do so (Petrie, 1992; Engelhardt, 2019).

In light of the above, it is not surprising that Sachunterricht is considered one of the most difficult subjects to teach (if not the most difficult one) for primary school teachers in Germany (Klafki, 1992). There is a consensus among teacher educators that university-based teacher education should enable pre-service primary school teachers to acquire the proper professionalization for teaching Sachunterricht, which empowers them to cope with the various requirements mentioned above in their future careers (Gesellschaft für Didaktik des Sachunterrichts, 2003; 2019; Marquardt-Mau et al., 1996). The extent to which university-based teacher education succeeds in this endeavour is one of the major issues addressed within research on teaching Sachunterricht.

On one hand, the development of pre-service primary school teachers' knowledge about pedagogy, teaching Sachunterricht, and the various content areas of Sachunterricht has already been thoroughly investigated by previous research (e.g., Appleton, 2003; Hartmann, 2018; Kirsch, 2022; Lange et al., 2012; Meschede et al., 2017; Niermann, 2017; Schmidt, 2015; Sothayapetch et al., 2013). The results of existing studies confirm that pre-service primary school teachers' knowledge and skills within these fields can be improved by university-based teacher education (Kleickmann, 2015; Meschede et al., 2020). Furthermore, in line with international research, it was shown that pre-service primary school teachers' content knowledge in natural and social sciences is tendentially lower as compared to that of pre-service

secondary school teachers (for a literature review, see Niermann, 2017).

On the other hand, studies (conducted in Germany) that address pre-service primary teachers' interest, academic self-concept, and sense of belonging regarding natural and/or social sciences are scarce (e.g., Beudels et al., 2021; Feser & Plotz, 2023; Lenzgeiger, 2022; Reichhardt, 2018; Wang & Sneed, 2019). Studies focusing on natural sciences have found that these affective-motivational orientations, although distinguishable, tend to be substantially correlated with one another (Feser & Plotz, 2023; Wang & Sneed, 2019). An analogous finding looms for social sciences. For example, the study of Reichhardt (2018) showed that primary school teachers' interest and academic self-concept regarding political sciences are strongly correlated. Additionally, a pattern in research seems to be emerging indicating that pre-service primary school teachers' interest, academic self-concept, and sense of belonging regarding natural and/or social sciences are relatively stable constructs that only change—if at all—over extended time periods within university-based teacher education (Beudels et al., 2021; Lenzgeiger, 2022).

Beyond that, thus far, the role of pre-service primary school teachers' interdisciplinary competence in teaching Sachunterricht has been addressed within educational research primarily from a theoretical perspective (e.g., Kalcsics, 2021; Künzli David et al., 2016; Wilhelm & Brühwiler, 2016). Although there is a growing body of research that empirically investigates university students' interdisciplinary competence, especially within higher education (e.g., Braßler & Dettmers, 2017; Engelhard 2019; Mansilla & Duraising, 2007) and engineering education (e.g., Richter & Paretti, 2009; Tormey & Laperrouza, 2023), to the best of our knowledge, such research has not yet been conducted with respect to the interdisciplinary competence of pre-service primary school teachers. In particular, there is a lack of empirical research, in Germany and internationally, about whether pre-service primary school teachers' interdisciplinary competence changes over their university-based teacher education and to what extent interdisciplinary competence is correlated with the domain-specific orientations of pre-service primary school teachers (e.g., their interest, academic self-concept, and sense of belonging regarding natural and/or social sciences).

### *Aim of the present study*

With the present study, we aim to add to previous research and address the research gaps detailed above. More precisely, we conducted a longitudinal study at a German university in which we surveyed pre-service primary school teachers over a period of 2 years during their undergraduate studies in natural and social sciences (for details, see below). This longitudinal study was guided by the following research questions:

(RQ1a) Do pre-service primary school teachers' interest, academic self-concept, and sense of belonging regarding natural sciences change during their undergraduate studies in natural sciences, and if so, to what extent?

(RQ1b) Do pre-service primary school teachers' interest, academic self-concept, and sense of belonging regarding social sciences change during their undergraduate studies in social sciences, and if so, to what extent?

(RQ2) Does pre-service primary school teachers' (self-evaluated) interdisciplinary competence change during their undergraduate studies in natural and social sciences, and if so, to what extent?

(RQ3) To what extent do pre-service primary school teachers' (self-evaluated) interdisciplinary competence and their interest, self-concept, and sense of belonging regarding natural and social sciences correlate with one another during their undergraduate studies?

Regarding research question RQ1, within the present study we expect at most moderate changes in pre-service primary school teachers' interest, academic self-concept, and sense of belonging regarding natural and/or social sciences, because previous research indicates that these constructs change only over extended time periods within university-based teacher education, if they change at all (see Introduction). Similarly, based on previous research findings, we expect substantial correlations between primary teachers' interest, academic self-concept, and sense of belonging regarding the natural and/or social sciences (research question RQ3). Conversely, due to the lack of previous research, it was not possible for us to draw up elaborate expectations regarding the change of pre-service primary school teachers' interdisciplinary competence and its correlative relationships with other constructs within the present study (research questions RQ2 and RQ3).

Below, we describe the context and design of our longitudinal study. Subsequently, we report and discuss the results of our data analysis regarding research questions RQ1 to RQ3.

## Method

### *Context of the study*

The present study was conducted within the university-based teacher education program "Sachunterricht"

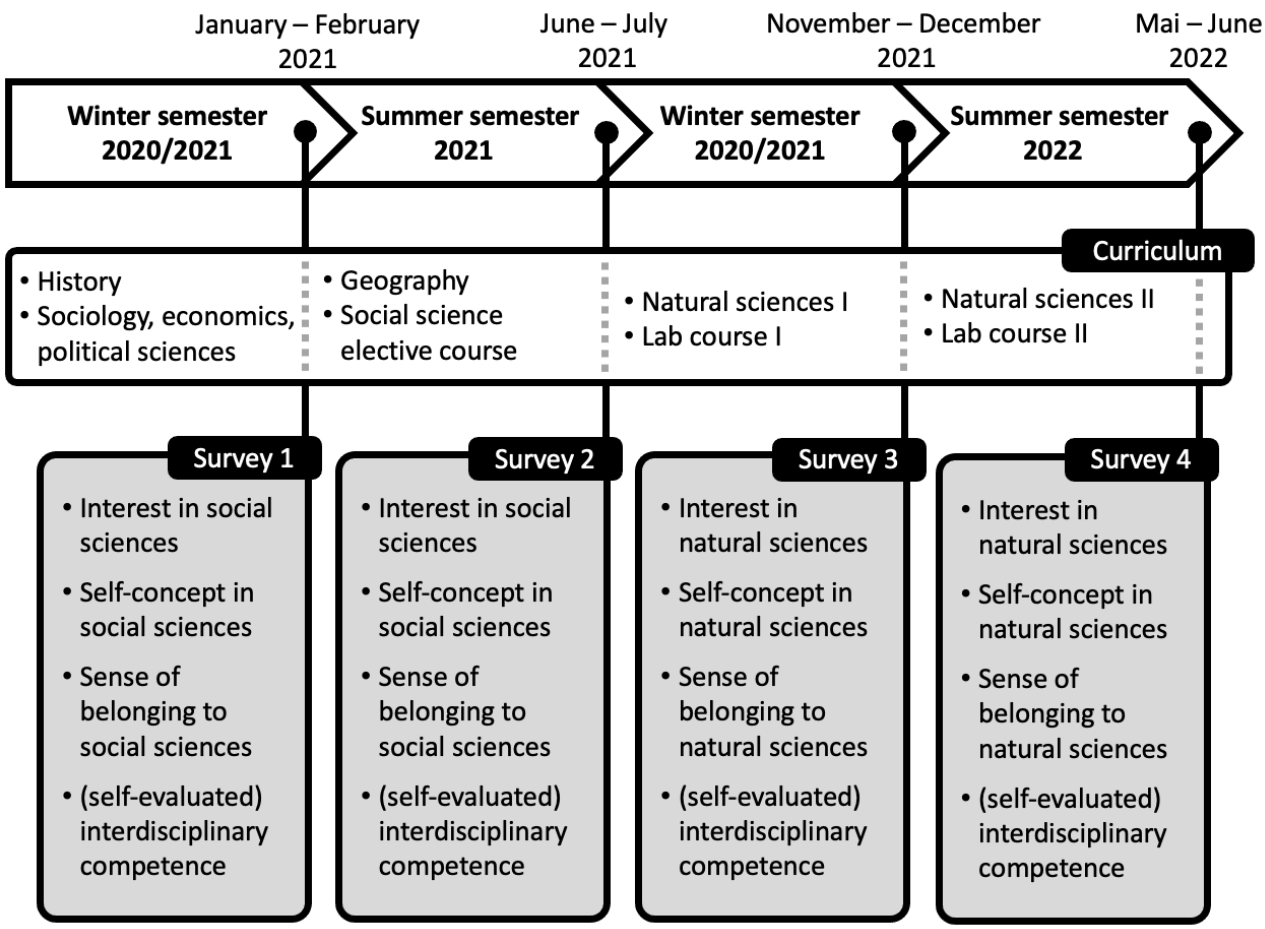
at the Universität Hamburg, Germany (Universität Hamburg, 2022). Pre-service teachers studying the "Primary Education" or "Primary Education for students with special education needs" degree programs at the Universität Hamburg can take this teacher education program as an elective. One major goal of this teacher education program is to provide pre-service primary school teachers a solid knowledge base within the content areas of Sachunterricht. Accordingly, the curriculum of this program comprises courses on both the natural and social sciences (see Figure 1). In their first two semesters, the pre-service primary school teachers study social sciences within various courses on economics, political sciences, sociology, history, and geography. Subsequently, they attend a two-semester course in which biology, chemistry and physics are taught in an integrative manner, together with an accompanying natural sciences laboratory class.

### Data collection

Pre-service primary school teachers who started the teacher education program "Sachunterricht" in the winter semester (April - September) of 2020/2021 were invited to participate in the present study ($N$ = 164). Incentives for participation were not offered. Data collection was conducted longitudinally via four online surveys spanning from the winter semester of 2020/2021 to the summer semester (September - March) of 2022 (Survey 1 to 4; see Figure 1). In accordance with the ethical and legal standards for educational research in Germany (Watteler & Ebel, 2019), the surveys were conducted anonymously and voluntarily. To match participants' responses across the four online surveys, subject-generated identification codes were used (Yurek et al., 2008).

Within each of the four surveys, participants' (self-evaluated) interdisciplinary competence was surveyed (see Figure 1). Furthermore, in line with the curricular order of the teacher education program "Sachunterricht" (see Context of the study), the focus of the four questionnaires differed (see Figure 1): the online questionnaires used in the winter semester of 2020/2021 (Survey 1) and summer semester of 2021 (Survey 2) surveyed participants' interest, academic self-concept, and sense of belonging regarding social sciences. Analogously, the online questionnaires used in the winter semester of 2021/2021 (Survey 3) and summer semester of 2022 (Survey 4) surveyed participants' interest, academic self-concept, and sense of belonging regarding natural sciences.

**Figure 1**
*Design of the longitudinal data collection*



## Measures

Table 1 summarizes the descriptive statics and psychometric properties of the instruments we used within the four online surveys of the present study. In the following sub-sections, these instruments are described in more detail. Further information, as well as an overall presentation of the questionnaires, are published in the technical report of the present study (see Feser & Michalik, 2023).

### *Interest in natural and social sciences*

To assess participants' interest in natural and social sciences, we used two scales developed by Rau (2017). Each of these scales consists of five items, which participants are asked to response to on a 4-point Likert scale. The items of both scales are nearly identical in wording; they only differ in whether they refer to natural or social sciences (e.g., "I am interested in learning something new in natural sciences" versus "I am interested in learning something new in social sciences").

### *Academic self-concept in natural and social sciences*

We used adapted versions of the instrument developed by Elsholz (2019) to assess participants' academic self-concept in natural and social sciences (number of items: five; bipolar response scale with scores ranging from 1 to 7). Analogously to the interest scales, the items of these scales differ only in whether they address natural or social sciences (e.g., "For me, learning new things in my natural science studies is easy/difficult" versus "For me, learning new things in my social science studies is easy/difficult").

### *Sense of belonging to natural and social sciences*

To assess the participants' sense of belonging to natural sciences, we used the instrument developed by Feser (2020). This instrument consists of 21 items (e.g., "I feel connected to the natural sciences community"), which require responses on a 5-point Likert scale. To assess participants' sense of belonging to social sciences, we adapted this instrument and changed the wording of the items to indicate social sciences (e.g., "I feel connected to the social sciences community").

### *(Self-evaluated) Interdisciplinary competence*

While knowledge about pedagogy, teaching Sachunterricht, and the various content areas of Sachunterricht is usually surveyed using paper-and-pencil or vignette tests, it is quite difficult to survey pre-service primary school teachers' interdisciplinary competence. Interdisciplinary competence is domain-unspecific and defined in terms of an individual's proficiency to cope within multidisciplinary contexts (see Introduction). For this reason, within empirical research, interdisciplinary competence is typically surveyed based on self-evaluations on the part of the participants (e.g., Claus & Wiese, 2019; Engelhardt, 2019; Lattuca et al., 2013; Tormey & Laperrouza, 2023). Therefore, to survey pre-service primary school teachers' interdisciplinary competence within the present study, we used an adapted version of the self-evaluation scale developed by Engelhardt (2019). This adapted version consists of six items, the wording of which was specified on the university-based teacher education program "Sachunterricht" and which participants are asked to response to on a 5-point

**Table 1**

*Descriptive statics and psychometric properties for the instruments used based on our surveyed sample*

| Instrument | Survey | $N_{item}$ | *N* | *M* | *SD* | Min | Max | α | $r_{it}$ |
|---|---|---|---|---|---|---|---|---|---|
| SB-S | $t_1$ | 21 | 75 | 3.36 | .57 | 2.18 | 4.38 | .75 | .36 to .64 |
| | $t_2$ | | 44 | 3.18 | .63 | 1.72 | 4.50 | .83 | .44 to .72 |
| SC-S | $t_1$ | 5 | 67 | 4.81 | .99 | 2.80 | 7.00 | .92 | .74 to .86 |
| | $t_2$ | | 44 | 4.60 | 1.15 | 1.00 | 7.00 | .93 | .72 to .88 |
| IN-S | $t_1$ | 5 | 81 | 3.12 | .53 | 1.60 | 4.00 | .82 | .54 to .66 |
| | $t_2$ | | 46 | 2.85 | .57 | 1.60 | 3.80 | .85 | .62 to .71 |
| SB-N | $t_3$ | 21 | 48 | 3.19 | .55 | 1.40 | 4.04 | .67 | .36 to .51 |
| | $t_4$ | | 44 | 3.05 | .68 | 1.57 | 4.92 | .78 | .46 to .71 |
| SC-N | $t_3$ | 5 | 46 | 4.06 | 1.12 | 1.00 | 7.00 | .92 | .72 to .86 |
| | $t_4$ | | 45 | 4.14 | .97 | 2.20 | 6.80 | .87 | .58 to .80 |
| IN-N | $t_3$ | 5 | 48 | 2.81 | .54 | 1.60 | 4.00 | .84 | .55 to .72 |
| | $t_4$ | | 46 | 2.91 | .49 | 1.80 | 4.00 | .81 | .50 to .73 |
| IDC | $t_1$ | 6 | 78 | 3.57 | .63 | 1.17 | 5.00 | .77 | .38 to .71 |
| | $t_2$ | | 45 | 3.49 | .66 | 1.83 | 4.83 | .76 | .25 to .74 |
| | $t_3$ | | 47 | 3.33 | .56 | 2.33 | 4.83 | .73 | .26 to .71 |
| | $t_4$ | | 43 | 3.33 | .67 | 1.50 | 4.83 | .80 | .30 to .72 |

Note. SB-S = sense of belonging to social sciences; SC-S = self-concept in social sciences; IN-S = interest in social sciences; SB-N = sense of belonging to natural sciences; SC-N = self-concept in natural sciences; IN-N = interest in natural sciences; IDC = self-evaluated interdisciplinary competence; $t_1$ = Survey 1; $t_2$ = Survey 2; $t_3$ = Survey 3; $t_4$ = Survey 4; N = number of participants; $N_{item}$ = number of items; M = arithmetic means; SD = standard deviation; min = minimum score; max = maximum score; α = Cronbach's alpha; $r_{it}$ = range of the items' selectivity coefficient.

Likert scale (e.g., "I can interconnect subject-specific concepts of different fields of the teacher education program 'Sachunterricht' and, as a result, achieve a better understanding").

### Participants and data analysis

Table 2 summarizes the participation rate across the four online surveys of the present study. In the first survey, the participation rate was 50.0 %; in the following three surveys, this rate dropped to between 28.0 and 29.9 %. In total, 103 (62.8 %) of the 164 pre-service primary school teachers who started the teacher education program "Sachunterricht" in the winter semester of 2020/2021 participated in at least one online survey. Only 23 (14.0 %) pre-service primary school teachers participated in all four surveys.

As a result, like many other longitudinal studies, the present study suffered from a substantial non-participation rate, as well as a declining re-participation rate. Thus, to decrease potential biases and reductions in test strength due to missing data, we utilized multiple imputation within our data analysis (Cox et al., 2014; Nissen et al., 2019) based on the assumption that it is plausible to predict the levels of a participant's attribute at a given point in time by using the levels of that attribute at a different point in time (e.g., predicting participants' interest in social science in the summer semester of 2021 based on their interest in social science in the winter semester of 2020/2021). For this multiple imputation, we used the data derived from the 60 pre-service primary school teachers (36.6 %) who participated in two or more surveys. Table 3 summarizes the central descriptive statistics of this sub-sample as compared to all pre-service primary school teachers who were invited to participate in the present study. The percentage share of missing values within this dataset ranged from 8.3 %, for participants' interest in social science in the winter semester of 2020/2021, to 38.3 %, for participants' self-evaluated interdisciplinary competence in the summer semester of 2022.

**Table 2**

*Participation rate across the four surveys*

| Survey | *N* | % | $N_p$ | *N* | % |
|---|---|---|---|---|---|
| $t_1$ | 82 | 50.0 | 1 | 43 | 26.2 |
| $t_2$ | 47 | 28.7 | ≥ 1 | 103 | 62.8 |
| $t_3$ | 49 | 29.9 | ≥ 2 | 60 | 36.6 |
| $t_4$ | 46 | 28.0 | ≥ 3 | 38 | 23.2 |
| $t_1$-$t_4$ | 23 | 14.0 | 4 | 23 | 14.0 |

Note. $t_1$ = Survey 1 (winter semester 2020/2021); $t_2$ = Survey 2 (summer semester 2021); $t_3$ = Survey 3 (winter semester 2021/2022); $t_4$ = Survey 4 (summer semester 2022); $t_1$-$t_4$ = Survey 1 to Survey 4; N = number of participants; % = percentage based on the total number of pre-service primary school teachers invited to participate in this study (N = 164); $N_p$ = number of surveys in which the same pre-service primary school teachers participated.

**Table 3**

*Descriptive statistics of the analyzed sample*

| | | Invited to participate | Participated in ≥ 2 surveys |
|---|---|---|---|
| Total | | 164 (100 %) | 60 (100 %) |
| Age | *M* | No information available | 22.56 |
| | *SD* | | 4.84 |
| High school graduation grade point average | *M* | No information available | 2.18 |
| | *SD* | | .57 |
| Gender | Female | 143 (87.20 %) | 55 (91.67 %) |
| | Male | 21 (12.80 %) | 5 (8.33 %) |
| Training in special education | Yes | 128 (78.05 %) | 43 (71.67 %) |
| | No | 36 (21.95 %) | 17 (28.33 %) |

Note. This table is based on the data derived from Survey 1; the German school grade scale (high school graduation grade point average) ranges from 1 = very good to 6 = insufficient; M = arithmetic means; SD = standard deviation; training in special education = pre-service primary school teachers who do (not) participate in a teacher education program about teaching students with special educational needs.

Data analysis was conducted using R Version 4.2.1 (The R Foundation for Statistical Computing, 2023). Via the "mice" package Version 3.14 (van Buuren & Groothuis-Oudshoorn, 2011), multiple imputation was performed (imputation method: linear regression through prediction; number of imputations: 200; number of iterations: 10). To examine research question RQ3, we performed pooled correlation analyses (Pearson's r) between participants' interdisciplinary competence and their interest, self-concept, and sense of belonging regarding natural and social sciences for each of the four online surveys (Heymans & Eekhout, 2019). Analogously, to address research questions RQ1 and RQ2, we used pooled paired-sample t-tests (Heymans & Eekhout, 2019); for research question RQ2, we also conducted a pooled Page's trend test beforehand (Page, 1963). Within these data analyses, we estimated the parameters of interest for each imputed dataset separately and combined them following Rubin's rules (Rubin, 1987).

### *Limitations*

Given our study design and methodological approach, the present study has some limitations that should be considered. First, due to a substantial non-participation rate and declining re-participation in our study, we utilized multiple imputation within our data analysis to decrease potential biases and reductions in test strength. Nevertheless, it is possible that the results of the present study still may exhibit some bias (Heymans & Eekhout, 2019), especially because our analyzed sample comprises only 36.6 % of all pre-

service primary school teachers who were invited to participate. Additionally, the participants in the present study were solely recruited within the university-based teacher education program "Sachunterricht" at the Universität Hamburg, Germany (see Context of the study). Consequently, it is conceivable that further results regarding RQ1 to RQ3 may emerge if the present study is replicated with pre-service primary school teachers' participating in university-based teacher education programs significantly differing from the programs in Hamburg. Moreover, our data collection was carried out during the COVID-19 pandemic. Because research indicates that pre-service teachers' professionalization within university-based teacher education was heavily influenced by the COVID-19 pandemic (e.g., An & Zakaria, 2022; Kan et al., 2022; Tekel et al., 2022), it is reasonable to assume that the results of the present study reflect this influence as well.

**Results**

*Changes in interest, academic self-concept, sense of belonging, and interdisciplinary competence*

In summary, the following pattern regarding research question RQ1 emerged within our data analysis (see Table 4): first, neither participants' academic self-concept in natural science nor their academic self-concept in social sciences showed a statistically significant change over time. The descriptive level hints, if at all, that there may be a very small decrease in participants' academic self-concept in social sciences ($p$ = .118; $d$ = -.22). Second, both participants' sense of belonging to natural sciences and their sense of belonging to social sciences showed a statistically significant decrease over time (p < .001). Furthermore, in both cases, the magnitude of this decrease was identical and moderate (d = -.67). Third and finally, participants' interest in natural and social sciences showed different changes over time. On one hand, participants' interest in social sciences decreased significantly between Surveys 1 and 2 ($p$ = .045; $d$ = -.32). On the other hand, between Surveys 3 and 4, their interest in natural sciences significantly increased ($p$ = .036; $d$ = .35). The magnitudes of these changes in interest are small.

Regarding the change in participants' (self-evaluated) interdisciplinary competence during their undergraduate studies (research question RQ2), the results of the pooled Page's trend test indicate a small but significant decrease over time ($L$ = 1558.49; ρ = .195[c]; $p$ = .004). However, as illustrated in Table 4, a more detailed analysis using pooled paired-sample t-tests revealed a significant, moderate decrease in participants' (self-evaluated) interdisciplinary competence only between Surveys 2 and 3 ($p$ = .003; $d$ = -.53).

**Table 4**

*Mean differences regarding participants' interest, self-concept, sense of belonging, and self-evaluated interdisciplinary competence across the four surveys*

Surveys 1 and 2 (winter semester 2020/2021 to summer semester 2021)

| Instrument | $M_1$ | $SD_1$ | $M_2$ | $SD_2$ | $t$ | $p$ | $d$ |
|---|---|---|---|---|---|---|---|
| SB-S | 3.33 | .59 | 3.17 | .60 | -3.68 | *** | -.67 |
| SC-S | 4.73 | 1.07 | 4.62 | 1.12 | -1.20 | .12 | -.22 |
| IN-S | 3.01 | .53 | 2.91 | .55 | -1.73 | * | -.32 |
| IDC | 3.54 | .66 | 3.55 | .58 | .17 | .43 | .03 |

Surveys 2 and 3 (summer semester 2021 to winter semester 2021/2022)

| Instrument | $M_2$ | $SD_2$ | $M_3$ | $SD_3$ | $t$ | $p$ | $d$ |
|---|---|---|---|---|---|---|---|
| IDC | 3.55 | .58 | 3.39 | .55 | -2.91 | ** | -.53 |

Surveys 3 and 4 (winter semester 2021/2022 to summer semester 2022)

| Instrument | $M_3$ | $SD_3$ | $M_4$ | $SD_4$ | $t$ | $p$ | $d$ |
|---|---|---|---|---|---|---|---|
| SB-N | 3.20 | .55 | 3.03 | .62 | -3.64 | *** | -0.67 |
| SC-N | 4.14 | 1.02 | 4.11 | .89 | -0.52 | .30 | -.10 |
| IN-N | 2.85 | .51 | 2.91 | .44 | 1.83 | * | .35 |
| IDC | 3.39 | .55 | 3.43 | .57 | .850 | .20 | .16 |

Note. SB-S = sense of belonging to social sciences; SC-S = self-concept in social sciences; IN-S = interest in social sciences; SB-N = sense of belonging to natural sciences; SC-S = self-concept in natural sciences; IDC = self-evaluated interdisciplinary competence; M = pooled arithmetic means; SD = pooled standard deviation; t = t-value (one-tailed pooled paired sample t-test); p = p-value; d = pooled Cohen's d with correlation adjustment for paired samples (see Cohen, 1988, p. 49); *p < .05; **p < .01; ***p < .001.

### Correlation analyses

Table 5 summarizes the results of our pooled correlation analyses. Our data analysis revealed that pre-service primary teachers' interest, academic self-concept, and sense of belonging regarding natural sciences correlate with one another at a moderate to high level (Survey 3: .57 ≤ r ≤ .61; Survey 4: .54 ≤ r ≤ .60). We also found an analogous pattern for participants' interest, academic self-concept, and sense of belonging regarding social sciences (Survey 1: .41 ≤ r ≤ .56; Survey 2: .51 ≤ r ≤ .57). However, within the social sciences, the correlations are smaller than within the natural sciences. Moreover, the correlations between these constructs hardly changed over time. Only the correlation between participants' academic self-concept and sense of belonging regarding social sciences increased noticeably (r = .41 in Survey 1 versus r = .57 in Survey 2).

Beyond that, our data analysis revealed that pre-service primary school teachers' (self-evaluated) interdisciplinary competence significantly correlates with their interest (.36 ≤ r ≤ .58), academic self-concept (.34 ≤ r ≤ .47), and sense of belonging regarding natural and social sciences (.35 ≤ r ≤ .41). Overall, these correlations range from weak to moderate. Furthermore, except for the correlation between participants' (self-evaluated) interdisciplinary competence and their interest in natural sciences, these correlations also hardly change over time (r = .44 in Survey 3 versus r = .58 in Survey 4).

## Discussion

### *Summary*

Within the present study, we investigated whether and to what extent pre-service primary school teachers' (self-evaluated) interdisciplinary competence, as well as their interest, academic self-concept, and sense of belonging regarding natural and social sciences, change during their undergraduate studies. Furthermore, we analyzed the extent to which these constructs correlate with one another. The results of our data analysis show that participants' academic self-concept regarding natural and social sciences did not change significantly over time, but their sense of belonging to both domains decreased significantly. Participants' interest in social sciences decreased significantly, while their interest in natural sciences increased significantly over time. Page's trend test indicates a small decrease in pre-service primary school teachers' (self-evaluated) interdisciplinary

**Table 5**

*Pooled correlations (Person's r) between participants' interest, self-concept, sense of belonging, and self-evaluated interdisciplinary competence*

| Survey 1 (winter semester 2020/2021) | | | | |
|---|---|---|---|---|
| | SB-S | SC-S | IN-S | IDC |
| SB-S | | .41*** | .56*** | .36** |
| SC-S | .41*** | | .51*** | .34** |
| IN-S | .56*** | .51*** | | .36** |
| IDC | .36** | .34** | .36** | |

| Survey 2 (summer semester 2021) | | | | |
|---|---|---|---|---|
| | SB-S | SC-S | IN-S | IDC |
| SB-S | | .57*** | .53*** | .35** |
| SC-S | .57*** | | .51*** | .43*** |
| IN-S | .53*** | .51*** | | .40** |
| IDC | .35** | .43*** | .40** | |

| Survey 3 (winter semester 2021/2022) | | | | |
|---|---|---|---|---|
| | SB-N | SC-N | IN-N | IDC |
| SB-N | | .57*** | .61*** | .41*** |
| SC-N | .57*** | | .60*** | .43*** |
| IN-N | .61*** | .60*** | | .44*** |
| IDC | .41*** | .43*** | .44*** | |

| Survey 4 (summer semester 2022) | | | | |
|---|---|---|---|---|
| | SB-N | SC-N | IN-N | IDC |
| SB-N | | .54*** | .60*** | .39** |
| SC-N | .54*** | | .60*** | .47*** |
| IN-N | .60*** | .60*** | | .58*** |
| IDC | .39** | .47*** | .58*** | |

Note. SB-S = sense of belonging to social sciences; SC-S = self-concept in social sciences; IN-S = interest in social sciences; SB-N = sense of belonging to natural sciences; SC-N = self-concept in natural sciences; IN-N = interest in natural sciences; IDC = self-evaluated interdisciplinary competence; **p < .01; ***p < .001.

competence across the four surveys of our study; particularly between Surveys 2 and 3, a significant decrease is evident. Beyond that, the correlations between pre-service primary school teachers' interest, academic self-concept, and sense of belonging regarding natural and social sciences were moderate to high and remained relatively stable over time. Participants' (self-evaluated) interdisciplinary competence correlated with the above-mentioned constructs at a weak to moderate level, which—with very few exceptions—also did not change over time.

*Conclusion and suggestions for future research*

With the limitations stated at the end of the Method section in mind, the results of the present study are, on one hand, in line with those of previous research (see Introduction). Consistent with the findings of other studies, we found that pre-service primary school teachers' interest, academic self-concept, and sense of belonging regarding natural sciences, although distinguishable, are substantially correlated with one another. Correspondingly, our data analysis also revealed a similar result for pre-service primary school teachers' interest, academic self-concept, and sense of belonging regarding social sciences. Additionally, within the present study, participants' interest, academic self-concept, and sense of belonging regarding natural and social sciences changed over time only weakly to moderately, if they changed at all. These results are also in line with the results of previous research indicating that these constructs are relatively stable and only change within university-based teacher education over extended time periods. An educational implication that derives from these results is that short-term interventions that aim to promote pre-service primary school teachers' interest, academic self-concept, or sense of belonging regarding natural and/or social sciences may show their intended outcomes—if at all—only to a very limited extent. Conversely, implementing the promotion of these motivational orientations as a cross-cutting and longitudinal component of university-based teacher education may be a more promising approach.

On the other hand, the present study, first and foremost, extends the current state of the research and provides starting points for potential future research. Our study revealed a significant decrease in pre-service primary school teachers' interest in social sciences and sense of belonging to natural and social sciences. One potential cause of these declines could be negative socialization experiences that pre-service primary teachers' may encounter during their undergraduate studies in natural and social sciences. Especially based on the results of previous research, it seems reasonable to assume that negative proficiency expectations on the part of lecturers toward pre-service primary school teachers (Muenks et al., 2020) and/or the labeling of pre-service primary school teachers as "second-class students" within faculties that predominantly award social status based on students' subject-specific knowledge and skills (Oevermann, 2010) are likely to negatively influence the development of pre-service primary school teachers' interest in and sense of belonging to university-based teacher education. Furthermore, a decrease in interest in social sciences and sense of belonging to natural and social sciences may also be influenced by cultural and societal expectations regarding primary teachers that have been explicitly and/or implicitly passed on to the participants (Feser & Haak, 2022). For example, societal messages such as "primary school teachers are generalist and not subject teachers like secondary school teachers" may shift pre-service primary teachers' focus to subject-unspecific or purely pedagogical aspects of their undergraduate studies, leading to a decrease in their sense of belonging to natural and social sciences and/or their interest in social sciences. As a result, future research should investigate whether the decreases in interest and sense of belonging encountered in the present study are replicable in other contexts of pre-service primary school teacher education. It would also be fruitful for future research to further examine the causes of such decreases.

In addition to a significant decrease in participants' interest in social sciences, the present study also revealed a significant increase in their interest in natural sciences. A potential cause for these changes in interest may be the different levels of satisfaction that pre-service primary school teachers have with their undergraduate studies in natural and social sciences. Adding to the above, another cause of the decrease in participants' interest in social sciences could be the fact that courses in their social science studies are not sufficiently aligned with the learning requirements and/or expectations of pre-service primary school teachers. As a result, participants may have developed the notion that social sciences, as academic disciplines, are less relevant or applicable to their future teaching practice, which, in turn, may have led to decreased interest in the social sciences. Conversely, the undergraduate natural science courses participants attended may have increased participants' interest in natural sciences by offering learning opportunities for developing knowledge and skills in natural science that they perceive as meaningful for their future teaching careers, or these courses may have provided participants with a solid understanding of the relevance of the natural science to primary science education (Marquant-Mau, 2001; Thomas & Durant, 1987). This interpretation is supported by the results of an evaluative survey accompanying the present study (for details, see Feser & Michalik, 2023). This survey revealed that pre-service primary school

teachers perceive the content of natural sciences courses within the university-based teacher education program "Sachunterricht" as more relevant to their future teaching careers than the content of social science courses. Additionally, several previous studies indicate that university-based teacher education programs whose content and structure are aligned with the professional development needs of pre-service primary school teachers positively influence their interest (e.g., Appleton, 1995; Beudels et al., 2021; Jarret,1999; Novak & Wisdom, 2018). Nevertheless, further research concerning the development of pre-service primary school teachers' interest should be conducted, especially because studies addressing their interest in social sciences have hardly been conducted thus far (for an exception, see Lenzgeiger, 2022).

The results of the present study also provide important insights into pre-service primary school teachers' (self-evaluated) interdisciplinary competence. On one hand, our data analysis revealed weak to moderate correlations between participants' (self-evaluated) interdisciplinary competence and their interest, academic self-concept, and sense of belonging regarding natural and social sciences. A similar result was found in previous studies by Scott et al. (2014; 2015), in which potential correlations between university students' (self-evaluated) interdisciplinary competence and their interest in physics, feeling of being recognized as a physics person, and physics identity were investigated. Therefore, it seems reasonable to assume that there may be limited but existing associations between pre-service primary school teachers' (self-evaluated) interdisciplinary competence and their interest, academic self-concept, and sense of belonging regarding natural and social sciences. However, the results of the present study certainly do not reveal any information about causal associations, specifically whether pre-service primary school teachers' interest, academic self-concept, and sense of belonging regarding natural and social sciences are impacted by their interdisciplinary competence, and/or vice versa. Nor can it be ruled out that the correlations just mentioned are spurious due to additional influencing or moderating factors (i.e., previous research hints that university students' curiosity is associated with both their interdisciplinary competence (Spelt et al., 2009) and their domain-specific affective-motivational orientations (e.g., Herpratiwi et al., 2018; Knecht, 2022; Yalız Solmaz, 2017)). Consequently, future research should be conducted to further clarify the associations between pre-service primary school teachers' (self-evaluated) interdisciplinary competence and their affective-motivational orientations related to natural and social sciences, such as their interest, academic self-concept, and sense of belonging.

Finally, on the other hand, within the present study, a small decrease in pre-service primary school teachers' (self-evaluated) interdisciplinary competence during their undergraduate studies in natural and social sciences was evident. Specifically, we found a moderate decrease between the two surveys in which the participants' curriculum moved from the social to the natural sciences (Surveys 2 and 3). One reasonable cause for this decline could be that the undergraduate courses participants attended solely promoted the acquisition of subject-specific knowledge and skills, rather than an interdisciplinary linking of such knowledge and skills. This, in turn, may have improved participants' knowledge and skills in natural and/or social sciences but, at the same time, may also counteracted any development regarding their interdisciplinary competence. Relatedly, especially when participants were shifting from social science to natural science courses, they may have encountered scientific concepts and methods that they perceived as new and unfamiliar. This potentially constituted a challenge to their interdisciplinary thinking and, accordingly, could explain the decrease in their (self-evaluated) interdisciplinary competence. However, an alternative interpretation of the observed decrease is based on the fact that, in the present study, participants' interdisciplinary competence was surveyed via self-evaluations and that, in educational research, it is well known that students may tend to overestimate their own competencies, especially when their competencies are low (e.g., Lindsey & Nagel, 2015; Mahmood, 2016; Kruger & Dunning, 1999). In all four surveys of the present study, the mean of the participants' (self-evaluated) interdisciplinary competence was notably higher than the theoretical mean on a five-point Likert scale ($3.39 \leq M_{1/2/3/4} \leq 3.55$ versus $M_{theoretical} = 3.00$; see Table 4). In other words, participants self-evaluated their interdisciplinary competence as fairly high, and the significant decrease in their self-evaluated interdisciplinary competence describes a shift toward the medium level. Thus, assuming that, at the beginning of their studies, pre-service primary school teachers' interdisciplinary competence was low rather than high and developed positively during their undergraduate studies, it may also be plausible that participants initially overestimated their interdisciplinary competence but, over time, gradually adopted a more realistic self-evaluation. This development, in turn, may have manifested itself in the empirical data as a decrease in participants' (self-evaluated) interdisciplinary competence. In our view, both these interpretations of the decrease in participants' (self-evaluated) interdisciplinary within the present study are quite reasonable. In order to clarify the extent to which one of these interpretations more accurately describes the change in primary school teachers' (self-evaluated) interdisciplinary competence during their undergraduate studies, further research is needed. In this regard, future research that aims to replicate the

results of the present study are needed, particularly studies that survey primary school teachers' interdisciplinary competence using measures other than self-evaluations.

**Acknowledgements**

We thank all participating pre-service primary school teachers and the ProfaLe project of the Universität Hamburg for supporting this research.

**Funding**

The present study is part of the ProfaLe project of the Universität Hamburg. The ProfaLe project is funded by Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research): [Grant Number 01JA1811].

**Footnotes**

[a]An approximate translation for Sachunterricht is natural sciences and social studies in primary education. Literally, the term Sachunterricht translates as education (German: Unterricht) on factual matters (German: Sache). For a comparison of Sachunterricht and the current primary school curricula in other European countries, see Blaseio (2021).

[b]More precisely, a further and very important domain integrated within Sachunterricht is technology education (Gesellschaft für Didaktik des Sachunterrichts, 2013). However, because technology education within Sachunterricht is not addressed in the present study, it is not discussed further in this paper.

[c]$\rho$ is the overall rank correlation between Surveys 1 to 4 (in reversed ordering) and the data (Page, 1963). Due to the reversed ordering of Surveys 1 to 4 within its calculation, $\rho$ is positive. Interpreted as an effect size, $\rho$ indicates an overall small decrease of participants' (self-evaluated) interdisciplinary competence over time.

**References**

An, B. G., & Zakaria, A. R. (2022). A scoping review of teacher training during COVID-19 pandemic. *International Education Studies*, *15*(2), 102–112. https://doi.org/10.5539/ies.v15n2p102

Appleton, K. (1995). Student teachers' confidence to teach science: Is more science knowledge necessary to improve self-confidence? *International Journal of Science Education*, *17*(3), 357–369. https://doi.org/10.1080/0950069950170307

Appleton, K. (2003). How do beginning primary school teachers cope with science? Toward an understanding of science teaching practice. *Research in Science Education*, *33*(1), 1–25. https://doi.org/10.1023/A:1023666618800

Beudels, M., Schroeder, N., & Preisfeld, A. (2021). „Ich traue mir zu ...". Effekte einer interdisziplinären Lehrveranstaltung auf motivationale Orientierungen angehender Sachunterrichtslehrkräfte [Effects of an interdisciplinary lecture course on pre-service primary school teachers' motivational orientations]. *PraxisForschungLehrer*innenBildung 3*(1), 188–220. https://doi.org/10.11576/PFLB-4845

Blaseio, B. (2021). Sachunterricht in Europa – Fachstrukturen für das geschichtliche, geographische und naturwissenschaftliche Lernen in der Grundschule [Sachunterricht in Europe – subject structures for learning history, geography and natural sciences in primary school]. *GDSU-Journal, 12*, 9–25. https://gdsu.de/sites/default/files/gdsu-info/files/9_25_12.pdf

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies. Competence viewed as a continuum. *Zeitschrift Für Psychologie*, *223*(1), 3–13.

Brassler, M., & Dettmers, J. (2017). How to enhance interdisciplinary competence—interdisciplinary problem-based learning versus interdisciplinary project-based learning. *Interdisciplinary Journal of Problem-Based Learning*, *11*(2). https://doi.org/10.7771/1541-5015.1686

Claus, A. M., & Wiese, B. S. (2019). Development and test of a model of interdisciplinary competencies. *European Journal of Work and Organizational Psychology*, *28*(2), 191–205. https://doi.org/10.1080/1359432X.2019.1567491

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates Publishers.

Cox, B. E., McIntosh, K., Reason, R. D., & Terenzini, P. T. (2014). Working with missing data in higher education research: A primer and real-world example. *The Review of Higher Education*, *37*(3), 377–402. https://doi.org/10.1353/rhe.2014.0026

Dickhäuser, O., Schöne, C., Spinath, B., & Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept [The Academic Self Concept Scale]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *23*(4), 393–405. https://doi.org/10.1024//0170-1789.23.4.393

Elsholz, M. (2019). *Das akademische Selbstkonzept angehender Physiklehrkräfte als Teil ihrer professionellen Identität [The academic self-concept of pre-service physics teachers as part of their professional identity]*. Logos.

Engelhardt, P. M. (2019). *Evaluation interdisziplinärer Kompetenzen von Studierenden [Evaluation of university students' interdisciplinary competence]* [Doctoral dissertation, Universität zu Köln]. http://nbn-resolving.de/urn:nbn:de:hbz:38-104692

Feser, M. S. (2020). Sense of belonging to science—Entwicklung eines Erhebungsinstruments für Lehramtsstudierende [Development of an instrument measuring per-service teachers' sense of belonging to natural science]. *Progress in Science Education*, *3*(2), 10–21. https://doi.org/10.25321/prise.2020.968

Feser, M. S. (2021). Students' and student teachers' sense of belonging to science: What do we know so far? *Journal of Baltic Science Education*, *20*(3), 340–343. https://doi.org/10.33225/jbse/21.20.340

Feser, M. S., & Haak, I. (2022). Key features of teacher identity: A systematic meta-review study with special focus on teachers of science or science-related subjects. *Studies in Science Education*, Advance online publication. https://doi.org/10.1080/03057267.2022.2108644

Feser, M. S., & Michalik, K. (2023). *Längsschnittliche Evaluation des natur- und gesellschaftswissenschaftlichen Sachunterrichtsstudiums der Universität Hamburg: Dokumentation der deskriptiven Ergebnisse und der Erhebungsinstrumente [Longitudinal evaluation of the* university-based teacher education program "Sachunterricht" at the Universität Hamburg. Technical report*]*. BoD – Books on Demand.

Feser, M. S., & Plotz, T. (2023). *Exploring the validity of a single-item instrument in assessing pre-service primary school teachers' sense of belonging to science*. EdarXiv. https://doi.org/10.35542/osf.io/djs4b

Gabriel-Busse, K., Kastens, C. P., & Kucharz, D. (2018). Fachspezifisch oder nicht? – Eine Studie zur Analyse der Binnenstruktur des Selbstkonzepts Sachunterricht [Subject-specific or not? – A study to analyze the internal structure of students' self-concept regarding Sachunterricht]. *Zeitschrift für Grundschulforschung*, *11*(2), 333–348. https://doi.org/10.1007/s42278-018-0017-1

Gesellschaft für Didaktik des Sachunterrichts (2003). *Perspectives Framework for General Studies in Primary Education*. Verlag Julius Klinkhardt.

Gesellschaft für Didaktik des Sachunterrichts (2013). *Perspektivrahmen Sachunterricht. Vollständig überarbeitete und erweiterte Ausgabe [Perspectives Framework for General Studies in Primary Education. Comprehensively revised and extended edition]*. Verlag Julius Klinkhardt.

Gesellschaft für Didaktik des Sachunterrichts (2019). *Qualitätsrahmen Lehrerbildung Sachunterricht und seine Didatik [Quality framework for teacher education on teaching Sachunterricht]*. Verlag Julius Klinkhardt.

Hartmann, C. (2018). Zur Bedeutsamkeit formaler Lerngelegenheiten für das fachdidaktische Wissen im Bereich der historischen Perspektive des Sachunterrichts [The importance of formal learning opportunities for pedagogical content knowledge about teaching history in primary school]. *Zeitschrift für Grundschulforschung*, *11*(2), 269–284. https://doi.org/10.1007/s42278-018-0016-2

Hasni, A., Lenoir, Y., & Froelich, A. (2015). The case of science, technology, and mathematics teachers in Quebec. *Issues in Interdisciplinary Studies*, *33*, 144–180. https://interdisciplinarystudies.org/docs/Vol33_2015/06_Vol_33_pp_144-180.pdf

Herpratiwi, Darsono, Sasmiati, & Pujiyatli. (2018). Cooperative learning and learning achievement in social science subjects for sociable students. *Eurasian Journal of Educational Research*, *75*, 81–98. https://doi.org/10.14689/ejer.2018.75.5

Heymans, M.W., & Eekhout I. (2019). *Applied missing data analysis using SPSS and (R)Studio.* https://bookdown.org/mwheymans/bookmi/

Jarrett, O. S. (1999). Science interest and confidence among preservice elementary teachers. *Journal of Elementary Science Education*, *11*(1), 49–59. https://doi.org/10.1007/BF03173790

Page, E. B. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association*, *58*(301), 216–230.

Pharo, E. J., Davison, A., Warr, K., Nursey-Bray, M., Beswick, K., Wapstra, E., & Jones, C. (2012). Can teacher collaboration overcome barriers to interdisciplinary learning in a disciplinary university? A case study using climate change. *Teaching in Higher Education*, *17*(5), 497–507. https://doi.org/10.1080/13562517.2012.658560

Petrie, H. G. (1992). Interdisciplinary education: Are we faced with insurmountable opportunities? *Review of Research in Education*, *18*(1), 299–333. https://doi.org/10.3102/0091732X018001299

Kahlert, J. (2014). Zwischen didaktischer Vernunft und inhaltlicher Beliebigkeit—Vielperspektivität im Sachunterricht [Multiperspectivity within Sachunterricht]. *Zeitschrift für Grundschulforschung*, *7*(2), 9–21.

Kalcsics, K. (2021). Es geht nur im Team. Plädoyer für enge fachliche Zusammenarbeit im Sachunterricht [A plea for close professional collaboration when teaching Sachunterricht]. *PraxisForschungLehrer*innenBildung, 3*(5), 123–130. https://doi.org/10.11576/PFLB-4781

Kan, L., Degotardi, S., & Li, H. (2022). The impact of COVID-19 on the delivery of preservice teacher education: A scoping review. *Asia Pacific Education Review*. https://doi.org/10.1007/s12564-022-09816-y

Kirsch, A. (2022). Messung von Unterrichtsplanungskompetenz im Fach Sachunterricht: Empirische Befunde zur Kompetenzentwicklung angehender Grundschullehrkräfte auf Grundlage einer mehrdimensionalen Kompetenzstruktur [Measuring lesson planning competence in the subject of Sachunterricht: Empirical findings on the competence development of prospective primary school teachers on the basis of a multidimensional competence structure]. *Zeitschrift für Erziehungswissenschaft*, *25*(4), 815–842. https://doi.org/10.1007/s11618-022-01116-w

Klafki, W. (1992). Allgemeinbildung in der Grundschule und Bildungsauftrag des Sachunterrichts [General education in primary school and educational mandate of Sachunterricht]. In R. Lauterbach, W. Köhnlein, K. Spreckelsen, & E. Klewitz (Eds.), *Brennpunkte des Sachunterrichts [Focus of Sachunterricht]* (pp. 11–31). Institut für Pädagogik der Naturwissenschaften und Gesellschaft für Didaktik des Sachunterrichts.

Kleickmann, T. (2015). *Professionelle Kompetenz von Primarschullehrkräften im Bereich des naturwissenschaftlichen Sachunterrichts [Primary school teachers' professional competence for teaching natural sciences]. Zeitschrift für Grundschulforschung*, 8(1), 7–22.

Knecht, M. (2022). Investigating nontraditional first-year students' epistemic curiosity during the research process: An exploratory, mixed-methods study. *College & Research Libraries*, *83*(6). https://doi.org/10.5860/crl.83.6.874

Köhnlein, W. (2015). Aufgaben und Ziele des Sachunterrichts [Aims and goals of Sachunterricht]. In J. Kahlert, M. Fölling-Albers, M. Götz, S. Miller, & S. Wittkowske (Eds.), *Handbuch Didaktik des Sachunterrichts [Handbook of Education in Sachunterricht]* (pp. 88–95). Verlag Julius Klinkhardt.

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, *33*(1), 27–50. https://doi.org/10.1080/09500693.2010.518645

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Künzli David, C., Gysin, S., & Bertschy, F. (2016). Sachunterricht als inter- und transdisziplinär konstituiertes Fach – Ansprüche an die Unterrichtsgestaltung und Überlegungen im Hinblick auf die Lehrerinnen- und Lehrerbildung [Sachunterricht as an inter- and transdisciplinary subject - demands on lesson design and considerations regarding teacher education]. *Beiträge zur Lehrerinnen- und Lehrerbildung*, *34*(3), 305–316. https://doi.org/10.36950/bzl.34.2016.9493

Kultusministerkonferenz (2015). *Empfehlung zur Arbeit in der Grundschule (Beschluss der Kultusministerkonferenz vom 02.07.1970 i. d. F. vom 11.06.2015) [Recommendation on the work in primary school (Decision by the Standing Conference of the Ministers of Education and Cultural Affairs from July 2, 1970 as amended on June 11, 2015)]*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1970/1970_07_02_Empfehlungen_Grundschule.pdf

Lange, K. (2015). Professionelle Kompetenzen von Lehrkräften im Sachunterricht [Professional competencies of teachers teaching Sachunterricht]. In J. Kahlert, M. Fölling-Albers, M. Götz, S. Miller, & S. Wittkowske (Eds.), *Handbuch Didaktik des Sachunterrichts [Handbook of Education in Sachunterricht]* (pp. 82–86). Verlag Julius Klinkhardt.

Lange, K., Kleickmann, T., Tröbst, S., & Möller, K. (2012). Fachdidaktisches Wissen von Lehrkräften und multiple Ziele im naturwissenschaftlichen Sachunterricht [Teachers' pedagogical content knowledge and multiple goals in science education]. *Zeitschrift für Erziehungswissenschaft*, *15*(1), 55–75. https://doi.org/10.1007/s11618-012-0258-z

Lattuca, L. R., Knight, D., & Bergom, I. (2013). Developing a measure of interdisciplinary competence. *International Journal of Engineering Education*, *29*(3), 726–739. https://www.ijee.ie/latestissues/Vol29-3/19_ijee2724ns.pdf

Lenoir, Y., & Hasni, A. (2016). Interdisciplinarity in primary and secondary school: Issues and perspectives. *Creative Education*, *07*(16), 2433–2458. https://doi.org/10.4236/ce.2016.716233

Lenzgeiger, B. (2022). Veränderung von Interesse und Selbstwirksamkeitsüberzeugungen Grundschullehramtsstudierender im Bereich der politischen Bildung [Changeability of teacher interest and teacher self-efficacy among primary school teachers-in-training in the field of political education]. *Zeitschrift für Grundschulforschung*, *15*(1), 205–220. https://doi.org/10.1007/s42278-021-00136-5

Lindsey, B. A., & Nagel, M. L. (2015). Do students know what they know? Exploring the accuracy of students' self-assessments. *Physical Review Special Topics - Physics Education Research*, *11*(2), Article 020103. https://doi.org/10.1103/PhysRevSTPER.11.020103

Mahmood, K. (2016). Do people overestimate their information literacy skills? A systematic review of empirical evidence on the Dunning-Kruger effect. *Communications in Information Literacy*, *10*(2), 199–213. https://doi.org/10.15760/comminfolit.2016.10.2.24

Mansilla, V. B., & Duraising, E. D. (2007). Targeted assessment of students' interdisciplinary work: An empirically grounded framework proposed. *The Journal of Higher Education*, *78*(2), 215–237. https://doi.org/10.1080/00221546.2007.11780874

Marquant-Mau, B. (2001). Scientific Literacy im Sachunterricht? [Scientific Literacy within Sachunterricht?] In D. Cech, W. Köhnlein, & H. Schreier (Eds.), *Die Aktualität der Pädagogik Martin Wagenscheins für den Sachunterricht [The topicality of Martin Wagenschein's pedagogy for teaching Sachunterricht]* (pp. 185–201). Verlag Julius Klinkhardt.

Marquardt-Mau, B., Köhnlein, W., Cech, D., & Lauterbach, R. (1996). *Lehrerbildung Sachunterricht [Teacher education on teaching Sachunterricht]*. Verlag Julius Klinkhardt.

Meschede, N., Fiebranz, A., Möller, K., & Steffensky, M. (2017). Teachers' professional vision, pedagogical content knowledge and beliefs: On its relation and differences between pre-service and in-service teachers. *Teaching and Teacher Education*, *66*, 158–170. https://doi.org/10.1016/j.tate.2017.04.010

Meschede, N., Hartinger, A., & Möller, K. (2020). Sachunterricht in der Lehrerinnen- und Lehrerbildung [Sachunterricht in teacher education]. In C. Cramer, J. König, M. Rothland, & S. Blömeke, *Handbuch Lehrerinnen- und Lehrerbildung [Handbook on teacher education]*. Verlag Julius Klinkhardt.

Muenks, K., Canning, E. A., LaCosse, J., Green, D. J., Zirkel, S., Garcia, J. A., & Murphy, M. C. (2020). Does my professor think my ability can change? Students' perceptions of their STEM professors' mindset beliefs predict their psychological vulnerability, engagement, and performance in class. *Journal of Experimental Psychology: General, 149*(11), 2119–2144. https://doi.org/10.1037/xge0000763

Niermann, Anne. (2017). *Professionswissen von Lehrerinnen und Lehrern des Mathematik- und Sachunterrichts [Professional knowledge of teachers teaching mathematics and Sachunterricht]*. Verlag Julius Klinkhardt.

Nissen, J., Donatello, R., & Van Dusen, B. (2019). Missing data and bias in physics education research: A case for using multiple imputation. *Physical Review Physics Education Research*, *15*(2), 020106. https://doi.org/10.1103/PhysRevPhysEducRes.15.020106

Novak, E., & Wisdom, S. (2018). Effects of 3d printing project-based learning on preservice elementary teachers' science attitudes, science content knowledge, and anxiety about teaching science. *Journal of Science Education and Technology*, *27*(5), 412–432. https://doi.org/10.1007/s10956-018-9733-5

Oevermann, U. (2010). Der Gegenbegriff zur Gesellschaft ist nicht Natur sondern Kultur [The counter concept to society is not nature but culture]. In F. Herrschaft & K. Lichtblau (Eds.), *Soziologie in Frankfurt. Eine Zwischenbilanz [Sociology in Frankfurt. An interim report]* (pp. 369–406). VS Verlag für Sozialwissenschaften.

Pendergast, D., Ronksley-Pavia, M., & Dewhurst, Y. (2020). A systematic quantitative literature review of pre-service teachers' sense of belonging during school-based experience. *International Journal of Home Economics*, *13*(2), 59–79.

Rau, S. (2017). *Durchführung von Sachunterricht im Vorbereitungsdienst: Eine längsschnittliche, videobasierte Unterrichtsanalyse [Teaching Sachunterricht by pre-service teachers during their traineeship: a longitudinal, video-based study]*. Logos.

Reichhart, B. (2018). *Lehrerprofessionalität im Bereich der politischen Bildung: Eine Studie zu motivationalen Orientierungen und Überzeugungen im Sachunterricht [Teacher professionalism in political science education: a study on motivational orientations and beliefs of teachers teaching Sachunterricht]*. Springer

Richter, D. M., & Paretti, M. C. (2009). Identifying barriers to and outcomes of interdisciplinarity in the engineering classroom. *European Journal of Engineering Education*, *34*(1), 29–45. https://doi.org/10.1080/03043790802710185

Rubin, D. B. (Ed.). (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons, Inc.

Schmidt, M. (2015). *Professionswissen von Sachunterrichtslehrkräften [Professional knowledge of teachers teaching Sachunterricht]*. Logos.

Scott, T. D., Hazari, Z., & Potvin, G. (2014). Interdisciplinary thinking and physics identity. *2013 Physics Education Research Conference Proceedings*, 329–332. https://doi.org/10.1119/perc.2013.pr.070

Scott, T. D., Hazari, Z., Potvin, G., Sadler, P. M., & Sonnert, G. (2015). Interdisciplinary affinity: Definitions and connections to physics identity. *2014 Physics Education Research Conference Proceedings*, 235–238. https://doi.org/10.1119/perc.2014.pr.055

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: validation of construct interpretations. *Review of Educational Research*, *46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Spelt, E. J. H., Biemans, H. J. A., Tobi, H., Luning, P. A., & Mulder, M. (2009). Teaching and learning in interdisciplinary higher education: A systematic review. *Educational Psychology Review*, *21*(4), 365–378. https://doi.org/10.1007/s10648-009-9113-z

Sothayapetch, P., Lavonen, J., & Juuti, K. (2013). Primary school teachers' interviews regarding pedagogical content knowledge (PCK) and general pedagogical knowledge (GPK). *European Journal of Science and Mathematics Education*, *1*(2), 84–105. https://doi.org/10.30935/scimath/9390

Tänzer, S. (2014). Konzeptionen und Positionen der Didaktik des Sachunterrichts [Concepts and Positions of Education in Sachunterricht]. In Gesellschaft für Didaktik des Sachunterrichts (Eds.), *Die Didaktik des Sachunterrichts und ihre Fachgesellschaft GDSU e. V. [Education in Sachunterricht and its academic society GDSU e. V.]* (pp. 57–74). Verlag Julius Klinkhardt.

Tekel, E., Öztekin Bayır, Ö., & Dulay, S. (2022). Teaching practicum during the covid-19 pandemic: A comparison of the practices in different countries. *International Journal of Progressive Education*, *18*(2), 71–86. https://doi.org/10.29329/ijpe.2022.431.5

The R Foundation for Statistical Computing (2023). The R project for statistical computing. https://www.r-project.org/

Thiel, S. (2003). Die Zusammenarbeit von Natur- und Sozialwissenschaften im Sachunterricht [Collaboration between natural and social sciences within Sachunterricht]. In H.-W. Kuhn (Ed.), *Sozialwissenschaftlicher Sachunterricht [Social sciences within Sachunterricht]* (pp. 287–303). Centaurus Verlag.

Thomas, G., & Durant, J. (1987). Why should we promote the public understanding of science? *Scientific Literacy Papers: A Journal of Research in Science, Education and Research, 1*, 1–14.

Tormey, R., & Laperrouza, M. (2023). The development, validation and use of an interprofessional project management questionnaire in engineering education. *European Journal of Engineering Education*. Advanced online publication. https://doi.org/10.1080/03043797.2023.2171854

Universität Hamburg (2022, September 20). *Unterrichtsfach Sachunterricht [The school subject Sachunterricht]*. https://www.ew.uni-hamburg.de/studium/unterrichtsfach-sachunterricht.html

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in *R*. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Wang, J., & Sneed, S. (2019). Exploring the design of scaffolding pedagogical instruction for elementary preservice teacher education. *Journal of Science Teacher Education*, *30*(5), 483–506. https://doi.org/10.1080/1046560X.2019.1583035

Watteler, O., & Ebel, T. (2019). Datenschutz im Forschungsdatenmagement [Data protection in data management for research]. In U. Jensen, S. Netscher, & K. Weller (Eds.), *Forschungsdatenmanagement sozialwissenschaftlicher Umfragedaten: Grundlagen und praktische Lösungen für den Umgang mit quantitativen Forschungsdaten [Data management in social science surveys: Principles and practical solutions for handling quantitative research data.]*. Verlag Barbara Budrich.

Wilhelm, M., & Brühwiler, C. (2016). Professionelle Kompetenzen für das Unterrichten interdisziplinärer Schulfächer – Desiderata in der Lehrerinnen- und Lehrerbildung [Professional competencies for teaching interdisciplinary school subjects - desiderata in teacher education]. *Beiträge zur Lehrerinnen- und Lehrerbildung*, *34*(3), 284–293. https://doi.org/10.36950/bzl.34.2016.9491

Yalız Solmaz, D. (2017). Relationship between lifelong learning levels and information literacy skills in teacher candidates. *Universal Journal of Educational Research*, *5*(6), 939–946. https://doi.org/10.13189/ujer.2017.050605

Yurek, L. A., Vasey, J., & Sullivan Havens, D. (2008). The use of self-generated identification codes in longitudinal research. *Evaluation Review*, *32*(5), 435–452. https://doi.org/10.1177/0193841X08316676

iejee

International Electronic Journal of Elementary Education
June 2023, Volume 15, Issue 5, 399-407

# Moodle-Based Development of Primary Education under Martial Law

Marianna Ostrovska[a,*], Olha Tretiak[b], Halyna Smolnykova[c], Olena Studenets[d], Maria Ivanchuk[e]

Received : 18 March 2023
Revised : 14 June 2023
Accepted : 20 June 2023
DOI : 10.26822/iejee.2023.308

[a] **Corresponding Author:** Marianna Ostrovska, Department of Pedagogy and Psychology, Ferenc Rákóczi II Transcarpathian Hungarian College of Higher Education, Beregovo, Ukraine.
E-mail: pedagog.m23@gmail.com
ORCID: https://orcid.org/0000-0001-8810-3810

[b] Olha Tretiak, Department of Science and Mathematics Education and Technologies, Institute of In-Service Education, Borys Hrinchenko Kyiv University, Kyiv, Ukraine.
E-mail: ottretiak@gmail.com
ORCID: https://orcid.org/0000-0002-1160-055X

[c] Halyna Smolnykova, Department of Science and Mathematics Education and Technologies, Institute of In-Service Education, Borys Hrinchenko Kyiv University, Kyiv, Ukraine.
E-mail: galasmolnikova@gmail.com
ORCID: https://orcid.org/0000-0001-5853-146X

[d] Olena Studenets, Department of Theory and Methods of Preschool and Primary Education, Khmelnytskyi Institute of Post-Graduate Pedagogical Education, Khmelnytskyi, Ukraine.
E-mail: olenastud@gmail.com
ORCID: https://orcid.org/0000-0003-2871-0621

[e] Maria Ivanchuk, Department of Pedagogic and Methodics of Primary School Teaching, Faculty of Pedagogy, Psychology and Social Work, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine.
E-mail: mariaivanchuuuk@gmail.com
ORCID: https://orcid.org/0000-0002-8104-9724



## Abstract

A specific feature of the organization of education during wartime is the creation of a modern educational environment for the harmonious development of students in conditions of war. The substantiation of ways of developing qualities, values and competences of primary school students is important for solving a number of contradictions between the quality training of students to be competitive in the process of building labour relations at age stages, and the insufficient level of their real preparation for productive activities. There is also a contradiction between a significant volume of promising information in the conditions of rapidly changing technologies and unadjusted educational environment to innovative changes in terms of its analysis and use. The aim of the study was to determine the impact of the Moodle platform on the academic performance of primary schoolers under martial law. The aim was achieved through the use of the following methods: questionnaire survey, testing, methods of statistical analysis. It was found that the introduction of the modernized curriculum in the Moodle platform had a positive effect on the academic performance of the students of the experimental group (EG) compared to the students who used the traditional platform. The crosswords, interactive assignments, as well as animated videos in case of correct completion of tests were also added to encourage students to complete additional assignments after the lesson. This also affected the positive result of the final stage of testing. A promising direction for further research may be the development of interactive tasks for primary schoolers to better assimilate the material they have learned.

**Keywords:**

Learning Management System, Moodle, Primary School, Distance Education, Martial Law

## Introduction

February 2022 divided the life of every Ukrainian into the periods before and after the war. All spheres of everyday life of society, including the educational sector, underwent global changes during the martial law. But, despite the dangerous situation in the country, educational institutions continue to provide educational services to students using online learning. The ability to obtain permanent access to education played an important role in the lives of students, parents and teachers. The necessary transition to an unusual format for obtaining educational services has contributed

to the development of the latest platforms and tools for online education, as well as online courses. This demonstrated that certain face-to-face educational elements can be transferred to an online environment, and some courses can be completely transferred to an online format. This effectively promotes the modernization of online learning, which has already been introduced and described as a form of flexible learning for part-time students or students who are unable to attend offline classes because of geographical distance (Adedoyin & Soykan, 2020). However, distance education takes away the desired experience of collaborative learning, socialization, and live communication (Putri et al., 2020; Tang & Braver 2020).

The COVID-19 pandemic gave an impetus to the understanding of the need to turn to a distance learning. In recent years, there has been a tendency to increasing transition of educational institutions to online learning. It is influenced by many economic, social and political factors, and their interaction on an international scale (Yazgan, 2022). However, researchers identified a low level of performance of most distance online environments in motivating students to learn, which is later reflected in their results during annual assessments, while there is almost no consistent strategy for e-learning (Abernathy, 2019).

It was determined that most studies on the impact of COVID-19 on education are focused on higher education only and, in most cases, depend on the academic subject. For example, there have been studies on the impact of limitations in education caused by the COVID-19 pandemic on the performance of medical students (Owusu-Fordjour et al., 2020), management students (Brammer & Clark, 2020), and economics students (Roman & Plopeanu, 2021). Therefore, it was decided to study the effectiveness of the introduction of the Moodle educational platform into the educational system of primary school students in order to enhance their motivation to study and effectively consolidate the acquired knowledge during distance education.

So, the aim of the study was to determine the impact of the Moodle platform on the development of primary education under martial law.

The aim was achieved by fulfilling the following research objectives:

1. Determining the effectiveness of traditional teaching and learning tools;
2. Determining the level of understanding of the material learnt by primary schoolers using the Moodle platform;
3. Conducting a survey of teachers, students and parents regarding the usability of the implemented Moodle platform.

Research hypothesis: If both teachers and parents perceive the use of the Moodle platform as a positive aspect of supplementing learning, an increased students' ability to develop more effective learning skills can be expected.

**Literature Review**

The most common topic of scientific discussions is the adaptation of educational services to distance learning. During the invasion of the Russian Federation on the territory of Ukraine, the education system faced unprecedented anthropological and social dynamics in combination with technological development. The problem of such dynamics (Laurent et al., 2020) has led to significant transformations in teaching and learning.

Many prospects have emerged during the wartime period and after the COVID-19 pandemic, and an analysis of recent research and publications reveals the current vectors of the digital transformation of education (Wang, 2021; Jones & McNulty, 2022). Digital transformation of education during martial law is a measure aimed at distance learning for the possibility of continuous training of learners. The effectiveness of the distance platform chosen by the teacher will have a significant impact on the quality of education from kindergarten to higher educational institution. Considering the issue in this context, the teacher needs to solve a number of organizational issues, taking into account the students' abilities and the conditions of martial law: find the key elements of the distance course, choose the right strategy for learning, create a high-quality video for asynchronous learning, develop distance learning assessments, set up remote group work, properly select synchronous virtual class tools and asynchronous platforms.

At the beginning of the military invasion, distance learning in Ukraine underwent changes that could have developed for years at other times in a usual environment. Burde et al. (2017) analysed the deployment of pedagogical engineering in distance education in emergency settings. Wang (2021) analysed the transformation of distance learning over the past five years. The study is about "emergency" distance learning using the example of 2020. As a result of the pandemic, all educational institutions were closed, offline classes were cancelled, and in a few days the whole world was forced to switch to distance education.

Engzell et al. (2021) determined whether school closures caused by the pandemic affected primary schoolers' performance, and identified a learning loss equivalent to one-fifth of the school year. The loss was higher than for students from less educated families. In the study of Pozo-Rico et al. (2020) analysed whether a teacher training programme is effective during the

COVID-19 period. It was found during the experiment that teachers who completed the programme coped with stress and information and communication technologies better. However, it does not examine the direct impact of distance learning on the aspects of primary education.

Moorhouse and Beaumont (2020a) studied teachers' use of a video conferencing system (VCS) to deliver synchronous online lessons to primary schoolers in Hong Kong. Teachers were given several suggestions for synchronous learning, such as giving students time to familiarize themselves with the technology using different VCS features (gallery view and screen sharing), and incorporating synchronous lessons into learning. It was suggested to include a pre-task to prepare students and a post-task to extend learning beyond the classroom.

The analysis gives grounds to state that most of the conducted studies were focused on the integration or mixing of digital technologies in offline learning. For example, Tay et al. (2017) analysed the implementation of digital technologies in primary schools in Singapore. It was determined that successful integration depends on the teaching approaches adopted by teachers. Bond (2020) conducted a systematic review of 110 academic publications on the effectiveness of flipped classrooms in primary and secondary school. The author found a positive relationship between flipped classrooms, student engagement and learning, with teacher-made videos, and the use of shared digital technologies (for example, LMS). However, the author also reported the need for further research to identify teachers' perceptions. Pulham and Graham (2018) conducted a review of educational documents to identify the competencies of primary and secondary school teachers required to deliver blended learning, along with technological knowledge such as the use of LMS, the ability to be flexible and personalize teaching content. Agreeing with Bond (2020), Pulham and Graham (2018) identified the lack of published research on blended learning in primary and secondary school.

However, it is important to note the difference between distance education during COVID-19 and wartime. During the pandemic, the main problem in the education system was the transition to a distance learning, because it was almost never used before. After the adaptation period, the students were able to obtain the necessary knowledge without hindrance and attend classes at the time specially allocated for this purpose. However, the war in Ukraine made adjustments to the educational environment, which led to the destabilization of the educational process at all levels. Under martial law, children switched to blended learning or to distance learning. But because of the lack of electricity, schools lost the opportunity to carry out the educational process systematically, which led to an imbalance in the equal acquisition of knowledge by all learners. Some of the schoolchildren were forced to leave the country in search of safety, and did not have the opportunity to attend distance classes, or they moved to safer cities in Ukraine without access to the Internet.

Therefore, it is very important that the educational process be organized so that all educators have the opportunity to solve the key problems of education that arose during the war and to guarantee maximum safety for the participants of the educational process with the simultaneous assimilation of the necessary high-quality knowledge.

Summarizing the above, we can conclude that recent global changes have been constantly affecting all forms of social life, and most of all, the educational system. Therefore, it gives impetus to the development of digital technologies.

Therefore, it can be concluded that global changes have influenced the development of digital technologies. At the beginning, the impact of digital technologies on education was almost invisible because the general technologies had limited utility in the classroom or the work of teachers (Cuban, 2001). In recent years, digital technologies specifically designed for learning, such as presentation software, Learning Management System (LMS), Student Response System (SRS), and game-based learning platforms (GBLP) have begun to play a significant role in education (Moorhouse & Beaumont, 2020b; Voogt et al., 2018). These changes have coincided with the development of Internet-ready portable devices such as tablets and Wi-Fi connectivity, meaning that teachers no longer need to use desktop computers and also take students to computer labs (Hockly & Dudeney, 2018).

Learning Management System is software that enables applying educational content and administer the learning and teaching process. An LMS uses human and technological resources that appropriately enhance learning into an environment for developing learning content designed for teachers as course leaders. Given that our aim in the research was to determine whether the use of e-learning systems can improve the educational process in primary school, the Moodle system was chosen as the learning environment. Moodle is a learning management system, but it is also perceived as a virtual learning environment.

The acronym Moodle is a modular object-oriented dynamic learning environment. Modularity means that the system accommodates smaller modules, activities that together form a whole, as in any e-learning system. The Moodle learning environment is dynamic

and the student takes an active part in the learning process. It is an open source system that allows users of the system to make changes to the programs and adapt them to their own needs. The Moodle toolset includes the following features: creation of a large number of courses using one system, course planning, activity schedule, calendar, user management, management of user roles and user groups in the course, working with existing files and educational content, testing and evaluating users, monitoring user activities, numerous tools for communication and user interaction, system management — backups, statistics, logs, advanced reference system.

The interaction of technology and education shapes e-learning, and their connection is strongly supported by those who develop educational content in e-learning systems. Well-developed and applied educational content in the e-learning system enables implementing various learning, teaching and knowledge testing scenarios.

## Materials and Methods

### *Design*

The research methodology was based on a combination of qualitative and quantitative empirical methods and provided for an interpretive approach to the analysis of the obtained data. The experimental part of the study lasted from September 2022 to the end of December 2022 (Table 1).

### Participants

For the reliability of the conducted experiment, the research was conducted at 4 secondary schools (2 urban and 2 rural schools) in the city of Uzhhorod. The experimental part of the study was designed for 4$^{th}$ grade students. This is justified by the fact that students will be able to independently answer questions in Google Forms and propose what they lacked on the Moodle platform.

The total of 816 students aged 9-11 years agreed to participate in the experiment. Of them, 210 people lived in another city or abroad.

Through random selection, students were divided into control (CG) and experimental (EG) groups — 408 students in each group.

A total of 42 teachers (14 supervising teachers and 28 English language teachers) aged 25 to 50 took part in the study. Teaching experience ranged from 5 to 35 years.

### Instruments

Author questionnaires were developed to conduct the survey (data collection was organized using Google Forms), the questionnaires used the Likert scale principle (1-5 points, where 1 is absolutely no, 2 — disagree / do not support, 3 — neutral, 4 — rather agree/support, 5 — absolutely agree/support). Cronbach's Alpha coefficient was used to check reliability, which was 0.792, indicating acceptable reliability. The Pearson correlation coefficient was used to validate the questionnaire. The obtained values were in the range of 0.7-0.8 and considered acceptable for research. Statistical data processing was carried out using the Microsoft Office.

### Data Collection

The following methods were used to conduct the experimental part of the study:

1. The method of systematic monitoring students' progress by teachers in order to identify the effectiveness of the Moodle platform in the educational process. Methods of testing, surveys and analysis of academic performance were also used.

**Table 1**

*Stages of the study to determine the impact of the Moodle platform on the development of primary schools under martial law*

| Research stage | Timing | Description of the research stage |
|---|---|---|
| Preparatory | May - August 2022 | 1) Selection of educational institutions for the experiment<br>2) Obtaining permission to conduct an experiment<br>3) Conducting a survey of teachers, students and parents to determine the advantages and disadvantages of traditional educational platforms<br>4) Adaptation of the curriculum to the Moodle platform |
| Initial | September-November 2022 | 1) Conducting input testing to determine student performance.<br>2) Division of students into control (CG) and experimental groups (EG).<br>3) Implementation of the Moodle platform in the educational process.<br>4) Conducting final testing to determine the success of students after the implementation of the Moodle platform.<br>5) Comparison of the obtained results for CG and EG. |
| Final | December 2022 | 1) Conducting a survey of teachers, students and parents regarding the convenience of the used platform.<br>2) Interpretation of the obtained results. |

Source: own author's development.

2. The study also involved a classification model of measurement scales: nominal, ordinal, interval, and relationship scale. For the study, the number of correctly completed assignments during training using Moodle as an educational platform was taken as a characteristic of students' performance.

3. The method of expert evaluations was used to analyse the results of the educational activities of students and teachers. A group of experts studied the state of cognitive independence of EG and CG students.

**Analysis of Data**

The Cronbach's alpha coefficient, which indicates the internal consistency of the test assignments, was used to check and interpret the data obtained at the end of the experimental part of the study. A comparison of the results of the CG and EG test assignments was also carried out.

**Ethical Criteria**

The chosen methods are verified for conducting research. The conducted experiment corresponds to the academic principles of professionalism, integrity, and respect for general human rights. All participants of the experiment (teachers, parents and students) consented to personal data processing. It was agreed to preserve the names of participants and names of educational institutions confidential. The instruments and techniques were tested for reliability and validity.

**Results**

The Moodle platform included the following educational elements: a lesson, a test, as well as means of communication and collaboration. The educational activity was designed as a set of pages on which educational content is presented in written or graphic form. Links, animations, multimedia and images were added to the learning content. The pages are united not only by content, but also by questions, with the help of which students personally checked their achievements. As this is electronic learning, students continue their studies after answering the questions correctly. If they make a mistake, they are returned to a page where the content must be repeated.

This way of learning obliges students to answer questions correctly, as one of the problems of e-learning is the loss of motivation, which can lead to skipping classes.

Testing in the Moodle system includes settings and a set of questions. The set of questions in the system is supplemented by selecting individual types of questions, forming a base from which questions are combined into a test. The test includes three types of questions: recall, completion, and calculation. When completing the test, students receive feedback on the correctness of the answer after each relevant question.

Students complete the assignment within the allocated time, and the task can be completed in three ways: submitting answers in writing when the grade is entered into the system; by recording the answers in the system itself and submitting the answers in the form of a created document that is transferred to the system.

The obtained average scores of students in each subject were analysed in order to select educational subjects (Table 2).

**Table 2**

*Percentage of scores obtained after control testing (%)*

| Percentage of correct answers | Elementary level | Medium level | Sufficient level | High |
|---|---|---|---|---|
| English language | 21 | 39 | 30 | 10 |
| Ukrainian language | 19 | 30 | 40 | 11 |
| Mathematics | 42 | 32 | 19 | 7 |
| Musical Art | 6 | 10 | 29 | 55 |
| I Explore the World | 0 | 6 | 24 | 70 |
| Literary Reading | 10 | 44 | 20 | 26 |

Source: author's own development

Taking into account the obtained data, it was decided to implement an experimental programme in subjects with different levels of student performance, namely: with a low level of Mathematics, with a medium level for English Language, and with the highest indicator for I Explore the World.

During remote classes, EG students used the Moodle platform, while CG continued their studies using the usual platforms.

A special feature of the Moodle platform when teaching Mathematics was that EG students were offered to pass a test after each lesson to better memorize the material they had learned. Training assignments were also offered for those willing to try. The difficulty of assignments gradually increased with each level, provided that the previous level was correctly completed. The main advantage was that the content of the assignments was close to the students' interests and also complemented the content of the textbooks without duplicating them.

When learning a foreign language, students expanded their vocabulary with the help of a built-in glossary. After each lesson, students could improve their listening, reading, speaking and writing skills by

completing test assignments. After passing the test, the student got access to an interesting short video with dialogues of the main characters of modern children's films.

In a similar way, the content of I Explore the World was updated.

At the end of the academic semester, CG and EG students were tested to determine the effectiveness of the chosen platform (Table 3).

**Table 3**

*Percentage ratio of obtained CG and EG scores at the end of the experiment*

| At the end of the experiment | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Subject | CG | EG | CG | EG | CG | EG | CG | EG |
| English Language | 10 | 8 | 39 | 39 | 38 | 38 | 13 | 15 |
| Mathematics | 39 | 39 | 31 | 30 | 18 | 19 | 12 | 12 |
| I Explore the World | 0 | 0 | 10 | 10 | 22 | 20 | 68 | 70 |

Source: author's own development

The input and output tests demonstrate the results of measuring the level of academic performance. If we compare the calculated empirical values of the Emp criterion with the critical value of 0.05=1.96 at the significance level of 0.05, we can see a tendency to $T_{emp} \geq Tcr$ for comparing the experimental and control groups at the significance level of 0.05. The sample mean and sample variance were used when finding $E_{mp}$ (Table 4).

**Table 4**

*Values of sample means, sample variances at the beginning and at the end of the experiment*

| The number of completed assignments | CG | | EG | |
|---|---|---|---|---|
| | Before | After | Before | After |
| Sample mean | 5.881 | 5.760 | 5.790 | 6.803 |
| Sample variance | 1.110 | 2.452 | 1.289 | 1.558 |

Source: author's own development

Therefore, it can be concluded that the reliability of the differences in the indicators of the compared samples is 95%.

A qualitative assessment of the level of involvement of EG students in using the Moodle platform was provided by independent works, which were evaluated by experts in terms of content, effectiveness, and independence. The level of students' independence in additional activities was determined by the completed assignments: fully completed, partially completed, and no reaction (Figure 1).

**Figure 1.**

*Obtained results of completing additional assignments in the CG (%)*

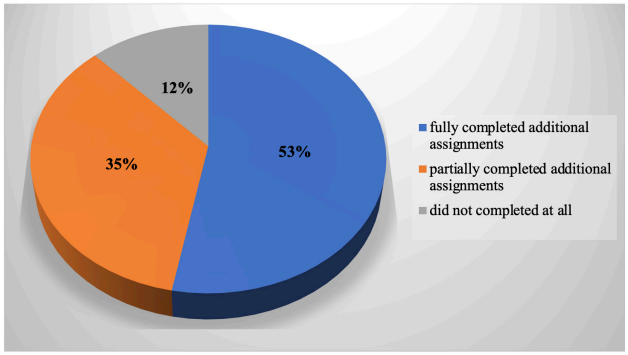



Figure 1 shows that 53% of students decided to complete all the proposed additional assignments. This is explained by the fact that the assignments were developed with elements of interactivity and in a game format (interesting crosswords, tests with a pleasant bonus in the form of a cartoon, etc.)

Teachers also paid attention to the ease of use of the studied Moodle platform. They drew attention to the fact that the platform plays presentation videos without delays, convenience in controlling the log, adding and controlling the material for classes.

## Discussion

The conducted research confirmed the importance of distance learning under martial law, especially for primary schoolers. The Learning Management System gives students the opportunity to continue their studies, communicate with classmates, and complete homework. The conducted study confirmed the initial hypothesis that if both teachers and parents perceive the use of the Moodle platform as a positive aspect of supplementing learning, one can expect an increased ability of students to develop more effective learning skills. Indeed, the obtained results were higher than the results in EG after the control testing of CG.

As not much time has passed since the beginning of the full-scale invasion of Ukraine, researchers have not yet had time to consider the issue of introducing the Moodle platform into distance education in primary education in context of military operations. Therefore, it was decided to compare the obtained results with researchers who considered the implementation of the Moodle platform among secondary school and higher school students during the COVID-19 pandemic. It was also planned to analyse the results obtained after the implementation of the Learning Management System. For example, Bulić and Blažević (2020) considered the field of biology. The test was designed for fifth, sixth, seventh and eighth grade students, and tested the knowledge of students learning on Moodle and students studying in a traditional classroom. The obtained results showed the absence of a statistically

significant difference in the arithmetic mean of the final test results between the two groups. On the contrary, the conducted research found positive results for the final testing. This is explained by the fact that CG students were interested in completing additional interactive assignments, which contributed to an increased level at the end of the experiment. In our opinion, students were able to distract themselves from the psychological stress caused by the state of war in the country through the use of interesting interactive assignments.

In their studies, Astriani et al. (2021), Marikar and Jayarathne (2016) determined the impact of using the Moodle platform to improve the learning of Mathematics for fifth grade students. It was found that implementing the curriculum using Moodle was effective for students' mathematical communication and self-efficacy. In general, students who used Moodle in e-learning performed better than those who received traditional education. In the current study, the results of students who studied using Moodle are also better than those who continued to study using conventional platforms.

Many studies are related to the use of Moodle in higher education. For example, Guillén-Gámez et al. (2022), Desnelita et al. (2021) found that Moodle has a significant impact on the development of virtual learning assignments. Garcia-Murillo et al. (2022) found a high degree of technological satisfaction with Moodle by students of higher educational institutions.

Holiver et al. (2020), Soub and Amarin (2021) determined the positive attitude of teachers and students towards the use of the platform in their practice. Our study confirms the results obtained by the researchers. The teachers, parents and students noted that the Moodle platform helps to use difficult content without delays in playback, facilitates monitoring of planned and completed assignments. In the event of an airstrike, teachers can stop the lesson and post the rest of materials on Moodle so that children and parents can work on them at any time after the danger is over.

We agree with the results of Bayramova and Aliyev (2019), Florjancic and Wiechetek (2022) that the development and implementation of modern technologies, namely Moodle, in education have brought many new opportunities and innovations supported by various educational platforms in the educational environment.

The findings of the study contribute to the understanding of the importance of modern LMS use and learning behaviours for the performance and satisfaction of primary schoolers during martial law.

**Conclusions**

Digitalization of the educational system is a current topic of the research. At the time of the invasion of the Russian Federation on the territory of Ukraine the educational system faced social and anthropological dynamics in combination with the rapid technological acceleration. The dynamics led to significant educational changes.

The conducted research emphasized and confirmed the importance of modernization and adaptation of educational platforms for better understanding of educational material by primary schoolers. Crosswords, interactive assignments, as well as watching animated videos in case of correct completion of test assignments were added to encourage students to complete additional assignments after the lesson. The main advantage of the Moodle platform was that the children were able to distract themselves from the war and concentrate on completing educational interactive assignments.

In the event of an airstrike, teachers can stop the lesson and post raw materials to Moodle so that children, with the help of their parents or independently, can continue learning at any time after the danger is over. This facilitates continuing systematic learning and unhindered learning of educational material.

A promising direction for further research may be the development of interactive assignments for primary schoolers to better memorize the material they have learned.

**References**

Abernathy, D. (2019). ADDIE in Action: A transformational course redesign process. *Journal for the Advancement of Educational Research*, *13*(1), 8.

Adedoyin, O. B., & Soykan, E. (2023). Covid-19 pandemic and online learning: The challenges and opportunities. *Interactive Learning Environments, 31(2), 863-875.* https://doi.org/10.1080/10494820.2020.1813180

Astriani, L., Setiyaningsih, D., Misriandi, M., & Yuliani, A. (2021). An implementation of e-learning based Moodle to develop mathematical communication skill and self-efficacy. *Kalamatika: Jurnal Pendidikan Matematika, 6*(2), 207-218. https://doi.org/10.22236/KALAMATIKA.vol6no2.2021pp207-218

Bayramova, U., & Aliyev, A. (2019). Innovations in Higher Education: MOODLE in Azerbaijan Universities. *1st International Informatics and Software Engineering Conference: Innovative Technologies for Digital Transformation, IISEC 2019 - Proceedings*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/UBMYK48245.2019.8965535

Bond, M. (2020). Facilitating student engagement through the flipped learning approach in K-12: A systematic review. *Computers & Education, 151*, 103819. https://doi.org/10.1016/j.compedu.2020.103819

Brammer, S., & Clark, T. (2020). COVID-19 and management education: Reflections on challenges, opportunities, and potential futures. *British Journal of Management, 31*, 453-456. https://doi.org/10.1111/1467-8551.12425

Bulić, M., & Blažević, I. (2020). The impact of online learning on student motivation in science and biology classes. *Journal of Elementary Education, 13*(1), 73-87. https://doi.org/10.18690/rei.13.1.73-87.2020

Burde, D., Kapit, A., Wahl, R. L., Guven, O., & Skarpeteig, M. I. (2017). Education in emergencies: A review of theory and research. *Review of Educational Research, 87*(3), 619-58.

Cuban, L. (2001). *Oversold and underused*. Harvard University Press.

Desnelita, Y., Susanti, W., Noratama Putri, R., Rizal, F., & Novaliendry, D. (2021). Application of a collaborative project-based learning model using Moodle in higher education. *RIGEO, 11*(9), 2141-2149.

Engzell, P., Frey, A., & Verhagen, M. (2021). Learning inequality during the Covid-19 pandemic. *Proceedings of the National Academy of Sciences*, *118*(17).

Florjancic, V., & Wiechetek, L. (2022). Using Moodle and MS Teams in higher education - a comparative study. *International Journal of Innovation and Learning*, *31*(2), 264–286. https://doi.org/10.1504/IJIL.2022.120649

Garcia-Murillo, G., Novoa-Hernandez, P., & Rodriguez, R. S. (2020). Technological satisfaction about Moodle in higher education - a meta-analysis. *Revista Iberoamericana de Tecnologias Del Aprendizaje*, *15*(4), 281–290. https://doi.org/10.1109/RITA.2020.3033201

Guillén-Gámez, F. D., Linde-Valenzuela, T., Ramos, M., & Mayorga-Fernandez, M. J. (2022). Identifying predictors of digital competence of educators and their impact on online guidance. *Research and Practice in Technology Enhanced Learning*, *17*(1). https://doi.org/10.1186/s41039-022-00197-9

Hockly, N., & Dudeney, G. (2018). Current and future digital trends in ELT. *RELC Journal, 49*(2), 164–178.

Holiver, N., Kurbatova, T., & Bondar, I. (2020). Blended learning for sustainable education: Moodle-based English for Specific Purposes teaching at Kryvyi Rih National University. *E3S Web of Conferences* (Vol. 166). EDP Sciences. https://doi.org/10.1051/e3sconf/202016610006

Jones, M., & McNulty, C. (2022). 'Everybody is sick and there won't be a cure': The Power of Privileging Student Voices Through COVID-19 Drawings. *International Electronic Journal of Elementary Education, 15*(2), 75-85. https://doi.org/10.26822/iejee.2023.280

Laurent, R., Dessus, P., & Vaufreydaz, D. (2020). Design spatial sociotechnique. *Distances et Médiations des Savoirs, 30*. https://doi.org/10.4000/dms.5228

Marikar, F., & Jayarathne, N. (2016). Effectiveness of MOODLE in education system in Sri Lankan University. *International Journal of Modern Education and Computer Science*, *8*(2), 54–58. https://doi.org/10.5815/ijmecs.2016.02.07

Moorhouse, B. L., & Beaumont, A. M. (2020a). Involving parents in their children's school-based English language writing using digital learning. *RELC Journal*, *51*(2), 259–267. https://doi.org/10.1177/0033688219859937

Moorhouse, B. L., & Beaumont, A. M. (2020b). Utilizing video conferencing software to teach young language learners in Hong Kong during the COVID-19 class suspensions. *TESOL Journal*, *11*(3). https://doi.org/10.1002/tesj.545

Owusu-Fordjour, C., Koomson, C. K., & Hanson, D. (2020). The impact of Covid-19 on learning - the perspective of the Ghanaian student. *European Journal of Education Studies*, *0*(0). https://doi.org/10.46827/ejes.v0i0.3000

Pozo-Rico, T., Gilar-Corbí, R., Izquierdo, A., & Castejón, J. L. (2020). Teacher training can make a difference: tools to overcome the impact of COVID-19 on primary schools. An experimental study. *International Journal of Environmental Research and Public Health, 17*(22), 1-23. https://doi.org/10.3390/ijerph17228633

Pulham, E., & Graham, C. R. (2018). Comparing K-12 online and blended teaching competencies: A literature review. *Distance Education, 39*(3), 411–432. https://doi.org/10.1080/01587919.2018.1476840

Putri, R. S., Purwanto, A., Pramono, R., Asbari, M., Wijayanti, L. M., & Hyun, C. C. (2020). Impact of the COVID-19 pandemic on online home learning: An explorative study of primary schools in Indonesia. *International Journal of Advanced Science and Technology, 29*, 4809–4818. https://doi.org/10.1371/journal.pone.0259546

Roman M., & Plopeanu A.P. (2021). The effectiveness of the emergency e-learning during COVID-19 pandemic. The case of higher education in economics in Romania. *International Review of Economics Education*, *37*, Article 100218. https://doi.org/10.1016/j.iree.2021.100218

Soub, T. F. A., & Amarin, N. Z. (2021). The reality of using Moodle in a distance education program. *Cypriot Journal of Educational Sciences*, *16*(5), 2173–2192. https://doi.org/10.18844/cjes.v16i5.6237

Tang, R., & Braver, T. S. (2020). Towards an individual differences perspective in mindfulness training research: Theoretical and empirical considerations. *Frontiers in Psychology, 11,* Article 818. https://doi.org/10.3389/fpsyg.2020.00818

Tay, L. Y., Melwani, M., Ong, J. L., & Ng, K. R. (2017). A case study of designing technology-enhanced learning in an elementary school in Singapore. *Learning: Research and Practice, 3*(2), 98–113. https://doi.org/10.1080/23735082.2017.1350737

Voogt, J., Knezek, G., Christensen, R., & Lai, K. W. (Eds.). (2018). *Second handbook of information technology in primary and secondary education*. New York: Springer.

Wang, C. X. (2021). CAFE: An instructional design model to assist K-12 teachers to teach remotely during and beyond the COVID-19 pandemic. *TechTrends, 65*, 8–16. https://doi.org/10.1007/s11528-020-00555-8

Yazgan, Ç. Ü. (2022). Attitudes and interaction practices towards distance education during the pandemic. *Education and Information Technologies*, *27*(4), 5349–5364. https://doi.org/10.1007/s10639-021-10843-2

This page is intentionally left blank.

www.iejee.com

iejee

International Electronic Journal of Elementary Education
June 2023, Volume 15, Issue 5, 409-424

# Comparing a Modified Structured Mix with a Modified Random Rotation Procedure to Teach Auditory-Visual Conditional Discriminations to Children with Autism

Brittany Marie DiSanti[a,*], Svein Eikeseth[b], Sigmund Eldevik[c]

Received : 15 January 2023
Revised : 2 June 2023
Accepted : 18 June 2023
DOI : 10.26822/iejee.2023.309

[a,*] **Corresponding Author:** Brittany Marie DiSanti, Department of Behavioral Science, Oslo Metropolitan University, Oslo, Norway.
E-mail: brittmarie13@gmail.com
ORCID: https://orcid.org/0000-0002-8333-3294

[b] Svein Eikeseth, Department of Behavioral Science, Oslo Metropolitan University, Oslo, Norway.
E-mail: svein.eikeseth@oslomet.no
ORCID: https://orcid.org/0000-0001-6404-9071

[c] Sigmund Eldevik, Department of Behavioral Science, Oslo Metropolitan University, Oslo, Norway.
E-mail: sigmund.eldevik@oslomet.no
ORCID: https://orcid.org/0000-0001-7029-1665



## Abstract

We evaluated two procedures to teach auditory-visual conditional discriminations (receptive labeling) to children with autism. The procedures evaluated a modified Structured Mix (SM) procedure and a modified Counterbalanced Random Rotation (RR) procedure. The modified SM procedure was based on the logic of simplifying the task by breaking it down into smaller, successive steps and by requiring mastery of each step before introducing the next. Compared to previous studies, the modified SM procedure contained fewer steps, less prompting, and a less stringent mastery criterion. The modified RR procedure targeted all three stimuli simultaneously by presenting them across consecutive trials, both during training and error correction. Sample stimuli were presented in a counterbalanced random order and the comparison stimuli were presented in counterbalanced random positions. Participants were nine children with autism. An adapted alternating treatments design was used. Results showed that the modified SM procedure was more efficient for four of the nine participants, the modified RR procedure was more efficient for one of the nine participants, both procedures were equally efficient for two participants, and neither procedure was effective for two of the nine participants. The modified SM procedure appeared more efficient than the SM procedure employed in previous studies. Despite results, further research is warranted to examine within subject comparisons between original discrimination training procedures and modified procedures.

**Keywords:**

Auditory-Visual Conditional Discrimination, Autism, Conditional Discrimination, Receptive Language, Stimulus Control

## Introduction

Children with neurodevelopmental disorders such as autism may demonstrate a limited repertoire of receptive language such as auditory-visual conditional discriminations. An auditory-visual conditional discrimination consists of hearing an auditory sample stimulus (e.g., teacher says "touch doll") and selecting the comparison

stimulus corresponding to the correct sample stimulus (e.g., learner touches the doll). This has been called an "if ... then" relation because there is a conditional relation between the sample and comparison stimuli (Sidman & Tailby, 1982). Although an auditory-visual conditional discrimination is a rather basic listener skill, it cannot be assumed to develop naturally in children with developmental delays (Iversen et al., 1986). An auditory-visual conditional discrimination requires that the visual comparison stimuli are observed and that they are discriminated from each other. Secondly, the auditory-verbal sample stimuli, which are presented successively (e.g., the child hears "touch doll" on one trial and "touch car" on the subsequent trial), are observed and discriminated from each other. Finally, and perhaps most challenging, the discrimination of the visual comparison stimulus and its reversal must be brought under instructional control of the verbal sample stimulus.

A discrimination training procedure described in several teaching manuals (Lovaas, 1977, 1981, 2003) has been used to teach auditory-visual conditional discriminations for more than four decades. This procedure was based on the logic of simplifying the task by breaking it down into smaller, successive steps and by requiring mastery of each step before introducing the next. This method has been referred to as the simple-conditional method (Grow et al., 2011) or a structured mix (SM) procedure (DiSanti et al., 2019). It has been questioned whether this procedure is optimal because the reinforcement contingency during the initial steps of training, where each sample stimuli is presented in blocks of trials, does not require discrimination of (or responding to) the verbal sample stimulus (Green, 2001). Rather, aspects of the comparison stimuli may be established as a controlling antecedent stimulus, which may interfere with subsequent control by the sample stimulus. This is because in a particular block of trials selecting the same comparison stimulus across consecutive trials is reinforced. Under this reinforcement contingency, the only stimulus controlling the child's response may be the object itself (i.e., comparison stimulus) and not the verbal sample stimulus. Indeed, the child does not need to discriminate the verbal sample stimulus since correct responding is achieved by selecting the same stimulus that produced reinforcement on previous trials (Sidman & Stoddard, 1966). Whenever selecting this stimulus is no longer reinforced (which would be the case when the reinforcement contingency has been reversed) responding to the other comparison is reinforced. Whenever an incorrect response occurs, the child may simply start responding to the previously trained comparison stimulus until another incorrect response occurs. As a result, an alternative source of stimulus control may be established that competes with the desired type of stimulus control required to establish a conditional discrimination (Green, 2001). In this way, the child may learn "not" to listen or attend to the teacher's instruction and subsequently this learning history may make it more difficult to get the child to discriminate the sample stimulus, which is necessary to establish a conditional discrimination.

To avoid these potential problems Green (2001) proposed that within a session or a block of trials: (a) a different sample stimulus should be presented on each trial, (b) there should be a minimum of three comparison stimuli on each trial, (c) each sample stimulus should be presented equally often, and (d) the position of the comparison stimuli should vary unsystematically across each trial.

Subsequent studies have suggested that the procedure proposed by Green (2001) is more efficient for most participants as compared to the SM procedure (or simple-conditional) (DiSanti et al., 2019; Grow et al., 2011; Grow et al., 2014; Grow & Van Der Hijde, 2017; Gutierrez et al., 2009; Holmes et al., 2015; Lin & Zhu, 2019; Vedora & Grandelski, 2015). For some learners with a limited verbal repertoire, however, the SM procedure has been found to be more efficient (DiSanti et al., 2019; Lin & Zhu, 2019). Likely, in learners with a limited verbal repertoire, some forms of deficits exist in the way verbal and nonverbal stimuli control listener and speaker behavior (Michael et al., 2011; Sundberg, 2016). Perhaps the SM procedure helps remediate some of these deficits by presenting these elements in a sequential order, introducing the next element only after the previous elements have been mastered, and establishing stimulus control. However, for those children with a more advanced verbal repertoire the SM procedure may interfere with the acquisition of conditional control. This may be due to reinforcing irrelevant sources of stimulus control which compete with the desired type of stimulus control required to eventually establish conditional discriminations.

Following this logic, the SM procedure employed in previous studies may have been less optimal because it contained unnecessary or redundant training steps, had a rigorous mastery criterion, and/or included a rigorous prompting procedure. This may have resulted in a type of overtraining and/or an increased possibility of establishing faulty stimulus control. For the purpose of the present study, we designed a modified version of the SM procedure (i.e., the modified SM procedure). First, training participants to select particular stimuli through mass trials with the inclusion of neutral distractors was eliminated and instead, stimuli that served as the discriminative stimulus (S+) on later steps were included as S-delta (S-). In previous studies, steps of mass trial teaching have involved training with neutral distractors (i.e., S- which later did not serve as S+; Holmes et al., 2015), or training stimuli in isolation (Grow et al., 2011; Grow et al., 2014; Gutierrez et al., 2009; Vedora et al., 2015). In past studies, as part of an errorless teaching procedure, prompts were gradually faded across a large number of trials. In the current study, errorless learning procedures were also included, but prompting was discontinued after one

correct response using a zero-second prompt delay. Modifications in the prompting procedure were made to reduce the potential for prompting an unnecessary number of trials and to minimize the potential for prompt dependency. Finally, the mastery criterion for each step was reduced to avoid overtraining. In previous studies, the mastery criterion for each step involved correct responding across a specified number of trials or across consecutive sessions. In the current study, the mastery criterion was less stringent; stimuli trained through mass trials required four-out-of-four consecutive correct responses. Steps that required discrimination between two or three stimuli required nine-out-of-nine consecutive correct responses.

In addition to the changes made to the modified SM procedure, we also made modifications to the RR procedure (i.e., modified RR procedure). The RR procedure employed by DiSanti et al. (2019) built on a "conditional only" method validated by Grow et al. (2011, 2014). This procedure was designed to teach responding to both the sample and comparison stimuli from the onset of training by presenting a different sample stimulus on each trial (random, but counterbalanced presentation), presenting each sample stimulus equally often and varying the position of the comparison stimuli unsystematically, but evenly across trials. However, typically following an incorrect response, error correction was implemented systematically; thus, not following the random, counterbalanced presentation of stimuli that was essential to the procedure. The same sample stimulus was presented on consecutive prompted trials and the comparison stimuli were placed in fixed positions across these trials, inadvertently diverting the procedure into a SM procedure, which the conditional-only method had been designed to avoid. Hence, the success of the RR procedure may be partly due to the fact that training during trials where error correction was implemented diverted from a RR procedure to more of a SM procedure. In the present study, the modified RR procedure presented trials during error correction, which were consistent with the RR procedure. More specifically, changes made to the RR in the current study were as follows: First, to make the mastery criterion for the modified RR condition comparable to the mastery criterion for the modified SM condition, the mastery criterion for the modified RR procedure was less stringent compared to previous studies. In the current study, mastery was defined as nine-out-of-nine consecutive correct responses, which was identical to the mastery criterion for the final step in the modified SM procedure. Second, in previous studies error correction typically entailed repeating the same trial while a prompt was being faded. That is, the exact position of the comparison stimuli remained the same across consecutive trials, while the particular sample stimulus (which was responded to incorrectly) was re-presented. In the current study, the error correction procedure for the modified RR procedure followed the logic of semi-random presentation of sample and comparison stimuli. That is, a different sample stimulus was presented on each trial, each sample stimulus was presented equally often, and the position of the comparison stimuli varied unsystematically across trials.

For some of the participants with limited and advanced verbal repertoires we trained more than one stimulus set, across both training procedures. This was done to examine whether the pattern of responding on the initial stimulus set could be replicated on additional stimulus sets within participants.

The present study was designed to examine the effectiveness of the modified SM procedure by comparing the number of trials to mastery for three auditory-visual conditional discriminations (i.e., a stimulus set) using the modified SM procedure to the number of trials to mastery using: (a) the modified RR procedure, (b) the original SM procedure (DiSanti et al., 2019), and (c) the original RR procedure (DiSanti et al., 2019). Data from the original SM and RR procedure were abstracted from the DiSanti et al. (2019) study.

## Method

### *Participants, Setting and Training Personnel*

Participants were nine, (seven males) three-to-thirteen-year-old children, diagnosed with autism spectrum disorder (ASD). The diagnosis was set by a licensed clinical psychologist. Inclusion criteria were as follows: a repertoire of at least 10 receptive labels, 10 motor imitations, and 10 visual-visual identical matching-to-sample discriminations. The purpose of these inclusion criteria was to ensure the participants were able to scan the visual array, attend to the instructor, and to acquire conditional discriminations via discrete trial teaching (Green, 2001). The number of auditory-visual conditional discriminations was reported by the participant's behavioral therapist before the start of the study.

All participants had previous exposure to both training procedures although the procedure may have appeared less strict in the applied setting compared to the current study (i.e., a less rigorous mastery criterion, flexible error correction procedures, implementation of mastered targets during acquisition training, combination of both training procedures). Table 1 exhibits participant characteristics, and scores from behavioral and developmental assessments.

Training sessions for all participants were conducted by the experimenter (all participants), lead behavior specialist (Participant 5, 6, and 7), or a registered behavior technician (Participant 4, 8, and 9). Sessions were conducted in the participants' school or clinic setting.

*Training Sets and Materials*

Each training set included three auditory sample stimuli in the form of spoken words. Three corresponding visual stimuli were used in the form of picture templates. A template with framed boxes for stimulus cards to be placed on the right, middle, and left was placed on an A3 laminated paper sheet (29.7 cm x 42.0 cm). A standard-sized clipboard was used for three of the participants due to lack of appropriate responding to stimuli when laminated templates were placed on the tabletop. Individual picture cards were made for the stimuli taught in each training condition. Picture cards were rotated on trials based on the sample stimulus, comparison stimuli available, and position required on the data collection sheet. Data collection sheets containing 30 trials were made for all steps and were randomized for position and stimuli.

For participants 1, 2, 3, 7, and 8 one set of stimuli for each condition was administered. For remaining participants, between two and three stimulus sets were taught to assess the extent to which results could be replicated across participant stimulus sets. Participants were not exposed to the training stimuli outside of the study. Furthermore, instructors who conducted the one-to-one training sessions did not include stimuli or related stimuli within the participant's other therapeutic treatment programs.

*Dependent Measures and Data Collection*

The dependent variable was the number of trials to mastery for three auditory-visual conditional discriminations (i.e., a stimulus set) in each training condition. One procedure was considered more efficient than the other procedure if the difference in number of trials to mastery exceeded ten percent (Ledford et al., 2019). If the difference in number of trials to mastery were ten percent or less, they were considered equally efficient. A correct response was defined as the participant pointing to the correct visual comparison stimulus within 5 seconds of the presentation of the auditory sample stimulus. A prompted response was defined as the participant selecting the correct comparison stimulus within 5 seconds of the experimenter providing a prompt. An incorrect response was scored if the participant selected the incorrect comparison stimulus or did not respond within 5 seconds after the presentation of the auditory sample stimulus. Data was also collected on the number of prompts required, number of errors that occurred, and maintenance at 4 and 6 weeks.

*Interobserver Agreement (IOA)*

To measure IOA, a second independent observer recorded participant responses for each training condition. Trial-by-trial agreement was calculated, and an agreement was scored if both the primary and secondary observer recorded (a) a correct response, (b) a prompted response, (c) an incorrect response, and (d) the position of the visual comparison stimulus. Interobserver agreement was calculated by taking the number of trials in agreement divided by the total number of trials in the session, multiplied by 100. Interobserver agreement was collected across all participants for a mean 39.8 % of the trials (range, 33% to 40%). Mean scores for interobserver agreement were 99.9% (range, 98.6% to 100%).

*Preference Assessment*

To identify putative reinforcers to be used during training, stimuli were chosen based on a teacher report of 20 preferred items. The twenty items were used in a Multiple-Stimulus-Without- Replacement-Preference assessment (MSWO) to identify the top five preferred items for each participant (DeLeon & Iwata, 1996). Before each training session, a brief MSWO using the preferred five items was conducted to identify preference. For two of the participants, the establishment of a token economy system had previously been in place for discrete trial training sessions. Participants who used a token economy system were exposed to a brief MSWO before training sessions in order to identify back-up reinforcers.

*Pretest*

Six target stimuli were identified through pretests for each stimulus set to ensure targets were unknown

**Table 1.**

*Participant Characteristics*

| Participant | Age | Receptive Labels | CARS | Developmental Age (Months) | Vineland ABC | Vineland Communication |
|---|---|---|---|---|---|---|
| 1 | 9:7 | >200 | 36.5* | 44* | 73 | 76 |
| 2 | 8:11 | >200 | 29.5* | 56* | 88 | 86 |
| 3 | 11:8 | >200 | 30* | 41* | 77 | 73 |
| 4 | 3:5 | >100 | 29 | 80 | 69 | 76 |
| 5 | 13 | <50 | 46.5 | 19 | 20 | 20 |
| 6 | 11:7 | <50 | 33 | 34 | 45 | 43 |
| 7 | 13:3 | <50 | 29.5 | 40-42 | 62 | 57 |
| 8 | 4:5 | <50 | 33.5 | 25 | 69 | 70 |
| 9 | 4:5 | <50 | 45 | 21 | 57 | 48 |

Note. Receptive Labels = receptive labels in the participant's repertoire before training; CARS = Childhood Autism Rating Scale Standard Version or High-Functioning Version (denoted with an *; Schopler et al., 2010); Developmental Age = developmental age as derived from Bayley Scales of Infant Development (Bayley, 2005) or Stanford Binet-Intelligence Scale (denoted with an *) (Roid, 2003); ABC = Adaptive Behavior Composite Score and Vineland Communication from Vineland Adaptive Behavior Scales-3 (Sparrow et al., 2016).

to the participants. All target stimuli were nouns, except one stimulus set for Participant 4, which was verbs. Three stimuli were presented on the template in front of the child. Pretest data sheets were created to ensure each stimulus was presented semi-randomly as the sample stimulus three times each, and each stimulus was presented in the right, middle, and left position three times each. The experimenter asked the participant to select one of the comparison stimuli. Stimuli were included in one of the training conditions if the participant responded less than or equal to 33% correct during the nine-trial probe. During the pretest, reinforcement was given approximately every 10 seconds for proper sitting and attending to the experimenter. Reinforcement for proper sitting and attending was not given immediately following a trial. No consequences were provided for incorrect or correct responding. When stimuli had been identified, three stimuli were randomly assigned to each discrimination training condition. Randomization of stimuli across conditions was assessed based on the initial sounds; that is, stimuli with similar, first sounds were not included within the same stimulus set (Wolery et al., 2014). Also, those stimuli that were part of a category (i.e., planets, continents, states, etc.) formed a stimulus set (e.g., Random Rotation: Mercury, Neptune, Jupiter). For one stimulus set (participant 1 modified structured mix condition), two stimuli began with the letter A, but the overlapping sounds were not the same (e.g., Asia and Africa).

*Teaching Procedure*

Training sessions were conducted in the morning and afternoon, five days a week. Participants received two sessions of each condition, daily. Each condition was counterbalanced during morning and afternoon training sessions. Morning training sessions took place between school arrival and lunch, and afternoon training sessions took place between lunch and school dismissal. For three participants, only one session of each condition was conducted, daily. The presentation of training conditions for these three participants were counterbalanced across days (i.e., Day 1: Random Rotation, Structured Mix; Day 2: Structured Mix, Random Rotation). Sessions consisted of 30 trials. Sessions were discontinued if the participant engaged in challenging behaviors (e.g., self-injury or aggression towards the trainer), or did not respond to instructions or prompts.

Both teaching conditions utilized a discrete trial teaching format (Eikeseth et al., 2014). The trainer presented the antecedent stimulus similarly to the comparison-first procedure (Grow et al., 2011; Grow et al., 2014; Kodak et al., 2015). Each trial began by placing the paper template containing the comparison stimuli in front of the participant on the tabletop before presenting the auditory sample stimulus. Next, the trainer presented an auditory sample stimulus (e.g., "cherry tree"). If the participant responded correctly by pointing to the correct comparison stimulus this was reinforced with verbal praise and a tangible or edible item. If the participant responded incorrectly or did not respond, they were told "no" or "try again," and the template was pulled away from the center of the table. On the next trial, the trainer presented the template and initiated a zero-second prompt delay by pointing to the correct comparison stimulus after the presentation of the auditory sample stimulus. If the child had not acquired the stimulus set within 500 training trials, training for that condition was discontinued.

*Modified Structured Mix (SM) Condition*

The modified SM condition included five steps. For all steps that included the presentation of a new sample stimulus the trainer initiated a zero-second prompt delay on the first trial only (i.e., Step 1, 2, and 4). On steps where two comparison stimuli were available (i.e., Step 1, 2, and 3) one position (right, middle, or left) was blank across the session. If the participant responded correctly with a zero-second prompt on the initial trial, the experimenter moved to the next trial on the pre-made data collection sheet and provided no prompt. If the participant responded correctly on the next trial, this was scored as correct and counted towards the mastery criterion. Next, the trainer continued to rotate the position of the comparison stimuli for the remaining trials; position of comparison stimuli and sample stimulus were dependent on the pre-made data collection sheet. If an incorrect response occurred during the session, error correction procedures were almost identical to the errorless learning procedures (implemented during Steps 1, 2, and 4). That is, following an incorrect response, the position of the comparison-stimuli was the same and the sample stimulus that was incorrect was re-presented with a zero-second prompt delay. Following a prompted response, the position of the comparison-stimuli remained the same and the sample stimulus was presented without a prompt. If the participant responded correctly, this trial was scored as correct, and the trainer rotated the position of the comparison stimuli for the next trial dependent on the data collection sheet. It should be noted that although the participant responded correctly without a prompt, this trial was not counted towards the mastery criterion. Prompted responses were reinforced with verbal praise. See Appendix for a detailed description of each step in the modified SM procedure.

**Appendix**

*Detailed Description of Each Step in the Modified SM Condition*

Step 1: Sample Stimulus 1. In Step 1, the trainer labeled stimulus 1 (e.g., "point to 'Africa'") while stimulus 1 and 2

served as the comparison stimuli. The position of both comparison stimuli rotated between the left, middle, and right position semi-randomly across trials. The mastery criterion was four consecutive trials correct within a 30-trial session.

Step 2: Sample Stimulus 2. In Step 2, the trainer labeled stimulus 2 while stimulus 1 and 2 served as the comparison stimuli. The position of both comparison stimuli rotated between the left, middle, and right position semi-randomly across trials. The mastery criterion was four consecutive trials correct within a 30-trial session.

Step 3: Sample Stimulus 1 and Sample Stimulus 2 Structured Mix. In Step 3, stimulus 1 and stimulus 2 were alternated semi-randomly as the sample stimulus across trials. Stimulus 1 and stimulus 2 served as the comparison stimuli and were rotated between the left, middle, and right positions semi-randomly across trials. The mastery criterion was nine out consecutive trials correct within a 30-trial session.

Step 4: Sample Stimulus 3. In Step 4, stimulus 3 was presented as the sample stimulus on all trials, while stimulus 1, 2, and 3 served as the comparison stimuli. The position of the comparison stimuli rotated between the left, middle, and right position semi-randomly across trials. The mastery criterion was four consecutive trials correct within a 30-trial session.

Step 5: Counterbalanced Random Rotation (Random Rotation). Within each block of nine trials, all three stimuli served as the sample stimulus three times each and the comparison stimuli appeared in each position three times each. The mastery criterion was nine consecutive trials correct within a 30-trial session.

*Modified Counterbalanced Random Rotation (RR) Condition*

The modified RR condition was identical to Step 5 of the modified SM condition except for some differences in the prompting procedure. At the onset of training, a zero-second prompt delay was provided for the first nine trials, during which each sample stimulus was presented three times: once in the left, middle, and right position. If the participant responded correctly across the nine prompted trials, prompting was discontinued. Subsequently, whenever an incorrect response occurred during sessions, a zero-second prompt delay was provided on the next three consecutive trials following an incorrect response. The comparison stimuli and sample stimuli were kept semi-random after an incorrect response, for example, if an incorrect response occurred on stimulus 1 in the middle position (Trial 11), but Trial 12 on the data collection sheet listed stimulus 2 as the sample stimulus in the left position, a zero-second prompt delay was initiated for stimulus 2. The next two prompted responses followed a semi-random presentation based on the position of the comparison stimuli and the sample stimulus presented on the pre-made data collection sheet. Following prompted responses, trials were presented without a prompt. The reason for the difference in the prompt procedure was that within a block of trials for the modified RR procedure a different sample stimulus should be presented on each trial, such that each sample stimulus is presented equally as often. Additionally, the position of the comparison stimuli should vary unsystematically across each trial. Nine consecutive prompted trials would ensure that each stimulus served as the sample stimulus at least once in each position. Prompted responses were reinforced with verbal praise.

*Error Analysis and Additional Error Correction Procedures*

If after a 30-trial session, a participant had not acquired the discrimination targeted during that session (i.e., fulfilled the mastery criterion of 4/4 or 9/9), error analysis was conducted to identify the extent to which errors occurred due to, for example, win-stay, lose-shift, win-shift, or position bias. A win-stay strategy may occur between different training steps (molar win-stay) of a discrimination training procedure or between consecutive trials (molecular win-stay) (Grow et al., 2011; Lovaas, 2003). That is, rather than attending to the change in auditory-sample stimulus the learner may respond to the stimulus which received reinforcement on the previous trial (molecular win-stay) or the previous step (molar win-stay). A win-shift strategy may occur when intermixing two sample stimuli, where the learner is reinforced for correct responding on trial 1, but the learner may shift to the other comparison stimulus on trial 2 to receive reinforcement. A lose-shift strategy may occur when the learner responds incorrectly to one of the comparison stimuli; thus, shifting to the other comparison stimulus due to a loss of reinforcement on the previous trial. A position bias may occur depending on how the comparison stimuli are arranged; that is, correct responding to the sample stimulus in a particular position may reinforce a higher percentage of responses to that specific position. Error analysis was taken across both training conditions, but additional procedures to correct for errors were only implemented for the modified SM condition. The reasoning for implementing error analysis and additional error correction procedures was from past research suggesting that the SM training condition could lead to faulty stimulus control (Green, 2001; Grow et al., 2011). Additional error correction procedures, following error analysis, are described in Table 2.

**Maintenance**

Maintenance tests were conducted four and six weeks after mastery for each condition. Maintenance tests followed the same format as the pretest.

**Procedural Integrity**

Procedural integrity was conducted by the lead classroom teacher or the experimenter. Data were collected on sections labeled: (a) preparing the session, (b) presentation of the sample and comparison stimuli, (c) prompting, (d) reinforcement, and (e) session structure. Preparing the session was scored as correct if necessary teaching materials and reinforcers were available and a positive relationship (i.e., child was smiling and interacting with the teacher) with the child had been established. During the presentation of the sample and comparison stimuli, the observer recorded whether clear instructions were provided to the child (i.e., presentation of only one sample stimulus indicated on the data collection sheet); if the instruction was appropriate (i.e., instruction followed the data collection sheet); and if the child was attentive during the presentation of the instruction (i.e., hands still, sitting upright in chair, looking at the instructor, looking at the visual array). Correct prompting included collecting data on whether the prompt led to a correct response, whether the correct prompt delay was provided and whether the correct number of responses were prompted. Reinforcement was scored as correct if praise and tangible reinforcers were used when correct responses occurred, whether reinforcement occurred within 2 seconds of a correct response, whether every correct response was reinforced during acquisition phases, whether praise was used for prompted responses, and whether incorrect responses were followed by "no" or "try again." The session structure was assessed correct if the trainer followed the correct discrete trial format (i.e., instruction, response, consequence); the inter-trial interval was no longer than 5 seconds; criterion levels were achieved before moving to different steps; and sessions ended on a correct response or with a task the child could perform correctly (i.e., did not have to be a receptive language task). Additional items listed within session structure assessed if the appropriate amount of time (a minimum of 60 minutes) was left between the two training procedures, and that 30 trials were conducted for both conditions in the morning and afternoon unless unethical to do so. The individual conducting the training was scored on whether they completed the tasks required for each of the five sections listed above. Procedural integrity was collected across all participants for a mean of 39.8% (range, 33% to 40%) of the sessions. Mean scores for procedural integrity were 98.2 % (range, 86% to 100%).

**Design**

An adapted alternating treatments design (Sindelar et al., 1985) was used to compare the SM condition to the RR condition. To counterbalance sequence effects, the order of training conditions was alternated semi-randomly so that each condition occurred an equal number of times in the morning and in the afternoon.

**Results**

Figure 1 shows individual data across blocks of trials for all participants (e.g., blocks of 4 trials for steps 1, 2, and 4 for modified SM; or 9 trials for modified RR and 9 trials for modified SM steps 3 & 5). Table 3 shows the number of trials to mastery, number of prompts, number of errors, and percentage correct at maintenance tests conducted four and six weeks after training.

Participant 1 was taught one stimulus set in each condition and acquired the stimulus set in each condition in less than 45 trials, although with 7 trials fewer in the modified RR condition. Six prompts and 3 errors occurred in the modified SM condition

**Table 2.**

*Error Correction Procedures for Participants 5, 8, and 9*

| Participant | Error pattern | Stimulus set and training step | Error correction procedure | Results |
|---|---|---|---|---|
| 5 | Bias towards comparison stimulus reinforced on previous trials. | Stimulus Set 2 and 3, training Step 3 (i.e., Discrimination between Stimulus 1 and Stimulus 2). | 1. Retraining from Step 1.<br><br>2. Prompting after two incorrect responses was started. | Neither retraining from Step 1 nor prompting after two incorrect responses resulted in acquisition of Step 3. Training on Step 3 was terminated because training had reached 500 trials. |
| 8 | Bias towards comparison stimuli not yet trained. | Stimulus Set 1, training Step 1. | Replacing the comparison distractor with a sample stimulus that would serve as stimulus 3. | Resulted in the acquisition of Step 1. Step 2 was subsequently mastered without procedural modifications. Training on Step 3 was terminated because training had reached 500 trials. |
| 9 | Bias towards comparison stimulus reinforced on previous trials.<br><br>Bias towards comparison stimuli not yet trained. | Stimulus Set 1, training Step 3 (i.e., Discrimination between Stimulus 1 and Stimulus 2).<br><br>Stimulus Set 2, training Step 2. | 1. Retraining from Step 1.<br><br>2. Replacing the comparison stimulus (distractor) with a neutral distractor that was not intended to be trained. | Retraining from Step 1 did not result in acquisition of Step 3. Training on Set 1 was terminated because training had reached 500 trials.<br><br>Resulted in the acquisition of Step 2. Steps 3 and 4 were subsequently mastered without procedural modifications. Training on Step 5 was terminated because training had reached 500 trials. |

compared to 16 prompts and 2 errors in the modified RR condition. Maintenance was high for both conditions.

Participant 2 was taught one stimulus set in each condition and acquired both in fewer than 40 trials. There was a two-trial difference in the number of trials to mastery across conditions. Six prompts and 1 error occurred in the modified SM condition, compared to 15 prompts and 2 errors in the modified RR condition. Maintenance for both conditions was low.

Participant 3 was taught one stimulus set in each condition and acquired the stimulus set in the modified SM condition in considerably fewer trials (74), than in the modified RR condition (128). The modified SM condition had fewer prompts and fewer errors, but maintenance was higher in the modified RR condition.

Participant 4 was taught two stimulus sets for each condition. Both training procedures were equally effective. Across both stimulus sets, fewer prompts were required for the modified SM condition, but in one stimulus set more errors occurred for the modified SM condition. Maintenance was 100% at four- and six-week follow-up, for both conditions and across both stimulus sets.

Participant 5 was trained on three stimulus sets and mastered stimulus set 1 with considerably fewer trials in the modified SM condition (93), with fewer prompts and errors, compared to the modified RR condition (252). Participant 5 failed to acquire the next two stimulus sets in either condition. For this participant, error analysis was conducted for stimulus set 2 and 3, and the error analysis and additional error correction procedures that were implemented are shown in Table 2. However, these procedures did not lead to acquisition of the discriminations in the modified SM condition. Maintenance was higher in the modified RR condition for stimulus set 1.

Participant 6 was taught three stimulus sets and acquired the first stimulus set with fewer trials in the modified SM condition, but acquired the two next stimulus sets with fewer trials in the modified RR condition. Participant 6 acquired stimulus set 1 in the modified SM condition in 33 trials, which was the minimum number of trials required to master a stimulus set without making any errors. Moreover, in the modified RR condition, stimulus set 3 was acquired in 18 trials, which was the minimum number of trials required to master a stimulus set without making any errors in this condition. In total, the number of prompts and number of errors were higher in the modified RR condition compared to the modified SM condition. For stimulus set 1, maintenance was higher in the modified SM condition at four- and six-week follow-up (67%, 78%), compared to the modified RR condition (44%, 0%). For stimulus set 2, maintenance was higher in the modified RR condition at four- and six-week follow-up (67%, 89%), compared to the modified SM condition (33%, 22%). For stimulus set 3, maintenance was higher in the modified RR condition (55% at four- and six-week follow-up), compared to the modified SM condition (44% at four- and six-week follow-up).

Participant 7 was taught one stimulus set and acquired the stimulus set in both conditions, but with fewer trials in the modified SM condition. The number of prompts and errors were fewer in the modified SM condition. Maintenance was not collected for this participant.

Participant 8 and 9 both failed to acquire any discriminations in both training conditions for one stimulus set (Participant 8) and two stimulus sets (Participant 9). Error analysis and additional error correction procedures were implemented but did not result in acquisition (see Table 2). The number of prompts were fewer in the modified SM condition, but also more errors occurred in this condition.

Considering participants as a group, the modified SM procedure was more efficient for four of the nine participants, the modified RR procedure was more efficient for one of the nine participants, both procedures were equally efficient for two participants, and finally, neither procedure was effective for two of the nine participants. The mean number of trials needed to acquire each stimulus set across participants in the modified SM condition was 65 trials compared to 105 trials per stimulus set in the modified RR condition. When tallying the total number of teaching trials required, a total of 654 trials were conducted in the modified SM condition compared to 1047 in the modified RR condition.

To compare the results from the present study to the results of DiSanti et al. (2019), we computed the mean number of trials to mastery for a stimulus set for each condition across participants in both studies. In the current study, seven participants were taught a total of 10 stimulus sets in the modified SM condition and 11 stimulus sets in the modified RR condition. In the DiSanti et al. (2019) study (experiment 2), four participants were taught a total of eight stimulus sets in the SM condition and another eight stimulus sets in the RR condition. Participants in the DiSanti et al. (2019) study were comparable in language skills to the participants in the present study. Table 4 exhibits mean number trials (and range) to acquire three auditory-visual conditional discriminations (i.e., a stimulus set) across the two different conditions, across both studies. The table also shows mean cognitive score and mean Vineland Adaptive Behavior Scale (VABS) scores for the participants in both studies. On average, the participants were similar in functioning across the two studies. Results show that a stimulus set was acquired in fewer trials, on average, in the modified SM condition (65 trials), followed by the RR condition (89 trials), the modified RR condition (119 trials), and the SM condition (132 trials).

**Table 3.**

*Number of Trials to Mastery, Number of Prompts, and Number of Incorrect Responses for each Participant*

| Participants | Stimulus Sets Taught | Trials to Mastery | | Number of Prompts | | Number of Errors | | Percentage Maintenance 4 Weeks | | Percentage Maintenance 6 Weeks | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | SM | RR | SM | RR | SM | RR | SM | RR | SM | RR |
| 1 | 1 | 44 | 37 | 6 | 16 | 3 | 2 | 100 | 100 | 100 | 100 |
| 2 | 1 | 38 | 36 | 6 | 15 | 1 | 2 | 33 | 22 | 22 | 33 |
| 3 | 1 | 74 | 128 | 11 | 54 | 7 | 15 | 33 | 35 | 55 | 89 |
| 4 | 2 | 192 | 189 | 36 | 92 | 28 | 24 | 100 | 100 | 100 | 100 |
| 5 | 3 | 93* | 252* | 331 | 780 | 351 | 281 | 33 | 44 | 33 | 55 |
| 6 | 3 | 134 | 246 | 19 | 136 | 8 | 37 | 47 | 55 | 48 | 48 |
| 7 | 1 | 79 | 159 | 10 | 73 | 7 | 22 | ** | ** | ** | ** |
| 8 | 1 | * | * | 161 | 197 | 166 | 117 | *** | *** | *** | *** |
| 9 | 2 | * | * | 326 | 641 | 302 | 222 | *** | *** | *** | *** |

Note. SM = modified structured mix condition; RR = counterbalanced random rotation condition. Bolded numbers in either the SM or RR condition under Trials to Mastery represent the most efficient condition. *= Mastery was not achieved for one or more stimulus sets. **= Not assessed. ***= Not assessed because the discriminations were not acquired. The 500-trial limit includes the sum of the number of trials to mastery, number of prompts and number of errors for one condition.

**Table 4.**

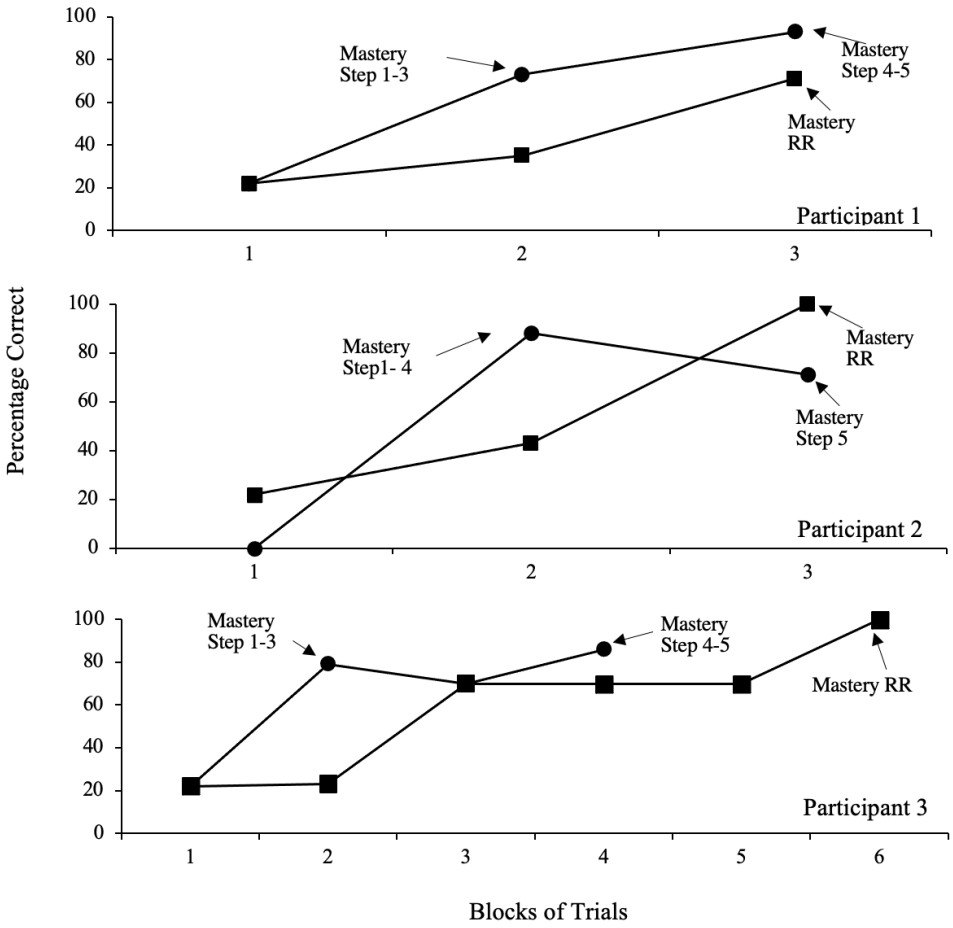
*Comparison of Results from The Present Study and DiSanti et al. (2019)*

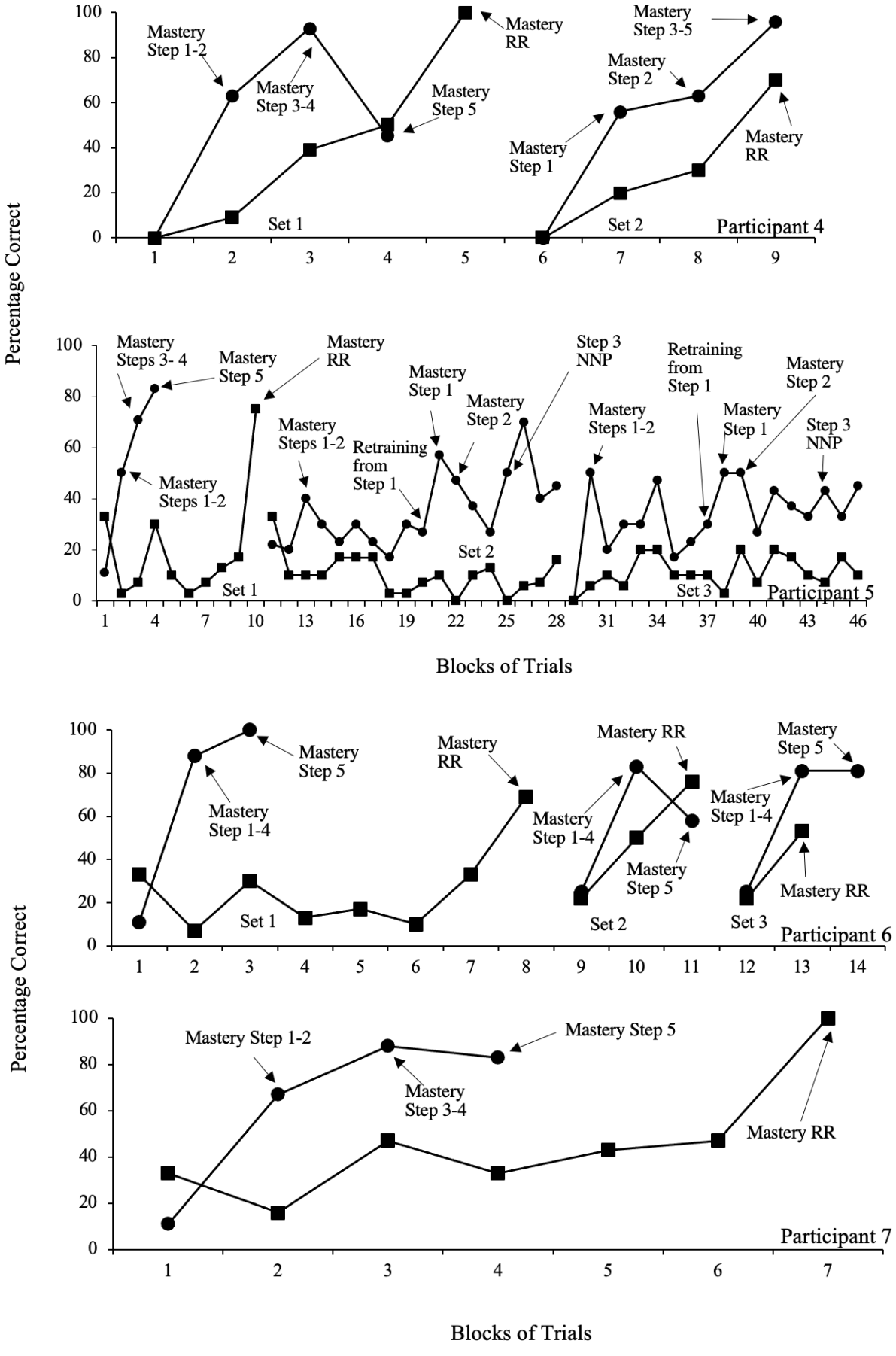
| Condition | Mean number of trials to mastery per stimulus set | Range | N | Number of stimulus sets mastered | Mean Developmental age (months) | Mean VABS ABC | Mean VABS Communication |
|---|---|---|---|---|---|---|---|
| Modified SM | 65 | 33-107 | 7 | 10 | 45 | 62 | 62 |
| RR* | 89 | 57-173 | 4 | 8 | 41 | 67 | 64 |
| Modified RR | 119 | 18-263 | 7 | 11 | 45 | 62 | 62 |
| SM* | 132 | 119-147 | 4 | 8 | 41 | 67 | 64 |

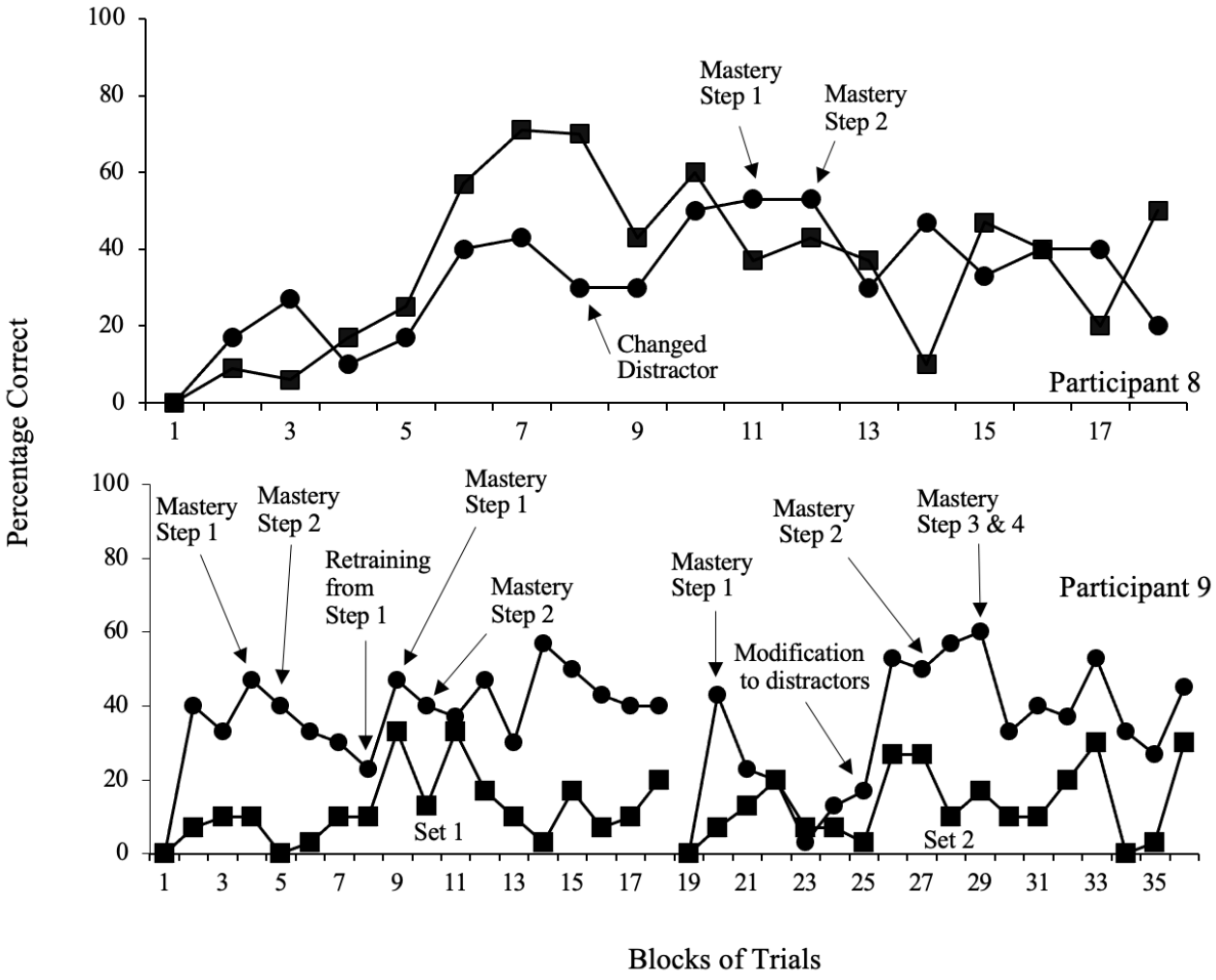
Note: * Data extracted from DiSanti et al. (2019), Experiment 2.

**Figure 1.**

*Individual data across modified SM and RR conditions.*




Percentage Correct
Mastery Step 1-2
Mastery Step 3-4
Mastery RR
Mastery Step 5
Mastery Step 3-5
Mastery Step 2
Mastery Step 1
Mastery RR
Set 1
Set 2
Participant 4
Mastery Steps 3- 4
Mastery Step 5
Mastery RR
Mastery Steps 1-2
Mastery Steps 1-2
Retraining from Step 1
Mastery Step 1
Mastery Step 2
Step 3 NNP
Mastery Steps 1-2
Retraining from Step 1
Mastery Step 1
Mastery Step 2
Step 3 NNP
Set 1
Set 2
Set 3
Participant 5
Blocks of Trials
Mastery Step 5
Mastery Step 1-4
Mastery RR
Mastery RR
Mastery Step 1-4
Mastery Step 5
Mastery Step 5
Mastery Step 1-4
Mastery RR
Set 1
Set 2
Set 3
Participant 6
Mastery Step 1-2
Mastery Step 5
Mastery Step 3-4
Mastery RR
Participant 7




Note. Percentage of correct responses per session for each participant across both conditions (Modified Structured Mix and Modified Random Rotation). Individual data points are represented across blocks of trials (e.g., blocks of 4 trials for Steps 1, 2, and 4 for modified SM; or 9 trials for modified RR and 9 trials for modified SM Steps 3 & 5). Sessions contained a maximum of 30 trials. The numbered arrows represent where each step of the Structured Mix condition was mastered. The arrows also represent where error correction procedures were conducted in the Structured Mix condition (Participants 5, 8, and 9).

## Discussion

We evaluated the efficiency of a modified SM procedure and a modified RR procedure for teaching auditory-visual conditional discriminations (receptive labels) to children with autism. The modified SM procedure contained fewer training steps, fewer prompts, and a less stringent mastery criteria compared to the SM method used in previous studies (DiSanti et al., 2019; Grow, et al., 2011; Grow et al., 2014; Grow & Van Der Hijde, 2017; Gutierrez et al., 2009; Holmes et al., 2015; Lin & Zhu, 2019; Vedora & Grandelski, 2015). Refer to the introduction section for a detailed description of the modifications made to the training steps, prompt procedures, and mastery criteria for the SM and RR procedures.

The efficiency of the two procedures were assessed by comparing the number of trials to mastery for establishing three auditory-visual conditional discriminations (i.e., a stimulus set) using the two different discrimination training procedures. Results showed that the modified SM procedure was more efficient for four of the nine participants, the modified RR procedure was more efficient for one of the nine participants, both procedures were equally efficient for two participants, and neither procedure was effective for two of the nine participants. For the stimulus sets which were mastered, the number of errors across the two conditions were comparable, except for two participants who were characterized as having limited verbal repertoires.

Maintenance was assessed four and six weeks after training for nine stimulus sets (Table 3). Across four of the stimulus sets, better maintenance scores were observed in the modified RR condition. For three stimulus sets, better maintenance was observed in the modified SM condition. Maintenance was equally effective for two stimulus sets (participant 4, stimulus set 1 & 2).

For one third of the participants, the modified RR was either equally efficient (two participants) or more efficient (one participant). These three participants had the highest scores on the communication subscale of the Vineland Adaptive Behavior Scale. This may suggest that for participants with an advanced verbal repertoire the modified RR procedure may be most efficient, which could be a topic for further research. For example, future research could compare the two procedures when teaching auditory-visual conditional discriminations to individuals with an intact verbal repertoire.

Based on available data in the present study, there is no clear-cut answer to the question of which approach is most efficient when teaching auditory-visual conditional discriminations to children with autism. Results from the current study suggest that a structured mixed procedure, in general, is more efficient when training steps are reduced, and prompt and mastery criterion are less stringent (i.e., the modified SM procedure used in the current study). Preliminary data suggest that the modified SM procedure is more efficient compared to the RR procedure, but this finding warrants replication, considering there was no intra-participant replication for most participants, except for two (P4 and P6). The modified SM procedure appeared to be more efficient for children with a limited verbal repertoire (DiSanti et al., 2019; Lin & Zhu, 2019); whereas preliminary data may suggest the RR procedure is more efficient for children with a more advanced verbal repertoire. This observation warrants further research.

For those children with an advanced or intact verbal repertoire, perhaps the SM procedure interferes with the acquisition of conditional control due to the reinforcement contingencies involved; that is, by reinforcing irrelevant sources of stimulus control that competes with the desired type of stimulus control required to eventually establish the conditional discrimination. This could be to select the comparison stimulus that was the S+ on the previous trial, select the comparison stimulus which has a stronger reinforcement history across trials or sessions, or to select the stimulus occurring in a particular position; all of which may interfere with sample-stimulus control. However, some children with a more limited verbal repertoire may not have learned to be affected by such reinforcement contingencies. Consistently selecting the S+ and not S- across successive trials and responding to reversal of S+ and S- functions may likely be a prerequisite for learning conditional discriminations (McIlvane et al., 1990). The SM procedure and the modified SM procedure may facilitate the establishment of these prerequisites for some participants.

Another reason why the SM procedure may be efficient is that a two-choice format is used when the first two stimuli are intermixed. This may have simplified the discrimination task and somehow facilitated learning of the sample-comparison relation. Correct responding during this part of training cannot occur unless the sample stimuli exert some type of control over comparison selection. The desired type of stimulus control would be to select S1 when hearing the name of S1 and select S2 when hearing the name of S2 (when S1 and S2 comprise the comparison array). However, it cannot be inferred from the data whether this type of stimulus control had emerged. Alternatively, the learner may have selected S1 and not S2 when hearing the name of S1, and excluded or rejected S1, hence selecting S2 by default when hearing the name of S2. In this case, a reject-relation has emerged (e.g., when the learner hears the name of S2, the learner rejects the previously reinforced S1 and selects the other comparison stimulus; Johnson & Sidman, 1993). Stimulus control of S2 selection when hearing the name of S2 may subsequently transfer from rejecting S1 to selecting S2, consequently resulting in the desired type of stimulus control. To ensure that a conditional discrimination has been established and that the child has learned receptive labeling, it is necessary to increase the stimulus array from two stimuli to three or more stimuli (Carrigan & Sidman, 1992; Sidman, 1987), but perhaps only after the discriminations in the two-choice format has been established. This could be a topic for future research.

Comparing the results from the modified RR procedure (current study) with results from the RR procedure (DiSanti et al., 2019), suggests that the RR procedure was more efficient than the modified RR procedure (Table 4). The mean number of trials were greater for the modified RR condition (119, compared to 89 in the RR condition), the maximum number of trials required for the modified RR condition were greater (263, compared to 173 in the RR condition), and the modified RR procedure had a higher range of trials (245, compared to 116 in the RR condition). Though preliminary, this could be because the error correction procedure in the RR procedure contained elements of a SM procedure. It is possible that the modified RR error correction procedure did not provide meaningful influence on learning and inflated the number of prompts required for the modified RR condition; thus, providing fewer opportunities for the participant to respond independently in the modified RR condition. Future research may explore this topic further.

Individual tailoring should be used when deciding which procedure to use for individual children. The current study provides some guidelines for which procedure to use for which children, but as we have seen variability in data exists. Hence, for this reason, which procedure to use for individual children must be

determined empirically. In addition, which procedure is most efficient may gradually change over time. For example, a particular child may initially benefit more from using the SM procedure when learning their first receptive labels. Later, the modified SM procedure may be more beneficial once initial labels are established. Finally, after becoming a prominent learner of receptive labels, the learner may not require the systematic approach offered by the SM or modified SM procedure, and training with the RR procedure will be more beneficial. This could be a topic for future research.

A problem with studying conditional discriminations is that it cannot be observed directly when and how the sample starts exerting comparison control. This process can only be inferred from manipulating various training procedures and observing the extent to which conditional relations result from these manipulations. For example, during mass trial teaching with the modified SM procedure, it is possible that the participants discriminated the sample stimuli to some extent and came to relate them with their corresponding comparison stimuli, even though this was not required by the reinforcement contingency (McIlvane et al., 1990). Unfortunately, it is difficult to verify this experimentally in an auditory-visual conditional discrimination task. This question is more available for experimental examination when using an arbitrary visual-visual conditional discrimination task. Consider the following experiment: In the presence of the visual sample stimulus A1, selection of the visual comparison stimulus A2 and not B2 is reinforced. This training will continue until performance is stable. Next, in the presence of B1, selection of B2 and not A2 is reinforced, and training is continued until performance is stable. So far, A and B stimuli have been trained using mass trials. To assess whether the participants came to discriminate the sample stimuli (A1 and B1) after mass trials, participants can be presented with a choice task, under extinction conditions, where A1 and B1 is presented together with two other stimuli (say X1 and Y1) which have not been part of any previous training. If A1 and B1 have not been discriminated between as part of the baseline training, participants should choose the four stimuli, on average, equally often. Conversely, if there is a bias towards selecting A1 and B1, these stimuli likely are selected because they previously have been discriminated between and associated with reinforcement. Unfortunately, it is difficult to conceive how this experiment would be carried out using auditory-visual conditional discrimination, and indeed, this may be the reason why there is more research available examining visual-visual conditional discriminations compared to auditory-visual conditional discriminations. This is unfortunate, since understanding auditory-visual conditional discriminations is important for understanding the acquisition of listener behavior.

Whenever ineffective, neither procedure appeared to have advantages over the other. Three participants failed to acquire auditory-visual conditional discriminations and every time a participant failed to acquire a stimulus set this happened concurrently in both the modified SM and modified RR conditions. Hence, when a participant did not acquire a stimulus set, one procedure did not produce advantages over the other. Training novel auditory-visual discriminations in isolation before systematically progressing to a conditional discrimination could have produced better outcomes for these learners (Lovaas, 1977). A lack of motivation could also account for these findings. A brief MSWO was conducted before each training session, but the participant could have become satiated on all five stimuli tested in the MSWO. If so, the preference assessment assessed preference between stimuli that were no longer effective as reinforcers. At this stage in training a MSWO identifying 20 new putative reinforcers could have been performed.

Error analysis was conducted across both the modified SM and RR conditions, but modifications were only made to the modified SM procedure; this could be interpreted as a limitation. In the modified SM condition, two types of faulty stimulus control were observed, and both were related to a stimulus bias. One type of stimulus bias was to select the comparison stimulus that had previously been associated with reinforcement (i.e., molar win-stay or molecular win-stay), which in individuals with developmental disabilities is the most typical error pattern observed (McIlvane & Stoddard, 1981). Another type of stimulus bias included selecting the novel comparison stimulus, that is the stimulus not yet trained (i.e., comparison stimulus bias) and selecting a different comparison stimulus after failing to contact reinforcement (i.e., lose-shift). Similar error patterns were observed in the modified RR condition.

Additional error correction procedures were attempted for three of the participants (5, 8, and 9). One error correction procedure consisted of retraining previously mastered steps. This was done because we were concerned that the mastery criterion of four consecutive correct responses was insufficient to maintain correct responding. This error correction procedure was included for two participants (5 and 9) but was not effective. This suggests that the progressive mastery criterion was not the problem. A second type of error correction procedure consisted of prompting after two incorrect responses (rather than after one incorrect response). We wondered whether corrective feedback alone (e.g., "try again") could facilitate learning and reduce the potential for prompt dependency. This procedure was employed for one participant (5) but did not improve performance. A third type of error correction procedure consisted of replacing the S- with a different S- when participants

showed a preference for comparison stimuli not yet trained. This happened for two participants (8 and 9), and for both participants this resulted in the acquisition of the target step. Additional procedural modifications were not made with the modified RR procedure when learners were failing to acquire the discriminations, and this is a limitation of the study.

Future research could compare the modified SM with the original SM procedure, the modified RR procedure with the original RR procedure, and the modified SM with the original RR procedure. Comparisons could also be made where error correction procedures were not included for either procedure, or the modified SM (or RR) procedure with two different error correction procedures. In addition, future research could consider additional error correction procedures, such as manipulation of the sample stimulus to facilitate its discrimination (Saunders & Spradlin, 1989). Examples include: requiring the participants to echo the verbal sample stimulus before presenting the comparison stimuli (providing that the learner has an echoic repertoire); requiring a manual sample stimulus response, such as requiring the participants to point to the sample stimulus before presenting the comparison stimuli; or re-presenting the auditory sample stimulus every two seconds until a comparison response is emitted (Green, 2001). Research indicates that effectiveness of different error correction procedures may vary across individuals (Carroll et al., 2015; Carroll et al., 2018; Leaf et al., 2020). Furthermore, future research may continue to compare the effectiveness of comparison-first or sample-first presentation, as this has been found to be idiosyncratic across literature (Cubicciotti et al., 2019).

For ethical reasons, training was terminated if a stimulus set was not acquired within 500 training trials (DiSanti et al., 2019). It is possible that some participants would have acquired the auditory-visual conditional discriminations if more training had been conducted. For example, one participant reached the final step in the modified SM condition, but training was discontinued after relatively few training trials because the 500-trial-limit was reached. Rather than discontinuing training for this participant, teaching could have continued if progress was made across sessions. Further evaluations should also be conducted to ensure the difficulty between receptive labels trained is comparable across conditions.

In addition to the limitations already mentioned, some additional limitations should be considered. Maintenance scores were relatively low across some participants who acquired the auditory-visual conditional discriminations in either both or one of the conditions. This could be due to the participants acquiring the auditory-visual conditional discriminations within very few sessions and that no maintenance training was conducted before four- and six-week follow-up. Future research might consider conducting additional maintenance training and follow-up assessments after the discriminations have been mastered.

## Declarations

### *Ethics approval*

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Norwegian Centre for Research Data (NSD).

Consent to participate

Informed consent was obtained from legal guardians.

Conflict of interest: none.

## References

Bayley, N. (2005). *Bayley scales of infant and toddler development (3rd ed.).* HarcourtAssessment, Inc.

Carrigan P. F. Jr. & Sidman M. (1982). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 183–204. https://doi.org/10.1901/jeab.1992.58-183.

Carroll, R. A., Joachim, B. T., St Peter, C. C. & Robinson, N. (2015). A comparison of error correction procedures on skill acquisition during discrete-trial instruction. *Journal of Applied Behavior Analysis, 48,* 257-273. https://doi.org/10.1002/jaba.205.

Carroll, R. A., Owsiany, J. & Cheatham, J. M. (2018). Using an abbreviated assessment to identify effective error correction procedures for individual learners during discrete-trial Instruction. *Journal of Applied Behavior Analysis, 51,* 482-501. https://doi.org/10.1002/jaba.460.

Cubicciotti, J. E., Vladescu, J. C., Reeve, K. F., Carrol, R. A. & Schnell, L. K. (2019). Effects of stimulus presentation order during auditory-visual conditional discrimination training for children with autism spectrum disorder. *Journal of Applied Behavior Analysis, 52,* 541-556. https://doi.org/10.1002/jaba.530.

DeLeon, I. G. & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforce preferences. *Journal of Applied Behavior Analysis*, *29*, 519-533.https://doi.org/10.1901/jaba.1996.29-519.

DiSanti, B. M., Eikeseth, S., Eldevik, S., Conrad, J. M. & Cotter-Fisher, K. (2019). Comparing structured mix and random rotation procedures to teach receptive labeling to Children with Autism. *Behavioral Interventions, 35*, 38-56. https://doi.org/10.1002/bin.1694.

Eikeseth, S., Smith, D. & Klintwall, L. (2014). Discrete trial teaching and discrimination training. I n P. Sturmey, J. Tarbox, D. Dixon & Matson (Eds.), *Handbook of early intervention and autism spectrum disorders: Research, practice, and policy* (pp. 229- 253). Springer.

Green, G. (2001). Behavior analytic instruction for learners with autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities*, *16*, 72-85. https://doi.org/10.1177%2F108835760101600203.

Grow, L. L., Carr, J. E., Kodak, T. M., Jostad, C. M. & Kisamore, A. N. (2011). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders. *Journal of Applied Behavior Analysis*, *44*, 475-498. https://doi.org/10.1901/jaba.2011.44475.

Grow, L. L., Kodak, T. M. & Carr, J. E. (2014). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders: A systematic replication. *Journal of Applied Behavior Analysis*, *47*, 600- 605. https://doi.org/10.1002/jaba.141.

Grow, L.L. & Van Der Hijde, R. (2017). A comparison of procedures for teaching receptive labeling of sight words to a child with autism spectrum disorder. *Journal of Behavior Analysis Practice*, *10*, 62-66. https://doi.org/10.1007/s40617-016-0133-0.

Gutierrez, A., Hale, M. N., O'Brien, H. A., Fischer, A. J., Drocher, J. S. & Alessandri, M. (2009). Evaluating the effectiveness of two commonly used discrete trial procedures for teaching receptive discrimination to young children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 11,* 1-12. https://doi.org/10.1016/j.rasd.2008.12.005.

Holmes, E. J., Eikeseth, S. & Schulze, K. A. (2015). Teaching individuals with autism receptive labeling skills involving conditional discriminations: A comparison of mass trial and intermixing before random rotation, random rotation only, and combined blocking. *Research in Autism Spectrum Disorders*, *11*, 1-12. https://doi.org/10.1016/j.rasd.2014.11.013.

Iversen, I. H., Sidman, M. & Carrigan, P. (1986). Stimulus definition in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, *45*, 297-304. https://doi.org/10.1901/jeab.1986.45-297.

Johnson, C. & Sidman, M. (1993). Conditional discrimination and equivalence relation: Control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 59,* 333-347.

Kodak, T., Clements, A., Paden, A., LeBlanc, B., Mintz, J. & Toussaint, K. A. (2015). Examination of the relation between an assessment of skills and performance on auditory visual conditional discriminations for children with autism spectrum disorder. *Journal of Applied Behavior Analysis, 48,* 52- 70. https://doi.org/10.1002/jaba.160.

Leaf, J. B., Cihon, J. H., Ferguson, J. L, Milne, C. M., Leaf, R. & McEachin, J. (2020). Comparing error correction to errorless learning: A randomized clinical trial. *The Analysis of Verbal Behavior, 36*, 1-20. https://doi.org/10.1007/s40616-019-00124-y.

Ledford, J. R., Chazin, K. T., Gagnon, K. L., Lord, A. K., Turner, V. R., & Zimmerman, K. N (2019). A systematic review of instructional comparisons in single-case research. *Remedial and Special Education, 1-14.* https://doi.org/10.1177/0741932519855059.

Lin, F. Y. & Zhu, J. (2019). Comparison of two discrimination methods in teaching Chinese children with autism. *Journal of Applied Behavior Analysis, 53,* 1145-1152. https://doi.org/10.1002/jaba.652.

Lovaas, O. I., Ackerman, A., Alexander, D., Firestone, P., Perkins, M., Young, D. B., Carr, E. G & Newsome, C. (1981). *Teaching developmentally disabled children: The ME book*. PRO-ED.

Lovaas, O. I. (1977). *The autistic child*. Irvington Publishers, Inc.

Lovaas, O. I. (2003). *Teaching individuals with developmental delays: Basic intervention techniques*. Pro-Ed.

Michael, J., Palmer, D. C. & Sundberg, M. L. (2011). The multiple control of verbal behavior. *The Analysis of Verbal Behavior, 27,* 3-22. https://doi.org/10.1007/BF03393089.

McIlvane, W. J. & Stoddard, T. (1981). Acquisition of matching-to-sample performances in severe retardation: Learning by exclusion. *Journal of Mental Deficiency Research, 25,* 33- 4 8 . https://doi.org/10.1111/j.1365-2788.1981.tb00091.x.

McIlvane, W. J., Dube, W. V., Kledaras, J. B., Iennaco, F. M. & Stoddard, L. T. (1990). Teaching relational discriminations to individuals with mental retardation: Some problems and possible solutions. *American Journal of Mental Retardation, 95*, 283-296.

Roid, G. H. (2003). *Stanford Binet intelligence scale (5th ed.).* Riverside Publishing.

Saunders K. J. & Spradlin J. E. (1989). Conditional discrimination in mentally retarded adults: The effect of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior, 52,* 1–12. https://doi.org/10.1901/jeab.1989.52-1.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J. & Love, S. R. (2010). *Childhood Autism Rating Scale (2nd ed.).* Western Psychological Services.

Sidman, M. (1987). Two choices are not enough. *Behavior Analysis, 22*, 11-18.

Sidman, M. & Stoddard, L. T. (1966). Programming perception and learning for retarded children. *International Review of Research in Mental Retardation, 2,* 151-208. https://doi.org/10.1016/S0074-7750(08)60206-2.

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*, 5–22. https://doi.org/10.1901/jeab.1982.37-5.

Sindelar, P., Rosenberg, M. & Wilson, R. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children, 8*, 67-76.

Sparrow, S. S. Cicchetti, D.V. & Saulnier, C. A. (2016). *Vineland adaptive behavior scales (3rd ed.).* Pearson.

Sundberg, M. (2016). Verbal stimulus control and the intraverbal relation. *Analysis of Verbal*

*Behavior, 32,* 107-124. https://doi.org/10.1007/s40616-016-0065-3.

Vedora, J. & Grandelski, K. (2015). A comparison of methods for teaching receptive language to toddlers with autism. *Journal of Applied Behavior Analysis, 48,* 188-193. https://doi.org/10.1002/jaba.167.

Wolery, M., Gast, D. L., & Ledford, J. R. (2014). Comparison designs. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (pp. 297-345). Routledge.