University of Foggia

A thesis submitted for the degree of Doctor of Philosophy in
NEUROSCIENCE AND EDUCATION (Curriculum: Education) – 37th Cycle

THE SELF-REGULATION OF ENGLISH LEARNING: A LONGITUDINAL STUDY ON L2 LEARNING AMONG ITALIAN PRIMARY SCHOOL STUDENTS

Field of Study: Psychology (M-PSI/01)

PhD Student

Eugenio Trotta

Supervisor

Paola Palladino, PhD

Academic year 2023 / 2024

CONTENTS

Abstract.....6

1. General introduction.....7

1.1. Aim of the doctoral thesis.....9

1.2. Theoretical framework.....10

1.2.1. MASRL Model.....11

1.2.2. The Control-Value Theory.....11

1.2.3. A comparative analysis of MASRL Model and CVT.....12

2. Study 1: the effects of COVID-19 pandemic on Italian primary school children's learning: a systematic review beyond the academic performance of Italian primary school students.....15

2.1. Method.....18

2.1.1. Literature search and search strategy.....18

2.1.2. Inclusion and exclusion criteria.....18

2.1.3. Screening process.....19

2.1.4. Quality assessment.....19

2.2. Results.....19

2.2.1. Data extraction.....19

2.2.2. Risk-of-bias assessment.....21

2.3. An overview of the effects of the school closure due to the COVID-19 pandemic.....25
2.3.1. Parents' opinion.....31
2.3.2. Teachers' opinion.....34
2.3.3. Students' opinion.....38

2.4.Discussion.....41

3. Study 2: investigating the interplay between cognitive, metacognitive, and emotional variables in L2 learning overtime.....48

3.1.Method.....49
3.1.1. Sample and procedure.....49
3.1.2. Measures.....50
3.1.2.1. Collective tests.....50
3.1.2.2. Individual tests.....51
3.1.3. Data analysis.....54

3.2.Results.....57
3.2.1. Relations between measures time after time.....57
3.2.2. Relations between measures overtime.....59
3.2.3. Growth Curve Analysis (GCA).....61
3.2.4. Multilevel model for repeated measures (MLM).....64

3.3.Discussion.....66

3.4.Conclusion.....68

4. Study 3: an investigation on emotions related to L2 learning in Italian primary school.....69

4.1. Method.....71

4.1.1. Participants.....71

4.1.2. Instruments and administration procedure.....71

4.1.3. Data analysis.....71

4.2. Results.....72

4.2.1. Analysis on the total sample.....72

4.2.1.1. Correlations.....72

4.2.1.2. Multiple linear regression.....73

4.2.2. Analysis on the partial sample.....74

4.3. Discussion.....77

4.4. Conclusion.....78

5. Study 4: analysis of psychometric properties of English Metacognitive Questionnaire for Primary School (EMQ–PS) 80

5.1. Introducing the original questionnaire.....82

5.2. Method.....83

5.2.1. Sample and procedure.....83

5.2.2. Teachers' perception.....85

5.2.3. Data analysis.....85

5.3. Results *87*

5.3.1. Preliminary analysis 87

5.3.2. Exploratory Factor Analysis 90

5.3.3. Principal Component Analysis 93

5.3.4. Evaluation of the identified dimensions 97

5.3.5. Qualitative analysis 99

5.3.6. Individual cases 100

5.3.7. Comparison of student self-assessments and teachers' perceptions 101

5.4. Discussion and conclusion *102*

6. General discussion 106

6.1. Study 1 ..... *106*

6.2. Study 2 ..... *107*

6.3. Study 3 ..... *108*

6.4. Study 4 ..... *109*

6.5. Limitation and considerations for the future *110*

7. General conclusion 111

8. Reference 112

Abstract

Self-regulated learning refers to the ability to regulate one's learning behavior, emotions and thoughts in a goal-directed manner. Due to its ability to include several variables that influence student learning, it is one of the most studied areas of research in educational psychology. However, learning changed radically in recent years. COVID-19 significantly affected all territories and contexts, not sparing primary school students. To date, there are countless questions about the effects of school closures on the learning of students. Through a systematic review (Study 1), this doctoral thesis contributes to shedding light on the learning conditions of such an affected population, highlighting changes in the cognitive, emotional and metacognitive fields in self-regulated learning contexts. Furthermore, despite being one of the fundamental subjects of the national surveys (INVALSI), as well as the first gateway to the world, there is no information about the learning of the second language in Italy, namely English. Therefore, this doctoral thesis continues with longitudinal research (Study 2) on the role of basic cognitive components (including L1 performance – Study 3), achievement emotions, and metacognition in L2 learning by Italian primary school post-pandemic students. The results highlighted a key role of working memory and metacognitive skills, as well as the constant presence of anxiety regarding the L2 over time. The doctoral thesis concludes with the psychometric analysis of a tool for investigating metacognitive, motivational and emotional skills in relation to L2: the English Metacognitive Questionnaire for Primary School (EMQ-PS – Study 4). This latter was developed to help professionals identifying the most at risk students and those in need of attention for an effective self-regulated learning.

Keywords:

L2 Learning; Achievement Emotions; Metacognition; Cognition; COVID-19.

1. General introduction

Learning is a complex and multidimensional process that involves various interconnected aspects. Undoubtedly, cognition is the strongest predictor of learning, as recognized by several studies (e.g., Deary et al., 2007; Floyd et al., 2003; Feraco et al., 2023). According to the Cattel-Horn-Carroll (CHC) theory of cognitive abilities (Schneider & McGrew, 2018), the human intelligence is multidimensional, divided into three strata: general intelligence (*g* – stratum III); broad abilities (stratum II); narrow abilities (stratum I). The stratum II (broad abilities) is an intermediate level that includes general cognitive abilities, such as Fluid Reasoning (Gf), Short-term Memory (Gsm), and Long-term Retrieval (Glr); the most specific level is the stratum III (narrow abilities), comprising specialized cognitive abilities, such as working memory – subsumed by Gsm (Floyd et al., 2003). According to the recent models, the CHC theory includes 16 broad abilities and almost 100 narrow abilities (Schneider & McGrew, 2018). Several meta-analyses highlighted how over 50% of academic achievement is explained uniquely by *g* (see for example, Roth et al., 2015; Kriegbaum et al., 2018; Zaboski II et al., 2018). Stratum by stratum, further studies recognized the role of Glr in learning, recognized as one of the most important abilities for scholastic achievement (Kaufman et al., 2012; Avitia et al., 2019). Furthermore, there are evidence that learning is related to Gsm, in particular with working memory (Bull et al., 2008; Schneider & Niklas, 2017; Quílez-Robres et al., 2021). Alloway and Alloway (2010) investigated whether working memory was merely a mediator of *g* or if it had a predictive role in pupils' academic achievement. Their results highlighted that working memory represents not a surrogate of general intelligence, but a dissociable ability that predicts academic achievement from early schooling, regardless of the learning domain (Floyd et al., 2003; Alloway & Passolunghi, 2011; Swanson & Alloway, 2012) or student's scholastic difficulties (Alloway, 2009). However, despite the relevance of cognitive abilities is unquestionable, attributing an exclusive role to them in learning can be a restriction.

Besides cognition, metacognitive factors play a role in self-regulated learning. The term metacognition indicates the knowledge and awareness of one's own cognitive processes as well as the self-reflective ability to regulate these processes to achieve a specific goal (Flavell, 1979). According to Nelson and Narens (1990), the cognitive processes related to a specific task (object-level) are influenced by a dynamic model, known as meta-level. The two levels

are linked by metacognitive monitoring and control. The first one informs the meta-level by the object-level: in other words, the metacognitive level observes and evaluates the processes and performance to be active on the task, based on one's knowledge and skills. The metacognitive control, instead, modifies the actions on the task (object-level) based on evaluations made. To give an example, imagine a student preparing for a test. The student is cognitively busied to read chapters from a textbook, trying to understand as much information as possible (object-level). The student evaluates (meta-level) his/her understanding of the material (monitoring) and decides whether to change study strategy, perhaps moving from reading to creating schemes or solving practical exercises (control). Depending on the timing, metacognitive skills – both monitoring and control – can be distinguished as prospective and retrospective, which means such as confidence in past choices and a prediction of success, respectively (Fleming et al., 2016; Destan et al., 2014; Siedlecka et al., 2016; Bayard et al., 2021).

Certainly, a relevant role in self-regulated learning is played by emotions (Boekaerts, 2007). Several studies highlighted learning is strongly connected to students' emotions directly linked to activities and outcomes, also known as achievement emotions (Pekrun, 2006; Artino & Jones, 2012; Putwain et al., 2013; Wu et al., 2021). Two recent systematic reviews (Tan et al., 2021; Wu & Yu, 2022) investigated the relationship between (positive/negative) achievement emotions and learning (or online learning), finding both that positive achievement emotions improved the learning effects more than negative ones, although the latter (in particular, anxiety and boredom) were also could be improve learning strategies (see also Boekaerts, 2007). The Control-Value Theory (CVT; Pekrun, 2000, 2006) demonstrates how achievement emotions are part of students' everyday life. These emotions are influenced by individual characteristics (appraisal related to control and value experiences) and by the context experienced at both proximal (e.g., task difficulty) and distal levels (e.g., educational systems). In turn, achievement emotions have a consequent impact in students' motivation, learning strategies, consideration of their own cognitive abilities, and ability to self-regulate their study, thereby influencing their academic success (McCombs, 2013; Pekrun & Linnenbrink-Gracia, 2012).

Therefore, it is indisputable the active and joint role of cognitive, metacognitive and emotional factors on self-regulated learning. Although it primarily discusses "affect" rather than specifically "achievement emotion", all the aforementioned factors can be found within

the metacognitive and affective model of self-regulated learning (MASRL – Efklides, 2011). The self-regulated learning (SRL) is a process through which students take control of their own learning, managing their personal resources to achieve a task (Zimmerman & Schunk, 2011). In the last decades, several models have been realized to describe the function of SRL, emphasizing different components of it (for a review, see Panadero, 2017). The MASRL model considers SRL as an iterative and dynamic process where metacognitive, cognitive and emotional factors interact between themselves. These interactions happen mainly at a microlevel known as the Task*Person level, underlining the secondary relevance of the individual characteristics of the learner with respect to the combination between them and the characteristics of the task processing. This condition makes the student "aware of the ongoing thinking, feelings, emotions, or physiological states denoting effort exertion during task" (Efklides, ibid.). Hence, the MASRL identifies the specific learning task as a central element. However, little is still known about the role these factors play (alone and together) in second language learning (from now on, L2 learning).

1.1. Aim of the doctoral thesis

Certainly, several studies shown the involvement of cognitive (Palladino & Cornoldi, 2004; Palladino & Ferrari, 2008; Rankin, 2017; Artuso & Palladino, 2019), metacognitive (Anderson, 2002; Raoofi et al., 2014; Aripin & Rahmat, 2021), and emotional factor (Shao et al., 2019; Teimouri et al., 2019; Li & Wei, 2023; Wang, 2024) on L2 learning, individually. However, less is known about the interconnected role of these factors. Furthermore, the limited evidence of their joint impact on L2 learning focused on an adult sample (Ghulamuddin et al., 2022). Therefore, this doctoral thesis aimed to address the following research questions:

#1) Is there a relationship between the achievement emotions, cognitive, metacognitive variables, and L2 learning?

#2) Is there a time effect on L2 learning, achievement emotions, cognitive, and metacognitive variables?

#3) How achievement emotions, cognitive and metacognitive variables can jointly influence English as a L2 performance in post-pandemic primary school students?

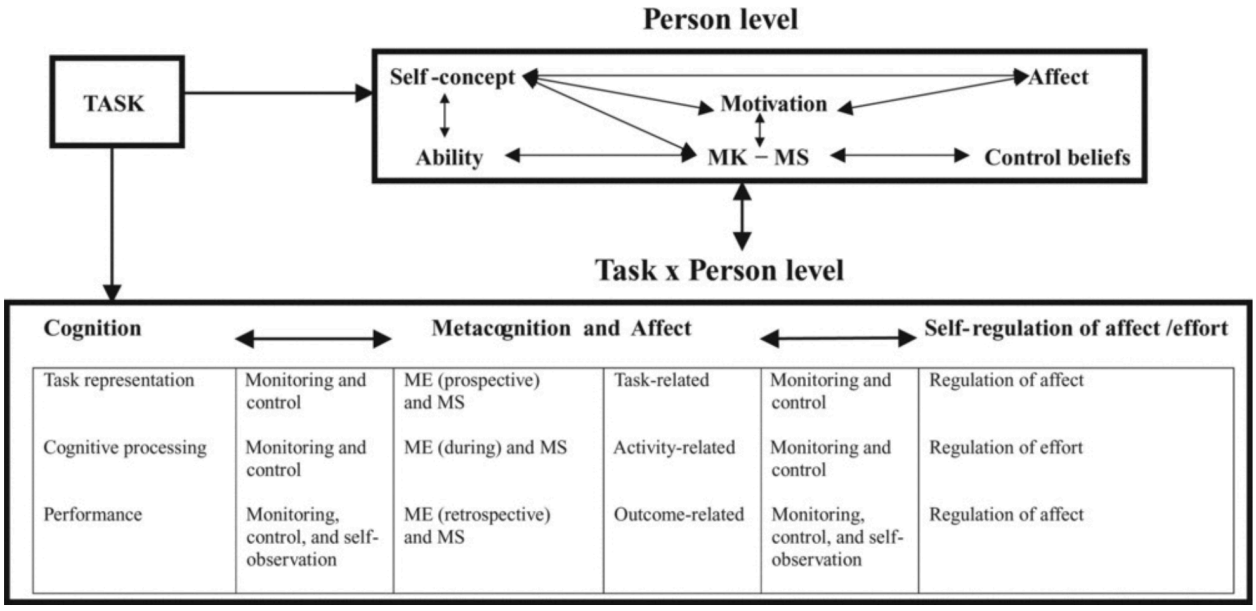
We expected to find a positive relation between positive achievement emotions, such as enjoyment with English learning performance, and an opposite relationship with negative achievement emotions, in particular with anxiety, increased during the school closure for COVID-19. Moreover, we expected to find a positive relation between cognitive and metacognitive measures and L2 performance. Finally, performance in the cognitive, metacognitive, and emotional measures as well as their relation to L2 performance were expected to change over time. To address these goals, we first conducted a systematic review (from now on, Study 1) to collect the opinions of parents, teachers, and students on: other dimensions of Italian primary school students affected by school closures, beyond academic performance; hypothetical agreement between the opinions of parents, teachers, and students regarding the different effects of school closures on Italian primary school students; possible differences between the effects of school closures on Italian primary school students and the students in other countries. After that, a longitudinal study (from now on, Study 2) was conducted to evaluate the role of achievement emotions, cognitive and metacognitive variables (singly or jointly) over time. Specifically, we focused on retrospective metacognition in Study 2, as several studies highlighted a predominant role of it, particularly the confidence judgment (CJ – retrospective monitoring), over prospective metacognition in self-regulated learning (Destan et al., 2014; Siedlecka et al., 2016; Bayard et al., 2021). Subsequently, we conducted two further studies: from now on, Study 3 and Study 4. Study 3 aimed to address the role of achievement emotions experienced in L2 learning context together with prior competence in the mother tongue (hence, Italian – L1). Study 4, instead, evaluated the reliability and psychometric properties of an instrument related to general metacognitive competence in L2 by Italian primary school students.

1.2. Theoretical framework

In our work, we considered it relevant both cognitive Models: The Metacognition with Motivation and Affect in Self-Regulated Learning (MASRL) Model for metacognitive competence (Efklides, 2011) and the Control-Value Theory (CVT) for achievement emotions (Pekrun, 2000). Despite both being relevant for learning, to our knowledge, no one considered them together. In this section, we analyzed the similarities, differences, and implications for understanding learning. First, we introduced a short explanation of both Models, starting from the MASRL one.

1.2.1. MASRL Model

Focused on the Task, the Model by Efklides considered two different levels: a Person Level and a Task*Person Level (see Figure 1). The first one is also called the "Macrolevel", representing a general top-down level of the student's stable characteristics regarding a specific Task (e.g.: self-concept; ability; motivation; metacognitive knowledge). The Task*Person Level is called the "Microlevel", with a bottom-up approach, where the student's focus shifts to the detailed processes involved in Task performance: cognition; metacognition and affect; self-regulation of affect/effort.



Note: ME = Metacognitive Experiences; MK = Metacognitive Knowledge; MS = Metacognitive Skills.

Figure 1. The MASRL Model.

1.2.2. The Control-Value Theory

The Control-Value Theory considered a series of complex interactions among emotions, perceptions of control, and values attributed to activities (Figure 2). The theory's approach is multidimensional, aiming to explain how these factors interact and influence individuals' behavior in educational contexts. The Theory by Pekrun does not explicitly include a concept of a "top-down" or "bottom-up" level in its conceptual framework. However, some aspects of the theory can be interpreted in terms of influences that emerge from the top or the bottom,

despite the factors being relevant for performing (positive or negative) achievement emotions. This Model considered four different levels: Environment (Distal and Proximal); Person; Emotion.

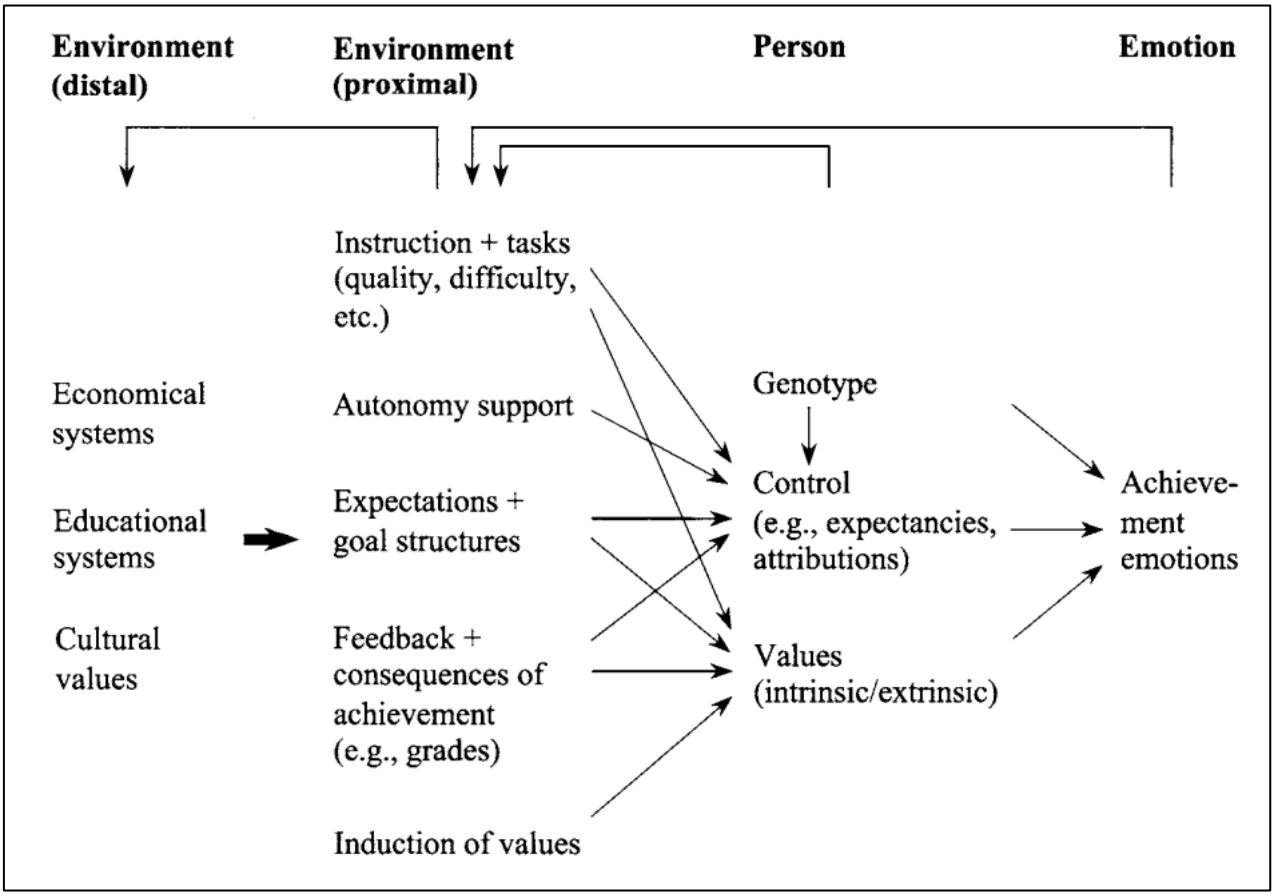


Figure 2. Control-Value Theory of achievement emotions.

1.2.3. A comparative analysis of MASRL Model and CVT

As first intuitive similarity, both Models share a Person level. The Person level by Pekrun considered genotype, control, and values factors. Hence, the theory considered crucial the perception of cognitive control (having or not having control over one's cognitive process during learning). This aspect could be interpreted in terms of a kind of "top-down" control that originates from the individual and influences how they approach cognitive challenges, in line with the Person level by Efklides. The achievement emotions themselves might be identified as the "bottom-up" level in the Control-Value Theory. In a broad view, emotions can be perceived as automatic and immediate responses to situational stimuli (Leventhal & Scherer, 1987). So, they could be considered a "bottom-up" aspect of the theory, as they can spontaneously emerge in response to specific situations.

We could say the two Models share the focus on learning (although the self-regulated learning in the MASRL Model by Efklides – for a review, see Panadero, 2017), considering how several different factors (e.g.: metacognition; motivation; achievement emotions) influence the learning process. The MASRL Model considered relevant the role of motivation for learning, covering a central role on the Person level. The motivation was not explicated in the structure of the Control-Value Theory, although it was postulated as a consequence of the perceived achievement emotions (Artino Jr, Holmboe, & Durning, 2012).

Another point of agreement between the Models is the role of the Task, which is central to the MASRL Model. According to Efklides (2011), the Task is linked to both levels (Person and Task*Person). Executing a specific Task, the students carry with them their characteristics (e.g.: cognitive ability; Metacognitive Task). Consequently, the evaluation of the objective Task characteristics starts from subjective experience. At the Task*Person level "the effect of general person characteristics progressively decreases and information coming from the monitoring of features of online Task processing receives precedence" (Efklides, 2011, p.7). In Pekrun's Control-Value Theory, the Task – located in the Proximal Environment level – is connected to the dimensions of the value attributed to the activity and perceived control (Person level). Hence, if a specific Task is valued as interesting, it can generate positive achievement emotions. On the contrary, if a Task is considered unimportant, negative emotions may arise. The Task can also influence the perception of control, as some Tasks may be seen as more controllable than others, in terms of expectancies or attributions, modulating the achievement emotions associated with the Task. Regarding the perception of control, even Efklides' Model considered it, in line with the CVT Theory: Pekrun referred to the student's expectancies and attribution, directly linked to the achievement emotions perceived; Efklides considered the perception of control as the students' beliefs as an agent, influencing (and influenced from) their metacognitive knowledge and skills at the Person level.

An interesting difference between the models is the consideration of the Environment. For Pekrun, it is the point of start, where he divided between Distal and Proximal Environment. The Distal Environment refers to broader aspects of the environment (cultural, social, or family factors) that can influence achievement emotions, exerting an indirect influence on learning. On the other hand, the Proximal Environment refers to more immediate aspects of the environment during the learning activity, such as the nature of the Task, as we reported.

The MASRL Model did not explicit the role of the Environment, particularly the Distal one. However, the Task itself could be considered as the (Proximal) Environment, since it changes based on its nature and, in turn, the attitude of the students regarding itself.

As Pekrun did not consider the concept of metacognition in his Model, Efklides did not consider the achievement emotions. However, Efklides' Model highlighted the role of affect, a distinct yet interconnected concept with emotion. As known, emotions are complex, brief experiences that involve a physiological, cognitive, and behavioral response to a stimulus or situation (Lazarus, 1991). The concept of affect refers to a general state of emotional or affective experience, including both specific emotions and a more general emotional state (Barrett & Russell, 1999). Therefore, emotion refers to specific and circumscribed emotional experiences, while affect is a broader term that can indicate a general emotional state that can include a range of specific emotions. According to Efklides (2011), affect and its regulation are relevant at both Levels in the perception of students' motivation and metacognitive competence towards learning.

In conclusion, both Models considered relevant aspects of learning, finding some similarities and differences. The MASRL Model explains in detail the interconnection of metacognitive, motivational, and affective competence and their interaction with the Microlevel and Macrolevel for self-regulated learning. The Control-Value Theory emphasizes the factors that realize positive/negative achievement emotions, relevant to explaining the mood of how a student may be approached at a specific Task.

Therefore, we started from the evaluation of the environment, which means in our case the school closure condition for COVID-19 in Italian primary schools (Distal Environment). Subsequently, we conducted the aforementioned studies to evaluate the role of several variables on the specific task of second language learning in a such particular and unknown population: post-pandemic primary school students.

2. Study 1: the effects of COVID-19 pandemic on Italian primary school children's learning: a systematic review beyond the academic performance of Italian primary school students.

This chapter refers to the study of Trotta et al. (2024). Here below a list of tables and figure included in this chapter.

List of tables included in Study 1.

Table 1	**Newcastle-Ottawa scale adapted for Cohort studies (NOS-C).**	22
Table 2	**Newcastle-Ottawa scale adapted for Case-control studies (NOS-CC).**	22
Table 3	**Newcastle-Ottawa scale adapted for Cross-sectional studies (NOS-CS).**	24
Table 4	**Synthesis of studies included.**	26
Table 5	**Descriptive synthesis of the included studies.**	30
Table 6	**Parents' opinion.**	34
Table 7	**Teachers' opinion.**	37
Table 8	**Students' opinion.**	39

List of figures included in Study 1.

Figure 3	**PRISMA flowchart of the literature search and screening process of Study 1.**	20
Figure 4	**ROBINS-I tool.**	21

9th March 2020 represents an unforgettable date in recent Italian history. On that date, the former Italian Prime Minister declared the start of the national lockdown (Galasso & Varriale, 2021). The pandemic has dramatically changed people's life worldwide, impacting all sectors, with repercussions also for the school sector. Indeed, different countries closed educational institutions to prevent the spread of COVID-19 (World Health Organization, 2023) causing an interruption of regular learning. The same has been done in Italy, where schools remained closed for almost 40 weeks (Gavosto & Romano, 2021; Hale et al., 2020). After some days/weeks of total school interruption, most countries, including Italy, started to adopt distance learning (Sector, 2020; Zhang et al., 2020). However, distance learning is a complex

process that needs a careful design to lead to effective learning (Bozkurt & Sharma, 2020). Due to the unpreparedness that led to the implementation of the distance learning, several studies revealed some achievement gap as an effect of school closure on the students' learning quality at different educational levels (Hammerstein et al., 2021; Thorn & Vincent-Lancrin, 2021).

König and Frey (2022) conducted a meta-analysis including 18 studies from different countries on primary and secondary school students, showing that school learning tended to be negatively affected by school closures, with a particular impact on primary school students. These results can also be found in Betthäuser et al. (2023) and Di Pietro (2023), comprising respectively 42 and 39 studies from different geographical areas on primary and secondary education students–even tertiary in Di Pietro (2023). Both meta-analyses agreed that students were unable to recover from the learning loss caused by school closures due to COVID-19. Several studies agreed that school closures mainly affected the most marginalized students, both scholastically (Cinquin et al., 2019; Grewenig et al., 2021) and emotionally (Banerjee et al., 2021; Scheer & Laubenstein, 2021). These students are identified as vulnerable due to disadvantaged educational, environmental, and social factors, including socio-economic status (SES), cultural differences, learning disabilities and special education needs (SEN). Also, the students included in these categories are accompanied by additional scholastic support compared to their classmates, which is challenging to dispense online, and may thus represent an additional source of difficulties in times of pandemic. This shows how the learning environment is essential to the student's well-being (Immordino-Yang, 2016; Cozzolino et al., 2022; García Bacete, 2014). As a consequence of the online transition, students experienced a period of forced isolation at home, away from other places of interest and their peers, with negative psychological consequences (Elharake et al., 2022). Several studies highlighted a fluctuation in the emotional behavior of students during the pandemic, with a significant increase in anxiety, depression, sadness (Chaabane et al., 2021; Meherali et al., 2021), suicide attempts (Saulle et al., 2022), and behavioral problems such as restlessness and inattention (Viner et al., 2022).

The debate on the quality of distance learning was also accentuated by the presence of problems related to teaching and assessment. Among the difficulties related to teaching, there are technological complications, such as the presence of adequate support tools, computers, good quality broadband. Another challenge is creating an ideal study environment (Goudeau

et al., 2021; Pokhrel & Chhetri, 2021; Yusuf & Ahmad, 2020). Furthermore, ensuring the proficiency adequate of both teachers and students, along with their parents, in the use of e-learning instruments and practices is essential (Panagouli et al., 2021; Sulla et al., 2022). Engzell et al. (2021) reported that the Netherlands–whose schools were closed for eight weeks–did not experience such technological problems. However, the study refers to a country with an excellent school performance according to the Organization for Economic Cooperation and Development (OECD) average, high academic autonomy, and the greatest presence of broadband among European countries. Despite the lack of such technological problems, learning losses were also found in Dutch students, highlighting that low socio-economic background and parental education/support negatively influenced students' learning (Bol, 2020).

Based on the information previously mentioned, it appears that the closure of schools due to COVID-19 and the subsequent distance learning affected students across educational levels worldwide, with implications extending beyond academic performance. It is plausible that similar circumstances have also transpired in Italy. Indeed, according to the Italian school calendar (Decreto Legislativo 16 aprile 1994, n. 297), Italian students lost over a year of regular learning during the pandemic. The national institute for the evaluation of the education and training system (INVALSI) annually tests the learning outcomes in language, mathematics, and English of Italian 2nd and 5th graders. The 2022 INVALSI data, compared with 2019 and 2021 cohorts, revealed a containment of the pandemic effects on learning of Italian language and a general slight worsening in mathematics (Rapporto INVALSI, 2022). Moreover, these data show that students with a linguistic disadvantage achieved lower results compared to an average student of the same age, highlighting the post-pandemic difficulties of students with low SES and a migrant background. However, although these data also consider socio-economic, cultural, and migratory aspects, they do not consider further aspects that may have influenced students' learning outcomes during the pandemic. Several international studies highlighted how the transition to distance learning led to serious emotional consequences and behavioral changes on students (Banerjee et al., 2021; Scheer & Laubenstein, 2021; Elsalem et al., 2020; Kee, 2021). Assessment of distance learning is still a widely debated topic, as digital skills among teachers and parents are heterogeneous–both in terms of the ability to use and of the availability of adequate tools (Elfirdoussi et al., 2020; Lassoued et al., 2020). There is also international agreement that distance learning has further

affected students already on the margins of school community, such as SEN students or those with cultural and economic difficulties, increasing social inequalities (Chaabane et al., 2021; Kawaoka et al., 2022; Pfefferbaum, 2021). For all these reasons, the Italian situation may represent a good case to be studied to understand the effects of long-time school closures. We decided to focus our investigation on primary school level, as it is a crucial period for pupils' learning outcomes.

2.1. Method

2.1.1. Literature search and search strategy

The articles were searched according to the PRISMA guidelines (Page et al., 2021). The study was preregistered on the Open Science Framework (https://doi.org/10.17605/OSF.IO/P4FGM –March 17th, 2023). On February 2nd, 2023; we searched on 4 different databases: Web of Science; Pubmed; EBSCOHost (including PsycArticles, PsycInfo, Psychology and Behavioral Sciences Collection); Scopus. We examined the preprint servers PsyArXiv, Open Science Framework, and PROSPERO, searching papers, preprints, or preregistered works on the same topics. Limiting the identification to the topic section (title, abstract, keywords), they used the following search string: *(COVID OR CORONA OR "SARS-COV-2") AND (ITAL*) AND (SCHOOL OR "PRIMARY SCHOOL" OR "ELEMENTARY SCHOOL" OR SCHOLASTIC) AND (LEARN*)*. Next, we used both the backward (the works cited in the selected articles) and forward (the studies that cited the considered articles) methods to identify any further studies.

2.1.2. Inclusion and exclusion criteria

We retained any article published between March 9th, 2020 (first COVID-19 lockdown in Italy) and February 2nd, 2023 (article search date); written in English or Italian. Study participants had to be Italian primary school students and/or their teachers/parents, but always with a specific focus on Italian primary school students' leaning during COVID-19 lockdown. We excluded studies that did not satisfy those criteria. Furthermore, we excluded non-journal

papers, editorials, dissertations, letters to authors, comments on published articles, and grey literature in general.

2.1.3. Screening process

Eligible studies were screened using Rayyan (Ouzzani et al., 2016). After removing duplicates, we assessed the titles and abstracts of all identified articles compared to the predefined criteria, discussing any disagreements later with the last two authors. Two authors reviewed the text of the selected publications and performed reliability using the Cohen's Kappa (κ) to measure agreement between two reviewers. Cohen's Kappa was interpreted using Landis and Koch (1977) convention as Poor (< 0.00), Slight (0.00–0.20), Fair (0.21–0.40), Moderate (0.41–0.60), Substantial (0.61–0.80), and Almost Perfect (0.81–1.00) Agreement.

2.1.4. Quality assessment

Given the heterogeneity of the studies included, we used different tools to assess the quality of the evidence and the risk of bias. We used the Risk of Bias in Non-randomized Studies–of Interventions tool (ROBINS-I – Sterne et al., 2016) for the assessment of the studies that used an experimental methodology (Baschenis et al., 2021). All the other studies used an observational methodology, therefore we used the Newcastle-Ottawa Scale (NOS – Peterson et al., 2011) in three different versions, depending on the type of study: cohort studies (NOS–C); case-control studies (NOS–CC); cross-sectional studies (NOS–CS). Our systematic review includes: 3 cohort studies (Bazoli et al., 2022; Borgonovi & Ferrara, 2023; Contini et al., 2022); 1 case-control study (Termine at el., 2022). The remaining 29 are cross-sectional studies. Two authors conducted an independent assessment of study quality, calculating percent agreement and Cohen's K (Cohen, 1960). Any concerns were discussed subsequently with the other authors.

2.2. Results

2.2.1. Data extraction

380 papers potentially relevant were found: 176 papers from Scopus, 116 from Web of Science, 61 from PubMed, and 27 from EBSCOhost. See Figure 3 for a PRISMA flowchart of the literature search process.

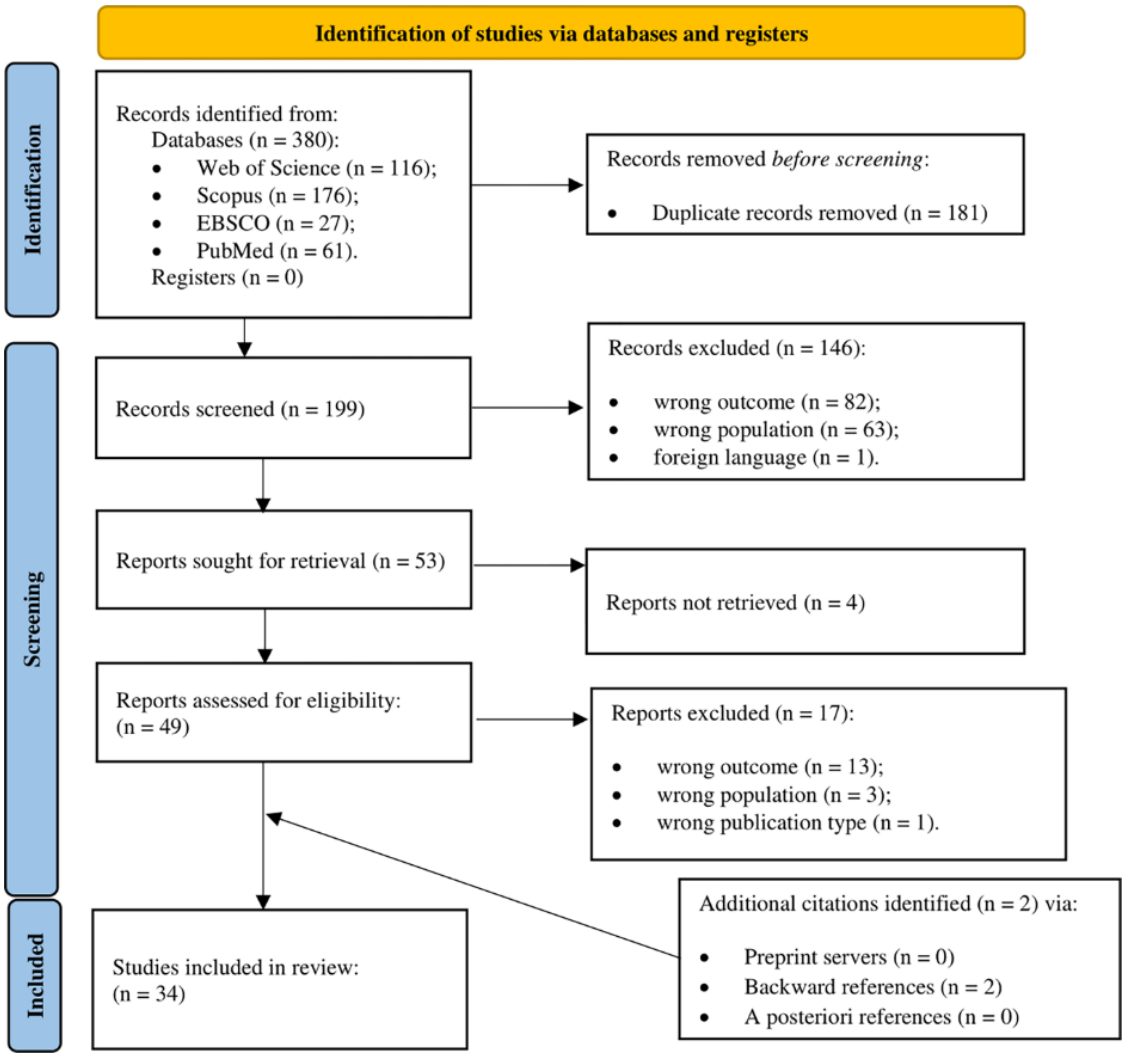


Figure 3. PRISMA flowchart of the literature search and screening process of Study 1.

First, we detected and removed 181 duplicates. We screened the titles and abstracts of the remaining 199 articles, removing 146 papers: 82 because of the wrong outcome (no reference to student learning); 63 for the wrong population examined (pre-school, middle-school, high-school, university, or others); 1 for the paper's language (different to English or Italian). Four records were excluded as the full text was not available. Two author independently reviewed the remaining 49 eligible studies, according to the criteria defined in advance. This process led to the exclusion of 17 papers: 13 for a different outcome (not focused on Italian primary school students' learning); 3 for the population (different from an Italian sample); 1 for the

type of publication (dissertation). Then, we included 2 articles consulting the papers' references via backward and forward search. We computed a Cohen's Kappa (κ) of 0.80 –a substantial agreement level (Landis & Koch, 1977). We discussed emerging discrepancies with the last two authors. We retained 34 articles suitable for inclusion.

2.2.2. Risk-of-bias assessment

Two authors independently assessed the methodological quality of the studies with ROBINS-I (Sterne et al., 2016), and NOS (Peterson et al., 2011) in three different versions (cohort, case-control, and cross-sectional studies). When the independent evaluation was finished, the two authors obtained a percentage of agreement of P(a) = 95% and a Cohen's K of 0.64 –a substantial agreement level (Cohen, 1960). All differences were discussed subsequently. Regarding the single study with a non-randomized experimental methodology (Baschenis et al., 2021), no concerns and/or high risk of bias were highlighted, and the study was judged to be at low risk of bias for all domains (see Figure 4).

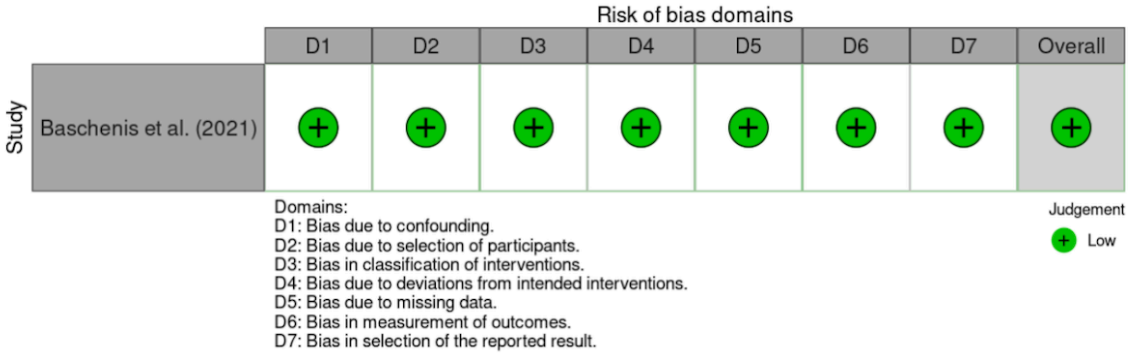


Figure 4. ROBINS-I tool.

The same was true for the three studies evaluated through the NOS-C (Bazoli et al., 2022; Borgonovi & Ferrara, 2023; Contini et al., 2022), that obtained a score of 7 (good score level– see Table 1).

Cohort studies (n=3)	Selection			Comparability		Outcome			Total (0-9) *
	Representativeness of the exposed sample	*Selection of the non-exposed cohort*	*Ascertainment of exposure*	*Demonstration that outcome of interest was not present at start of study*	*Comparability of cohorts on the basis of the design or analysis controlled for confounders*	*Assessment of outcome*	*Was follow-up long enough for outcomes to occur?*	*Adequacy of follow-up of cohorts*	
Bazoli et al. (2022)	1	1	1	1	1	0	1	1	7
Borgonovi & Ferrara (2023)	1	1	1	1	1	0	1	1	7
Contini et al. (2022)	1	1	1	1	1	0	1	1	7

Note: * 9 = Very Good; 7–8 = Good; 5–6 = Satisfactory; 0– 4 = Unsatisfactory.

Table 1. Newcastle-Ottawa scale adapted for Cohort studies (NOS-C).

All these studies presented an unsatisfactory value only in the Outcome section, due to the lack of information about the method of evaluating the results (e.g., independent blind assessment). Only one study (Termine et al., 2022) was evaluated through the NOS-CC, and obtained a total score of 5 (satisfactory – see Table 2).

Case-control studies (n=1)	Selection				Comparability	Exposure			Total (0-9) *
	Is the case definition adequate?	*Representativeness of the cases*	*Selection of controls*	*Definition of controls*	*Comparability of cases and controls on the basis of the design or analysis*	*Ascertainment of exposure*	*Same method of ascertainment for cases and controls?*	*Non-response rate*	
Termine et al. (2022)	0	1	1	1	1	0	1	0	5

Note: * 9 = Very Good; 7–8 = Good; 5–6 = Satisfactory; 0– 4 = Unsatisfactory.

Table 2. Newcastle-Ottawa scale adapted for Case-control studies (NOS-CC).

The unsatisfactory scores are mainly due to the lack of information about the rate of non-responders and the self-report data collection method (survey). 29 studies were evaluated through the NOS-CS (Table 3). Only one study achieved a good level of evaluation (Doz et al., 2023) and only 2 a satisfactory criterion (Bertoletti et al., 2023; Crisci et al., 2021). There is still much debate within the scientific community on the inclusion of studies highly vulnerable to bias within systematic reviews and meta-analyses (Jüni et al., 2001; Wood et al., 2008). We decided to maintain studies with high risk of bias for different reasons. First, assessing the risk of bias does not necessarily imply the possibility that a study is biased, but

only that it is susceptible to bias (Harvey & Dijkers, 2019). Second, we decided to preserve our research by Tipe II error, thus the possibility to exclude informative papers because of a too stringent methodology impossible to adopt in consideration of the pandemic period and the covered topic. Third, it should be noted that these high risks of bias levels are mainly due to the choice of sampling methods (snowball) and the evaluation methods used (surveys or structured interviews). This choice is to be considered binding as the pandemic period in Italy consisted in both the closure of research institutions and the impossibility of researchers to conduct experimental studies involving schools.

Cross-sectional studies (n=29)	Selection				Comparability	Outcome		Total (0-10) *
	Representativeness of the sample	*Sample size*	*Non-respondents*	*Ascertainment of the exposure*	*Adjustment of the outcome*	*Assessment of the outcome*	*Statistical test*	
Bazzoli et al. (2021)	0	0	0	0	0	1	1	2
Benigno et al. (2020)	0	0	0	0	0	0	0	0
Bertoletti et al. (2023)	1	1	0	2	1	0	1	6
Canals-Botines et al. (2021)	0	0	0	0	0	0	0	0
Capperucci et al. (2022)	0	0	0	0	0	0	0	0
Capurso & Roy Boco (2021)	0	0	0	0	0	0	0	0
Champeaux et al. (2022)	0	0	0	0	0	0	1	1
Colombo & Santagati (2022)	0	0	0	0	0	0	0	0
Crescenza et al. (2021)	0	0	0	0	0	0	0	0
Crisci et al. (2021)	1	0	1	2	0	0	1	5
Decarli et al. (2022)	0	0	0	0	0	0	1	1
Doz et al. (2022)	1	1	1	1	1	2	1	8
Ferretti, et al. (2021)	0	0	0	0	0	0	0	0
Gaggi et al. (2020, September)	0	0	0	1	0	0	0	1
Gentile et al. (2021)	0	0	0	0	1	0	1	2
Guzzo et al. (2022)	0	0	0	0	0	0	1	1
Ianes & Bellacicco (2020)	0	0	0	0	0	0	1	1
Inguscio et al. (2023)	0	0	0	2	0	0	1	3
Mangiavacchi et al. (2021)	0	0	0	0	0	0	1	1
Marchese et al. (2022)	0	0	0	0	0	0	0	0
Picca et al. (2021)	0	0	0	0	0	0	1	1
Ranieri et al. (2020)	0	0	0	0	0	0	0	0
Scarpellini et al. (2021)	0	0	0	0	0	0	1	1
Segre et al. (2021)	0	0	0	0	0	0	1	1
Szpunar et al. (2021)	0	0	0	0	0	0	0	0
Termine et al. (2021)	0	0	0	2	0	0	1	3
Thorell et al. (2021)	0	0	0	0	0	0	1	1
Thorell et al. (2022)	0	0	0	2	0	0	1	3
Zaccoletti et al. (2020)	0	0	0	2	1	0	1	4

Note: * 9 = Very Good; 7–8 = Good; 5–6 = Satisfactory; 0– 4 = Unsatisfactory

Table 3. Newcastle-Ottawa scale adapted for Cross-sectional studies (NOS-CS).

2.3. An overview of the effects of the school closure due to the COVID-19 pandemic

In addition to the scholastic achievement, several studies included in our systematic review found that the school closure due to Covid-19 affected Italian primary school students on an emotional level (see, for example, Mangiavacchi et al., 2021), particularly for those who were already on the margins of the educational system prior to the COVID-19 pandemic, due to socio-cultural, socio-economic difficulties, and/or difficulties due to one or more diagnostic conditions (e.g., Crisci et al., 2021). According to the studies included in our work, some students may have encountered difficulties during distance learning, from a technical-practical point of view, such as: difficulties in following the online lesson, technical problems or competence of teachers/students with the new learning methodology (for example, Bazzoli et al., 2021). Hence, we categorized the articles according to the specific dimensions investigated, namely: emotional aspects of the learning context; marginality-inequality effects; didactic problems; learning quality. See Table 4 for a synthesis of the studies included and the investigated dimensions.

Authors	**Investigated dimension(s)**			
	Emotional aspects of the learning context	Marginality-inequality effects	Didactic problems	Learning quality
Baschenis et al., 2021		✓		
Bazoli et al., 2022		✓		✓
Bazzoli et al., 2021		✓	✓	
Benigno et al., 2020		✓		
Bertoletti et al., 2023			✓	✓
Borgonovi & Ferrara, 2023		✓		✓
Canals-Botines et al., 2021			✓	
Capperucci et al., 2022			✓	✓
Capurso & Roy Boco, 2023		✓		
Champeaux et al., 2022	✓			
Colombo & Santagati, 2022		✓		
Contini et al., 2022		✓		✓
Crescenza et al., 2021	✓	✓		
Crisci et al., 2021		✓		
Decarli et al., 2022		✓	✓	✓
Doz et al., 2023	✓		✓	
Ferretti et al., 2021			✓	
Gaggi et al., 2020			✓	
Gentile et al., 2021				✓
Guzzo et al., 2022			✓	
Ianes & Bellacicco, 2020		✓		
Inguscio et al., 2023		✓		
Mangiavacchi et al., 2021	✓			✓
Marchese et al., 2022		✓		

Authors	Investigated dimension(s)			
	Emotional aspects of the learning context	Marginality-inequality effects	Didactic problems	Learning quality
Picca et al., 2021	✓			
Ranieri et al., 2020		✓	✓	
Scarpellini et al., 2021	✓	✓	✓	✓
Segre et al., 2021	✓			
Szpunar et al., 2021	✓			
Termine et al., 2021		✓		
Termine et al., 2022		✓		
Thorell et al., 202	✓	✓		
Thorell et al., 2022		✓		
Zaccoletti et al., 2020	✓			
TOTAL	**10**	**20**	**11**	**9**

Table 4. Synthesis of studies included.

The section called "Emotional aspects of the learning context" includes studies investigating emotional aspects experienced by students in distance learning during the pandemic period. By the general term "emotional" we mean studies that considered both short-term (e.g., fear, joy, sadness, anger) and long-term (e.g., anxiety, frustration, serenity) emotional states. We also include cognitive processes linked to emotional responses, such as attention and motivation to learn (Kim & Pekrun, 2014; Taylor & Fragopanagos, 2005). We divided the studies according to the involved sample's opinion: parents; teachers; students. The study by Crescenza et al. (2021) was the only one that investigated these aspects from the perspectives of both teachers and parents. Therefore, we included the study in both sub-categories. In summary, 2 studies collected information from a sample of teachers, 8 from parents, and 1 from students. We repeated the same distinction also in the other sections.

Within the section "Marginality-inequality effects", we included the papers that have at-risk students within their sample: students with specific learning disorders (SLDs), neurodevelopmental disorders (e.g., autism spectrum disorder, ASD; attention deficit and hyperactivity disorder, ADHD), or neurodiverse students in general. We also included students with SEN, linguistic-cultural difficulties, and participants with low SES. Of the 20 papers included, 7 collected information from a sample of parents, 6 from teachers, and 7 from students. However, Baschenis et al. (2021) and Inguscio et al. (2023) collected data through questionnaires to parents and tested children through psychometric tests; therefore, we placed them in both subsections. Consequently, the section includes 7 papers that collected information from a sample of parents, 6 from teachers, and 9 from students.

The new teaching approach adopted during distance learning has brought difficulties in evaluating student performance, as well as didactic problems encountered by parents or teachers themselves. The category called “Didactic problems” includes studies that considered problems encountered during distance learning, such as the provision of the teaching itself, technical problems, or competence of the teachers/pupils in the use of the new learning tools, issues with the online assessment of students’ achievements. This section includes 11 studies that dealt with these issues in the pandemic period: 2 according to the parents’ opinion; 9 to the teachers’ opinion. No articles evaluated this topic according to students’ opinion.

The category called “Learning quality” includes studies that compared academic performance during the pandemic with pre-pandemic performances. Of the 9 papers included, 3 collected information from a sample of parents, 3 from teachers, and 3 from students.

Therefore, this work includes 10 articles dealing with emotional aspects of the learning context, 20 related to marginality-inequality effects, 11 on didactic problems, 9 on learning quality. Some articles investigated two or more areas of learning; therefore, they were included in each dimension investigated. See Table 5 for sample characteristics and demographic information about the studies included.

ID	Authors	Sample				Italian location(s)
		Type	N	Sex	Age	
					Mean / Sd	
1	Baschenis et al., 2021	P/S	63 parents' children with dyslexia (EG). 38 parents' children without dyslexia (CG)	NA	NA / NA	North (Pavia, Lombardy)
			65 with dyslexia (EG); 52 without dyslexia (CG)	F = 20 with dyslexia	10.64 with dyslexia / 1.60 with dyslexia	
				F = 20 without dyslexia	9.80 without dyslexia / 1.57 without dyslexia	
2	Bazoli et al., 2022	S	'18/'19 samples: 24.781 students for 5th grade; 29.675 8th grade. '20/'21 samples: 16.631 for 5th grade; 9.708 8th grade.	NA	NA / NA	NA
3	Bazzoli et al., 2021	T	3121 (39.4% primary school teachers)	NA	51 / NA	48% North, 22% Center, 30% South*
4	Benigno et al., 2020	T	12	F = 12	NA / NA	8 children's hospitals in Italy (4 located in North, 1 in Center, and 3 in South)
5	Bertoletti et al., 2023	T	1407 (51% primary school teachers)	F= 93%	48.94 / 0.64	52% from North
6	Borgonovi & Ferrara, 2023	S	<1mn (985.909 primary school students)	NA	NA / NA	NA
7	Canals-Botines et al., 2021	T	31 (Italian sample)	(Italian sample) F= 29 M = 2	NA / NA	Center (Lazio)
8	Capperucci et al., 2022	T	11.828 (47.5% primary school teachers)	NA	NA / NA	Center (Tuscany and Umbria)
9	Capurso & Roy Boco, 2023	T	21 (10 primary school teachers)	F= 90%	43.6 / 8.8	NA
10	Champeaux et al., 2022	P	2.455 Italian and 1.838 French families (3.769 Italian and 3.183 French children. The 47.1% of Italian children were primary school students)	NA	Italian children F+M = 8.3 (between 2-18 years old) / Italian Children F+M = 3.5	54.9% North, 25.7% Center, 19.4% South*
11	Colombo & Santagati, 2022	T	145 (53% Italian primary school students)	F = 95% of the entire sample	NA / NA	For the entire sample: 50% North, 15% Center, 35% South
12	Contini et al., 2022	S	1539	NA	NA / NA	North (Torino)

ID	Authors	Sample					Italian location(s)
		Type	N	Sex	Age		
					Mean	Sd	
13	Crescenza et al., 2021	P/T	991 parents (57% parents' primary school students)	F = 826 (83.4%)	NA	NA	North, Center, South (Lombardy, Lazio, Puglia)
			794 teachers (39% primary school teachers)	F = 674 (84.9%)	NA	NA	
14	Crisci et al., 2021	P	637 parents (339 primary school students)	Parents F = 92%	NA	NA	NA
				Children M = 48%	Children F+M = 10.8	Children F+M = 3.24	
15	Decarli et al., 2022	T	120 (59 primary school teachers)	F = 105	49.66	9.5	For the entire sample: 14 schools from North, 10 from Center, 21 from South
16	Doz et al., 2023	T	270 (135 of primary schools)	F = 239	49.5	9.74	NA
17	Ferretti et al., 2021	T	244 (122 primary school teachers)	NA	NA	NA	NA
18	Gaggi et al., 2020	P	490 families	NA	NA	NA	17 out of 20 Italian regions
19	Gentile et al., 2021	P	19.527 families (almost 40% of 32.000 children were from primary schools)	F = 86% (the responding was generally the mother)	30.8% of them were 45-49 years old.	NA	68.9% from Center
20	Guzzo et al., 2022	T	977 (315 primary school teachers)	F = 792 (81%)	Less than 28 (1%) – more than 60 (15%). 50-59 (44%)	NA	48% North, 12% Center, 40% South*
21	Ianes & Bellacicco, 2020	T	3.291 (40,9% primary school students)	NA	NA	NA	For the entire sample: 42.3% North, 15.2% Center, 42.4% South
22	Inguscio et al., 2023	P/S	61 parents	F = 51 (mothers – 83.6%)	F = 44.09 M = 49.66	F = ±6.23 M = ±4.69	Center (Polyclinic Umberto I, Rome)
			61 children	F= 36 (daughters – 59.02%)	F = 12.19 M = 12.33	F= ±2.96 M = ±3.01	
23	Mangiavacchi et al., 2021	P	2101 families (3619 Children under 16, 40.3% primary school students)	Parents F = 93% of respondents were mothers	Parents F = 41 M = 44	NA	53.6% North, 26.7% Center, 19.7% South
				Children M = 50.7%	Children F+M = 7	NA	
24	Marchese et al., 2022	S	33 (15 primary school students)	Primary school students F= 6	Primary school students = 7	NA	South (Barletta)

ID	Authors	Sample				Italian location(s)
		Type	N	Sex	Age	
					Mean / Sd	
25	Picca et al., 2021	P	Parents of 3392 children [1704 (50.2%) Italian primary school students]	Italian primary school students M = 53%	Italian primary school students: between 6 and 10 / NA	NA
26	Ranieri et al., 2020	S/T	820	NA	NA / NA	Different Italian regions
27	Scarpellini et al., 2021	P	F = 1.601 mothers (1.138 mothers of primary school students)	Mothers' primary school students F = All	Mothers' primary school students F = 42,5 / NA	70.2% from North
			Primary school students F+M = 1.148	Primary school students F = 545 (48.2%)	Primary school students F+M = 8 / NA	
28	Segre et al., 2021	S	82 (54.9% primary school students)	F = 46.3%	10.4 / NA	North Italy (Milan)
29	Szpunar et al., 2021	P	5022 parents	Parents F= 4452	Parents = 43.51 / Parents =5.99	For the entire sample: 25.3% North, 64.2% Center, 10.5% South*
			5822 children (38% primary school students)	Children = F= 2830	Children = 9.3 / Children = 4	
30	Termine et al., 2021	S	1362 with Neurodevelopment disorders (NDD – 455 primary school students)	M= 861 with NDD	primary school students = 9 / NA	North (Varese)
			6943 without NDD (3318 primary school students)	M= 3280 without NDD		
31	Termine et al., 2022	S	with tic disorder =49	M= 39	6–18 years / NA	North (Varese)
			Control group = 245	M= 125	NA	
32	Thorell et al., 2021	P	6720 parents (794 from Italy)	NA	NA / NA	NA
33	Thorell et al., 2022	P	1010 from different European Countries	NA	NA / NA	NA
34	Zaccoletti et al., 2020	P	567 (Italian's parents = 173)	F = 89% (Italian's mother)	42,90 (Italian's parents) / 5,76 (Italian's parents)	NA
			Italian primary school students F+M = 72%	F = 46% (Italian's children)	9,65 (Italian's children) / 2,14 (Italian's children)	

Note: P = Parents. S = Students. T = Teachers. M = Male. F = Female. NA = Not available. *The southern area marked with an asterisk includes the major islands.

Table 5.Descriptive synthesis of the included studies.

2.3.1. Parents' opinion.

Fourteen studies have delved into the effects of distance learning during the COVID-19 pandemic from the perspective of parents (Baschenis et al., 2021; Crisci et al., 2021; Mangiavacchi et al., 2021; Champeaux et al., 2022; Crescenza et al., 2021; Gaggi et al., 2020; Gentile et al., 2021; Inguscio et al., 2023; Picca et al., 2021, Scarpellini et al., 2021; Szpunar et al., 2021; Thorell et al., 2021, 2022; Zaccoletti et al., 2020). Online questionnaires were the main method of data collection (see Table 6 for a description of the studies included in this section). Consensus among these studies is that distance learning has posed significant challenges for children. Concerns about the quality of learning during distance education were widespread among parents, with many expressing dissatisfactions with the effectiveness of online instruction (Mangiavacchi et al., 2021; Gentile et al., 2021; Scarpellini et al., 2021). Parents have reported difficulties in supporting their children academically (Crescenza et al., 2021), leading to disruptions in children's emotional well-being, including increased aggression and restlessness (Champeaux et al., 2022; Scarpellini et al., 2021). There has also been a noticeable decrease in attention, concentration, and engagement during home study sessions (Thorell et al., 2021; Zaccoletti et al., 2020). Additionally, concerns have been raised about increased screen time and decreased participation in school and extracurricular activities (Mangiavacchi et al., 2021; Szpunar et al., 2021; Zaccoletti et al., 2020). Despite these challenges, one study found a positive effect on children's sleep and family relationships due to parents' remote work (Picca et al., 2021). Furthermore, several parents reported amplified difficulties faced by marginalized and neurodiverse students during distance learning. Regarding to the marginality-inequality condition, parents of students with special educational needs (SEN) or neurodevelopmental disorders have expressed dissatisfaction with the support received and highlighted the negative impact on their children's academic performance and psychological well-being (Crisci et al., 2021; Thorell et al., 2021, 2022), along with higher levels of fidgetiness (Crescenza et al., 2021). Similarly, parents of dyslexic children have reported significant difficulties in online learning compared to parents of neurotypical children (Baschenis et al., 2021). Inguscio et al. (2023) investigated the perception of online learning by students (and their parents) with or without hearing loss (HL). The presence or absence of HL did not influence the parents' estimation of the difficulties with the distance learning. Over 50% of the parental population defined the distance learning

as quite useful, although with lower scores than the direct relationship (in class) between child-teacher and schoolmates. Furthermore, instructional problems have been noted, with some children unable to participate in distance learning due to technological constraints or perceived inefficacy of remote learning methods (Gaggi et al., 2020; Scarpellini et al., 2021). The investigation conducted by Gaggi et al. (2020) found that 4.5% of the children in their sample (22 out of 490) never participated in distance learning during the pandemic, due to family technological difficulties or school impreparation, or even because some teachers considered it a suboptimal method for teaching young students. More than 70% of the parental sample of Scarpellini et al. (2021) reported that they were not in favor of the learning methods imposed by the distance learning, and many of them (80%) found themselves in great difficulty in giving technological help to children.

Authors	Aim	Assessment tool	Investigated dimension(s)	Results
Baschenis et al., 2021	A cross-sectional comparison of reading skills between children with and without dyslexia evaluated pre- and post-lockdown	For parents, an ad hoc questionnaire. EG was assessed before and after the lockdown; CG only after.	Marginality-inequality effects	None of the CG parents reported difficulties in their children following online lessons, unlike the EG parents, who expressed difficulties in carrying out their homework. Both groups of parents reported worsening scholastically and relationally.
Champeaux et al., 2022	Evaluation of Italian/French parents on children's home-schooling process and emotional well-being	Online questionnaire	Emotional aspects of the learning context	For the Italian sample, there was a loss of learning (M= −5.138; SD=2.614) and a worsening of the emotional well-being (M=−0.670; SD=0.957) of their children.
Crescenza et al., 2021	Compare the opinion of all participants involved in school learning on the distance learning	Online questionnaire	Emotional aspects of the learning context	Parents report having had to apply several times for a work permit to support their children during the distance learning. Increased levels of restlessness among students with disabilities.
			Marginality-inequality effects	Increased levels of restlessness among students with disabilities.
Crisci et al., 2021	Investigate the role of background, child (executive functions), and parental (psychological well-being) factors in the perception of negative/positive effects of distance learning	Online questionnaire, based on: Childhood Executive Functioning Inventory (CHEXI); Positive Mental Health (PMH)	Marginality-inequality effects	The negative effects of distance learning were related to a younger age and a greater EF deficit. Parents' psychological well-being had no significant effect
Gaggi et al., 2020	Evaluate the technologies and methodologies adopted	Online questionnaire	Didactic Problems	Regarding to learning, 22 families (4.5%) reported that the school did not activate distance

Authors	Aim	Assessment tool	Investigated dimension(s)	Results
	during the COVID-19 period by teachers and the impact it has had on students and their families			learning at all, to be understood as any type of activity related to learning, from passive activities, such as sending/checking homework, to the same distance learning.
Gentile et al., 2021	Evaluate the technologies and methodologies adopted during the COVID-19 period	Online questionnaire	Learning Quality	Regarding to distance learning's efficacy, parents' primary school students reported, in a 10-point Likert, an average of 5 point.
Inguscio et al., 2023	Investigate the psychological characteristics of distance learning on Italian students with/out HL	Online questionnaire, based on: Questionnaire on school wellbeing (QBS); State-Trait Anxiety Inventory for adults (STAY-Y); Revised Children's Manifest Scale (RCMAS-2)	Marginality-inequality effects	No significant difference of anxiety between parents, reported as "not particularly problematic" according to the Italian norms.
Mangiavacchi et al., 2021	Investigate the parents' reactions to the lockdown and the role of these changes on children's emotional wellbeing, time with their parents and their educational outcomes	Online questionnaire	Emotional aspects of the learning context	Parents observed a significant reduction in their children's emotional wellbeing.
			Learning Quality	Parents observed a reduction in productive activities (school, homework, extracurricular activities): from around 60% of daily activity to less than 40%. Slight improvement in reading time.
Picca et al., 2021	Investigate the positive/negative effect of COVID-19 on children's behaviors, daily life, distance learning, and the use of digital device	Online questionnaire	Emotional aspects of the learning context	For the Italian primary school students, the improvement of family relationships and parents' remote working positively influenced sleep, emotional and behavioral disturbances; the same were negatively influenced by screen time and time spent watching TV
Scarpellini et al., 2021	Evaluation of mothers on children's home-schooling process and emotional well-being	Online questionnaire	Emotional aspects of the learning context	Mothers observed behavioral changes in their children, specifically restlessness (69.1%) and aggressiveness (33.3%)
			Marginality-inequality effects	Over the 50% of the children with not specified chronic disorders attended distance learning only once a week. However, almost the 15% of them did not receive any specific support.
			Didactic Problems	Although most experienced no difficulties with the technology (80.7%), mothers did not approve distance learning (72.2%).

Authors	Aim	Assessment tool	Investigated dimension(s)	Results
			Learning Quality	11.5% of elementary school students were either not evaluated, didn't obtain any grades (53.1%), or received lower grades compared to their prior academic performance (7.1%). Nevertheless, a majority of primary school students assessed, specifically 80%, received identical grades as they did before the COVID-19 pandemic.
Szpunar et al., 2021	Investigate the changes in students' lifestyle habits due to the pandemic, in particular the school activities.	Online questionnaire	Emotional aspects of the learning context	12.6% of primary school students spend more than 3 hours per day doing homework, more time than daily distance learning. The 56.1% of the same students do homework more than before of the lockdown (20.3% less)
Thorell et al., 2021	Comparison of parenting experiences on homeschooling from different European countries	Online questionnaire	Emotional aspects of the learning context	The results referring to Italy highlight a strong parental concern for the negative effects of the distance learning on their children. Also, not everyone has been able to stick with homeschooling fluently.
			Marginality-inequality effects	49.1% of parents of SEN students reported that the support they received during distance learning was not sufficient.
Thorell et al., 2022	Investigate the parents' perception about the role of ADHD/ASD's deficit or both on the positive/negative effects of distance learning	Online questionnaire	Marginality-inequality effects	Higher levels of negative effects of distance learning in children with ADHD and/or ASD.
Zaccoletti et al., 2020	Examine the opinion of Italian and Portuguese parents about impact of lockdown on students' academic motivation and their participation in extracurricular activities	Online questionnaire, based on AMOS 8-15	Emotional aspects of the learning context	Decreased academic motivation and extracurricular activities in both Countries

Table 6. Parents' opinion.

2.3.2. Teachers' opinion.

Fourteen studies have investigated teachers' opinions on emotional aspects, marginality-inequality effects, instructional issues, and the quality of their students' learning (Doz et al.,

2023; Bertoletti et al., 2023; Bazzoli et al., 2021; Benigno et al., 2020; Canals-Botines et al., 2021; Capperucci et al., 2022; Capurso & Roy Boco, 2023; Colombo & Santagati, 2022; Crescenza et al., 2021; Decarli et al., 2022; Ferretti et al., 2021; Guzzo et al., 2022; Ianes & Bellacicco, 2020; Ranieri et al., 2020). Only two studies investigated emotional status of students according to their teachers (Doz et al., 2023; Crescenza et al., 2021). The teachers' sample by Doz et al. (2023) observed a decline in students' attention and motivation, particularly among those lacking parental support. Over 60% of the surveyed teachers by Crescenza et al. (2021) acknowledged the importance of prioritizing the emotional well-being of students and families. However, they also expressed concerns about their perceived lack of proficiency in fulfilling this role. It is indisputable that the pandemic has had disruptive effects on student populations who already presented a difficult clinical-family picture; within the school context there was inevitably a loss or a slowdown in learning (Bazzoli et al., 2021; Benigno et al., 2020). Teachers reported that their students with special educational needs (SEN) have faced significant challenges, including changes in relationships, decreased autonomy, and passive participation (Capurso & Roy Boco, 2023; Decarli et al., 2022), or even no online lesson activities for these students, given the difficulties in delivering their personalized teaching plans online (Colombo & Santagati, 2022; Ianes & Bellacicco, 2020). While some teachers recognized the benefits of distance learning, such as innovative teaching methods, many encountered difficulties in building relationships with students and adapting to technological limitations (Bertoletti et al., 2023; Guzzo et al., 2022). As regards didactic problems, digital competence has emerged as a critical issue, with widespread resistance to digital methods among older teachers (Doz et al., 2023; Bertoletti et al., 2023; Canals-Botines et al., 2021; Decarli et al., 2022). This resistance has negatively influenced student performance, underscoring the importance of ongoing training and support for teachers (Guzzo et al., 2022). Additionally, assessing students' academic performance in a remote environment has posed significant challenges, necessitating adaptation of assessment methods. Capperucci et al. (2022) reported that the most common methods in primary school, during the pandemic period, were homework (78.7%), tests (63.3%), oral questions (62.0%), unspecified exercises (59.0%). Over 60% of teachers who responded to the survey by Ferretti et al. (2021) did not believe assessment was feasible in distance learning environments. See Table 7 for a description of the studies on teachers' opinion.

Authors	Aim	Assessment tool	Investigated dimension(s)	Results
Bazzoli et al., 2021	Teachers' perspective on distance learning's criticalities and territorial differences during the pandemic	Online questionnaire	Marginality-inequality effects	With greater difficulties in southern Italy, the teachers reported that: 7.7% of the students did not participate in distance learning; 10.3% with irregular participation; 21.8%, although participating regularly, had learning difficulties.
			Didactic Problems	58.1% of teachers expressed difficulties in assessing learning during the distance learning.
Benigno et al., 2020	Evaluate the impact of the pandemic on the italian SiHo services	Group interview	Marginality-inequality effects	Hospital teachers report the same difficulties as traditional schools, both from a relational point of view with the students and from an organization of the distance learning
Bertoletti et al., 2023	Identify subgroups of teachers based on the use of digital tools and how these subgroups differ in terms of satisfaction and students' performance.	Online questionnaire + students' standardized test scores (INVALSI)	Didactic Problems	A third of teachers proved resistant to digital methods, using available technologies in a rather limited and resistant way.
			Learning Quality	The teachers who had greater digital skills were more satisfied with their teaching, positively also associated with better student performance.
Canals-Botines et al., 2021	Spanish, Italian, Romanian teachers' experiences and perspective on distance learning before, during, and after COVID-19	Online questionnaire	Didactic Problems	Teachers were somewhat prepared to distance learning, although they had struggled to track children's progress and maintain grounding during school closures.
Capperucci et al., 2022	Comparison of national data with data from Tuscany and Umbria on teaching and assessment strategies used during the distance learning	Online questionnaire	Didactic Problems	Teachers who applied alternative methods of assessment provided a higher rating of the students' learning
			Learning Quality	Regarding the effectiveness of the distance learning for students learning, the average value of the Tuscan sample was lower than the national average, while the Umbrian sample was in line with that.
Capurso & Roy Boco, 2023	Evaluate the changes in roles, relationships and activities caused by school closures in SEN students, according to their teachers.	Semi-structured interviews (56min on average)	Marginality-inequality effects	Several teachers reported that, during the lockdown, SEN students lost the opportunity to relate to the entire class group, often stuck in a dual situation with the support teacher
Colombo & Santagati, 2022	Teachers' experiences and opinions about the strategies used during lockdown with students with disabilities (SD). Furthermore, their emotions on their work, and the relationship with students' parents	Online questionnaire	Marginality-inequality effects	Increased of inadequacy and stress of the majority of teachers (65%, 62% respectively). Most of them reported the difficult to work alone at home (53%). According to the survey responds, 27% of primary school teachers did no activity with SD in their class. Over 40% of teachers did not maintain direct contact with the parents of the pupils.
Crescenza et al., 2021	Compare the opinion of all participants involved in school learning on the distance learning	Online questionnaire	Emotional aspects of the learning context	Most teachers report the need to give more voice to students' emotions.

Authors	Aim	Assessment tool	Investigated dimension(s)	Results
Decarli et al., 2022	Evaluate the teachers' experience during distance learning and which type of strategies they used with SEN students	Online questionnaire	Marginality-inequality effects	Regarding SEN students, less involvement and greater attention difficulties were highlighted on their part, although teachers reported simplifying the lesson and/or individualizing it for their needs
			Didactic Problems	Most of the teachers reported feeling comfortable and sufficiently proficient with the technologies used during the distance learning.
Doz et al., 2023	Investigate the impact of the distance learning on teachers and their views on students' learning difficulties	Online questionnaire	Learning Quality	They highlighted a disagreement that distance learning helped students learn more efficiently.
			Emotional aspects of the learning context	The teachers reported students' problems in social relationships (33.3%), difficulty in respecting independence (17.4%), motivation (14, 4%), attention (11.8%), and problems with adherence routines (7%). For primary school, a significantly higher number of students relational difficulties (66%).
			Didactic Problems	The teachers reported students' problems in the use of technologies (53.7%), lack of support from families (7%) and problems with adherence routines (7%).
Ferretti et al., 2021	Evaluate teachers' relationship to new distance learning practices, with a main focus on remote evaluation methods	Online questionnaire	Didactic Problems	Most teachers have adapted to the new teaching methodologies; however, they consider that it is not possible to adequately evaluate in a remote way.
Guzzo et al., 2022	Evaluation of teachers' distance learning experience	Online questionnaire	Didactic Problems	A substantial difficulty in using the new method of delivering lessons was highlighted, trying to involve students as much as possible also through SMS and/or WhatsApp
Ianes & Bellacicco, 2020	One month after the pandemic, evaluate the activation of distance learning for all students and the inclusion of students with disabilities.	Online questionnaire	Marginality-inequality effects	One month after the pandemic, 11.20% of primary school teachers report that they have not activated the distance learning in their class or in some, especially in the south (10.00%). Still regarding primary school, 24.60% of students with disabilities were excluded from the distance learning due to its ineffectiveness, while 9.90% were excluded due to the objectives of the PEI, for which didactic interventions not in presence.
Ranieri et al., 2020	Explore teachers' response to distance learning challenges	Online questionnaire	Didactic Problems	Teachers used different methodologies to allow students to be assessed during the pandemic period, including observation of attitudes, and teamwork.

Table 7. Teachers' opinion.

2.3.3. Students' opinion.

As highlighted in the preceding sections, students have faced various emotional challenges and learning difficulties due to school closure during the pandemic period. This section provides a summary of the results of 10 studies that investigated students' own opinions (Baschenis et al., 2021; Bazoli et al., 2022; Borgonovi & Ferrara, 2023; Contini et al., 2022; Termine et al., 2021, 2022; Inguscio et al., 2023; Marchese et al., 2022; Ranieri et al., 2020; Segre et al., 2021). See Table 8 for a description of every study included in this section.

Authors	Aim	Assessment tool	Investigated dimension(s)	Results
Baschenis et al., 2021	A cross-sectional comparison of reading skills between children with and without dyslexia evaluated pre- and post-lockdown	Neuropsychological tests (DDE-2; MT-3-Clinic); ad hoc questionnaire (for children and parents). EG was assessed before and after the lockdown; CG only after.	Marginality-inequality effects	Progressive worsening of reading skills (speed and accuracy) and increase in errors of text, words, non-words. EG's 59%-63% did not reach the expected improvement in reading speed
Bazoli et al., 2022	Investigate the learning loss in math and reading achievement and assess whether educational inequalities are related to the social position of students' families, their geographical area of residence, their migrant status and their educational background	INVALSI ('18-'19 – '20-'21)	Marginality-inequality effects	There is no evidence that COVID-19 worsened the achievements of students in the lower SES more than those of upper-middle SES.
			Learning Quality	Except for reading among 5th grade students, the pandemic generated learning losses among the attendants of every school grade in both participants tested by INVALSI.
Borgonovi & Ferrara, 2023	Investigate the difference in math and reading achievement of students in '20-'21 (COVID-19) and '18-'19 (non-COVID-19)	INVALSI ('18-'19 – '20-'21) + data from the 2017–2018 and 2015–2016 tests	Marginality-inequality effects	Primary school students belonging to the lowest ESCS quartile (first quartile) achieved better reading scores and lower loss in mathematics in the COVID-19 compared to the non-COVID-19 sample. Those belonging to the second quartile reported the lowest reading gain and the greatest information loss in mathematics.
			Learning Quality	For primary school students, they showed a smaller reduction in math achievement and an increase

Authors	Aim	Assessment tool	Investigated dimension(s)	Results
				in reading achievement in the COVID-19 cohort
Contini et al., 2022	Estimate the effect of school closure on math skills in a primary school sample	INVALSI data: they compared the standardized tests in math and Italian at the end of grade 2 (pre-test) with a standardized assessment administered by the researchers at the end of grade 3 (post-test).	Marginality-inequality effects	The pandemic exacerbated the pre-existing disparities among various socio-economic groups.
			Learning Quality	Average loss of 0.19 SD in test scores
Inguscio et al., 2023	Investigate the psychological characteristics of distance learning on Italian students with/out HL	Online questionnaire, based on: Questionnaire on school wellbeing (QBS); State-Trait Anxiety Inventory for adults (STAY-Y); Revised Children's Manifest Scale (RCMAS-2)	Marginality-inequality effects	The HL factors did not affect the school wellbeing. They did not find significant difference between children with/out HL
Marchese et al., 2022	Investigate if 1:1 distance learning with a special education teacher was effective in helping SEN children during the pandemic	Online questionnaire	Marginality-inequality effects	Although with difficulties in pursuing and achieving outcomes, all SEN children who followed the platform 1:1 achieved the expected outcomes for their deficit.
Ranieri et al., 2020	Explore teachers' response to distance learning challenges	Online questionnaire	Marginality-inequality effects	Over 50% of students did not have proper technological access, due to the family unit not having the necessary devices. Furthermore, the lack of technical (44%) or linguistic skills (33%) on the part of the parents had a particular impact.
Segre et al., 2021	Investigate the students' psychological distress and changes in routine	Video-interviews	Emotional aspects of the learning context	79.3% of the sample reported difficulty adapting to distance learning, perceiving greater learning difficulty at home. Changes in eating and sleeping habits were also highlighted.
Termine et al., 2021	Evaluate the impact of social distancing and lifestyle changes on students with/out NDDs during COVID-19	Online questionnaire	Marginality-inequality effects	Major disadvantage in distance learning for children from low SES families. Worse learning outcomes during distance learning by NDD students
Termine et al., 2022	Investigate the impact of the pandemic on patients with tic disorders.	Online questionnaire	Marginality-inequality effects	The presence of tics was associated with greater difficulty with distance learning

Table 8. Students' opinion.

Segre et al. (2021) conducted interviews revealing increased levels of anxiety (over 75%) and concerns about family members contracting COVID-19 among almost 80% of students in their sample. Dyslexic students faced heightened difficulties with online lessons and homework compared to their peers without specific learning disorders (Baschenis et al., 2021), although such difficulties in following online lessons were not highlighted by the sample of SLDs students investigated in Marchese et al. (2022). However, their sample (21 students from a single primary school, in southern Italy) cannot be considered representative. Inguscio et al. (2023) did not find a significant impact of hearing impairment on academic well-being, while Contini et al. (2022) and Ranieri et al. (2020) highlighted exacerbated educational disparities, particularly among socioeconomically disadvantaged students. The pandemic has worsened existing disparities among students, especially those from lower socioeconomic backgrounds, as Borgonovi & Ferrara (2023) demonstrated a non-linear relationship between the pandemic's effects on learning and Economic, Social, and Cultural Status (ESCS – Campodifiori et al., 2010), with medium-low SES students experiencing the greatest setbacks. Termine et al. (2021, 2022) also noted significant challenges for students with neurodevelopmental disorders (NDDs), attributing the difficulties to the inadequacy of distance learning for their individualized educational needs. Termine et al. (2022) collected information on a sample of 49 patients between 6 and 18 years old affected by tic disorders (common hyperkinetic movement disorder in childhood). More than 50% of their sample reported an increase in the severity of tics during the pandemic, with an increase in restlessness in 30% of the participants and in perceived pain in 21% of them. According to the regression analyses carried out by Termine et al. (2022), the persistence and increase of tics was directly associated with a learning difficulty in the distance learning modality. In 2021, Termine et al. (2021) conducted a study also in a hospital setting, but with a larger sample size, including patients with tic disorders (n = 46) and not only: 781 with SLD, 125 with ADHD, 42 with ASD. The analyses carried out by Termine et al. (2021) reported a strongly negative effects of distance learning condition in students with Neurodevelopmental Disorders (NDD), adapting less than the control group at this new condition for several reason: distance learning difficulties were more pronounced for children with NDD; students with NDD felt more often reprimanded by their parents during the COVID-19 lockdown; they tended to feel less the absence of their peers, indicating a negative impact on the socialization.

It is interesting to note that no studies have been found that highlight whether Italian primary school students experienced difficulties in using the new distance learning methodology, or any technical didactic problems (e.g.: lack of adequate devices, quality of connection Internet). Regarding the quality of learning, three studies used INVALSI data, revealing adverse effects on academic performance, particularly in mathematics, among primary school students during the pandemic. Borgonovi and Ferrara (2023) and Bazoli et al. (2022) observed a decline in mathematical skills, while reading scores varied across regions. In fact, the results by Borgonovi and Ferrara (2023) were in line with those by Bazoli et al. (2022), except for the distribution of learning loss on the Italian territory. While Bazoli et al. (2022) divided the area of residence into regions, identifying a territorial homogeneity with respect to learning loss, Borgonovi & Ferrara (2023) divided the area into Italian provinces (thus a smaller division compared to Bazoli et al., 2022), identifying a somewhat heterogeneous condition. Regarding mathematics, there was more negative results in the provinces further north-west, and in the regions further south (Puglia and Basilicata, and a part of Sicily) compared to the entire national territory. For reading, the areas most affected by learning loss were the north and central north of Italy and the islands. Contini et al. (2022) also noted greater losses for students with pre-existing academic difficulties, low socioeconomic status, and female students, especially those with highly educated parents.

2.4.Discussion

The COVID-19 pandemic has had a considerable impact on learning, generating changes in many psychosocial characteristics of the learning context experienced by children. This systematic review would like to offer an overview of the effects of COVID-19 school closures on Italian primary school students. We did not focus exclusively on the loss in terms of academic performance, but also on the fundamental role that the learning context may have had on the emotions of pupils, according to them and their caregivers. Overall, the findings indicate that younger students struggled with learning outcomes during the pandemic, with a consensus among parents and teachers about the negative impact of distance learning on effective learning, especially in subjects such as mathematics, as demonstrated by the INVALSI measurements (Bazoli et al., 2022; Borgonovi & Ferrara, 2023). The drastic change in the learning setting was above all studded with technical-practical difficulties, which

contributed to increasing inequalities between students, disrupting the learning experience of students who were already considered at risk of marginalization. These difficulties include limited access to devices and reliable internet connection, lack of technology skills among students and teachers, and difficulty managing distance learning effectively.

To our knowledge, this is the first work that considers the opinion of all the relevant figures in school learning with a view that covers not only the individual academic performance, but also the emotional aspects, technological difficulties, and personal problems of Italian primary school students. We consider the identification of these aspects relevant because of several reasons: (1) the emblematic territory (Italy experienced one of the longest periods of school closure due to the pandemic) and sample (primary school students experienced the first years of education in an unprecedented way for the Italian territory); (2) the possibility of future disruptions to the regular school schedule for other exceptional events (e.g., the recent flooding in Emilia-Romagna); (3) the potential contribution to a deeper understanding of the nuanced impact of distance learning on primary school education; (4) the understanding of any consequences of distance learning not only from an observational point of view, but also to implement improvements on a national level, developing more targeted interventions and support systems for those students who may have struggled the most in this new learning environment. Also, by including studies that explore the opinion of teachers, parents and students themselves, this systematic review constitutes the opportunity to observe the same phenomenon by different perspectives which is a rarity in the field.

Our results highlight several similarities and differences among the perspectives of parents, teachers, and students regarding the effects of COVID-19 school closures on students' learning and well-being. A key similarity concerns the negative impact on learning and emotional well-being during the transition to remote learning. All three groups acknowledged a decline in students' attention, motivation, and engagement during online classes, as well as an increase in anxiety levels and concerns. Additionally, both parents and teachers highlighted the significant challenge in providing necessary support to students during remote learning, especially for students with special educational needs or neurodevelopmental disorders. This suggests an agreement among the perceptions of different stakeholders regarding the difficulties encountered by students during the pandemic. Such disadvantageous condition is not identified in students with hearing loss (Inguscio et al., 2023), neither according to the opinion of the students themselves nor their parents.

However, there are also differences in perspectives related to the main problems due to COVID-19. For example, parents expressed concerns about excessive screen time and decreased participation in extracurricular activities, while teachers mainly focused on the need to improve digital competence and the challenge of adapting assessment methods. Moreover, teachers acknowledged some advantages of remote learning, such as innovation in teaching methods, which parents seem not have directly noticed.

As previously mentioned, no study investigated potential difficulties among students in approaching the new instructional methodology. One possible explanation could be attributed to their adaptability to technology. It is plausible that these students, who are typically exposed to various digital devices and platforms from an early age, may have a propensity to navigate and engage with online learning tools. The early exposure to technology may have facilitated a smoother transition to distance learning, potentially reducing the perceived challenges or obstacles encountered during the adaptation process (Onofrei & Iancu, 2015). However, further research is needed about this point.

Students also highlighted specific difficulties related to their learning disorder and emphasized the negative impact on learning quality, thus supporting the conclusions of the other two groups (parents and teachers). It is interesting to note how both parents and teachers considered support for students crucial in this new mode of learning, although each stakeholder pointed to the other: parents were resentful of the teachers' lack of attention towards their pupils, especially those with previous difficulties; teachers were dissatisfied with the lack of attention and support from parents. Concerning the lack of support from some teachers, it is conceivable as expressed by some studies considered (Bertoletti et al., 2023; Decarli et al., 2022): teachers with more years of professional experience, accustomed to certain learning methodologies, encountered greater difficulties in reinventing themselves, especially towards technological novelty. The reported lack of parental support by teachers could be related with the level of education highlighted by Contini et al. (2022), which suggest that students who have been most affected by COVID-19 school closures are those with highly educated parents. It is possible to postulate retrospectively that, the higher the level of education attained, the greater the professional responsibilities achieved by parents, and therefore the greater the workload, reducing the daily available time and hindering careful instructional support for their children. However, it is important to note that the relationship between parents' level of education and their involvement in their children's learning is

complex and can vary depending on a range of factors, including cultural, socioeconomic, and individual context (Engzell et al., 2021; Davis-Kean, 2005).

Finally, we examined eventual differences between the effects of school closures on Italian primary school students and the international literature on the topic. This inquiry was thoroughly conducted through the analysis of the opinions of different populations (parents, teachers, students), while maintaining a focus on students' learning and well-being. This comparison could provide valuable insights into the specificities of the Italian experience and its relevance in the global context, thus contributing to delineating targeted strategies and interventions to address the effects of the pandemic on students' education.

Concerning the emotional aspects, our findings are in line with research conducted overseas. In their international systematic review, Viner et al. (2022) analyzed the association between school closure and the well-being of K19 students. Compared to primary school students, school closure and the obligation to stay at home drastically reduced healthy extracurricular activities, such as sports, increasing sedentary habits (more time spent in front of the screen). Furthermore, a relationship between school closure and an increase in emotional and behavioral difficulties is once again confirmed. Our results are also in line with a recent systematic review of Italian K12 students during school closure (Bonvino et al., 2023). Their results confirmed an increase in emotional dysfunction demonstrated by an increase in anxiety, anger, difficulty concentrating, and even sleep disturbances. However, it is interesting that Picca et al. (2021) was the only study that found a positive effect of the pandemic period, precisely on children's sleep. Further research on this matter is necessary. Still, Panchal et al. (2023) carried out an intercontinental systematic review on the mental health of children and adolescents (≤ 19 years) exposed to the pandemic lockdown. Like them, our national data indicate that lack of adequate extracurricular activities and increased time spent in front of electronic tools negatively impacted children's emotional health and psychological well-being, increasing perceptions of negative emotions associated with school closures for COVID-19 (feeling of loneliness, anger, irritability, fear, and stress). Furthermore, the authors also identified a higher risk of mental health problems and increased stress for students with special educational needs, although this also depended on space availability (e.g., having a garden or adequate space for studying at home) and family socio economic status. Further international research highlighted how students already on the margins of school community suffered scholastically (Cinquin et al., 2019; Grewenig et al.,

2021) and emotionally because of the school closure more than their peers (Banerjee et al., 2021; Scheer & Laubenstein, 2021). Even in this area, our results do not differ from international literature, although we highlighted that the condition differed based on the difficulty experienced by the students (e.g., students with special education needs vs. low socio-economic status vs. linguistic-cultural differences).

We did not find international reviews that specifically addressed technical-practical educational issues, such as the knowledge and competence of teachers in adapting to new teaching methodologies or the availability and quality of technological devices in the context of distance learning. However, a study conducted by Engzell et al. (2021) sheds some light on this matter. This study reported that Dutch schools were closed for "only" eight weeks during a period of pandemic-related disruptions, without experiencing complications in the continuation of education. One of the factors that contributed to the maintenance of an adequate education was the excellent quality of broadband in the country. As regards digital preparation of teachers in this new teaching method, our results are aligned with a survey conducted in Germany by König et al. (2020), about the importance of teachers possessing digital competence in achieving meaningful educational objectives. In summary, while technical-practical issues were not extensively explored in secondary studies, the primary evidence, and data available do underscore the significance of factors such as broadband quality and teachers' digital skills in addressing the challenges of distance learning effectively.

As regards academic achievement, our review showed a learning loss in mathematical skills, in particular for females and students with learning difficulties, and a slight improvement in reading skills (Hammerstein et al., 2021; König & Frey, 2022; Betthäuser et al., 2023; Di Pietro, 2023). A comparison with international review is needed due to the several differences and similarities between the Italian context and other countries. Contrary to Italian context, the international study by Betthäuser et al. (2023) did not find an effects across age (no difference between primary and older students in learning loss). Similarly to our results, they found a greater learning loss in math than reading, although we highlighted a slight improvement in reading skills of Italian primary students. Our data are in line with the Danish study by Birkelund and Karlson (2023), that found a leaning gain in reading performance of pupils in Grades 2 and 4, and a leaning loss in the same topic between older students (grade 8). It is important to note they lived a different time range of school closure: only eight weeks

for Danish primary school students, and 22 weeks for Danish Grade 8 students. Hence, it is conceivable that prolonged school closures may have a detrimental effect on children's learning (see, for example, Schult et al., 2022), raising long-term concerns for countries that experienced prolonged school closures, such as Italy. Our data are also in line with the US study by Kuhfeld, Lewis and Peltier (2023), where they highlighted a larger decline for primary school students compared to secondary school students, but not in the same school subjects as the ones that were analyzed in the current review: while Italian studies highlighted greater learning loss in mathematics and a slight improvement in reading, they found similar pre- and post-pandemic results for mathematics, and a greater loss in reading among younger students. Furthermore, our study highlighted a worsening in existing disparities among students, especially those with learning disabilities or special education needs (SEN), and from lower socioeconomic backgrounds. These results are aligned with several international studies. As regards students with previous scholastic difficulties, two US studies (Fuchs et al., 2023; Chatzoglou et al., 2023) highlighted that learning loss in reading and mathematics was substantially greater for students with prior difficulties than for the general population. As previously reported, it is conceivable that the structured education provided by the schools that these students require could not be adequately delivered in the new learning modality. This once again underlines the need to identify new, more inclusive learning methodologies. Compared to students with different family backgrounds, heterogeneity is highlighted at an international level. While the study by Kuhfeld, Lewis and Peltier (2023), in line with our findings, highlighted considerable differences between ethnic groups, with lower performance among students with high socioeconomic poverty, not all studies found an association between learning loss and socioeconomic status (Birkelund & Karlson, 2023). We might assume that this difference is due to the concept of "socio-economic status" varies considerably from country to country (in this case: USA vs. Denmark). This may mean that measures used to assess the socioeconomic level of study participants may not be perfectly comparable across different nations or regions. Furthermore, country-specific cultural, political, and economic factors could significantly influence how socioeconomic level affects student learning losses. Furthermore, it should be remembered that the sample of the Danish primary school of Birkelund and Karlson (2023) experienced a school closure of only 8 weeks, while the study of Kuhfeld, Lewis and Peltier (2023) referred to US students, where the school closure lasted several months. This would further confirm the hypothesis that the

rapid return to school reduced the risk of learning losses for students with previous difficulties, also in relation to the possible lack of adequate devices to follow distance learning (e.g., electronic devices; good quality or total absence of Wi-fi) within the family unit. The comparison between Italian and international studies highlighted an extremely heterogeneous picture, with numerous differences and similarities that showed a harmful effect of school closures due to COVID-19 based on different factors: territory; duration of school closure; pre-pandemic disadvantaged conditions. Further research on the topic is needed to highlight long-term post-pandemic effects on the quality of learning of primary school students, not only Italians.

3. Study 2: investigating the interplay between cognitive, metacognitive, and emotional variables in L2 learning overtime.

This chapter refers to the study of Palladino et al. (under review). Here below a list of tables and figure included in this chapter.

List of tables included in Study 2.

Table 9	**Pearson–Correlations between the various measures at T0.**	58
Table 10	**Pearson–Correlations between the various measures at T1.**	58
Table 11	**Pearson–Correlations between the various measures at T2.**	58
Table 12	**Pearson–Correlations between the various measures at T3.**	59
Table 13	**Pearson–Correlations between T0 with T1, T2, and T3.**	60
Table 14	**Pearson–Correlations between T1 with T2, and T3.**	61
Table 15	**Pearson–Correlations between T2 and T3.**	61
Table 16	**Parameter estimates of growth curve analysis.**	63
Table 17	**Multilevel parameter estimates for the MLM on English as L2.**	66

List of figures included in Study 2.

Figure 5	**Examples from the AEQ-ES.**	51
Figure 6	**Aggregate Growth Curve Analysis T0-T3.**	62
Figure 7	**Individual Growth Curve Analysis T0-T3.**	63

The current thesis therefore aims to investigate cognitive, achievement emotions and metacognition, in primary school children learning English as L2. The goal represents an original aim that may significantly contribute to better understanding the complexity of the learning process during early stages for several reasons. Besides the possibility to communicate with people from other countries and having insight in their cultures, or having more career chances, studying a L2 has been shown to improve cognitive abilities as well as influencing positively academic achievement in general, to laying the foundations of language skills, particularly phonology, and to generating a learning advantage, especially when L2

learning begins in primary school (Hakuta et al., 2003; Stewart, 2005). Despite the recognition of the great opportunity in starting it young, many studies on L2 learning at school focused on college and university students (e.g., Sparks & Ganschow, 2001), and only a few included children from the first steps of the learning process (e.g., Berguno & Bowler, 2004, for primary school students). Although the literature indicates that there is an extensive base of research on L2 learning and its relationship with cognitive, metacognitive, and emotional components, most of the studies examined a single variable within a specific research field (but isolated metacognition, emotions, the cognitive basis of learning, etc.,). We aimed to investigate, with a longitudinal design, cognition, metacognition, as well as achievement emotions related to learning English as L2 in primary school children to understand their role in L2 vocabulary acquisition.

3.1.Method

3.1.1. Sample and procedure

The present study included a total of 311 children (mother tongue is Italian or adequately fluent in Italian). Six children were excluded from the analyses because of teacher-reported psychological, behavioral, or neurological disorders. The final sample consisted of a total of 305 children aged between 6 and 9 years (M_{age} = 7.44 years; *SD* = 0.59; 148 females, 48.5%) attending the 2nd and 3rd grades (Academic year 2021/2022) of different public primary schools located in different regions of northern and southern Italy. Participation in the study was strictly dependent on oral and written informed consent from each student's parents and/or legal caretakers, as mandated by the APA's ethical principles. Participation was dependent on written informed consent from each student's parents and/or legal caretakers. All children had to be either native Italian speakers or sufficiently fluent in Italian. We involved in the activities all participants who had neurological, psychological, or behavioral problems and who had parental consent. However, these students were excluded from the data analysis. Children were assessed twice, always during school time, in two different days: in their own class for the collective tests, and, in a quiet room, for the individual testing.

Students were assessed: in November 2021 (T0); 6 months later, in May 2022 (T1); in November 2022 (T2). Students who were absent during one or both data collections were called a second time, so that the smallest number of students were excluded from the study. Both tests (collective and individual) had an average duration of 30 minutes each. The studies were carried out according to the ethical guidelines of the Italian Association of Psychology (AIP) and approved by the Ethics Committee of the Psychology section of University of Padova (prot. n. 34/19).

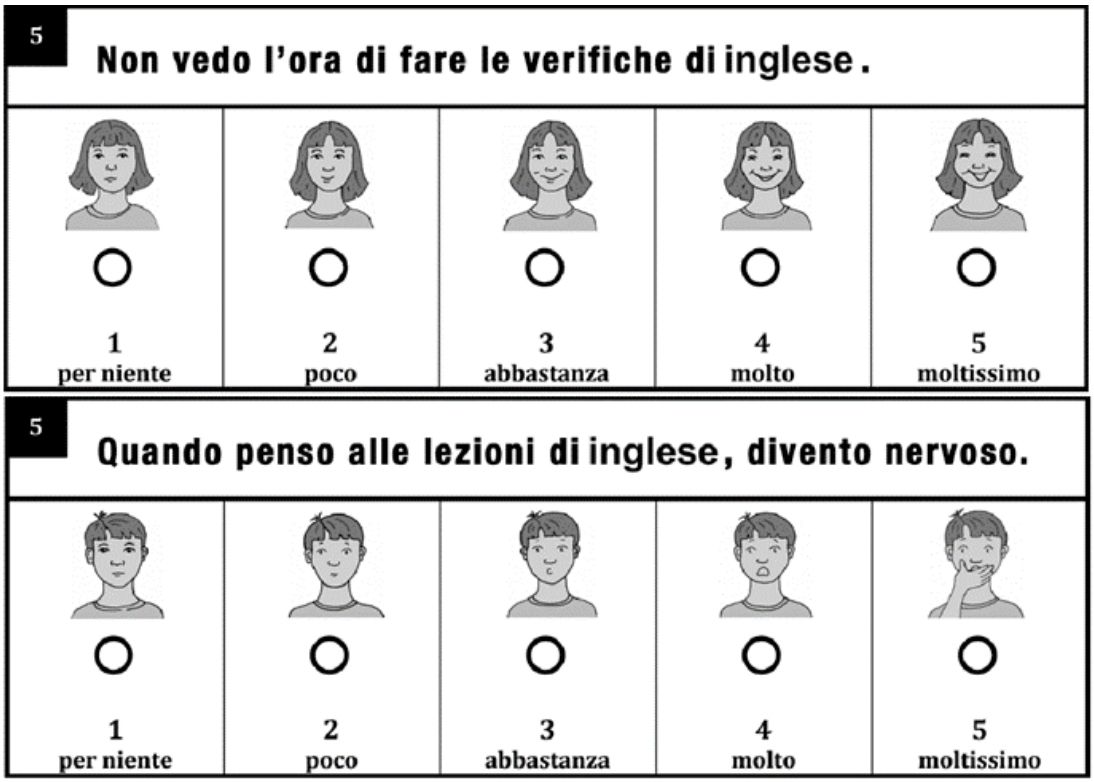
3.1.2. Measures

As reported, we administrated several tests, both collectively and individually. Below we explain the tests used for each of these administration conditions.

3.1.2.1. Collective tests

- Achievement emotions related to L2

We evaluated the achievement emotions perceived by students in second language learning using the Achievement Emotion Questionnaire (AEQ - Pekrun et al., 2011) in the primary school form (AEQ-ES - Lichtenfeld et al., 2012). We adapted the Italian version by Racanello et al. (2019), used to evaluate the achievement emotions perceived in math, for our main topic: English as a second language. The AEQ-ES consisted of a 5-point Likert scale from 1 (not at all) to 5 (very much). The aim of the questionnaire was to evaluate three achievement emotions (joy, boredom, anxiety) experienced during three different contexts of second language learning: lessons, homework, tests. Each response was accompanied by faces (adapted to gender) representing the intensity of the emotional experience (see Figure 5 for an example). In contrast to the other tests mentioned, the AEQ-ES was administered collectively in class. The Italian version (Racanello et al., 2019) revealed a reliability of each factor between .61 and .93 in terms of Omega values. We modified only the terms related to the mathematics domain with the English language domain, maintaining the original structure of the items. Following the experimentation, we found an internal reliability of $\alpha = .78$.



Note: From top to bottom: "I can't wait to do the English tests" (joy, question n.5 about tests in class, for girls); "when I think about English lessons, I get nervous" (anxiety, question n.5 about lessons, for boys).

Figure 5.Examples from the AEQ-ES.

3.1.2.2. Individual tests

- Immediate Non-Words Recall (INWR)

We used a list of ten disyllabic non-words as the first test to evaluate the memory span of the students (Marinelli et al., 2016). The list of terms was the same for 2nd and 3rd grade students. The researcher introduced them to the list of non-words, asking them to pay as much attention as possible, because they had to repeat as many terms as they could remember. The exercise was composed of five alternating repetitions: between a repetition of non-words and another by the researcher, there was a re-enactment by the student. The score was calculated adding the number of non-words correctly pronounced at each repetition.

- Digit Span forward and backward

We evaluated the verbal short-term memory of students using the Digit Span tests (Orsini et al., 2012), both forward and backward versions. Each test involved 21 numerical sequences,

ranging from a minimum of 3 numbers (for forward recall) and 2 numbers (for backward recall) to a maximum of 9 numbers (for forward recall) and 8 numbers (for backward recall), with zero excluded. These sequences were organized by length and mnemonic complexity. We used the same sequences for both 2nd and 3rd grade students. In the forward version, after the researcher pronounced a sequence, the student had to repeat it verbally. In the backward version, the student had to pronounce the sequence in reverse order (e.g., if the researcher said "3-8-6", the student would respond with "6-8-3"). The test was interrupted after three consecutive incorrect answers in the same sequence. However, if the student made one or two consecutive mistakes but correctly recalled the next sequence, the test continued.

- Delayed Non-Word Recall (DNWR)

In addition to assessing short-term memory, the Digit Span Tests served as an interval during which participants were engaged in other activities to prevent continuous rehearsal of non-words. Following this, we assessed each student's long-term memory (specifically, episodic memory) through the Delayed non-word recall (DNWR). During this test, we asked the student to recall all the non-words he/she could remember from the first exercise (Immediate non-words recall). There was no time limit for completing the test; we stopped whenever the student indicated that he/she did not remember any more non-words. We assigned 1 point for each correctly recalled non-word, without counting repetitions of the same non-word (e.g., if a participant repeated the target non-word "*FIPO*" twice, he/she received only 1 point).

- L2 performance

To evaluate the L2 performance (specifically, English), we introduced students to three different tools, administered individually: Vocabulary Entry; Vocabulary Output; Sentence Building. The proves had different items for 2nd and 3rd grade, created according to National indications for the school grade curriculum. For the Vocabulary Entry, we introduced them to six stimuli: school supplies for 2nd grade, clothing for 3rd grade. The researcher asked the student which one of those stimuli was called with a specific term in English, and the student had to point to the corresponding image, according to his/her opinion. For the Vocabulary Output, the request was overturned. Introducing the student to six new stimuli (animals for

2nd grade, rooms of the house for 3rd grade), the researcher pointed to a stimulus, and the student had to reply with the specific term of that. The score for both proves is calculated from a minimum of zero to a maximum of six, assigning one point for each correct answer, and zero points for each incorrect. For the Sentence Building, we introduced two different topics for each grade school. In the first one, we asked students to translate a simple Italian phrase in English (e.g.: "*Come si chiede ad un bambino chi è il suo cantante preferito, in inglese?*"). In the second one, we introduced a concept in Italian, translating ourselves in English, and asking students to reply in English (e.g.: "*mi chiedono a che ora comincio la scuola. When is your class starting?*"). The Sentence Building score was calculated by adding the number of words that made up the grammatically correct form of each answer. Any additional words (e.g. good morning; please; etc.) were not taken into consideration. The final score consisted of the sum of the correct answers in the Vocabulary Entry (maximum 6), in the Vocabulary Output (maximum 6) and in the Sentence Building (maximum 10 for the 2nd grade, and 11 for the 3rd grade), hence from 0 to 22 or 23.

- Metacognitive Monitoring (MM)

Monitoring relates to the ability to judge one's own performance and to express certainty or uncertainty about one's own responses (e.g., "How sure am I to have given the right answer to that question?"; see Roebers & Spiess, 2017). To measure metacognitive monitoring related to English learning as a L2, children's confidence judgements on the correctness of their answers on the English vocabulary test were collected (Roebers et al., 2012, 2017). After performing each vocabulary subtest, the children were presented again with each item of each task and asked how certain they were that they gave the correct answer (i.e., metacognitive monitoring). For each item, children had to indicate their level of confidence on a 5-point Likert scale ranging from "not sure at all" to "very sure", represented by smileys, ranging from a very sad face to a very happy face, respectively. Specifically, we asked them "How sure are you that you indicated the right picture?" (for English vocabulary comprehension) and "How sure are you that you named this picture correctly?" (for English vocabulary production). The metacognitive monitoring score was calculated separately for correct and incorrect answers by calculating the mean confidence rates. In addition to pure confidence judgements, the ability to monitor performance on the English vocabulary test was considered

by calculating a discrimination score by subtracting the mean confidence judgement score of incorrect responses from the mean confidence judgement score of correct responses. This value expresses the ability to differentiate between correct and incorrect responses, with a greater value indicating a better ability to discriminate between correct and incorrect answers (Roebers, 2002). To control for the typical confounding effect between monitoring and performance, we included pure confidence judgements in the model (see Roebers et al., 2012).

- Metacognitive Control (MC)

After performing the metacognitive monitoring task, we evaluated the children's metacognitive control ability, since accurate monitoring is an important prerequisite for effective control (i.e., monitoring-based control; see Dunlosky & Metcalfe, 2008). The children were presented again with each item of each subtest and asked to judge whether they had given the correct answer or not. Two boxes were used, one red box and one green box. For maintenance or withdrawal of their previously given answers, children were asked to put the items for which they believed that they had given the correct answer in the green box and to put those items for which they believed to have given the wrong answer in the red box (see Lavis & Mahy, 2021, for a similar approach). Children were given 1 point if they correctly put a correct item in the green box (i.e., maintenance of correct responses) or an incorrect item in the red box (i.e., withdrawal of incorrect responses), whereas they received 0 points if they put a correct item in the red box or an incorrect item in the green box. First, we calculated the proportion of correctly maintained items and the proportion of withdrawn incorrect items separately. To have a more appropriate measure of metacognitive control, considering high and low English vocabulary performance, within the model, we used a unique measure including both control measures of correct and incorrect responses (Roebers & Spiess, 2017).

3.1.3. Data analysis

We conducted a post hoc power analysis to evaluate the sample size necessary to achieve the statistical power given that $\alpha = .05$ and moderate to high effect size (ES = .50–70).

Parametric Pearson r correlations were calculated between L2 learning, emotional, metacognitive, and cognitive measures included in the present study to highlight any relationship between them. A Growth Curve Analysis (GCA – Mirman, 2017) was carried out to determine significant differences for English vocabulary (L2), the achievement emotions, metacognition, and cognition overtime. We opted for a GCA rather than a Repeated Measures ANOVA for several reasons. First, Repeated Measures ANOVA assume a linear and uniform trend over time, which can limit the ability to capture more complex changes or nonlinear variations, whereas GCA allows for the modeling of both linear and nonlinear changes over time, allowing the use of growth terms (e.g., linear, quadratic, cubic) to represent complex growth curves. Furthermore, a common problem in longitudinal studies is the handling of missing data. GCA is much more robust to missing data than ANOVA, which generally requires complete data at each time point: this often leads to the need to drop cases with missing measures, reducing statistical power (Krueger & Tian, 2004; Mirman, 2017).

Next, a multilevel model for repeated measures (MLM) was applied to the explanatory model to analyze whether L2 learning outcomes within a repeated measure design could be predicted by a specific child's school, age and achievement emotion, metacognition, and cognition. The assumptions of multilevel models (e.g., residuals, outliers, and collinearity) were checked and addressed before conducting the primary analyses. Since the outcome was conceived as a longitudinal continuous variable, a planned comparison within a restricted maximum likelihood (REML)-based mixed-effects model with Satterthwaite adjustment was applied to address model misspecification and its unbiasedness for data missing completely at random or at random (Cnaan et al., 1997).

The data for this study could be conceived having a hierarchical structure: 305 students, who were tested several times, and who were strictly nested within three schools. Hierarchical linear models, especially in the education studies literature, are useful for analyzing repeated measures or longitudinal data (Snijders & Bosker, 2012). Despite the potentially nested nature of our data, a common role in performing multilevel analysis suggested that at least five levels of the grouping variable represent the minimum number of factor levels should be for random effect (Gelman & Hill, 2006). Substantially, having few levels makes the estimation of the variance of random effects terms difficult, and multilevel models typically add little over classical models (Gomes, 2022). Based on these remarks, we first, tested the potential contribution of school (≤3 levels) as a random effect. Otherwise, we considered schools in

the model as fixed effect and retained the students' ID as unique random effect in the model. Emotional, metacognitive and cognitive measures as well as age were also included in the model as covariates.

Upon these considerations, the explanatory variables were added stepwise into the model. First, we estimated the unconditional null model (Model 0 or "the intercept model") without any predictor variables, which provided the overall L2 learning for all the students. It divides the variance of the initial score as well as the variance of the learning L2 into between-students components. As mentioned above, the dependent variable learning L2 was measured at the individual student level. To test the common role of >5 levels for a random variable in the model, we estimated an additional intercept model including both participants and schools at a random level. An improvement in the effect size in this latter model over the previous one will justify the use of a grouping variable with few levels (in our case, schools with three levels). Next, we estimated a new intermediate model (Model #1) adding fixed effect variables to the model and a final model (Model #2) including both fixed and random effects. In the fixed effect models, the school variable has been introduced as a pairwise comparison (e.g., school 2 – school 1; school 3 – school 1), as well as the age variable. Covariate variables were included in the last model as random effects. Cross-level interactions were also allowed and were subsequently removed from the models if they were not significant.

R-squared marginal and conditional indices, as well as -2Loglikelihood (-2LL), were used to assess model fit. The marginal R-squared (R^2M) considers only the variance of the fixed effects, whereas the conditional R-squared (R^2C) takes both the fixed and random effects into account. The likelihood ratio chi-square test (LRT) with the provided Akaike information criterion (AIC) and Bayesian information criterion (BIC) were also used to assess competing models. A significant LRT also provided evidence of an improvement in the present model over the null model. The model with lower AIC and BIC scores is expected to exhibit a superior balance between its ability to fit the dataset and its ability to avoid overfitting the dataset. All tests time and covariate effects were conducted using an alpha level of .05 and reported estimates, standard errors and 95% confidence intervals. Random components are reported in terms of variance, standard deviation, and intraclass correlation (ICC). Statistical analyses were performed via Statistical analyses were performed via R statistics, lme4 (Bates et al., 2009), ggplot2 packages (Wickham, 2011), Jamovi 2.6 (The jamovi project, 2024; R Core Team, 2024; Gallucci, 2019; Lüdecke et al., 2020), and G*power (Faul et al., 2007).

3.2. Results

According to the power analysis (α = .05 and ES = .50–.70), our recruited sample ensured good statistical power (1-β = 99%) for the difference between repeated measures (ANOVA). Correlations between L2 learning, cognitive, metacognitive, and emotional measures were computed. Since the consistent number of investigated variables, increased by the different time phases of our longitudinal study, we decided to propose the correlations in different tables, to facilitate the readers. First, we introduced a correlation analysis for each time, to highlight the transversal relationship between the different variables (e.g.: T0 vs. T0; T1 vs. T1; etc. See section 6.2.1.). Subsequently, the tables relating to the longitudinal comparison are reported, that is, the role of each variable with respect to itself and the others between the different time phases (T0 vs. T1 vs. T2 vs. T3. See section 6.2.2.).

3.2.1. Relations between measures time after time.

The results revealed that L2 learning at all time (Table 9; Table 10; Table 11;Table 12) was highly positively correlated with the respective metacognitive skills, both monitoring and control, except for metacognitive control at T3 (Table 12): the latter appears, surprisingly, slightly negatively correlated with L2 learning. As expected, L2 learning was all time highly negative correlated (with a time increment) with the anxiety, revealing the less anxiety students had about English as L2, the better they learned it. No significant correlations were found between L2 and joy or boredom, except for boredom at T3 ($r = —0.27$, $p < .01$). Regarding cognitive variables, all of them appeared correlated with L2 learning, in particular the working memory (evaluated with the Immediate Non-Word Recall – INWR). We noted no correlation between different types of variables, such as cognitive and emotional variables. Some sporadic trend toward significance was highlighted, e.g.: short-term memory (specifically, Digit Span Forward) and Metacognitive Control at T0 and T1 (respectively: $r = 0.16$, $p = .01$; $r = 0.25$, $p = .03$); the working memory (INWR) and Metacognitive Control at T1 ($r = 0.26$, $p = .02$); Anxiety with Digit Span Forward ($r = —0.35$, $p = .02$) and Delayed Non-Word Recall ($r = —0.30$, $p = .04$) at T1; Digit Span Backward with Joy at T1 ($r = —0.35$, $p = .02$) and Metacognitive Control at T2 ($r = 0.18$, $p = .01$).

	T0.I.N.W.R.	T0.Span.F	T0.Span.B	T0.D.N.W.R.	T0.L2	T0.M.M.	T0.M.C.	T0.Joy	T0.Boredom
T0.Span.F	0.24***								
T0.Span.B	0.15*	0.48***							
T0.D.N.W.R.	0.53***	0.1	0.06						
T0.L2	0.15*	0.22***	0.11	0.15*					
T0.M.M.	0.04	0.03	0.04	0.08	0.3***				
T0.M.C.	0.1	0.16*	0.1	0.01	0.37***	—0.10			
T0.Joy	—0.08	0.09	0.01	0.06	0.08	—0.01	—0.07		
T0.Boredom	0.06	—0.13	—0.07	—0.11	—0.1	—0.09	0.04	—0.64***	
T0.Anxiety	0.03	—0.13	0.01	—0.09	—0.17**	—0.11	—0.08	—0.44***	0.67***

* p < .05, ** p < .01, *** p < .001.

I.N.W.R.: Immediate Non-Word Recall; Span.F: Digit Span, Forward version; Span.B: Digit Span, Backward version; D.N.W.R.: Delayed Non-Word Recall; L2: English as 2nd language; M.M.: Metacognitive Monitoring; M.C.: Metacognitive Control.

Table 9. Pearson–Correlations between the various measures at T0.

	T1.I.N.W.R.	T1.Span.F	T1.Span.B	T1.D.N.W.R.	T1.L2	T1.M.M.	T1.M.C.	T1.Joy	T1.Boredom
T1.Span.F	0.35**								
T1.Span.B	0.11	0.36**							
T1.D.N.W.R.	0.51***	0.35**	0.21						
T1.L2	0.52***	0.34**	0.14	0.23*					
T1.M.M.	0.18	0.01	—0.02	0.17	0.3***				
T1.M.C.	0.26*	0.25*	0.04	0.02	0.33***	0.04			
T1.Joy	0.17	0.21	0.29*	0.25	—0.02	0.08	0.03		
T1.Boredom	—0.13	—0.21	—0.07	—0.15	—0.11	—0.06	—0.07	—0.70***	
T1.Anxiety	—0.14	—0.35*	—0.16	—0.30*	—0.24***	—0.08	—0.1	—0.49***	0.67***

* p < .05, ** p < .01, *** p < .001.

Note: I.N.W.R.: Immediate Non-Word Recall; Span.F: Digit Span, Forward version; Span.B: Digit Span, Backward version; D.N.W.R.: Delayed Non-Word Recall; L2: English as 2nd language; M.M.: Metacognitive Monitoring; M.C.: Metacognitive Control.

Table 10. Pearson–Correlations between the various measures at T1.

	T2.I.N.W.R.	T2.Span.F	T2.Span.B	T2.D.N.W.R.	T2.L2	T2.M.M.	T2.M.C.	T2.Joy	T2.Boredom
T2.Span.F	0.22**								
T2.Span.B	0.23**	0.42***							
T2.D.N.W.R.	0.61***	0.15*	0.19**						
T2.L2	0.17*	0.19**	0.28***	0.28***					
T2.M.M.	—0.03	—0.07	0.01	0.06	0.26***				
T2.M.C.	0.02	0.11	0.18**	—0.02	0.35***	—0.14*			
T2.Joy	—0.04	—0.12	—0.11	0.06	0.03	0.19**	—0.1		
T2.Boredom	0.05	0.05	0.04	—0.04	—0.08	—0.16*	0.06	—0.73***	
T2.Anxiety	0.05	—0.08	0.05	—0.09	—0.26***	—0.17*	—0.00	—0.50***	0.62***

* p < .05, ** p < .01, *** p < .001.

Note: I.N.W.R.: Immediate Non-Word Recall; Span.F: Digit Span, Forward version; Span.B: Digit Span, Backward version; D.N.W.R.: Delayed Non-Word Recall; L2: English as 2nd language; M.M.: Metacognitive Monitoring; M.C.: Metacognitive Control.

Table 11. Pearson–Correlations between the various measures at T2.

	T3.I.N.W.R.	T3.Span.F	T3.Span.B	T3.D.N.W.R.	T3.L2	T3.M.M.	T3.M.C.	T3.Joy	T3.Boredom
T3.Span.F	0.21*								
T3.Span.B	0.30***	0.46***							
T3.D.N.W.R.	0.76***	0.19*	0.27**						
T3.L2	0.32***	0.27**	0.34***	0.14					
T3.M.M.	0.30***	0.19*	0.23**	0.16	0.33***				
T3.M.C.	0.09	0.12	0.01	0.07	—0.18*	—0.03			
T3.Joy	0.17	—0.08	—0.06	0.21*	0.15	0.16	0.03		
T3.Boredom	—0.12	0.05	0.09	—0.05	—0.27**	—0.15	0.09	—0.78***	
T3.Anxiety	—0.27**	—0.23*	—0.1	—0.19*	—0.33***	—0.17	0.05	—0.36***	0.49***

* p < .05, ** p < .01, *** p < .001.

Note: I.N.W.R.: Immediate Non-Word Recall; Span.F: Digit Span, Forward version; Span.B: Digit Span, Backward version; D.N.W.R.: Delayed Non-Word Recall; L2: English as 2nd language; M.M.: Metacognitive Monitoring; M.C.: Metacognitive Control.

Table 12. Pearson–Correlations between the various measures at T3.

3.2.2. Relations between measures overtime.

We found several correlations between our variables and themselves overtime (see Table 13; Table 14; Table 15). It is interesting to note that L2 learning at T0 was significantly associated to the same task at T1 and T2, but not at T3. The latter was also not related to L2 at T1, and just a trend of significancy with T2. Regarding the other variables, the role of metacognitive and emotional skills (specifically, anxiety perceived) in predicting the learning of English as a second language overtime seems to be identifiable as constant. We conducted a growth curve to better highlight the role of time in each variable.

	T0.I.N.W.R.	T0.Span.F	T0.Span.B	T0.D.N.W.R.	T0.L2	T0.M.M.	T0.M.C.	T0.Joy	T0.Boredom	T0.Anxiety
T1.I.N.W.R.	0.53***	0.36**	0.12	0.49***	0.3**	0.2	0.28*	0.1	—0.02	0.12
T1.Span.F	0.41***	0.96***	0.31**	0.38***	0.32**	0.08	0.21	0.03	0.06	0.11
T1.Span.B	0.21	0.25*	0.98***	0.20	0.06	—0.09	—0.00	—0.03	0.18	0.13
T1.D.N.W.R.	0.54***	0.30**	0.19	0.95***	0.14	—0.04	0.18	0.05	0.04	0.07
T1.L2	0.24***	0.24***	0.14*	0.18**	0.54***	0.14*	0.27***	0.08	—0.14*	—0.17**
T1.M.M.	0.06	0.06	0.06	0.09	0.22***	0.24***	0.13*	—0.04	—0.02	—0.07
T1.M.C.	0.02	0.17**	—0.00	0.05	0.27***	0.09	0.27***	0.14*	—0.2**	—0.21***
T1.Joy	—0.16*	—0.04	0.08	0.04	0.10	0.05	—0.00	0.43***	—0.37***	—0.33***
T1.Boredom	0.1	—0.10	—0.03	—0.1	—0.1	—0.03	—0.01	—0.33***	0.46***	0.34***
T1.Anxiety	—0.04	—0.26***	—0.11	—0.10	—0.21**	—0.06	—0.05	—0.28***	0.38***	0.51***
T2.I.N.W.R.	0.13	0.20*	0.14	0.21*	0.11	—0.02	0.02	0.11	—0.05	—0.08
T2.Span.F	0.15*	0.49***	0.32***	0.18*	0.04	—0.15*	0.03	0.16*	—0.12	—0.13
T2.Span.B	0.34***	0.26***	0.37***	0.37***	0.13	0.05	0.04	0.01	—0.04	—0.02
T2.D.N.W.R.	0.18*	0.18*	0.12	0.37***	0.22**	—0.01	0.1	0.03	0.03	—0.01
T2.L2	0.22**	0.25**	0.14	0.25***	0.5***	0.13	0.4***	0.06	—0.13	—0.22**
T2.M.M.	0.05	0.06	—0.03	0.17*	0.24***	0.3***	0.06	0.09	—0.15*	—0.16*
T2.M.C.	0.04	0.04	0.20**	0.08	0.21**	—0.07	0.25***	0.07	—0.11	—0.18*
T2.Joy	—0.18*	—0.08	—0.06	—0.04	0.03	0.04	—0.02	0.35***	—0.21**	—0.17*
T2.Boredom	0.13	0.03	0.06	0.02	—0.06	0.03	—0.05	—0.28***	0.31***	0.18*
T2.Anxiety	0.07	—0.10	0.03	—0.09	—0.23***	—0.11	—0.15*	—0.13	0.17*	0.32***
T3.I.N.W.R.	0.26*	0.29**	0.18	0.30**	0.15	0.09	0.01	0.11	—0.04	—0.10
T3.Span.F	0.2	0.42***	0.35***	0.11	0.11	0.11	—0.08	—0.05	0.1	0.14
T3.Span.B	0.11	0.36***	0.46***	0.17	0.2*	0.02	—0.07	0.03	—0.05	0.09
T3.D.N.W.R.	0.26*	0.26*	0.12	0.3**	0.16	0.18	0.00	0.12	—0.01	—0.08
T3.L2	0.19	0.32**	0.20	0.19	—0.00	—0.16	—0.08	0.21*	—0.11	—0.01
T3.M.M.	0.18	0.17	0.20	0.19	0.02	0.06	0.02	0.25**	—0.2*	—0.2
T3.M.C.	—0.2	—0.06	0.11	—0.08	0.01	0.05	—0.07	—0.03	0.00	0.10
T3.Joy	0.02	0.17	—0.12	0.05	0.04	0.12	—0.15	0.29***	—0.23**	—0.14
T3.Boredom	—0.02	—0.12	0.03	—0.03	—0.02	—0.06	0.13	—0.30***	0.32***	0.16
T3.Anxiety	—0.25*	—0.17	—0.09	—0.27**	—0.2*	—0.07	—0.18*	—0.26**	0.19*	0.36***

* $p < .05$, ** $p < .01$, *** $p < .001$.

Note: I.N.W.R.: Immediate Non-Word Recall; Span.F: Digit Span, Forward version; Span.B: Digit Span, Backward version; D.N.W.R.: Delayed Non-Word Recall; L2: English as 2nd language; M.M.: Metacognitive Monitoring; M.C.: Metacognitive Control.

Table 13. Pearson–Correlations between T0 with T1, T2, and T3.

	T1.I.N.W.R.	T1.Span.F	T1.Span.B	T1.D.N.W.R.	T1.L2	T1.M.M.	T1.M.C.	T1.Joy	T1.Boredom	T1.Anxiety
T2.I.N.W.R.	0.49***	0.25*	0.13	0.31*	0.21**	—0.06	0.10	—0.04	0.1	—0.04
T2.Span.F	0.14	0.57***	0.5***	0.22	0.15*	—0.04	0.02	0.02	0.01	—0.12
T2.Span.B	0.27*	0.48***	0.48***	0.32*	0.29***	0.14*	0.07	—0.07	0.03	—0.03
T2.D.N.W.R.	0.55***	0.29*	0.24	0.39**	0.22***	0.04	0.12	—0.02	0.11	0.01
T2.L2	0.51***	0.26*	0.10	0.33**	0.67***	0.28***	0.23***	0.09	—0.09	—0.23**
T2.M.M.	0.26*	0.06	—0.20	0.16	0.14*	0.35***	0.17*	0.04	—0.01	—0.11
T2.M.C.	0.05	0.15	0.28*	0.15	0.33***	0.13*	0.24***	0.11	—0.07	—0.16*
T2.Joy	0.1	—0.07	—0.22	0.07	—0.14*	0.1	—0.11	0.65***	—0.50***	—0.44***
T2.Boredom	—0.18	—0.06	0.3*	—0.16	0.02	—0.13	0.04	—0.50***	0.5***	0.41***
T2.Anxiety	—0.33**	—0.09	0.25*	—0.27*	—0.1	—0.13	0.01	—0.42***	0.38***	0.52***
T3.I.N.W.R.	0.26	0.37*	0.25	0.23	0.30***	0.1	—0.01	0.15	—0.08	—0.17
T3.Span.F	0.10	0.61***	0.49**	0.22	0.18*	0.06	0.17	—0.05	0.09	0.04
T3.Span.B	—0.08	0.55***	0.48**	0.10	0.23**	0.07	0.1	—0.01	0.02	—0.01
T3.D.N.W.R.	0.22	0.43*	0.19	0.29	0.23*	0.13	0.02	0.11	—0.03	—0.07
T3.L2	0.09	0.36*	0.28	0.22	0.15	—0.02	—0.21*	0.14	—0.14	—0.07
T3.M.M.	—0.12	0.08	0.20	—0.13	0.18*	0.09	—0.14	0.06	—0.11	—0.10
T3.M.C.	—0.25	—0.07	0.08	—0.13	—0.08	0.02	0.1	0.08	—0.01	0.02
T3.Joy	0.28*	0.14	—0.22	0.21	0.03	0.12	0.01	0.37***	—0.34***	—0.27**
T3.Boredom	—0.17	—0.11	0.13	—0.22	0.01	—0.04	—0.01	—0.31***	0.43***	0.32***
T3.Anxiety	—0.28*	—0.35*	—0.23	—0.47**	—0.27***	—0.21*	—0.08	—0.21*	0.28**	0.49***

* p < .05, ** p < .01, *** p < .001.

Note: I.N.W.R.: Immediate Non-Word Recall; Span.F: Digit Span, Forward version; Span.B: Digit Span, Backward version; D.N.W.R.: Delayed Non-Word Recall; L2: English as 2nd language; M.M.: Metacognitive Monitoring; M.C.: Metacognitive Control.

Table 14. Pearson–Correlations between T1 with T2, and T3.

	T2.I.N.W.R.	T2.Span.F	T2.Span.B	T2.D.N.W.R.	T2.L2	T2.M.M.	T2.M.C.	T2.Joy	T2.Boredom	T2.Anxiety
T3.I.N.W.R.	0.37***	0.22*	0.26**	0.46***	0.28**	0.13	0.07	—0.04	0.01	—0.03
T3.Span.F	0.13	0.60***	0.31***	0.15	0.28**	0.09	0.02	—0.16	0.07	—0.09
T3.Span.B	0.11	0.38***	0.50***	0.23*	0.26**	0.08	0.09	—0.01	—0.08	—0.09
T3.D.N.W.R.	0.35***	0.14	0.16	0.47***	0.26**	0.10	0.05	—0.00	—0.01	—0.06
T3.L2	0.05	0.27**	0.31***	0.06	0.18*	—0.01	—0.12	0.18*	—0.19*	—0.13
T3.M.M.	—0.05	0.17	0.09	0.14	0.17	0.19*	—0.02	0.11	—0.07	—0.05
T3.M.C.	—0.10	0.08	—0.13	—0.08	—0.02	—0.11	0.06	—0.08	0.1	0.06
T3.Joy	0.05	—0.09	—0.16	0.22**	0.09	0.08	0.03	0.42***	—0.29***	—0.22**
T3.Boredom	0.03	0.06	0.13	—0.06	0.02	—0.03	0.05	—0.43***	0.49***	0.28***
T3.Anxiety	—0.07	—0.34***	—0.20*	—0.11	—0.34***	—0.12	—0.20*	—0.13	0.19*	0.47***

* p < .05, ** p < .01, *** p < .001.

Note: I.N.W.R.: Immediate Non-Word Recall; Span.F: Digit Span, Forward version; Span.B: Digit Span, Backward version; D.N.W.R.: Delayed Non-Word Recall; L2: English as 2nd language; M.M.: Metacognitive Monitoring; M.C.: Metacognitive Control.

Table 15. Pearson–Correlations between T2 and T3.

3.2.3. Growth Curve Analysis (GCA)

Table 16 shows parameter estimates of the growth curve analysis (GCA) conducted on our variables over time. In summary, the analysis suggests significant progress in learning and cognitive abilities, but this is accompanied by growing emotional stress, which may affect

student motivation and overall learning experience. In more detail, the effect of time was significant for nearly all variables (p<.001), showing substantial changes for each variable over the time points (T0-T3). Regarding our dependent variable, L2 learning, the GCA highlighted a significant improvement (β=1.39, p<.001), suggesting that students improved their English performance over time. For INWR, the effect of time was positive and significant (β=2.24, p<.001), indicating an improvement in non-word recall ability over time. Both digit span tasks also showed significant increases (p<.001) in short-term memory ability over time. Metacognitive Monitoring (β=−0.20, p<.001) and Joy (β=−0.35, p<.001) decreased, indicating a slight decline in students' ability to monitor their performance and less joy perceived in L2 learning as time progresses. Contrary to the positive emotion, Boredom and Anxiety tended to increase slightly over time (β= 0.19, p<.001 for Boredom; β= 0.06, p=0.003 for Anxiety), indicating a rise in negative emotions during L2 learning. See Figure 6 for an overall view of the average trends, and Figure 7 for detailed analyses of individual variability.

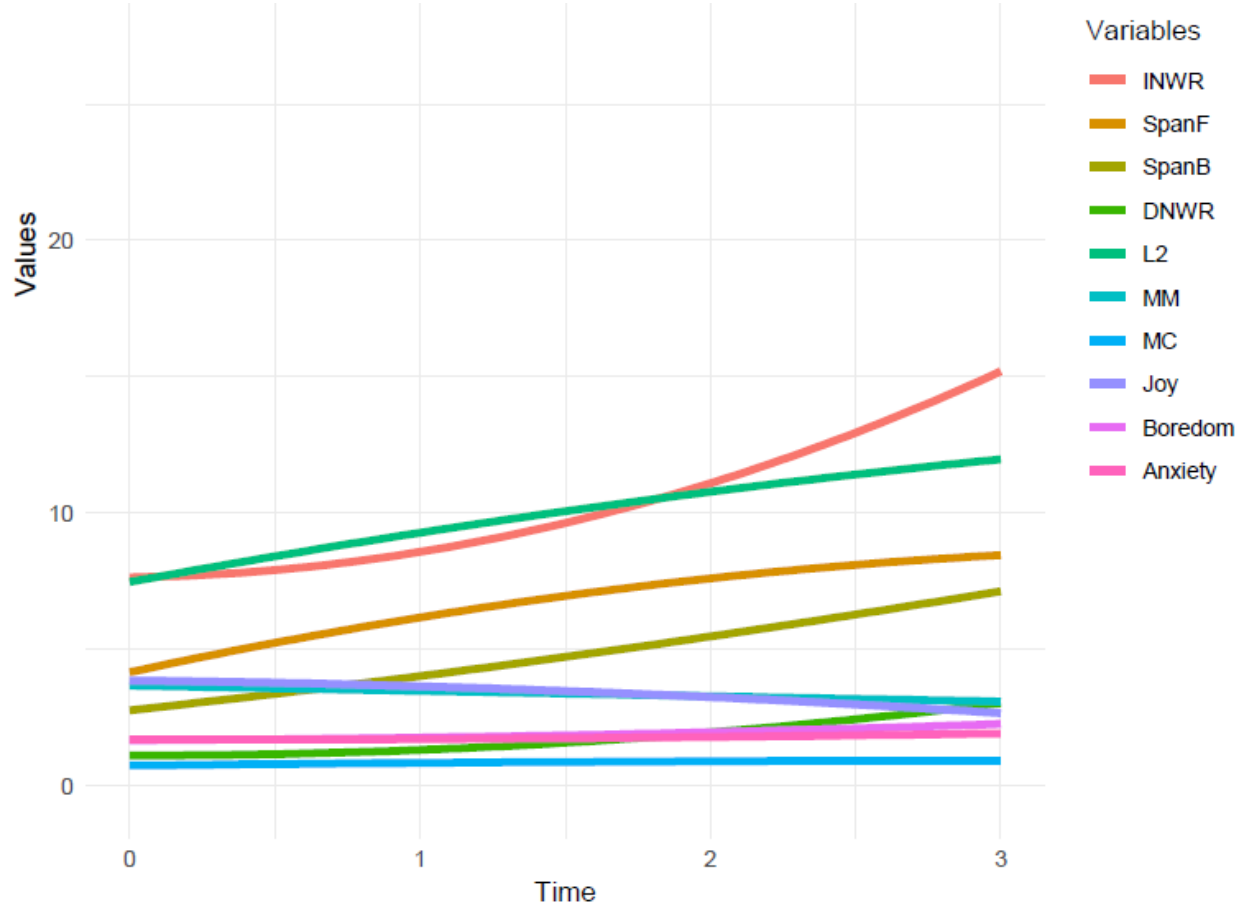


Figure 6. Aggregate Growth Curve Analysis T0-T3.

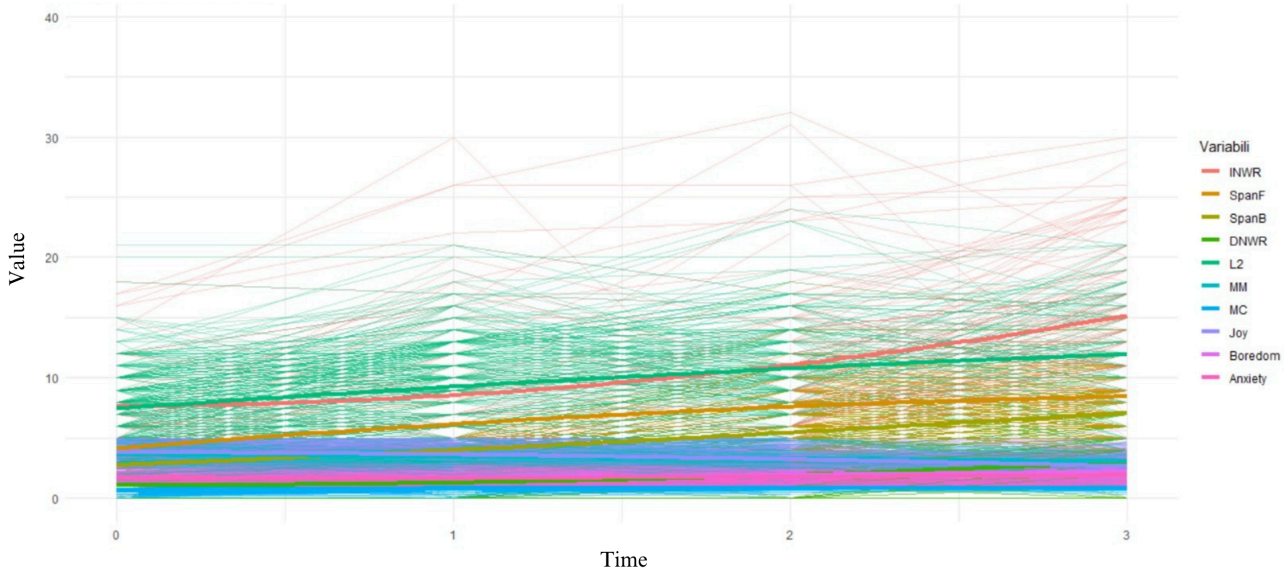


Figure 7. Individual Growth Curve Analysis T0-T3.

	INWR		**SpanF**		**SpanB**		**DNWR**		**L2**		**MM**		**MC**		**Joy**		**Boredom**		**Anxiety**	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	7.28	**<.001**	4.33	**<.001**	2.77	**<.001**	1.01	**<.001**	7.66	**<.001**	3.66	**<.001**	0.74	**<.001**	3.95	**<.001**	1.60	**<.001**	1.66	**<.001**
TIME	2.24	**<.001**	1.50	**<.001**	1.40	**<.001**	0.55	**<.001**	1.39	**<.001**	–0.20	**<.001**	0.06	**<.001**	–0.35	**<.001**	0.19	**<.001**	0.06	**0.003**
Random Effects																				
σ^2	23.77		2.96		1.99		1.47		10.37		0.55		0.04		0.71		0.59		0.34	
τ_{00}	8.57 $_{ID}$		1.37 $_{ID}$		0.97 $_{ID}$		0.92 $_{ID}$		7.60 $_{ID}$		0.19 $_{ID}$		0.01 $_{ID}$		0.54 $_{ID}$		0.45 $_{ID}$		0.27 $_{ID}$	
ICC	0.27		0.32		0.33		0.39		0.42		0.26		0.20		0.43		0.43		0.44	
Observations	656		698		698		698		979		943		979		907		907		907	
R^2M/ R^2C	0.181/0.398		0.415/0.600		0.475/0.647		0.148/0.477		0.103/0.482		0.054/0.298		0.075/0.262		0.104/0.491		0.037/0.453		0.006/0.446	

Note: INWR: Immediate Non–Word Recall; SpanF: Digit Span, Forward version; SpanB: Digit Span, Backward version; DNWR: Delayed Non–Word Recall; L2: English as 2nd language; MM: Metacognitive Monitoring; MC: Metacognitive Control. σ^2 = variance residual; τ_{00} = variance for random effects; R^2M = R^2 Marginal; R^2C = R^2 Conditional.

Table 16. Parameter estimates of growth curve analysis.

3.2.4. Multilevel model for repeated measures (MLM)

Neither univariate nor multivariate outliers were detected, as assessed by a Q–Q plot of the residuals. All primary study variables were deemed to be normally distributed (Kolmogorov–Smirnov test = .045, p = .25), and the residual errors were approximately normally distributed. No multicollinearity issues were detected among the predictors. Table 17 presents model fits and parameter estimate of the three random intercept models: the unconditional null model, an intermediate model, and a final model fitted via the REML and bobyqa optimizer (Bound Optimization by Quadratic Approximation).

The unconditional null model (Model 0) represents a simple variance component model that predicts the overall L2 learning for all students. As expected, the model intercept was significant at ($t(979) = 47.9$, $p < .001$). The null model contained no predictor variable; only the intercept was modelled as a combination of a fixed effect and a random deviation from it. The participants' ID was included in the model as a random effect to address observation dependency problems. The ICC of the dependent variable was .38; thus, approximately 38% of the variance in L2 learning was explained by interindividual differences. This also means that a larger part of the variance could be explained by situational characteristics. We subsequently estimated an alternative null model retaining both the ID and the school as random components. The model's intercept estimate of this latter model was close to that of the previous model ($t(979) = 48.86$, $p < .001$), with an ICC of .33. Thus, the inclusion of the school grouping variable as a random effect in the model explained only 5% more of the dependent variable than before. Concerning random effects, we have enough evidence to conclude that the participants' ID cluster contributed more than the school cluster did to the amount of variation in L2 learning. Hence, we shifted it as a fixed effect in the following models.

Model #1 accounts for 38% of the total variance (R^2C), and the 8% of the variance (R^2M) model was explained by the fixed effect (ICC of .33). The lowest AIC/BIC values and the significant likelihood ratio (LRT = 91.1, df = 1, $p < .001$) provided evidence for an improvement in the present model over the null model. The model's intercept, corresponding to schools and the first age level, was significant ($t(979) = 11.0$, $p = .007$). Among the predictor variables in the equation, Students in school 2 ($t(979) = 5.99$, $p = <.001$) had an average score higher than students in school 1. School 3, however, did not show a significant

difference compared to school 1 (t(979) = 0.66, p = .512). Significant fixed effect was found for age (t(979) = –.50, p = .078).

Next, the final model (Model 2), which includes random and fixed effects, demonstrated an improvement in model fit over previous models (see AIC and BIC values in Table 17), with a decreased ICC. Model 2 fit the data well which accounts for 51.1% of the total variance (R^2C). 41.2% of the variance model was explained by fixed and covariate effects (R^2M). The model's intercept, corresponding to schools and age, was significant (t(515) = 18.83, p = .003).

Among the cognitive variables as predictors in the equation, only Digit Span Backward (t(515) = 5.46, p <.001) and the Delayed Non-Word Recall (t(515) = 2.62, p = .01) had significant effects in determining L2 learning. Both Metacognitive Monitoring scores (t(515) = 4.48, p -.001) and Metacognitive Control (t(515) = 5.97, p = <.001) were found to predict English as L2. Among emotions, Anxiety was also found to predict L2 learning (t(515) = -3.07, p = 0.002). No significant fixed effect included Joy (t(515) = –1.13, p = 0.258) or Boredom (t(515) = –1.28, p = 0.221). No significant cross-level interaction was observed in either Model #1 or Model #2.

	Model 0			Model 1			Model 2		
Predictors	Estimates [95% CI]	*SE*	*p*	Estimates [95% CI]	*SE*	*p*	Estimates [95% CI]	*SE*	*p*
(Intercept)	9.28 [8.90; 9.66]	0.19	<.001	9.01 [8.65; 9.38]	0.18	<.001	9.39 [9.00; 9.77]	0.20	<.001
School 2 – School 1				2.63 [1.78; 3.50]	0.44	<.001	1.51 [0.68; 2.34]	0.42	<.001
School 3 – School 1				0.32 [–0.63; 1.26]	0.48	.512	0.67 [–1.03; 1.17]	0.56	0.905
Age				–0.59 [–1.06; 0.05]	0.29	.078	0.03 [–0.55; 0.62]	0.30	0.915
I.N.W.R							0.06 [–0.01 0.13]	0.04	0.137
Span.F							0.09 [–0.09; 0.28]	0.09	0.341
Span.B							0.57 [0.36; 0.77]	0.10	<.001
D.N.W.R.							.36 [0.10; 0.64]	0.14	.009
M.M.							0.91 [0.51; 1.31]	0.14	<.001
M.C.							4.52 [3.04; 6.01]	3.04	<.001
Joy							–0.25 [–0.68; 0.18]	0.22	.258
Boredom							–.34 [–0.87; 0.20]	0.27	.221
Anxiety							–0.92 [–1.51; –0.33]	0.30	.002

	Model 0			Model 1			Model 2		
Predictors	Estimates [95% CI]	*SE*	*p*	Estimates [95% CI]	*SE*	*p*	Estimates [95% CI]	*SE*	*p*
Random Effects									
σ2	12.64			12.69			11.59		
τ00	7.8 ID			6.13 ID			2.34 ID		
ICC	0.38			0.33			0.17		
Observations	979			979			515		
R^2M/R^2C	0.000/0.382			0.0825/0.3812			0.412/0.511		
Deviance	5598.338			5551.5392			2816.108		
AIC	5602.893			5561.4394			2826.599		
BIC	5618.998			5592.8585			2909.770		
log–Likelihood	–2799.169			–2775.7696			–1408.054		

Note: I.N.W.R.: Immediate Non-Word Recall; Span.F: Digit Span, Forward version; Span.B: Digit Span, Backward version; D.N.W.R.: Delayed Non-Word Recall; M.M.: Metacognitive Monitoring; M.C.: Metacognitive Control.

σ2 = variance residual; τ00 = variance for random effects; ICC= intraclass correlation; R^2M = R^2 Marginal; R^2C = R^2 Conditional.

Table 17. Multilevel parameter estimates for the MLM on English as L2.

3.3. Discussion

The present study aimed to investigate within a longitudinal design the role of metacognitive and emotional variables on primary school children's English learning as L2. Thus, children attending the 2nd and 3rd grades of primary school were tested several times (November 2021, T0; 6 months later, May 2022, T1; November 2022, T2) on an English vocabulary test, metacognitive monitoring and control abilities related to performance on the English vocabulary test, as well as on achievement emotions related to learning English as L2.

The results were in line with our hypotheses according to which achievement emotions, metacognitive and cognitive variables are related to children's performance on the L2 learning. Correlational analyses revealed a constant role of cognitive variables on L2 learning at each specific time of evaluation, recognizing an interaction between basic cognitive skills and linguistic skills since the first years of approaching the L2. Regarding the metacognitive competence investigated, the decline in metacognitive control identified at T3 ($r = —0.18$, $p < .05$) could suggest an overestimated effect of students' self-efficacy (Schunk, 1985). During the more advanced stages of learning, in fact, students could develop greater self-efficacy and confidence in their abilities, reducing the need to resort to metacognitive control strategies for every aspect of language learning. This shift to a more autonomous approach could explain why lower levels of MC correlate with higher levels of L2 competence at T3, since students

who reach a good linguistic level tend to reduce the use of explicit control in favor of a more fluid and less structured management of the process. We cannot exclude the possibility that the difference in the significance of the effects that metacognitive monitoring and control had on L2 might also be due to the specificity of the measurement. Similarly, it would be more comprehensive to include both a general measure of metacognition (e.g., declarative metacognition) and a performance-related measure of metacognition (i.e., procedural metacognition). However, during the study presented, no further metacognitive skills beyond monitoring and control were investigated, including students' strategies. We investigated these metacognitive aspects in Study 4. Regarding the achievement emotions investigated, it seems that anxiety had a constant role (both for a specific time and over the years) in L2 performance. In line with the literature, in fact, a lower perceived anxiety corresponds to a better L2 performance. Although this might appear to be a comforting finding, it is necessary to report that no significant role was highlighted by positive emotions such as joy, highlighting how anxiety is the most experienced emotion in school.

Correlation analyses highlight the direct relationship of L2 with other variables at different time points, while Growth Curve Analysis and Multi-Level Model (MLM) describe the evolution over time. Specifically, Growth Curve Analysis provides a temporal view of changes in L2, while the MLM explores in detail how different variables influence L2 at the individual and school levels, also considering hierarchical and group effects. Growth Curve Analysis for L2 revealed a constant improvement over time, indicating that continuous and progressive learning is effective for participants. The significance of the time estimate ($p < .001$) indicated that temporal progression is a key factor for L2 learning, with a constant and positive trend at each measurement interval. Similarly, temporal improvements were evident for cognitive and metacognitive skills (in particular, metacognitive monitoring). As highlighted by the correlations, the analyses suggested that an increase in anxiety may slow down L2 progress or not follow the same positive trajectory as other variables. Both random effects and Intraclass Correlation Coefficients showed that there are significant differences between individuals, suggesting that some progress more quickly based on personal factors. For this reason, we calculated an MLM, which revealed a constant effect not only of the students' previous individual differences in L2, but also that belonging to one of the three schools evaluated in our study could be the source of disparities in L2 performance. A better performance levels were highlighted in students from School 2, the only one belonging to a

territory in Northern Italy. Future research will certainly have to consider not only the territorial position of the school, but also the socioeconomic condition of the child and the school he belongs to. Model #2 again highlighted the support of all cognitive variables (in particular, working memory), retrospective metacognitive competence (both monitoring and control) and the constant role of anxiety (in this case, as a significant obstacle) in learning English as a L2.

3.4. Conclusion

To reconnect with Pekrun's control-value theory (2006) and Efklides' model of SRL (Efklides, 2011), we believe that the two models are important in the explanation and prediction of learning outcomes and that their integration would shed more light on various aspects of L2 learning as well as the learning process and outcomes in general.

Taken together, the results of the present study support the view that cognition, metacognition and achievement emotions contribute to the increase in vocabulary learning in English as L2 of primary school children during their second and third grades. The study of cognition, metacognition and emotions in an early stage of learning may represent a way to better understand how to support a learning process that may sometimes be difficult and generate negative emotions and reduced motivation (Efklides, 2011). The role played by those variables in predicting English as L2 learning seems to suggest that promoting them in the learning context may favour the experience of fewer negative emotions during the learning process and therefore support better self-regulated achievement. Moreover, integrating these insights in the educational setting, such as in teaching and instruction, might contribute to a more positive and meaningful learning experience for children (Gregersen, 2013).

4. Study 3: an investigation on emotions related to L2 learning in Italian primary school

This chapter refers to the study of Trotta et al. (under review [A]). Here below a list of tables and figure included in this chapter.

List of tables included in Study 3.

Table 18	**Pearson correlation between L1, L2 measures and emotional variables related to L2.**	73
Table 19	**ANOVA regression models.**	73
Table 20	**Multiple linear regression with Vocabulary in L2 as a dependent variable.**	74
Table 21	**Frequency of performance bands at the PMA (Vocabulary in L1).**	75

List of figures included in Study 3.

Figure 8	**Differences between students with High or Low Level in L1 performance relative to L2 Vocabulary Entry.**	76
Figure 9	**Differences between students with High or Low Level in L1 performance relative to L2 Vocabulary Output.**	76
Figure 10	**Differences between students with High or Low Level in L1 performance relative to L2 Sentence Building.**	77

Despite L2 learning is recognized as a fundamental step in the student's personal and social growth, the latest INVALSI data (2023) revealed a decline in the positive trend that characterized the years prior to COVID-19 by Italian primary school students. There was a decrease of 7% compared to 2022 in the Reading test, and 4% in the Listening. Similarly, there was also a post-pandemic worsening in the mother tongue (L1), Italian. The same INVALSI data from 2023 revealed a worsening of 4% compared to 2022 for 2nd grade students, and 6% for those in the 5th grade. For over 60 years, the role of L1 competence (L1) on L2 performance is recognized (Lado, 1957; Bernhardt, 2000, Palladino & Cornoldi, 2023). Learners must achieve a certain level of competence in their L1 before positive effects on L2 learning can be observed (Lasagabaster, 2001). According to the Hypothesis of Linguistic Interdependence (Cummins, 1979), the level of language proficiency and literacy in L2 can be influenced by previous competence in L1. Therefore, if a student has a low level of

proficiency in L1, they are likely to have low proficiency in L2 as well. Dufva and Voeten (1999) conducted a three-year longitudinal study with over 150 Finnish primary school students, starting from the first primary (M_{Age}=7 years), with the aim of highlighting an effect of L1 proficiency on L2 (English language) performance. The results showed that previous literacy skills in the L1 (Finnish language) explained almost 60% of the variance in English proficiency and that phonological memory also played a significant role in L2 learning. Sparks and Ganschow (1993) developed the Linguistic Coding Deficit Hypothesis (LCDH), a theory that identifies phonological and syntax attitudes in L1 as the most relevant cognitive factors in L2 learning. Ganschow et al. (1991), in fact, highlighted a significant difference in spelling and writing performance of high school students, divided according to their academic skills in learning L2. Recent Italian studies found similar results in dyslexic primary school students, showing significantly poorer control in reading and writing English as L2 (Palladino et al., 2013, 2016). LCDH theory recognized a link between L1 performance and L2 learning difficulties, mainly attributable to phonological coding skills (ability to organize, analyze and combine the sounds of language) rather than to affective variables (Jeong-Won, 1996). However, Sparks himself also highlighted emotional and motivational characteristics of low interest in the subject and related emotions among the profiles of students with low L2 learning (Sparks & Ganschow, 1995). In a more recent longitudinal study, Sparks et al. (2009) followed over 50 monolingual U.S. primary school students for ten years to highlight which predictors of L2 performance (Spanish, French, or German) were. The model introduced by the researchers indicated that 66% of the variance in L2 competence depended not only on linguistic aptitude in L1, but also on the affective-motivational characteristics of learning. Regarding the emotional aspects involved in L2 learning, the role of anxiety is widely recognized (MacIntyre, 1995). A recent review conducted by Shao et al. (2019) reported that anxiety related to L2 learning is the most studied emotion, while only a few studies have also examined other emotions. Although a role of emotions perceived during the learning process is widely recognized for its success and failure (see for example: Linnenbrink & Pintrich, 2002; Pintrich & De Groot, 2003; Tzohar-Rozen & Kramarski, 2014), there is still little knowledge of the role of different emotions (boredom and joy, for example) and of the specific contribution to L2 learning. Moreover, most studies mainly included adolescent and adult participants (e.g., Lee, 2014), while less is available in younger student populations. Therefore, the present study focused on the early stages of approaching English as a second

language, i.e. learning English in primary school, with the intention of answering the following research questions: do L1 learning and perceived emotions regarding L2 learning, predict English as L2 in primary school? What achievement emotions influence the L2 performance of Italian primary school students? Do students with low vocabulary proficiency and potential fragility in their native language experience fewer positive emotions in learning L2 English and have lower levels of Proficiency in L2 vocabulary?

4.1.Method

4.1.1. Participants

A total of 182 Italian primary school students, native speakers or adequately fluent in Italian, were evaluated towards the end of the second term of the school year (April 2023), attending the 3rd and 4th grade of different primary schools, located in Southern (Puglia) and Northern Italy (Lombardy). The students were between 7 and 10 years old (M_{Age} = 8.45 years; SD = 0.55; 46.82% females). They were tested at school during school hours, in the classroom for collective tests and, in a reserved classroom of the same school, for individual tests.

4.1.2. Instruments and administration procedure

In addition to the tests specifically designed to evaluate L2 performance and the AEQ-ES for achievement emotions (both explained in Study 2), the verbal meaning section of Primary Mental Abilities (PMA – Rubini & Rossi, 1982) was used to assess vocabulary competence (L1). The test consists of 30 target words, each followed by 4 words. In the five minutes of the test, the students had to identify, for each target word, the identifiable word as a synonym among the four answer possibilities. The total score was calculated giving one point for each correct answer, and zero points for each incorrect or unassigned answer. Therefore, the final score ranged from a minimum of zero to a maximum of thirty.

The test was administered collectively in class.

4.1.3. Data analysis

Two different sets of analyses were conducted. First, considering the total sample (N = 182), correlational analyses and multiple linear regressions were carried out to investigate the relationship between the performance of English as L2, the performance in the L1 vocabulary (PMA) and the average score of emotions (joy, boredom, anxiety) perceived in the L2 learning contexts. Secondly, a second set of analyses was conducted considering a sub-sample (N=132) to verify whether conditions of low performance in the skill level related to L1, and/or conditions of risk of disorder could differently influence performance in L2 and the learning emotions associated with it. In particular, the 25th, 50th and 75th percentiles of the PMA-verbal meaning scores were calculated. Subsequently, only the data relating to the group of students with a score below the 25th percentile, renamed *Low*, and the one with a score above the 75th, renamed *High*, were used for subsequent t-test analyses for independent samples. All analyses, both those on the total sample (N = 182) and those on the sub-sample (N = 132) were performed using IBM SPSS Statistics software (version 29.0.1.0).

4.2. Results

4.2.1. Analysis on the total sample

4.2.1.1. Correlations

Table 18 reports the results of the Pearson correlation analysis between the PMA score indicative of vocabulary skills in L1 (M = 21.58, SD = 7.43), the vocabulary ability score in L2 (M = 10.69, SD = 4.94) and the total emotions of Joy (M = 3.05, SD = 1.08), Boredom (M = 1.98, SD = 1.00) and Anxiety (M = 1.82, SD = 0.74) experienced in L2 learning contexts. In line with expectations, a significant positive correlation was observed between scores relating to Italian vocabulary (L1) and English vocabulary (L2). In addition, L2-related scores had a significant positive correlation with Joy score and negative with Boredom and Anxiety scores. Finally, the score for L1 correlated negatively with Joy, but no significant associations with Boredom and Anxiety were found.

	L2 Vocabulary	Joy	Boredom	Anxiety
L1 Vocabulary	.475**	–.147*	.017	.023
L2 vocabulary		.126*	–.148*	–.283**

** p < .001; * p < .05. (one-tailed)

Table 18. Pearson correlation between L1, L2 measures and emotional variables related to L2.

4.2.1.2. Multiple linear regression

A multiple linear regression was conducted using gender, age, L1 Vocabulary, and the three achievement emotions as predictors to verify the role of L1 Vocabulary and L2-related emotions (Joy, Boredom and Anxiety) in predicting L2 Vocabulary performance (dependent variable). The results showed a significant positive effect of the L1 predictor (R^2 = .226), a significant negative effect of the Age predictor (R^2 = .090) and a significant effect of the Anxiety predictor (R^2 = .055), for a total variance explained by the model equal to 37% (Table 19).

Model	R	R^2		Sum of squares	df	Mean Square	F	*Sign.*
1	.475	.226	Regression	997.65	1	997.65	52.52	*.000*
			Residual	3419.11	180	18.99		
			Total	4416.76	181			
2	.562	.316	Regression	1393.55	2	696.78	41.25	*.000*
			Residual	3023.21	179	16.88		
			Total	4416.76	181			
3	.609	.371	Regression	1637.83	3	545.94	34.97	*.000*
			Residual	2778.93	178	16.61		
			Total	4416.76	181			

Table 19. ANOVA regression models.

Gender, as well as the level of Boredom and Joy experienced in L2 learning, did not play a significant role in predicting performance in L2 (see Table 20), therefore we removed those variables. The results of the ANOVA showed Model 3 ($F_{(3, 178)}$ = 34.97, p < .001) as the best

one, for which L2 performance was predicted negatively by Age and Anxiety and positively by L1 ability. Therefore, the results suggested a positive relationship between skills measured by vocabulary tests in L1 and performance in L2, as well as a negative relationship between the Anxiety experienced during L2 learning and performance in L2. Furthermore, the results indicates that with increasing age there was a worsening of the vocabulary in L2. A possible explanation for this paradoxical effect could be linked to the intrinsic differences in the tests administered, which were differentiated for the two class levels involved, and built based on the school program provided for each class level. Therefore, this effect should be further investigated, to better specify whether it could be associated with an actual cohort effect or whether it was due to a methodological nature.

	Variable	B	SE	β	t	*Sign.*	Lower limit 95% CI	Upper limit 95% CI
Model 1	(Constant)	3.87	.99		3.89	.000	1.91	5.83
	L1 vocabulary	.31	.04	.47	7.24	.000	.23	.40
Model 2	(Constant)	25.99	4.66		5.57	.000	16.78	35.18
	L1 vocabulary	.38	.04	.57	8.81	.000	.29	.46
	Age	–3.11	.64	–.31	–4.84	.000	–4.38	–1.84
Model 3	(Constant)	25.01	4.48		5.57	.000	16.14	33.86
	L1 vocabulary	.37	.04	.56	8.97	.000	.29	.45
	Age	–2.58	.63	–.26	–4.07	.000	–3.83	–1.33
	Anxiety	–1.59	.40	–.24	–3.95	.000	–2.38	–.79

Note: B = non-standardized beta coefficient; β = standardized beta coefficient; SE = standard error. CI = 95% Confidence Interval.

Table 20. Multiple linear regression with Vocabulary in L2 as a dependent variable.

4.2.2. Analysis on the partial sample

The total sample was divided into three groups to analyze possible differences between students who reported very low scores in the L1 Vocabulary test: *Low* (PMA score ≤ 17); *Medium* (average competence, 17 < PMA score > 24); *High* (PMA score ≥ 24). The data

relating to the 132 students who were in the extreme groups were used for the subsequent analyses (see Table 21).

PMA Performance	Frequency	Percentage
Low	64	35.2
Medium	50	27.5
High	68	37.4
Total	182	100.0

Table 21. Frequency of performance bands at the PMA (Vocabulary in L1).

A series of independent t-tests were conducted to verify the presence of differences between the *High* and *Low* score group in L1 regarding the scores obtained in the L2 tests (Vocabulary Entry, Vocabulary Output, Sentence Building and Total Score) and the emotions felt during L2 learning (Joy, Boredom, Anxiety).

The results showed that the two groups did not differ significantly regarding any of the emotions analyzed (Joy: $t = 1.31$, $p > .05$; Boredom: $t = –.714$, $p > .05$; Anxiety: $t = –.455$, $p > .05$) while the difference was statistically significant for all the other dependent variables. In particular, the *High* group reported better performance than the *Low* group in Vocabulary Entry scores ($M_{High} = 5.00$, $SD_{High} = 1.46$; $M_{Low} = 3.73$, $SD_{Low} = 1.75$; $t = –4.50$, $p < .001$, see Figure 8), Vocabulary Output ($M_{High} = 3.26$, $SD_{High} = 1.68$; $M_{Low} = 2.59$, $SD_{Low} = 1.89$; $t = –2.15$, $p < .05$, see Figure 9), Sentence Building ($M_{High} = 4.41$, $SD_{High} = 3.25$; $M_{Low} = 1.54$, $SD_{Low} = 2.23$; $t = –5.86$, $p < .001$, see Figure 10), and in the total score of Vocabulary at L2 ($M_{High} = 12.68$, $SD_{High} = 4.41$; $M_{Low} = 7.88$, $SD_{Low} = 4.22$; $t = –6.38$, $p < .001$).

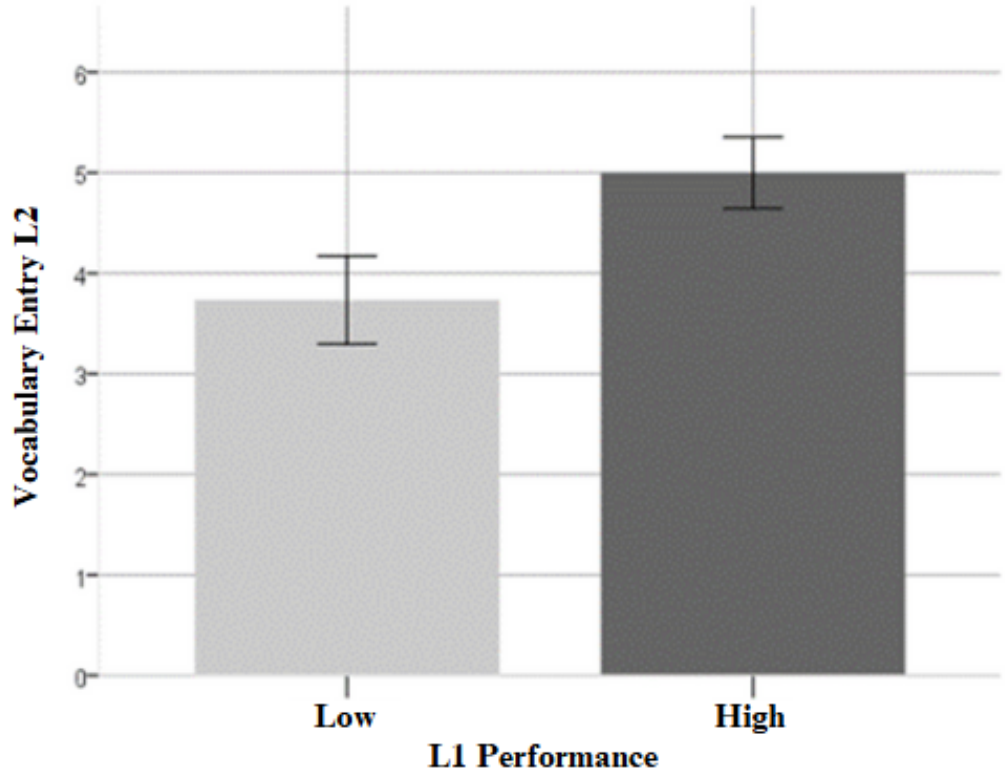


Figure 8. Differences between students with High or Low Level in L1 performance relative to L2 Vocabulary Entry.

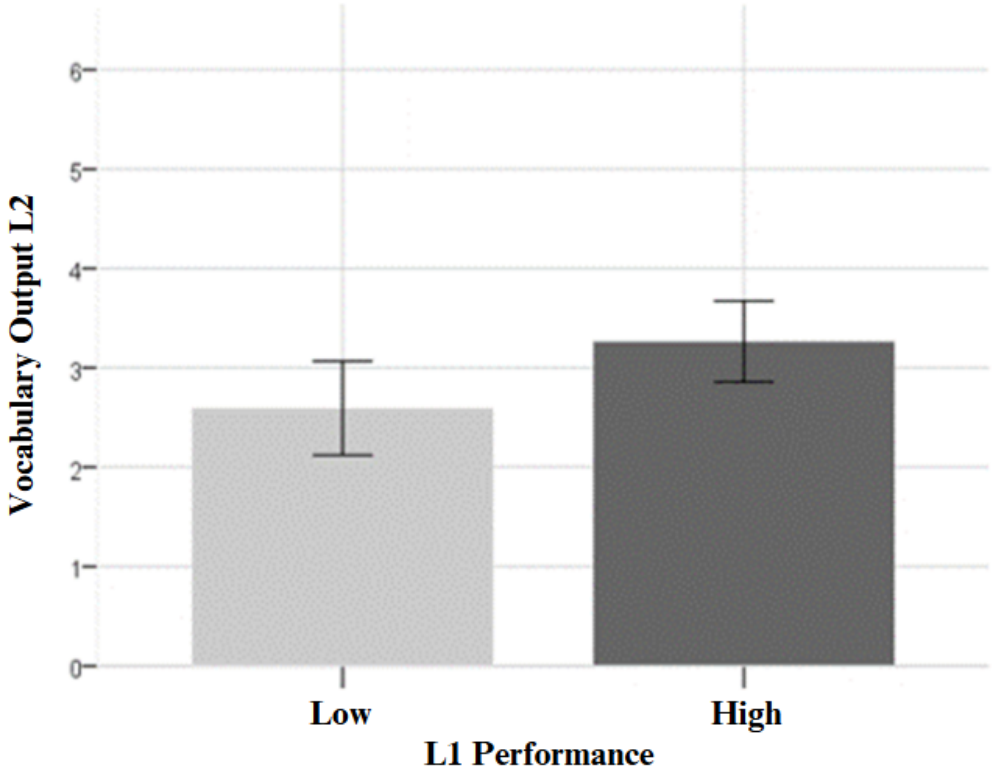


Figure 9. Differences between students with High or Low Level in L1 performance relative to L2 Vocabulary Output.

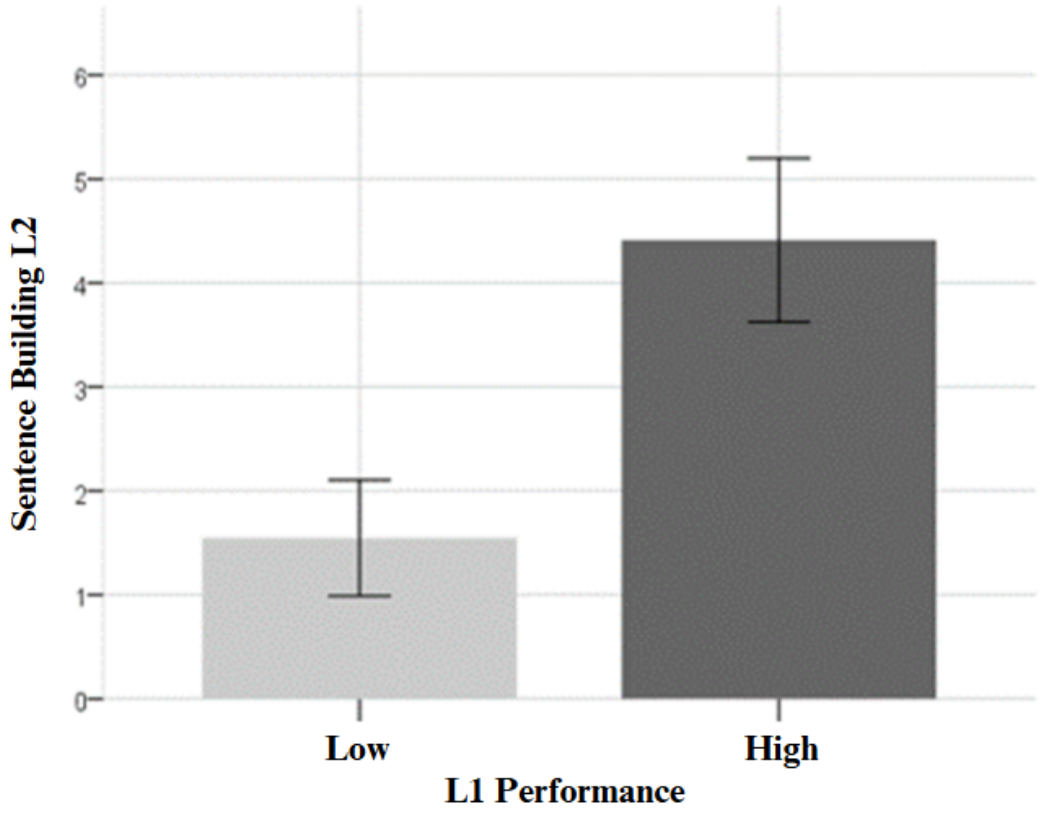


Figure 10. Differences between students with High or Low Level in L1 performance relative to L2 Sentence Building.

4.3.Discussion

The research conducted allows us to understand the role of competence in the mother tongue (the knowledge of vocabulary in Italian) and emotions experienced in L2 learning activities in primary school children. The results showed, in line with the existing literature referring mostly to secondary school students and young adults (Ferrari & Palladino, 2007, 2012; Palladino et al., 2013; Palladino et al., 2016; Sparks 2019), that from the early stages of studying L2 English to primary school, the significant role of language skills in L1 is found. This is evident both from regression analyses, which showed that vocabulary competence in L1 was a strong predictor of L2, and from the comparison between students with low and high vocabulary in L1. The comparison revealed rather limited performance in L2 of children with low vocabulary in the mother tongue, significantly lower than children with very good skills in the mother tongue. The analysis of performance in L2 English showed a very marked gap in the Sentence Building task in which children with low L1 vocabulary correctly produced only one and a half words and they were, therefore, in clear difficulty compared to children with high vocabulary. The latter could construct a short sentence (question or

answer) with about 4 correct words produced on average in the Sentence Building task, while peers with low performance in vocabulary in L1 were unable to construct a short sentence and respond with a single word or at most two. The mother tongue and, in particular, the vocabulary in Italian, was confirmed in this age group as a predictive competence of L2 learning. This result is in line with the results present in the literature for older populations and education (Abbasian et al., 2020; Harrison & Krol, 2007; Lee, 2014; Vulchanova et al., 2014), confirming the relationship between L1 and L2 even in primary school children, in a phase of strong learning development, in which a first basic competence of L2 is also started. These results are particularly relevant as the research has focused mainly on other age groups and on more advanced L2 skills and almost absent are the studies on the primary school group. Our study also collected the measurement of the emotional components of the learning process. This component, often neglected, is considered crucial in many models that represent learning from a perspective of self-regulation and relationship with the context (Boekaerts, 1996; Efklides, 2011; Puustinen & Pulkkinen, 2001). The study of emotions in English language learning at school highlights a significant role of anxiety as a determinant of learning performance. It seems worthy of further investigation a negative emotion predicted English performance in primary school children, naturally in a direction that corresponded to greater anxiety and less learning. However, this mechanism appears less clear than expected because the comparison between children more and less competent in L1 did not reveal a greater experience of anxiety about L2 learning in the group of less competent, although a significant difference in performance in L2 emerged. The anxiety experienced by children in L2 learning therefore appears to be an important factor but independent of the skills already acquired in the L1 vocabulary. It should also be noted that anxiety measured in this way, like other emotions, is the sum of the scores indicated by the participants in three different learning contexts: lessons, homework and classroom tests. The score is therefore representative of an emotion experienced across different activities and environmental and social learning contexts.

4.4. Conclusion

In conclusion, the present study highlights a link between vocabulary learning in the mother tongue and in L2 English in Italian primary school children in line with the literature on other

age groups and levels of schooling. The emotions experienced in L2 learning also seem to play a relevant role: anxiety about L2 appears to be predictive in a negative sense of performance in L2. However, the comparison between children with low and high vocabulary in L1 does not highlight different emotional experiences, although it highlights a different performance in L2. Future studies could help shed light on the role of the emotional component of L2 learning both in the direction of a greater understanding of the phenomenon and to evaluate teaching approaches that consider the emotions felt in learning contexts.

5. Study 4: Analysis of the psychometric properties of the English Metacognitive Questionnaire for Primary School (EMQ–PS)

This chapter refers to the study of Trotta et al. (under review [B]). Here below a list of tables and figure included in this chapter.

List of tables included in Study 4.

Table 22	**Descriptive statistics of the sample.**	84
Table 23	**Means, Standard Deviation, Skewness, and Kurtosis.**	88
Table 24	**Correlation matrix between individual items.**	89
Table 25	**Preliminary Exploratory Factor Analysis Models.**	91
Table 26	**Exploratory Factor Analysis: Comparison of Factor Solutions.**	91
Table 27	**Principal Component Analysis (Matrix of rotated components).**	96
Table 28	**Correlation matrix between the EMQ–PS total and the dimensions.**	97
Table 29	**Descriptive characteristics and correlations between the identified dimensions.**	98
Table 30	**EMQ–PS means and standard deviations (SD).**	98
Table 31	**EMQ–PS performance bands.**	99
Table 32	**Frequency distributions of attributive choices.**	100
Table 33	**Correlation matrix (Students*Teachers).**	102

List of figures included in Study 4.

Errore. L'origine riferimento non è stata trovata.	Errore. L'origine riferimento non è stata trovata.	**Errore. Il segnalibro non è definito.**
Figure 12	**Exploratory Factor Analysis Models: 3 and 4 Latent Factor Solutions.**	93
Figure 13	**Scree plot**	94

As reported, the role of metacognition in learning is widely recognized (for a review on the subject, see Wang, Haertel, & Walberg, 1990). Roebers et al. (2012) evaluated the

metacognitive monitoring and control skills of primary school students on their homework. The task consisted of writing certain words from images of objects and animals. Next, they were asked to rate a judgment of confidence about how confident they were that the word was spelled correctly (metacognitive monitoring) through a seven-point Likert scale. Finally, participants had the opportunity to delete previously written words if they felt they were spelled incorrectly. Metacognitive control was quantified using the number of words deleted or maintained adequately or incorrectly. To give an example of adequate control: words written correctly and maintained; misspelled and withdrawn words. Their results revealed that both metacognitive monitoring and control (as well as executive functions, which were also investigated) are significantly associated with students' academic achievement.

To measure metacognitive awareness in adults, Schraw and Dennison (1994) developed the Metacognitive Awareness Inventory (MAI), a 52-item instrument designed to investigate two specific factors, namely cognition knowledge and cognition regulation. Also designed for adult populations, Pedone et al. (2017) developed The Metacognition Self-Assessment Scale (MSAS), an 18-item self-report tool on a 5-point Likert scale, aimed at investigating five metacognitive skills: 1) monitoring; 2) differentiation; 3) integration; 4) decentralization and 5) mastery. Kreutzer et al. (1975) proposed a tool to assess children's metamemory, using a battery of semi-structured interviews to investigate their knowledge of everyday events. The scales presented are just a few examples of the countless tools available in the literature for the measurement of metacognitive functioning (for a review, see Smith et al., 2023). However, the currently available tools for measuring metacognitive skills have some disadvantages. First, they are mainly aimed at adolescent or adult populations (e.g., Pedone et al., 2017). When designed for younger populations, as in the case of primary school students, most existing scales are constructed to provide a hetero assessment of children's metacognitive abilities, referring to the opinions of parents and teachers (Gioia, 2000) or referring to video observations (Whitebread et al., 2009). Furthermore, the scales aimed directly at primary school students investigate one or two specific aspects of metacognition (metacognitive knowledge; monitoring; control; etc.), and they do not consider other components, such as metacognitive experiences, goals or strategies (e.g.: Amsterlaw, 2006; Meyer et al., 2010; Roebers et al., 2009).

To our knowledge, there is no validated tool for primary aimed at investigating the role of metacognitive skills in learning English as L2. The only instrument on this topic is the English

Metacognitive Questionnaire (EMQ), developed by Palladino et al. (2019) to investigate metacognitive abilities in Italian secondary school students. The EMQ evaluates metacognitive and motivational aspects of the individual student on L2, highlighting functional and dysfunctional aspects in the system of meta-knowledge, beliefs and values, and to conduct the setting of a possible enhancement intervention in the most appropriate direction. Having highlighted the role of metacognitive competences since primary school (see, for example: De Jager et al., 2005; Roebers et al., 2012; Williams & Atkins, 2009), and since there is currently no tool for the evaluation of the latter in the learning of English as L2, we analyzed the psychometric properties of EMQ in a sample of Italian primary school students (*English Metacognitive Questionnaire for Primary School - EMQ–PS*).

5.1. Introducing the original questionnaire

The EMQ (Palladino et al., 2019) is a self-assessment questionnaire consisting of 31 items. The first 29 items express, with respect to each statement, the degree of agreement of the individual student through a four-point Likert scale (1 = against; 2 = somewhat against; 3 = somewhat agree; 4 = agree). Of the 29 items, 13 are reverse, while the remaining 16 items have a formulation that follows the direction of the general construct. The 29 items investigated four specific areas of metacognitive abilities, as identified within the study by Palladino et al. (2019):

- Beliefs ("*Credenze"*), composed of eight items: it refers to the skills (general and personal) that students may or may not consider relevant for L2 learning (e.g.: "*Con l'impegno e l'esercizio tutti possono imparare l'inglese*"; "*Per leggere bene in inglese è utile conoscere il significato di molte parole*").
- Experienced ("*Vissuti"*), eight items: the emotional and affective experiences of students towards L2 learning (e.g.: "*Durante le verifiche di inglese mi sento agitato, provo malessere più che in altre materie*"; "*Durante i compiti di inglese mi sento tranquillo e fiducioso*").
- Strategies (*"Strategie"*) eight items: this area investigates the conscious attitude of students, understood as control and strategic aspects that they can use for L2 learning (e.g.: "*Vedere film, cartoni animati o ascoltare canzoni di inglese è utile per*

migliorare l'inglese"; "*Quando leggo in inglese penso a quello che so sull'argomento*").

- Utility (*"Utilità"*), five items: students' perception of usefulness in L2 learning (e.g., "*Sapere l'inglese è importante per ottenere un buon lavoro*"; "*Studiare l'inglese serve solo per ottenere buoni voti*").

The last two items (30 and 31) measure a dimension called *Attributions*, which investigates the causes of success and failure perceived by students in L2 learning. Both items have five answer options. The student is asked to choose, for each item, the two answers in order of importance that best indicate (in the student's opinion) the reasons related to their success and failure. Alternative answers refer to the following causes: commitment, skill, help, characteristics of the task, luck.

5.2. Method

5.2.1. Sample and procedure

We involved 292 primary school students (M_{Age} = 9.17 years; SD = .63; range = 7-11 years) from three different Italian regions, namely Lombardy, Puglia, and Calabria. Participation was strictly linked to obtaining parental consent. Twelve students were excluded because they did not receive such consent. In addition, twelve more students with a certified diagnosis of intellectual disabilities, neurological or language disorders, with specific learning disorders and/or other diagnoses were excluded from the data analysis. The final sample considered for the analyses is composed of 258 students ($N_{females}$ = 130; 46% of the sample) attending the 3rd ($N_{females}$ = 62) and 4th ($N_{females}$ = 68) primary school in the 2022/2023 academic year (Table 22). All students involved were native Italian speakers or bilingual, with fluent verbal ability in the Italian language. The study was conducted according to the ethical guidelines of the Italian Association of Psychology (AIP) and approved by the Ethics Committee of the Psychology Section of the University of Pavia (approval of project number 34/19, Year 2019).

Gender	N	Age (years)		Grade	
		Mean	SD	3rd	4th
Male	150	9.16	.64	70	80

Female	130	9.18	.63	62	68
Total / Mean	280	9.17	.63	132	148

Table 22. Descriptive statistics of the sample.

The content of the questionnaire, originally designed for lower secondary school students, has been adapted for primary school pupils. This adaptation facilitated primary school students, making each question more in line with their level of cognitive and linguistic development, simplifying the language, maintaining the original structure and content of the questionnaire. The administration procedure took place collectively within the class group, in self-assessment using paper format. The average compilation time of the entire battery was about twenty minutes, but no time limitation was imposed. Unlike the original version, three initial items were introduced, with content unrelated to the research topic, to allow each student to familiarize themselves with the type of statements and response methods they would face (e.g.: *"Mi piacciono i cani"* – see **Errore. L'origine riferimento non è stata trovata.**); these introductory items were excluded from the analyses. Once the questionnaire was delivered to each student, the researcher introduced the method of compilation by reading the initial familiarization items to the entire class group. Afterwards, students were told that there were no right or wrong answers, but personal ones, encouraging an honest answer. At the end of the familiarization with the test, the students proceeded independently, aware that they could ask for help from the researcher and/or reference teacher in case of non-understanding of one or more items.

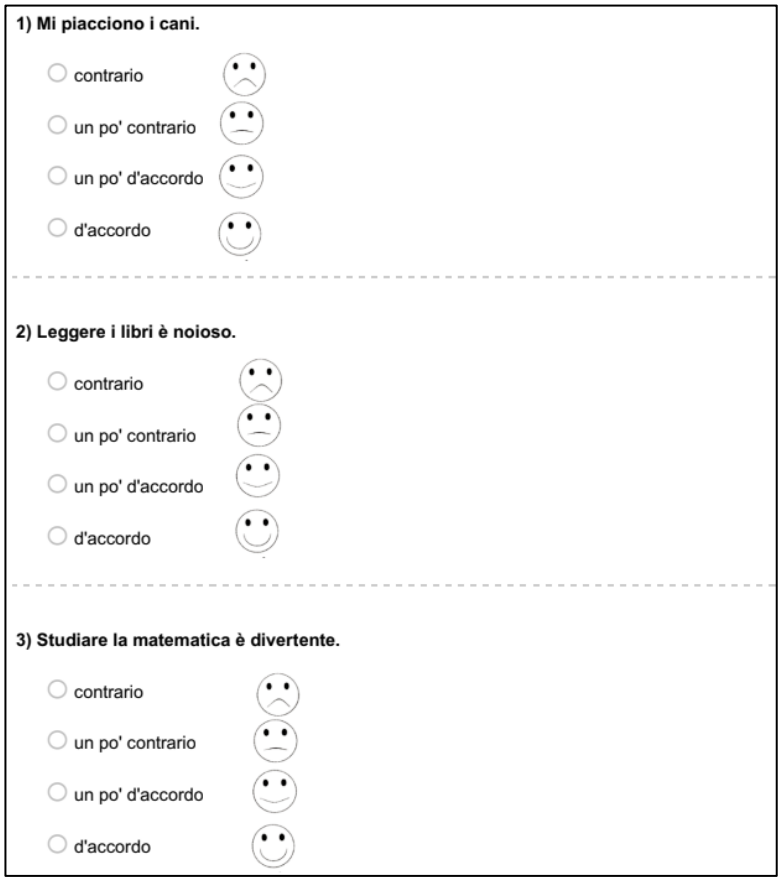


Figure 11. Familiarization item.

5.2.2. Teachers' perception

A version of the EMQ was administered to the teachers of the classes involved. This version, aimed to collect information from teachers about their opinion on the metacognitive skills of each student in their classes. These data were collected to compare them with the students' self-evaluations, to highlight the degree of awareness of teachers about the metacognitive skills of their students (Carr & Kurtz, 1991).

5.2.3. Data analysis

All analyses were carried out through IBM SPSS Statistics (version 28.0.1.0), Jamovi (version 2.4.0 – The jamovi project, 2023; R Core Team, 2022; Revelle, 2023) and M*Plus* 8.11 (Muthén & Muthén, 2017, 2023).

At the preliminary level, the descriptive statistics of the instrument were analyzed, verifying the normality of the frequency distributions of each item and each subscale, as well as

calculating the mean scores and standard deviations. The descriptive reliability of each dimension of the instrument was then analyzed, considering two distinct indices: the Cronbach's alpha through the Spearman-Brown Split-half method (Warren, 2015) and, since it is a scale with ordinal values (Likert scale), it was considered appropriate to also consider the Omega index (McDonald, 1999) on the polychoric matrix (McNeish, 2018).

Several steps were conducted for the analysis of the factorial structure of the instrument. First, an exploratory factor analysis (EFA; Tucker & MacCallum, 1997), to explore the multivariate structure of items, and reduce the size of the instrument by identifying the fewest non-directly observable latent variables (or latent factors) into which the instrument organizes itself (Watkins, 2018). A robust estimation method (MLR – Robust Maximum Likelihood Estimation; Maydeu-Olivares, 2017) and the following criteria to evaluate the goodness of the fit of the models: a) the χ^2 test, with the associated indices greater than .95 (Comparative-Fit Index – CFI, and Tucker-Lewis-Fit Index – TLI; Hu & Bentler, 1999); b) the value of RMSEA (Root Mean Square Error of Approximation) and its confidence intervals less than .05 (Kline, 2023); c) the value of SRMR (Standardized Root Mean Square Residual) less than .06 (Kline, 2023).

We intended to compare the structure of the EMQ-PS with the structure of the original EMQ for the lower secondary school sample (Palladino et al., 2019). Therefore, we conducted a confirmatory factor analysis (CFA; Watkins, 2018), in line with the procedure carried out by the comparison study. In line with the comparison study that presents the original version of the scale (Palladino et al, 2019), the Principal Component Analysis (PCA) method was used (Fabrigar et al., 1999).

We used both EFA and PCA for several reasons. First, we carried out the same analyses of the original scale (Palladino et al., 2019), to compare it with the present study. Hence, we used the principal component analysis. However, PCA and EFA have two distinct objectives. PCA is a multivariate technique that aims to reduce the dimensionality of a scale, i.e. to simplify a complex set of data in which observations are described by several interrelated quantitative dependent variables, preserving as much as possible the information contained in the original data, and analyzing the entire correlation matrix of the scale (Abdi & Williams, 2010; Watkins, 2018). EFA, on the other hand, aims to identify latent variables underlying the functioning of the scale, assuming that the identifiable latent factors are influenced by the basic macro-construct (Cudeck, 2000; Tucker & MacCallum, 1997).

5.3. Results

5.3.1. Preliminary analysis

A first step of descriptive analysis on the individual items was conducted, analyzing the mean values, standard deviations, skewness and kurtosis of the 29 quantitative items of the EMQ-PS, as reported in Table 23 (Orcan, 2020). This analysis indicates a normal distribution of the data, and it was therefore possible to proceed with the next steps.

Correlation matrix analysis is a critical step for the implementation of exploratory and/or confirmatory factor analysis, as well as for the implementation of principal component analysis (PCA), as the latter's task is to identify the fewest factors necessary to better explain the total variance observed between the elements of the correlation matrix (Grimm & Yarnold, 1995; Watkins, 2018). The correlations between the individual items were then analyzed, as reported Table 24. The results highlighted correlation indices suitable for the implementation of both exploratory and confirmatory factor analyses (Watkins, 2018).

Item		Mean	SD	Skewness	SE	Kurtosis	SE
1.S	Guardo film, cartoni animati o ascolto canzoni in inglese per migliorare il mio inglese	2.56	1.204	–0.0887	0.152	–1.539	0.303
2.S	Se sbaglio un esercizio, mi esercito di più per fare meglio la prossima volta	3.44	0.943	–1.5848	0.152	1.269	0.303
3.U	Sapere l'inglese è importante per il mio futuro	3.60	0.769	–2.0883	0.152	3.757	0.303
4.B	Con l'impegno e l'esercizio, tutti possono imparare l'inglese	3.64	0.743	–2.0667	0.152	3.416	0.303
5.E*	Durante le verifiche di inglese mi sento agitato/a	2.71	1.208	–0.1629	0.152	–1.575	0.303
6.S	Se scrivo in inglese, dopo rileggo e controllo gli errori	3.22	1.071	–1.0974	0.152	–0.217	0.303
7.U*	Sapere l'inglese non è importante	3.47	0.980	–1.6603	0.152	1.292	0.303
8.B	È più facile studiare l'inglese se penso che mi sarà utile in futuro	3.30	0.992	–1.2354	0.152	0.286	0.303
9.E*	Se devo parlare in inglese provo vergogna	2.82	1.115	–0.2463	0.152	–1.412	0.303
10.S	Quando la maestra parla inglese, sto attento/a ai gesti o alle espressioni del suo volto	3.29	1.005	–1.1844	0.152	0.103	0.303
11.E*	Se sbaglio un esercizio, ci resto molto male e penso che io non sia bravo/a con l'inglese	2.84	1.208	–0.4073	0.152	–1.445	0.303
12.U	Sapere l'inglese è utile per viaggiare	3.71	0.717	–2.5904	0.152	5.978	0.303
13.S*	Spesso faccio i compiti di inglese con la TV o la musica accesa	3.10	1.190	–0.8340	0.152	–0.971	0.303
14.B*	L'inglese si impara solo se si è bravi	3.09	1.172	–0.8177	0.152	–0.954	0.302
15.U	I miei genitori pensano che sia importante sapere l'inglese	3.67	0.782	–2.5306	0.152	5.528	0.302
16.E*	Prima di fare un compito di inglese so già che sarà difficile	2.88	1.127	–0.4921	0.152	–1.190	0.302
17.B*	Imparare a memoria non è sempre utile perché non tutti hanno una buona memoria	2.38	1.210	0.1672	0.152	–1.532	0.303
18.S	Quando leggo in inglese mi aiuto guardando il titolo e le immagini	3.25	1.011	–1.1830	0.152	0.169	0.302
19.E*	Quando la maestra parla in inglese sono poco attento/a, perché so che faccio fatica a capire	3.03	1.171	–0.6864	0.152	–1.128	0.303
20.E	Mi piace parlare in inglese	3.00	1.110	–0.7289	0.152	–0.870	0.303
21.E	Durante i compiti di inglese mi sento tranquillo/a e fiducioso/a	2.97	1.100	–0.6218	0.152	–1.008	0.303
22.B*	Se non ricordo è perché non so	3.02	1.049	–0.6239	0.152	–0.922	0.303
23.B	Per imparare l'inglese è utile ascoltarlo o leggerlo tutti i giorni	3.19	0.969	–0.9704	0.152	–0.160	0.303
24.B*	Rileggere in inglese è inutile, perché se non si capisce subito, rileggendo non si migliora	3.35	1.056	–1.3460	0.152	0.274	0.303
25.U*	Studiare l'inglese serve solo per ottenere buoni voti	3.19	1.089	–1.0437	0.152	–0.365	0.303
26.S	Quando leggo in inglese penso a quello che so sull'argomento	3.22	0.993	–1.0366	0.152	–0.128	0.303
27.S	Se non so come si dice una cosa in inglese mi aiuto con i gesti	2.61	1.198	–0.1795	0.152	–1.502	0.303
28.E*	Quando non riesco a fare un esercizio di inglese, mi prende l'ansia	2.57	1.228	–0.0734	0.152	–1.589	0.302
29.B	Per leggere bene in inglese, bisogna sapere il significato di molte parole	3.38	0.815	–1.3230	0.152	1.262	0.302

Note: N=258. B=Beliefs; E=Experienced; S=Strategies; U=Utilities. * = item reverse.

Table 23. Means, Standard Deviation, Skewness, and Kurtosis.

	1.S	2.S	3.U	4.B	5.E	6.S	7.U	8.B	9.E	10.S	11.E	12.U	13.S	14.B	15.U	16.E	17.B	18.S	19.E	20.E	21.E	22.B	23.B	24.B	25.U	26.S	27.S	28.E
1.S	—																											
2.S	0.21***	—																										
3.U	0.27***	0.2**	—																									
4.B	0.25***	0.28***	0.27***	—																								
5.E	0.09	—0.10	0.03	0.08	—																							
6.S	0.25***	0.38***	0.38***	0.24***	0.03	—																						
7.U	0.07	0.16*	0.37***	0.13*	0.06	0.17**	—																					
8.B	0.26***	0.26***	0.49***	0.34***	—0.02	0.35***	0.30***	—																				
9.E	0.05	0.08	0.05	0.14*	0.33***	—0.01	0.12	0.04	—																			
10.S	0.25***	0.23***	0.42***	0.33***	0.03	0.28***	0.25***	0.37***	0.09	—																		
11.E	0.10	0.10	0.03	0.14*	0.41***	0.05	0.18**	0.05	0.27***	0.04	—																	
12.U	0.13*	0.20**	0.36***	0.20**	0.03	0.32***	0.12*	0.32***	0.08	0.16*	0.06	—																
13.S	—0.05	—0.04	—0.08	0.04	0.13*	—0.10	—0.02	—0.13*	0.03	—0.10	0.10	0.03	—															
14.B	0.00	0.03	0.10	0.07	0.13*	0.03	0.18**	0.06	—0.06	0.01	0.11	0.09	0.12	—														
15.U	0.19**	0.20**	0.43***	0.21***	—0.04	0.37***	0.24***	0.28***	—0.01	0.32***	—0.02	0.24***	—0.05	0.08	—													
16.E	0.20**	0.04	0.23***	0.21***	0.41***	0.09	0.27***	0.13*	0.41***	0.19**	0.41***	0.05	0.07	0.10	0.13*	—												
17.B	—0.04	0.04	0.00	0.09	0.08	0.07	0.18**	0.01	0.20**	—0.01	0.11	0.02	—0.03	0.03	0.01	0.15*	—											
18.S	0.04	0.27***	0.07	0.17**	—0.07	0.16*	—0.01	0.14*	—0.11	0.18**	0.01	0.02	—0.11	—0.10	0.06	—0.08	0.05	—										
19.E	0.21***	0.20**	0.19**	0.23***	0.26***	0.24***	0.25***	0.15*	0.37***	0.21***	0.29***	0.15*	—0.01	0.10	0.10	0.47***	0.21***	—0.01	—									
20.E	0.44***	0.28***	0.42***	0.34***	0.25***	0.27***	0.31***	0.36***	0.31***	0.36***	0.24***	0.16*	—0.00	0.08	0.27***	0.38***	0.05	0.10	0.29***	—								
21.E	0.30***	0.17**	0.34***	0.29***	0.32***	0.27***	0.26***	0.28***	0.31***	0.35***	0.28***	0.14*	—0.02	0.13*	0.20**	0.45***	0.18**	0.03	0.42***	0.61***	—							
22.B	0.08	0.02	0.02	0.09	0.20**	0.04	0.22***	0.04	0.20**	0.12	0.30***	0.06	0.04	0.21***	0.06	0.35***	0.19**	—0.07	0.25***	0.18**	0.29***	—						
23.B	0.18**	0.10	0.35***	0.27***	—0.04	0.23***	0.07	0.38***	0.06	0.32***	—0.08	0.22***	—0.05	—0.14*	0.27***	0.06	—0.08	0.10	0.18**	0.31***	0.33***	—0.02	—					
24.B	0.14*	0.09	0.17**	0.19**	0.14*	0.16**	0.19**	0.09	0.12*	0.14*	0.19**	0.23***	0.05	0.17**	0.19**	0.18**	0.11	0.12	0.24***	0.29***	0.15*	0.10	0.05	—				
25.U	0.02	0.01	0.10	—0.06	0.12	—0.02	0.21***	0.04	0.11	0.04	0.15*	0.16**	0.04	0.40***	0.02	0.17**	0.10	—0.11	0.16*	0.033	0.03	0.18**	—0.14*	0.30***	—			
26.S	0.22***	0.30***	0.22***	0.22***	0.09	0.23***	0.08	0.24***	0.04	0.31***	0.01	0.17**	—0.12*	0.02	0.29***	0.12	—0.08	0.15*	0.22***	0.16*	0.25***	—0.00	0.25***	0.08	—0.02	—		
27.S	0.10	0.19**	0.17**	0.08	—0.14*	0.10	0.03	0.24***	—0.06	0.28***	—0.04	—0.04	— 0.24***	0.00	0.10	—0.05	—0.11	0.25***	—0.09	0.19**	0.15*	—0.03	0.23***	—0.05	—0.09	0.26***	—	
28.E	0.19**	0.02	0.08	0.13*	0.54***	0.15*	0.14*	0.06	0.43***	0.16*	0.45***	0.02	0.13*	0.07	0.07	0.51***	0.19**	—0.10	0.31***	0.35***	0.37***	0.24***	0.07	0.16*	0.16**	0.10	— 0.10	—
29.B	—0.04	0.10	0.19**	0.03	0.05	0.16**	0.08	0.11	—0.04	0.03	0.03	0.09	0.13*	—0.04	0.20**	0.01	0.03	0.03	0.06	0.05	0.08	—0.06	0.10	—0.01	—0.02	0.11	0.01	0.05

* p < .05, ** p < .01, *** p < .001

Table 24. Correlation matrix between individual items.

5.3.2. Exploratory Factor Analysis

As a first step in the exploration of the instrument structure, an exploratory factor analysis was conducted, with factorization of the main axis as the extraction method, orthogonal rotation according to the Varimax method to verify the dimensionality of the instrument, and robust estimation method for continuous variables (MLR; Maydeu-Olivares, 2017; Watkins, 2018). The 1- to 4-factor models were estimated, and each model was compared to the K-1 factor model, to preserve the most parsimonious model (Wikrama et al., 2016). Principal axis factoriality was chosen as it focuses exclusively on the common variance between the variables considered, leaving out the specific variance of each variable (Barbaranelli, 2003). Compared to the rotation method, the factors were correlated with each other with correlations less than |. 30|, confirming a non-oblique solution. The models obtained using all 29 items (Models 1-4; Table 25) showed an ambiguous factor structure, characterized by items without significant saturation on any factor (items 13, 17 and 29), and items with a saturation greater than .3 in two or more factors (items 7, 20, 21). Eliminating these items, the exploratory factor analysis showed that the 3-factor and 4-factor models represented a stronger factor structure, and a better fit (Table 25; Table 26).

		χ^2 (df)	CFI	TLI	RMSEA	SRMR	Models' comparison	χ^2 diff	Δ df
Models a 1-4	Model a1	959.662 (377) ***	.567	.534	.077 [.071-.083]	.097			
(29 items)	Model a2	539.456 (349) ***	.859	.836	.046 [.038-.053]	.055	a1 vs. a2	446.894 ***	28
	Model a3	443.715 (322) ***	.910	.886	.038 [.029-.047]	.045	a2 vs. a3	80.775 ***	27
	Model a4	413.995 (296) ***	.912	.880	.039 [.030-.048]	.041	a3 vs. a4	32.680 (n.s.)	26
Models b 1-4	Model b1	832.305 (299) ***	.581	.544	.083 [.076-.090]	.101			
(26 items)	Model b2	441.656 (274) ***	.868	.844	.049 [.040-.057]	.056	b1 vs. b2	432.855 ***	25
	Model b3	347.060 (250) ***	.924	.901	.039 [.028-.048]	.043	b2 vs. b3	81.865 ***	24
	Model b4	314.261 (227) ***	.931	.902	.039 [.028-.049]	.038	b3 vs. b4	32.405 (n.s.)	23
Models c 1-4	Model c1	789.568 (275) ***	.538	.496	.085 [.078-.092]	.101			
(25 items)	Model c2	381.164 (251) ***	.883	.860	.045 [.035-.054]	.055	c1 vs. c2	572.025 ***	24
	Model c3	289.209 (228) **	.945	.928	.032 [.019-.043]	.042	c2 vs. c3	79.738 ***	23
	Model c4	252.701 (206) *	.958	.939	.030 [.014-.041]	.036	c3 vs. c4	32.694 (n.s.)	22
Models d 1-4	Model d1	743.529 (252) ***	.511	.465	.087 [.080-.094]	.106			
(25 items)	Model d2	355.954 (229) ***	.874	.848	.046 [.037-.056]	.056	d1 vs. d2	505.034 ***	23
	Model d3	268.721 (207) ***	.939	.918	.034 [.021-.045]	.042	d2 vs. d3	76.643 ***	22
	Model d4	234.147 (186) ***	.952	.929	.032 [.017-.044]	.036	d3 vs. d4	31.134 (n.s.)	21
Models e 1-4	Model e1	706.926 (230) ***	.501	.451	.090 [.082-.097]	.108			
(25 items)	Model e2	319.712 (208) ***	.883	.858	.046 [.035-.055]	.054	e1 vs. e2	546.285 ***	22
	Model e3	244.017 (187) **	.940	.919	.034 [.021-.046]	.042	e2 vs. e3	66.763 ***	21
	Model e4	207.593 (167) *	.957	.936	.031 [.014-.043]	.036	e3 vs. e4	32.087 *	20

Note: * p < .05, ** p < .01, *** p < .001, n.s. = not significant.
Models b: items 13, 17 and 29 excluded. Models c: item 20 excluded. Models d: item 21 excluded. Models e: item 7 excluded.

Table 25. Preliminary Exploratory Factor Analysis Models.

	χ^2 (df)	CFI	TLI	RMSEA	SRMR	Models' comparison	χ^2 diff	Δ df
Model e1	706.926 (230) ***	.501	.451	.090 [.082-.097]	.108			
Model e2	319.712 (208) ***	.883	.858	.046 [.035-.055]	.054	e1 vs. e2	546.285 ***	22
Model e3	244.017 (187) **	.940	.919	.034 [.021-.046]	.042	e2 vs. e3	66.763 ***	21
Model e4	207.593 (167) *	.957	.936	.031 [.014-.043]	.036	e3 vs. e4	32.087 *	20

Table 26. Exploratory Factor Analysis: Comparison of Factor Solutions.

In particular, the 3-factor model (Model e3) had an adequate fit [χ^2 (187) = 244.017, p < .01; CFI = .940; RMSEA = .034 (.021 - .046)], and highlighted the presence of the following dimensions: Factor 1 – "*Beliefs and Strategies*", representative of each student's perception of their metacognitive skills in L2 learning (e.g., "*Per imparare l'inglese è utile ascoltarlo o leggerlo tutti i giorni*", or "*È più facile studiare l'inglese se penso che mi sarà utile in futuro*"), as well as the individual strategies put in place to learn the language (e.g., "*Se scrivo in inglese, dopo rileggo e controllo gli errori*", or "*Quando leggo in inglese mi aiuto guardando il titolo e le immagini*"); Factor 2 – "*Experienced*", representative of the emotional aspects related to performance and learning L2 (e.g., "*Durante le verifiche di inglese mi sento agitato"*, or *"Se devo parlare in inglese provo vergogna*"); Factor 3 – "*Attitudes*", representative of the aspects of belief related to L2 learning (e.g., *"L'inglese si impara solo se si è bravi"*, or *"Studiare l'inglese serve solo per ottenere buoni voti"*). In terms of correlations between factors: the second and third factors showed a moderate positive association (r = .240); a positive association was also found between the second and first factors, albeit to a lesser extent (r = .145); the first and third factors were not significantly associated with each other. The 4-factor model (Model e4) had an adequate fit [χ^2 (167) = 207.593, p < .05; CFI = .957; RMSEA = .031 (.014 - .043)] comparatively better than the 3-factor model [$\Delta\ \chi^2$ (20) = 32.087; p < .05]. In terms of identifying latent dimensions, this 4-factor model considered the factor structure of the tool in a more precise way, as the additional factor allowed a clearer division between the metacognitive components and those related to strategies (see Figure 12). In particular, this 4-factor structure identified the following dimensions: Factor 1 – "*Beliefs*", representative of the metacognitive aspects related to L2 learning (e.g., "*Sapere l'inglese è importante per il mio futuro*"); Factor 2 – "*Strategies*", uniquely representative of the strategies that the student implemented to learn the language (e.g., "*Se non so come si dice una cosa in inglese, mi aiuto con i gesti*"); Factor 3 – "*Experienced*", representative of the emotional aspects related to L2 learning, completely superimposable to the corresponding factor in the 3-way solution, to which is added an item relating to a metacognitive aspect ("*Con l'impegno e l'esercizio tutti possono imparare l'inglese*"); Factor 4 – "*Attitudes*", representative of the aspects of belief related to L2 learning, and overlapping with the corresponding factor in the solution to 3. Regarding the correlations between factors, in this case this solution confirmed the presence of a significant positive association of moderate magnitude between the factor related to emotional aspects

and that related to attitudes (r = .236), as well as the absence of a significant association between metacognitive aspects and belief aspects (Factor 1 and Factor 4). Between the two factors better defined within the 4-way solution (the one related to strategies and the one related to metacognitive abilities), a significant positive correlation of medium magnitude emerged (r = .405).

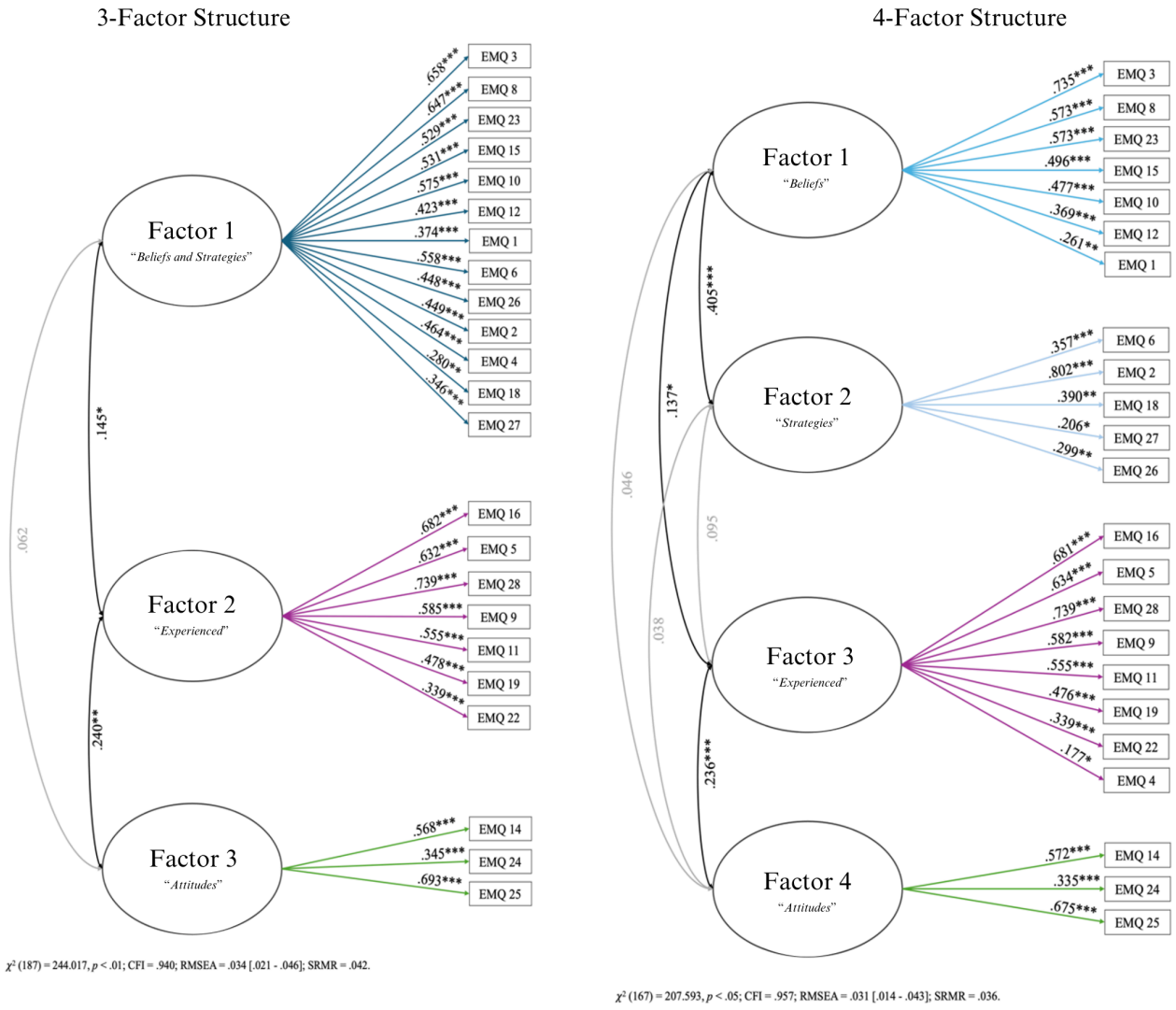


Figure 12. Exploratory Factor Analysis Models: 3 and 4 Latent Factor Solutions.

5.3.3. Principal Component Analysis

To confirm the factor structure that emerged from the EFA models, as well as to compare the factor structure of the instrument of the present study with the original one that considered older students, confirmatory factor analysis models were implemented, following the methods and procedures of the original scale (Palladino et al., 2019). A principal component analysis

(PCA) was carried out, which used the principal component method as the extraction method, and rotation of Varimax. Both the Kaiser-Meyer-Olkin adequacy index (KMO = .815) and the Bartlett sphericity test [χ^2 (406) = 1880.252; $p < .001$] confirmed the adequacy of the sample and the factorization of the correlation matrix. Regarding the number of components, the eigenvalue method equal to 1 (Kaiser, 1960) showed that the best factorial solution had 8 factors, and this model explained a cumulative variance of 56.9%. However, the information from the Scree test (Cattell, 1966), also called the "elbow method", suggested that the number of plausible factors was 4, identified as the inflection point of the curve (Figure 13). However, scree test did not always provide a clear solution to the number of factors. An examination of eigenvalues and their differences was performed (Cureton & D'Agostino, 2013; D'Agostino & Russell, 2005), again identifying the fourth factor as the last non-random factor, immediately preceding the substantially larger difference (<0.1 – in this case, from 2.355 to 0.185). The cumulative explained variance for a number of factors set at 4 was 41.3%.

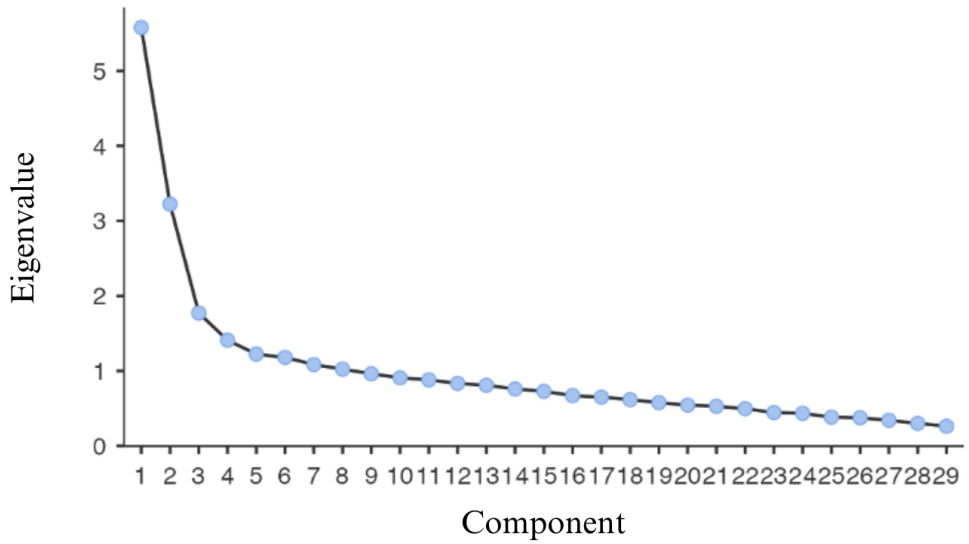


Figure 13. Scree plot.

As shown in Table 27, some items theoretically belonging to different dimensions saturated within the same factor (i.e. Factor 1). A similar situation was also found in the original scale (Palladino et al., 2019), where the authors renamed this dimension as "*Experienced*" (R^2 = 13%). In the original study, it showed that lower secondary students attached fundamental importance to their emotional and affective perceptions when learning English as a second language (Palladino et al., 2019). Individual perceptions on an emotional and affective level played a similar important role in our primary school sample, as reported by the variance

explained almost overlapping with that highlighted in the study by Palladino and colleagues (R^2 = 13%).

Compared to the identification of Factor 2 through PCA, most of the items that saturated within it referred to the role of students' previous, current and future knowledge on L2 learning, as well as to a series of thoughts that could influence the performance of the same learning (e.g.: *"Quando leggo in inglese penso a quello che so sull'argomento"; "Sapere l'inglese è importante per il mio futuro"; "Per leggere bene in inglese, bisogna sapere il significato di molte parole").* Therefore, this factor was called "*Metaknowledge*".

Regarding Factor 3, the original scale defined this dimension as "*Utility*" (R^2 = 6%; Palladino et al., 2019). The four items that in our PCA models saturated in this factor belonged to the theoretical area of *Beliefs* ("*L'inglese si impara solo se si è bravi"; Rileggere in inglese è inutile, perché se non si capisce subito, rileggendo non si migliora")* and the *Utilities ("Sapere l'inglese non è importante"; Studiare l'inglese serve solo per ottenere buoni voti")*. As previously reported, the tool allowed to investigate not only the metacognitive behavior, but also the motivational component. The items saturated in the third factor, in fact, investigated the intrinsic and extrinsic motivation of L2 learning. For this reason, this factor was renamed as "*Motivation*" (R^2 = 7%).

As a final factor, a defined factor of "*Strategies*" emerged (R^2 = 8%), which could be totally superimposed on the homonymous factor of the original version (R^2 = 6%). Within this last factor, 3 out of 4 items referred to the actions implemented by students to achieve their goal ("*Quando leggo in inglese mi aiuto guardando il titolo e le immagini"; "Se non so come si dice una cosa in inglese mi aiuto con i gesti"; "Per leggere bene in inglese, bisogna sapere il significato di molte parole").* The fourth item belonged to the *Beliefs area* ("*Spesso faccio i compiti di inglese con la TV o la musica accesa"),* however it included attention to study strategies that could influence the quality of learning, confirming the choice to keep the label of the dimension as "*Strategies*". It is relevant to pay further attention to two items, saturated in the latter factor: item 13 ("*Spesso faccio i compiti di inglese con la TV o la musica accesa")* and 29 ("*Per leggere bene in inglese, bisogna sapere il significato di molte parole").* Both items had a high level of saturation but negative (respectively: –0.526; –0.44) and had few (and scarce) significant associations in the correlation matrix with the additional items. The authors considered these values interesting, as primary school students did not see the

presence of distractors or the scarce amount of their English vocabulary as an obstacle to their English.

Original area	Item	Component			
		Metaknowledge (EV* 15,27%)	Experienced (EV 13,32%)	Motivation (EV 6,87%)	Strategies (EV 5,86%)
Beliefs	4	0,487			
Beliefs	8	0,636			
Beliefs	14			0,725	
Beliefs	17		0,259		
Beliefs	22		0,43	0,364	
Beliefs	23	0,56		–0,306	
Beliefs	24	0,252		0,431	
Beliefs	29	0,371			–0,44
Strategies	1	0,361	0,258		
Strategies	2	0,504			
Strategies	6	0,673			
Strategies	10	0,515			0,365
Strategies	13				–0,526
Strategies	18				0,391
Strategies	26	0,466			
Strategies	27				0,667
Utility	3	0,713			
Utility	7	0,334		0,447	
Utility	12	0,54			–0,309
Utility	15	0,635			
Utility	25			0,748	
Experienced	5		0,667		
Experienced	9		0,646		
Experienced	11		0,628		
Experienced	16		0,731		
Experienced	19	0,258	0,561		
Experienced	20	0,508	0,512		
Experienced	21	0,394	0,619		
Experienced	28		0,754		

Note: Extraction method: Principal Component Analysis. Rotation method: Varimax with Kaiser normalization. Convergence for rotation performed in 10 iterations. *EV =Explained Variance.

Table 27. Principal Component Analysis (Matrix of rotated components).

5.3.4. Evaluation of the identified dimensions

Once the four factors of the EMQ-PS tool were identified, some basic descriptive analyses were conducted on the emerged dimensions, such as the analysis of the correlations between the dimensions, as well as the analysis of the reliability of the emerged dimensions. Table 28 shows that all correlations were statistically significant ($p < .001$), suggesting a consistent relationship between the size of the instrument and the total.

	Beliefs	**Experienced**	**Strategies**	**Attitudes**	**Total**
Beliefs	—				
Experienced	0.47***	—			
Strategies	0.38***	0.27***	—		
Attitudes	0.56***	0.31***	0.34***	—	
Total	0.78***	0.81***	0.64***	0.66***	—
*** p < .001					

Table 28. Correlation matrix between the EMQ–PS total and the dimensions.

The Split-half analysis revealed an acceptable level of reliability of the total scale ($\alpha = .707$), confirmed by the analysis of the level of reliability calculated on the polychoric matrix ($\omega = .898$). Regarding the four dimensions identified by exploratory and confirmatory factor analysis, the dimensions of *Beliefs* and *Experienced* demonstrated good reliability (respectively, for the dimension of *Beliefs,* $\alpha = .730$ and $\omega = .750$; for the dimension of *Experienced,* $\alpha = .787$ and $\omega = .786$). The *Strategies* and *Attitudes* dimensions, on the other hand, showed lower levels of reliability than the previous dimensions (for the *Strategies,* $\alpha = .589$ and $\omega = .703$; for the *Attitudes*, $\alpha = .554$ and $\omega = .598$). See Table 29 for more information on the confidence levels for each identified.

	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1 – Beliefs	–			
Factor 2 – Strategies	.380***	–		
Factor 3 – Experienced	.468***	.270***	–	
Factor 4 – Attitudes	.556***	.337***	.311***	–
Mean (SD)	3.33 (.58)	3.15 (.64)	2.94 (.71)	3.21 (.80)
Cronbach's α	.730	.589	.787	.554
McDonald's ω	.750	.703	.786	.598

*** $p < .001$

Table 29. Descriptive characteristics and correlations between the identified dimensions.

The mean values and standard deviations of the dimensions identified are shown below, divided by 3rd and 4th grade (Table 30).

Area	Grade	Mean	SD
Metaknowledge	3rd	36.1	6.48
	4th	37.2	5.47
Experienced	3rd	28.8	6.94
	4th	27.5	7.18
Motivation	3rd	12.8	3.04
	4th	13.3	2.67
Strategies	3rd	12.1	2.06
	4th	12.1	2.14
Total	3rd	89.8	13.66
	4th	90.0	12.45

Table 30. EMQ–PS means and standard deviations (SD).

A t-test was carried out with independent samples for the two grades, which did not show significant differences between the answers of students at different school levels. Therefore, the percentile values and performance bands for the entire sample were reported (Table 31).

Area	N	Media	SD	Percentiles					Performance bands			
				10th	20th	30th	50th	70th	RTI	RTA	STP	TPA
Metaknowledge	258	36.7	5.97	28.70	32.0	35.0	38.0	41.0	<29	29-34	35-41	>41
Experienced	258	28.1	7.09	17.00	22.0	24.1	29.0	33.0	<17	17-23	24-33	>33
Motivation	258	13.1	2.85	9.00	10.0	12.0	14.0	15.0	<9	9-11	12-15	>15
Strategies	258	12.1	2.10	9.00	10.0	11.0	12.0	13.0	<9	9-10	11-13	>13
Total	258	89.9	13.00	71.00	77.4	84.0	92.0	99.0	<71	71-83	84-99	>99

Note: RTI = Request for Teaching Intervention; RTA = Request for Teaching Attention; STP = Sufficient Teaching Performance; TPA = Teaching Performance Achieved.

Table 31. EMQ–PS performance bands.

5.3.5. Qualitative analysis

The English Metacognitive Questionnaire (both the original and the Primary School version) had two qualitative items, namely 30 and 31 (respectively: *"Se sbagli un compito di inglese pensi"; "Se prendi un bel voto in un compito di inglese pensi").* These items investigated the area of *Attributions*, i.e. to which factors (internal and/or external) the student attributed his or her successes and failures in learning English as an L2. As reported previously, both questions had five answer options: the student had to choose, for each item, two answers in order of importance that best indicate (in the student's opinion) the reasons related to their success and failure. In our sample, 72% of participants (N ~ 185) answered the request correctly, and it was therefore not possible to consider the responses of 28% of participants. Although it was conceivable that age would influence the difficulty in following the indications relating to the two qualitative items (which require a different response mode from the quantitative items), subsequent analyses did not show disparities in the compilation of the items compared to the class attended. It should be noted, however, that these qualitative items were collected to qualitatively assess the student's attribution style of his/her academic successes and failures. Therefore, the lack of response to these items did not negatively affect the analyses conducted to explore the dimensionality and structure of the EMQ-PS instrument. The frequency distributions of responses by students who responded correctly by following the instructions are presented below (Table 32).

Attribution	Item. Response	% Total				
		Skill	Help	Task	Luck	Effort
Failure	30.1	6.5 %	7.1 %	14.7 %	5.4 %	66.3 %
	30.2	25.7 %	18.6 %	20.8 %	23.0 %	12.0 %
Success	31.1	6.0 %	3.3 %	5.4 %	22.8 %	62.5 %
	31.2	20.5 %	28.6 %	18.9 %	19.5 %	12.4 %

Table 32. Frequency distributions of attributive choices.

As shown, most primary school students (more than 60% of the sample) attributed both their successes and failures in L2 learning to personal effort, as the first and most important consequence. According to the Self-Theories (Dweck, 2013), students therefore attributed to themselves a personal and dynamic competence in achieving their goals. A rather heterogeneous situation was found in the second possibility of response in both conditions. In the case of failure, 25% of students attributed their failure to their abilities as a second option, i.e. aspects that are always personal although innate and static over time. About the success, more than 28% attributed as a second response option the failure to receive an external contribution, i.e. help from the teacher. Both responses were less adaptive than attribution to commitment (Moè, 2010).

5.3.6. Individual cases

We introduce two cases as examples of the use of the tool for profile insights and structuring of targeted interventions. The data of the questionnaire in question are compared with the performance of the English test carried out in the same academic year (2022/2023). We introduce the case of a male general participant, from the province of Lombardy, with Italian L1, without any diagnosis. Considering that the sample mean for English vocabulary was 12 (SD = 3.91), the participant scored about 1.76 standard deviations below the mean, achieving a score of 5. Examining the score achieved by the same student in the EMQ-PS, there was a sufficient performance in the Strategies applied to learning English as L2, and a performance fully achieved in the Experienced (see Table 32 for the performance bands). On the other hand, the fields of “Metaknowledge” and “Motivation” were deficient (respectively: request for didactic attention; request for immediate intervention). The EMQ-PS allows us to understand which metacognitive areas can be identified as insufficient and focus on them for the promotion of real and effective learning. In this case, it is considered necessary to start

from the promotion of motivational components (intrinsic and extrinsic) of the student for L2 learning, strengthening the awareness of one's previous skills and knowledge, and therefore of the value of one's Metaknowledge (a useful tool could be Botto et al., 2019). A further example is the score obtained by a student from the Calabrian territory, also with an Italian mother tongue. The student achieved an average score of English in the sample, for a total of 11. Moreover, it presents teaching performance sufficiently or fully achieved in all metacognitive areas investigated by the EMQ-PS, except for Metacognitive Strategies. The student in question, in fact, achieved a score of 8 in the latter area, which highlights a request for immediate intervention. After a comparison with the performance of the student in school subject, the EMQ-PS leads to a greater awareness of the study methodology, acquiring awareness on which areas to focus more on to promote real and effective learning. In this case, the metacognitive fall seems to suggest the investigation of the strategies and behaviors used for L2 learning, contributing to the awareness of the importance of acquiring new and more effective strategies.

5.3.7. Comparison of student self-assessments and teachers' perceptions

We developed and administered a scale to the English language teacher of each class to gather their opinions on the metacognitive and motivational knowledge of each student (e.g., "*Quando il/la mio/a alunno/a sbaglia un esercizio, si esercita di più per fare meglio la prossima volta"; "Quando non riesce a fare un esercizio di inglese, il/la mio/a alunno/a si lascia prendere dall'ansia*"). A correlation analysis was conducted between the dimensions of the EMQ-PS and the responses to a questionnaire distributed to the group of teachers to assess teachers' perceptions of the scale's areas. As shown in Table 33, no significant correlations were found between the students' self-reported data and their teachers' evaluations across any of the four dimensions analyzed. This information is particularly noteworthy, as it highlights a misalignment between teachers' assessments of students' motivations, emotional experiences, and metacognitive strategies for L2 learning, and what the students themselves reported. However, there is a trend toward significance in the relationship between the *Experienced* and the *Motivation* dimension reported by teachers ($r = .150$; $p < .05$).

Students' metacognition→ Teachers' opinion↓	**Metaknowledge**	**Experienced**	**Motivation**	**Strategies**	**Total**
Metaknowledge	–0.04	–0.08	–0.05	–0.03	–0.08
Experienced	–0.06	–0.02	–0.07	–0.03	–0.06
Motivation	–0.07	0.15*	0.10	–0.09	0.06
Strategies	–0.07	–0.10	–0.10	–0.05	–0.11
Total	–0.06	–0.05	–0.06	–0.04	–0.07

*p<.05

*Table 33. Correlation matrix (Students*Teachers).*

5.4. Discussion and conclusion

The aim of the present study was to contribute to scientific research on metacognitive competences through the adaptation and validation of the English Metacognitive Questionnaire (EMQ) by Palladino et al. (2019) in an Italian primary school sample. Regarding the internal consistency of the instrument, both the four dimensions and the entire questionnaire showed an adequate level of reliability. The results obtained through exploratory and confirmatory factor analyses confirmed a 4-factor structure, as identified by Palladino et al. (2019) and by the Flavell's Cognitive Monitoring Model (1979). However, some differences emerged between the original and the primary school version: this difference was conceivable and predictable, as these versions were designed and validated within samples of different ages (lower secondary school vs. primary school) and in two different historical periods (2019 vs. 2023), further characterized by a drastic change in the organization of teaching due to the pandemic period for COVID-19 between the aforementioned academic years which may have influenced the learning and emotional correlates of primary school students. The first factor identified ("*Metaknowledge")* included items related to a more general view of personal and parents' knowledge on the relevance of learning English, as well as awareness of learning strategies, monitoring of one's own knowledge and self-regulation. The second factor ("*Experienced"),* as in the original version, referred to the emotional and affective perceptions of students in L2 learning. The third factor ("*Motivation*") defined the importance of the student's motivational ability for real learning, in line with different studies that underlined the fundamental role of motivation for learning a second language (Dörnyei, 2001; Gardner, 2001; Oga-Baldwin et al., 2017). The fourth factor ("*Strategies")* collected items that allowed us to understand how much the student was

aware of being able to learn better, through different study strategies. The second and fourth factors ("*Experienced*"; "*Strategies*") remained unchanged in their interpretation compared to the original version (Palladino et al., 2019). Conversely, a distinction between the two scales concerns the absence of the dimensions of *Beliefs* and *Attitudes* in the EMQ–PS. It should be noted, however, that they are intrinsically present within the factors identified in the present study which have been defined as "*Metaknowledge*" and "*Motivation*". Compared to the "*Experienced*" factor, having shown a very similar explained variance between the two scales (13.00% vs. 13.32%), it is interesting to note that in the present study the "*Experienced*" area is identified as the second factor in PCA and 3-factor EFA, and as the third factor in 4-factor EFA, while in the version by Palladino et al. (2019) it was the factor with the highest number of saturations. This phenomenon may express the fact that primary school students gave less importance to their emotional and affective perceptions for effective English language learning than their secondary school counterpart, perhaps due to lack of experience or little "emotional maturation". Compared to the original scale, the size of the "*Strategies*" ultimately shifts dramatically, specifically from being the second factor (R^2 = 8%) of the original scale, to being the fourth and last factor identified of the EMQ–PS (R^2 = 6%). Again, this difference could be ascribed to a lack of experience on the part of the students to highlight the best action to take, depending on the study condition and the context. Although our main objective was not specifically the study of the latent structure of the instrument, 'this analysis was considered as an appropriate intermediate step, and allowed us to identify a series of similarities and differences with the analysis of the principal components. While the PCA had unequivocally identified a 4-factor solution, the EFA results suggest that plausible solutions may be either 3-factor or 4-factor. This difference could be mainly attributed to the fact that, in the case of PCA, we replicated the analyses carried out by Palladino and colleagues, also maintaining some items that were already ambiguous or with saturation indices that were not entirely adequate (Palladino et al., 2019). Despite this difference, the saturation of items in the first three factors is totally overlapping in both analyses. Regarding the "*Strategies*" factor, it is conceivable that primary school students did not fully master the approaches and actions that a person adopts to achieve their learning objectives. However, a comparison between our EFA models and Palladino et al. (2019) is not possible, as in this comparative study an analysis of the latent structure was not conducted, but only a confirmatory factor analysis based on the PCA method. No significant associations emerged between teachers' perception

and students' self-evaluations of their metacognitive abilities, except for the relationship between the *Experienced* and the *Motivation* reported by teachers. We can assume in retrospect that higher levels of positive emotions (e.g., joy) are recognized by teachers as more motivating students to learn English. The implementation of continuing education programs for teachers is considered useful. These programs could focus on improving teachers' awareness of students' emotional and motivational experiences and strategies to support them more effectively. Using the information collected, teachers may be able to adapt their teaching practices to better meet the individual needs of students, actively involving them in the assessment and feedback process, in a collaborative and inclusive learning environment. It is suggested the implementation of continuous monitoring systems of students' perceptions, allowing the timely identification of any changes in their motivations, emotions and metacognitive strategies, and prompt intervention by the teacher. Integrating these practical implications can foster a more stimulating and inclusive learning environment, which supports students in achieving their learning goals.

Certainly, a limitation of our study can be found in the small number of responses to qualitative items. Almost 100 students out of 258 of the total sample (i.e. 39% of the sample) did not fill out this part of the questionnaire or did so incorrectly. One explanation could be that the two qualitative items had a different type of response: hence, the student had to change the way he/she answered which may have been more difficult for someone. In the future, it is considered necessary to pay more attention to a correct explanation of the instructions by researchers for qualitative items to promote full understanding and correct compilation.

Despite the limitations, the EMQ-PS has good psychometric requirements, and it is considerable useful for identifying students' strengths and weaknesses with respect to L2 learning since primary school. The study of the different metacognitive skills by the students allows to encourage a comparison within the class group to promote interpersonal knowledge. In addition, it proves to be an efficient investigation tool for the understanding of the individual student with respect to metalinguistic skills, as well as his own consideration of himself and his successes and failures. The assessment of L2 metacognitive skills from primary school is important for several reasons related to self-regulated learning and emotions that contribute to the creation of contexts of well-being: for the awareness of learning itself, allowing students to be more aware of how they learn and to adapt their strategies according to specific challenges; to promote students' ability to self-assess their efforts and self-regulate

in learning the English language; for the role of motivation in learning, as those who develop metacognitive skills often maintain greater motivation in L2 learning (Cornoldi, 1995, 2007; Botto, Palladino & Cornoldi, 2019; Palladino & Cornoldi, 2023). In conclusion, assessing primary school students' metacognitive skills in L2 learning is crucial to help them become autonomous and motivated learners.

6. General discussion

The thesis presented aimed to investigate not only the current level of knowledge of L2 by Italian post-pandemic primary school students, but also the role of basic cognitive skills (including competence in the mother tongue), metacognitive skills and emotions perceived by students in learning L2. As highlighted, in fact, the role of these variables is widely recognized. However, they were examined – in relation to L2 – in isolation, and never together. Therefore, we aimed to contribute to the understanding of the factors that support the self-regulation of L2 learning in primary school, through a longitudinal study. This national project was born from the collaboration between the University of Foggia as leader with several universities (Free University of Bolzano; University of Pavia; University of Calabria), private institutions (British School Pistoia) and several schools in Italy. We conducted several studies to address the following research questions: #1) Is there a relationship between the achievement emotions, cognitive, metacognitive variables, and L2 learning? #2) Is there a time effect on L2 learning, achievement emotions, cognitive, and metacognitive variables? #3) How achievement emotions, cognitive and metacognitive variables can jointly influence English as a L2 performance in post-pandemic primary school students?

First, we considered necessary to evaluate the effect of school closure by COVID-19 pandemic on Italian primary students, not only on their academic performance, but other dimensions affected. Subsequently, we conducted a longitudinal study to investigate the role of our variables in L2 learning, and the mother tongue too. Finally, we considered relevant to develop an instrument that it could allow the early identification of children at risk in L2 learning, not only from a metacognitive but also motivational, and emotional, point of view.

6.1.Study 1

The COVID-19 pandemic had drastic consequences for many, not sparing students of all levels. It was presumable that the school closure due to the pandemic condition influenced the learning of students in general, as evidenced globally by the scientific literature. However, less evident was the information relating to primary school students, in particular from the Italian territory, one of the Western nations most affected by the pandemic phenomenon. Our

systematic review provided an overview of the effects of the school closure due to the COVID-19 pandemic on the learning outcomes of Italian primary school students (target population) according to their opinion, and the opinions of their teachers and/or parents. As shown, the school closure affected not only the academic achievement of students, but also several dimensions. Specifically, changes in the emotional, cognitive and metacognitive domains are evident. First, students showed a significant deterioration in emotional well-being during school closures. Parents and teachers reported an increase in anxiety, frustration and behavioral problems, such as agitation and aggression. Social distancing and distance learning led to a sense of isolation, with a negative impact on students' motivation and social interaction, which are essential for their emotional balance. Students with learning difficulties or those belonging to socio-economically disadvantaged groups suffered more from these negative emotional effects, which may have further affected their academic performance. Cognitively, students had difficulties in maintaining concentration and engagement in the context of distance learning. Technological challenges, the absence of direct interaction with teachers and the lack of adequate facilities for independent learning exacerbated cognitive difficulties, especially in subjects such as mathematics. On a metacognitive level, many students found it difficult to regulate and monitor their learning. The lack of immediate feedback and the disruption of the school routine reduced students' self-efficacy and ability to plan and reflect on their academic progress. Another critical point that emerged is the lack of specific data regarding the learning of English as a second language during the pandemic. Although it is a canonical subject of the INVALSI tests, the data at our disposal provide information on students' performance in Italian and (especially) mathematics, and therefore the emotional, cognitive and metacognitive condition experienced by students with respect to these subjects.

6.2. Study 2

As reported, our systematic review showed no studies focused on learning English as a second language in Italian primary school students. This lack of data highlights a fundamental knowledge gap of the pandemic effects on young students. Learning a new language from primary school, in fact, has different benefits and opportunities: in addition to the possibility of communicating with people from other countries and learning about their cultures, learning

English as a second language helps to lay the foundations of language skills, generating an advantage in learning and academic performance in general. Despite the proven benefits of learning L2, most studies focused on high school or university students, and only a few of them investigated children early in the learning process.

For over three years, therefore, over 300 primary school students in Italy were involved in a series of tests aimed at assessing their skills in L2 and in the variables exposed. Our results certainly highlighted a role of temporal maturity and individual characteristics, but above all a collaboration of cognitive, metacognitive and emotional skills in learning English as a second language by Italian primary school students. There was a particular role for working memory and the ability to metacognitively monitor one's performance on the specific task, in line with scientific literature (Baddeley, 2015; Kormos & Sáfár, 2008; Sohn & Doane, 2003; Taouki, Lallier, & Soto, 2022). Regarding achievement emotions, it is not surprising that the inverse role of anxiety was relevant: the lower the perceived anxiety towards the English language, the better the performance. It certainly makes us think that there was no significant role for positive emotions such as joy, highlighting that anxiety is not only the most experienced emotion among Italian primary school students, but it appears as the only emotion experienced during English lessons (and probably others as well).

6.3. Study 3

The individual and territorial differences highlighted in Study 2 led us to take a step back, introducing an evaluation of students' performance in their mother tongue, i.e. Italian. In line with the scientific literature (Hdstijn & Bossers, 1992; Prior et al., 2014; Sparks et al., 2012), this PhD thesis showed that lexical competence in L1 was a strong predictor of L2, revealed by the comparison between students with a rather limited performance in L1 and peers with higher skills. Considering the entire sample, significant correlations were found, positive between performance in Italian and the emotion of Joy, and negative between performance in English and the emotions of Boredom and Anxiety. Furthermore, in line with the individual differences highlighted in Study 2, regression analysis showed that age, performance level in mother tongue and anxiety significantly predicted performance in the second language. Considering, on the other hand, the sub-sample composed of participants with extreme performance in the mother tongue ($< 25^{th}$ and $>75^{th}$ percentile), it emerged that the two groups

did not differ for any of the measured emotions, but the mother tongue high-performance group reported significantly higher performance in all parameters related to English. Learning English as a second language in Italian primary school children appeared to be related to the competence achieved in the motor tongue, in agreement with the literature that has examined secondary school or university students and adults, and is also explained by the level of anxiety, which negatively predicts learning. However, the comparison between groups with high and low competence in Italian showed for the latter only worst performance in English but no differences in the emotions experienced in learning contexts. It is relevant to note that Study 3 referred to a cross-sectional study. Future studies should investigate such differences in the role played by emotions in the general population and in students with extreme performance in their mother tongue, considering also the time component. Despite that, Study 3 allowed us to highlight the role of L1 proficiency in L2 learning.

6.4. Study 4

As the mother tongue, we evaluated students' metacognitive competence for immediate performance in L2, leaving aside their metacognitive reflection skills regarding learning English as an L2 in general. Since there were no validated tools for the assessment of these skills in L2, we carried out the psychometric analysis of the English Metacognitive Questionnaire for Primary School (EMQ-PS). Our results showed not only that it was an adequate tool for investigating metacognitive skills from primary school, but also a useful aid to figures in the field of child development (psychologists, teachers, educators) to highlight students with motivational and metacognitive difficulties regarding English, allowing the early identification of students who could subsequently experience negative situations (disinterest, anxiety, inattention). Specifically, the development of this tool seemed as necessary as it was natural in its validation. It is the practical basis of the union between the two main models: the MASRL Model and the Control-Value Theory. Specifically, the EMQ-PS proves to be an effective tool for investigating the Macro Level of Efklides' model, and the homonymous level identified by Pekrun: the Person level. In fact, through its application, the professional can investigate the strategic abilities, metacognitive knowledge and control of the student's own beliefs. Furthermore, the tool is useful for identifying the motivation (therefore the intrinsic and extrinsic values) perceived by the student in L2 learning. The

motivation score in particular guides the professional's evaluation in understanding the achievement emotions experienced by the student, and in planning together with him/her the most metacognitively appropriate study method, based on the specific task and the student's personal skills.

6.5. Limitation and considerations for the future.

Certainly, further investigations are needed. Specifically, as highlighted, it is evident that individual differences play a fundamental role in L2 learning, as well as in the variables investigated. Similarly, a difference due to territorial belonging (schools that perform better than others) was highlighted. Future studies need to consider the family background and socio-economic condition to which each student belongs to identify not only individual skills, but also the family condition and the opinion of the parents themselves, as well as their teachers. It is relevant to ask, in fact, whether the opinion of reference figures (such as teachers) can influence the performance of the L2 of each student. Several studies focused on the accuracy of teachers' judgment, i.e. the ability of teachers to effectively assess students' attributes and calibrate learning needs and task demands accordingly (Artelt & Gräsel, 2009). Although it considers various aspects related to judgment (task/person), accuracy of judgment refers to the student's performance in a specific task or school subject, and not to the teacher's opinion of his or her student's abilities, which can convey the student's emotional (as well as cognitive and metacognitive) response in future learning. It is therefore necessary that future research consider the correspondence between teachers' opinions on students' learning ability and their real learning capacity. Another condition highlighted by our systematic review is the difficulty for teachers to adapt to new teaching methodologies (digital students and distance learning) while recognizing their value and learning opportunity. Future studies will certainly have to consider the opportunity to investigate the use of innovative teaching methodologies for the English language and beyond, such as the use of applications software (apps). Digital tools for learning L2, such as apps, are now within everyone's reach, offering a variety of teaching strategies with the aim of developing communication skills in all children.

7. General conclusion

In conclusion, the doctoral thesis presented leads to many answers, but paving the way for further and greater research questions. This PhD thesis explored the effects of the pandemic on self-regulated learning in primary school, with a focus on the acquisition of English as a second language (L2). The pandemic period highlighted educational gaps and challenges, especially for younger students, highlighting the complexity of remote teaching and the need to adapt to an ever-changing learning environment. “For the times they are a-changin” sang Bob Dylan, and we find it further among the school banks of young post-pandemic populations, indeed. The changes in the learning of Italian primary school students are tangible and need further investigation. The pandemic period accelerated a reflection on the approach to education in primary school, demonstrating how a flexible and innovative education system can better support the new generations in a constantly evolving world. This thesis represents a first step towards a deeper understanding of the factors that influence L2 learning and offers a basis for future research aimed at improving the educational experience of young Italian students.

8. Reference

Abbasian, R., Hadian, B., & Vaez-Dalili, M. (2020). Exploring the interplay between Iranian EFL learners' first language (L1) literacy resources and their performance on L2 receptive skills. *Current Psychology*, *39*(6), 1900-1909. https://doi.org/10.1007/s12144-018-0059-5

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, *2*(4), 433-459. https://doi.org/10.1002/wics.101

Alloway, T. P. (2009). Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. *European journal of psychological assessment*, *25*(2), 92-98. https://doi.org/10.1027/1015-5759.25.2.92

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of experimental child psychology*, *106*(1), 20-29. https://doi.org/10.1016/j.jecp.2009.11.003

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*(1), 133-137. https://doi.org/10.1016/j.lindif.2010.09.013

Amsterlaw, J. (2006). Children's beliefs about everyday reasoning. *Child Development, 77*(2), 443-464. https://doi.org/10.1111/j.1467-8624.2006.00881.x

Anderson, N. J. (2002). The Role of Metacognition in Second Language Teaching and Learning. ERIC Digest.

Aripin, N., & Rahmat, N. H. (2021). Metacognitive writing strategies model used by ESL writers in the writing process: A study across gender. *International Journal of Asian Social Science*, *11*(1), 1-9. https://doi.org/10.18488/journal.1.2021.111.1.9

Artelt, C., & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für pädagogische Psychologie*, *23*(34), 157-160. http://dx.doi.org/10.1024/1010-0652.23.34.157

Artino Jr, A. R., & Jones II, K. D. (2012). Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *The internet and higher education*, *15*(3), 170-175. https://doi.org/10.1016/j.iheduc.2012.01.006

Artino Jr, A. R., Holmboe, E. S., & Durning, S. J. (2012). Can achievement emotions be used to better understand motivation, learning, and performance in medical education?. *Medical Teacher*, *34*(3), 240-244. https://doi.org/10.3109/0142159X.2012.643265

Artuso, C., & Palladino, P. (2019). The role of memory updating in shallow native and opaque second language learning. *Second Language Research*, *35*(3), 377-396. https://doi.org/10.1177/0267658318777022

Avitia, M. J., Kaufman, A. S., Bray, M., & Kaufman, J. C. (2019). Relationship between reading and long-term storage and retrieval (Glr) in college students. *Applied Neuropsychology: Adult*, *26*(2), 111-123. https://doi.org/10.1080/23279095.2017.1362562

Baddeley, A. D. (2015). Working memory in second language learning. *Working memory in second language acquisition and processing*, 17-28. https://doi.org/10.21832/9781783093595-005

Banerjee, T., Khan, A., & Kesavan, P. (2021). Impact of lockdown and school closure on children in special schools: A single-centre survey. *BMJ paediatrics open*, *5*(1). https://doi.org/10.1136/bmjpo-2020-000981

Barbaranelli, C. (2003). Factor analysis: Exploratory. In *Encyclopedia of Psychological Assessment* (pp. 403-407). Sage.

Barrett, L. F., & Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current directions in psychological science*, *8*(1), 10-14. https://doi.org/10.1111/1467-8721.00003

Baschenis, I. M. C., Farinotti, L., Zavani, E., Grumi, S., Bernasconi, P., Rosso, E., ... & Chiappedi, M. (2021). Reading skills of children with dyslexia improved less than expected during the COVID-19 lockdown in Italy. *Children*, *8*(7), 560. https://doi.org/10.3390/children8070560

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., & others. (2009). Package 'lme4.' URL: https://lme4.r-forge.r-project.org/

Bayard, N. S., van Loon, M. H., Steiner, M., & Roebers, C. M. (2021). Developmental improvements and persisting difficulties in children's metacognitive monitoring and control skills: Cross-sectional and longitudinal perspectives. *Child development*, *92*(3), 1118-1136. https://doi.org/10.1111/cdev.13486

Bazoli, N., Marzadro, S., Schizzerotto, A., & Vergolini, L. (2022). Learning loss and students' social origins during the covid-19 pandemic in Italy. *FBK-IRVAPP Working Papers*, *3*, 2022.

Bazzoli, N., Dagnes, J., Carbone, D., & Barberis, E. (2021). La didattica a distanza nell'Italia diseguale: criticità e differenze territoriali durante la prima ondata Covid. *Rivista geografica italiana: CXXVIII, 3, 2021*, 36-58.

Benigno, V., Dagnino, F., & Fante, C. (2020). Exploring the impact of the COVID-19 pandemic on Italy's school-in-hospital (SiHo) services: The teachers' perspective. *Continuity in Education*, *1*(1), 136. DOI: https://doi.org/10.5334%2Fcie.26

Berguno, G., & Bowler, D. M. (2004). Communicative interactions, knowledge of a second language, and theory of mind in young children. *The Journal of Genetic Psychology*, *165*(3), 293-309. https://doi.org/10.3200/GNTP.165.3.293-309

Bernhardt, E. B. (2000). Second-language reading as a case study of reading scholarship in the 20th century. *Handbook of reading research*, *3*, 791-811.

Bertoletti, A., Soncin, M., Cannistrà, M., & Agasisti, T. (2023). The educational effects of emergency remote teaching practices—the case of covid-19 school closure in Italy. *PLoS One*, *18*(1), e0280494. https://doi.org/10.1371/journal.pone.0280494

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature human behaviour*, *7*(3), 375-385. https://doi.org/10.1038/s41562-022-01506-4

Birkelund, J. F., & Karlson, K. B. (2023). No evidence of a major learning slide 14 months into the COVID-19 pandemic in Denmark. *European Societies*, *25*(3), 468-488. https://doi.org/10.1080/14616696.2022.2129085

Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European psychologist*, *1*(2), 100-112. https://doi.org/10.1027/1016-9040.1.2.100

Boekaerts, M. (2007). Understanding students' affective processes in the classroom. In *Emotion in education* (pp. 37-56). Academic Press. https://doi.org/10.1016/B978-012372545-5/50004-6

Bol, T. (2020). Inequality in homeschooling during the Corona crisis in the Netherlands. First results from the LISS Panel. https://doi.org/10.31235/osf.io/hf32q

Bonvino, A., Calvio, A., Stallone, R., Marinelli, C. V., Quarto, T., Petito, A., ... & Monacis, L. (2023). Emotions in times of pandemic crisis among Italian children: A systematic review. *International Journal of Environmental Research and Public Health*, *20*(12), 6168. https://doi.org/10.3390/ijerph20126168

Borgonovi, F., & Ferrara, A. (2023). COVID-19 and inequalities in educational achievement in Italy. *Research in social stratification and mobility*, *83*, 100760. https://doi.org/10.1016/j.rssm.2023.100760

Botto, M., Palladino, P., & Cornoldi, C. (2019). *English is fun!*. Vol. 2. Trento: Erickson.

Bozkurt, A., & Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. *Asian journal of distance education*, *15*(1), i-vi. Retrieved from https://asianjde.com/ojs/index.php/AsianJDE/article/view/447

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental neuropsychology*, *33*(3), 205-228. https://doi.org/10.1080/87565640801982312

Campodifiori, E., Figura, E., Papini, M., & Ricci, R. (2010). *Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia*. Invalsi. Available from: https://www.provincia.bz.it/servizio-valutazione-italiano/download/escs_invalsi.pdf

Canals-Botines, M., Crișan, M. M., Leproni, R., Medina-Casanovas, N., Pujol Tubau, M., Raluy Alonso, A. (2021). The Challenge of Distant Teaching for Primary School Teachers of English during the Covid-19: the Cases of Spain, Italy and Romania. *Lea*, 10, 393-404. https://doi.org/10.13128/ LEA-1824-484x-12798

Capperucci, D., Scierri, I. D. M., Salvadori, I., Batini, F., Toti, G., Barbisoni, G., & Pera, E. (2022). Remote teaching during COVID-19 Emergency: Teaching and assessment strategies and the role of previous training. *Education Sciences*, *12*(10), 646. https://doi.org/10.3390/educsci12100646

Capurso, M., & Roy Boco, T. D. (2023). Covid-19 Lockdown in Italy: perspectives and observations from special education teachers of changes in teaching and the student experience. *Childcare in practice*, *29*(4), 405-419. https://doi.org/10.1080/13575279.2021.1975649

Carr, M., & Kurtz, B. E. (1991). Teachers' perceptions of Their Students' metacognition, Attributions, and Self-Concept. *British Journal of Educational Psychology*, *61*(2), 197-206.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245-276. http://dx.doi.org/10.1207/s15327906mbr0102_10

Chaabane, S., Doraiswamy, S., Chaabna, K., Mamtani, R., & Cheema, S. (2021). The impact of COVID-19 school closure on child and adolescent health: a rapid systematic review. *Children*, *8*(5), 415. https://doi.org/10.3390/children8050415

Champeaux, H., Mangiavacchi, L., Marchetta, F., & Piccoli, L. (2022). Child development and distance learning in the age of COVID-19. *Review of Economics of the Household*, *20*(3), 659-685. https://doi.org/10.1007/s11150-022-09606-w

Chatzoglou, E., Fishstrom, S., Payne, S. B., Andress, T. T., & Vaughn, S. (2023). The footprint of the COVID-19 pandemic in reading performance of students in the US with and without disabilities. *Research in Developmental Disabilities*, *140*, 104585. https://doi.org/10.1016/j.ridd.2023.104585

Cinquin, P. A., Guitton, P., & Sauzéon, H. (2019). Online e-learning and cognitive disabilities: A systematic review. *Computers & Education*, *130*, 152-167. https://doi.org/10.1016/j.compedu.2018.12.004

Cnaan, A., Laird, N. M., & Slasor, P. (1997). Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in medicine*, *16*(20), 2349-2380. https://doi.org/10.1002/(SICI)1097-0258(19971030)16:20%3C2349::AID-SIM667%3E3.0.CO;2-E

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, *20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Colombo, M., & Santagati, M. (2022). The inclusion of students with disabilities: Challenges for Italian teachers during the covid-19 pandemic. *Social Inclusion*, *10*(2), 195-205. https://doi.org/10.17645/si.v10i2.5035

Contini, D., Di Tommaso, M. L., Muratori, C., Piazzalunga, D., & Schiavon, L. (2022). Who lost the most? Mathematics achievement during the COVID-19 pandemic. *The BE Journal of Economic Analysis & Policy*, *22*(2), 399-408. https://doi.org/10.1515/bejeap-2021-0447

Cornoldi, C. (1995). *Matematica e metacognizione: atteggiamenti metacognitivi e processi di controllo* (Vol. 43). Edizioni Erickson.

Cornoldi, C. (2007). *Difficoltà e Disturbi Dell'apprendimento* (pp. 1-337). Il mulino.

Cozzolino, M., Vivo, D. R., & Celia, G. (2022). School-based mind–body interventions: a research review. *Human Arenas*, *5*(2), 262-278. https://doi.org/10.1007/s42087-020-00163-1

Crescenza, G., Fiorucci, M., Rossiello, M. C., & Stillo, L. (2021). Education and the pandemic: distance learning and the school-family relationship. *Research in Education and Learning Innovation Archives*, (26), 73-85. http://dx.doi.org/10.7203/realia.26.18078

Crisci, G., Mammarella, I. C., Moscardino, U. M., Roch, M., & Thorell, L. B. (2021). Distance learning effects among Italian children and parents during COVID-19 related school lockdown. *Frontiers in Psychiatry*, *12*, 782353. https://doi.org/10.3389/fpsyt.2021.782353

Cudeck, R. (2000). Exploratory factor analysis. In *Handbook of applied multivariate statistics and mathematical modeling* (pp. 265-296). Academic Press.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of educational research*, *49*(2), 222-251. https://doi.org/10.3102/00346543049002222

Cureton, E. E., & D'Agostino, R. B. (2013). Factor analysis: An applied approach. *Psychology press*. https://doi.org/10.4324/9781315799476

D'agostino Sr, R. B., & Russell, H. K. (2005). Scree test. *Encyclopedia of biostatistics*, 7. https://doi.org/10.1002/0470011815.b2a10082

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of family psychology*, *19*(2), 294. http://dx.doi.org/10.1037/0893-3200.19.2.294

De Jager, B., Jansen, M., & Reezigt, G. (2005). The development of metacognition in primary school learning environments. School effectiveness and school improvement, *16*(2), 179-196. https://doi.org/10.1080/09243450500114181

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13-21. https://doi.org/10.1016/j.intell.2006.02.001

Decarli, G., Surian, L., Vignoli, M., & Franchin, L. (2022). What Teachers Tell Us about the Impact of the COVID-19 Pandemic on Public Education in Italy. *European Journal of Educational Research*, *11*(2), 1103-1114. http://dx.doi.org/10.12973/eu-jer.11.2.1103

Decreto Legislativo 16 aprile 1994, n. 297. Testo Unico delle disposizioni legislative vigenti in materia di istruzione, relative alle scuole di ogni ordine e grado (s.o. G.U. n.115 del 19/5/1994). Available from: https://www.miur.gov.it/documents/20182/206879/Decreto+Legislativo+16+aprile+1994%2C+n.+297.pdf/4e92e3d8-2ada-43f2-9568-e273c3e61ad3?version=1.0&t=1495463422783

Destan, N., Hembacher, E., Ghetti, S., & Roebers, C. M. (2014). Early metacognitive abilities: The interplay of monitoring and control processes in 5-to 7-year-old children. *Journal of experimental child psychology*, *126*, 213-228. https://doi.org/10.1016/j.jecp.2014.04.001

Di Pietro, G. (2023). The impact of Covid-19 on student achievement: Evidence from a recent meta-analysis. *Educational Research Review*, *39*, 100530. https://doi.org/10.1016/j.edurev.2023.100530

Dörnyei, Z. (2001). New themes and approaches in second language motivation research. *Annual review of applied linguistics*, *21*, 43-59. https://doi.org/10.1017/S0267190501000034

Doz, E., Cuder, A., Caputi, M., Pellizzoni, S., & Passolunghi, M. C. (2023). Distance learning environment: perspective of Italian primary and secondary teachers during COVID-19 pandemic. *Learning Environments Research*, *26*(2), 555-571. https://doi.org/10.1007/s10984-022-09451-9

Dufva, M., & Voeten, M. J. (1999). Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Applied psycholinguistics*, *20*(3), 329-348. https://doi.org/10.1017/S014271649900301X

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.

Dweck, C. S. (2013). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology press.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational psychologist*, *46*(1), 6-25. https://doi.org/10.1080/00461520.2011.538645

Elfirdoussi, S., Lachgar, M., Kabaili, H., Rochdi, A., Goujdami, D., & El Firdoussi, L. (2020). Assessing distance learning in higher education during the COVID-19 pandemic. *Education Research International*, *2020*(1), 8890633. https://doi.org/10.1155/2020/8890633

Elharake, J. A., Akbar, F., Malik, A. A., Gilliam, W., & Omer, S. B. (2022). Mental health impact of COVID-19 among children and college students: A systematic review. *Child Psychiatry & Human Development*, 1-13. https://doi.org/10.1007/s10578-021-01297-1

Elsalem, L., Al-Azzam, N., Jum'ah, A. A., Obeidat, N., Sindiani, A. M., & Kheirallah, K. A. (2020). Stress and behavioral changes with remote E-exams during the Covid-19 pandemic: A cross-sectional study among undergraduates of medical sciences. *Annals of Medicine and Surgery*, *60*, 271-279. https://doi.org/10.1016/j.amsu.2020.10.058

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *118*(17), e2022376118. https://doi.org/10.1073/pnas.2022376118

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, *4*(3), 272. http://dx.doi.org/10.1037//1082-989X.4.3.272

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Feraco, T., Resnati, D., Fregonese, D., Spoto, A., & Meneghetti, C. (2023). An integrated model of school students' academic achievement and life satisfaction. Linking soft skills, extracurricular activities, self-regulated learning, motivation, and emotions. *European Journal of Psychology of Education*, *38*(1), 109-130. https://doi.org/10.1007/s10212-022-00601-4

Ferrari, M., & Palladino, P. (2007). Foreign language learning difficulties in Italian children: are they associated with other learning difficulties?. *Journal of Learning Disabilities*, *40*(3), 256-269.https://doi.org/10.1177/00222194070400030601

Ferrari, M., & Palladino, P. (2012). A longitudinal study of English as foreign language learning: L1 predictors in Italian students. *Applied Cognitive Psychology*, *26*(4), 616-625. https://doi.org/10.1002/acp.2839

Ferretti, F., Santi, G. R. P., Del Zozzo, A., Garzetti, M., & Bolondi, G. (2021). Assessment practices and beliefs: Teachers' perspectives on assessment during long distance learning. *Education Sciences*, *11*(6), 264. https://doi.org/10.3390/educsci11060264

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906. https://psycnet.apa.org/doi/10.1037/0003-066X.34.10.906

Fleming, S. M., Massoni, S., Gajdos, T., & Vergnaud, J. C. (2016). Metacognition about the past and future: quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neuroscience of Consciousness*, *2016*(1), niw018. https://doi.org/10.1093/nc/niw018

Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and mathematics achievement across the school-age years. *Psychology in the Schools*, *40*(2), 155-171. https://doi.org/10.1002/pits.10083

Fuchs, L. S., Seethaler, P. M., Fuchs, D., & Espinas, D. (2023). Severe pandemic learning loss and the promise of remotely delivered intervention in students with comorbid reading and mathematics learning difficulty. *Journal of learning disabilities*, *56*(4), 278-294. https://doi.org/10.1177/00222194231170313

Gaggi, O., Kolasinska, A. B., Mirri, S., & Prandi, C. (2020, September). The new classmate: an exploration of how CoVid-19 affected primary schools activities in Italy. In *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good* (pp. 36-41). https://doi.org/10.1145/3411170.3411239

Galasso, V., & Varriale, C. (2021). The Effectiveness of Leaders' Public Communication during COVID-19. *BAFFI CAREFIN Centre Research Paper*, (2021-162). https://dx.doi.org/10.2139/ssrn.3844502

Gallucci, M. (2019). *GAMLj: General analyses for linear models*. [jamovi module]. Retrieved from https://gamlj.github.io/.

Ganschow, L., Sparks, R. L., Javorsky, J., Pohlman, J., & Bishop-Marbury, A. (1991). Identifying native language difficulties among foreign language learners in college: a" foreign" language learning disability?. *Journal of Learning Disabilities*, *24*(9), 530-541. https://doi.org/10.1177/002221949102400905

García Bacete, F.J., Marande Perrin, G., Schneider, B.H., Blanchard, C. (2014). Effects of School on the Well-Being of Children and Adolescents. In: Ben-Arieh, A., Casas, F., Frønes, I., Korbin, J. (eds) Handbook of Child Well-Being. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-9063-8_149

Gardner, R. C. (2001). Language learning motivation: The student, the teacher, and the researcher. *Texas papers in foreign language education*, *6*(1), 1-18.

Gavosto, A., & Romano, B. (2021). COVID-19 e learning loss: quali misure senza misura. *Disponibile all'indirizzo: https://www.fondazioneagnelli.it/wp-content/uploads/2021/05/Learning-loss-e-perdita-di-capitale-umano.pdf*

Gelman, A. & Hill, J., (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge: Cambridge University Press.

Gentile, M., Benigno, V., Caruso, G., Chifari, A., Ferlino, L., Fulantelli, G., & Allegra, M. (2021). Italian parents' perception about learning practices and educational effectiveness of remote schooling during the first lockdown. *QWERTY-Interdisciplinary Journal of Technology, Culture and Education*, *16*(2), 87-108. https://doi.org/10.30557/QW000045

Ghulamuddin, N. J. A., Sharif, T. I. S. T., & Abd Rahman, Z. (2022). STEM STUDENTS' ENGLISH LANGUAGE LEARNING STRATEGIES AND UNDERLYING FACTORS PERTAINING TO OPEN DISTANCE LEARNING. *Journal of Nusantara Studies (JONUS)*, *7*(2), 81-102. https://doi.org/10.24200/jonus.vol7iss2pp81-102

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Test review behavior rating inventory of executive function. *Child Neuropsychology, 6*(3), 235-238. https://doi.org/10.1076/chin.6.3.235.3152

Gomes, D. G. (2022). Should I use fixed effects or random effects when I have fewer than five levels of a grouping factor in a mixed-effects model?. *PeerJ*, *10*, e12794. https://doi.org/10.7717/peerj.12794

Goudeau, S., Sanrey, C., Stanczak, A., Manstead, A., & Darnon, C. (2021). Why lockdown and distance learning during the COVID-19 pandemic are likely to increase the social class achievement gap. *Nature human behaviour*, *5*(10), 1273-1281. https://doi.org/10.1038/s41562-021-01212-7

Grewenig, E., Lergetporer, P., Werner, K., Woessmann, L., & Zierow, L. (2021). COVID-19 and educational inequality: How school closures affect low-and high-achieving students. *European economic review*, *140*, 103920. https://doi.org/10.1016/j.euroecorev.2021.103920

Grimm, L. G., & Yarnold, P. R. (1995). Reading and understanding multivariate statistics. *American Psychological Association*.

Guzzo, T., Boffo, S., Ferri, F., Gagliardi, F., & Grifoni, P. (2022). Towards quality digital learning: Lessons learned during COVID-19 and recommended actions—The teachers' perspective. *Sustainability*, *14*(14), 8438. https://doi.org/10.3390/su14148438

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological science*, *14*(1), 31-38. https://doi.org/10.1111/1467-9280.01415

Hale T, Webster S, Petherick A, Phillips T, Kira B. Oxford COVID-19 government response tracker (OxCGRT). Last Updat. 2020;8:30.

Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). Effects of COVID-19-related school closures on student achievement-a systematic review. *Frontiers in psychology*, *12*, 746289. https://doi.org/10.3389/fpsyg.2021.746289

Harrison, G. L., & Krol, L. (2007). Relationship between L1 and L2 word-level reading and phonological processing in adults learning English as a second language. *Journal of Research in Reading*, *30*(4), 379-393. https://doi.org/10.1111/j.1467-9817.2007.00351.x

Harvey, L. A., & Dijkers, M. P. (2019). Should trials that are highly vulnerable to bias be excluded from systematic reviews?. *Spinal Cord*, *57*(9), 715-716. https://doi.org/10.1038/s41393-019-0340-y

Hdstijn, J. H., & Bossers, B. (1992). Individual differences in L2 proficiency as a function of L1 proficiency. *European Journal of Cognitive Psychology*, *4*(4), 341-353. https://doi.org/10.1080/09541449208406192

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Ianes, D., & Bellacicco, R. (2020). Didattica a distanza durante il lockdown. L'impatto percepito dagli insegnanti sull'inclusione degli studenti con disabilità. *L'integrazione Scolastica e Sociale*, *19*(3), 25-47.

Immordino-Yang, M. H. (2016). Emotion, sociality, and the brain's default mode network: Insights for educational practice and policy. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 211-219. https://doi.org/10.1177/2372732216656869

Inguscio, B. M. S., Nicastri, M., Giallini, I., Greco, A., Babiloni, F., Cartocci, G., & Mancini, P. (2023). School wellbeing and psychological characteristics of online learning in families of children with and without hearing loss during the Covid-19 pandemic. *Psychology in the Schools*, *60*(1), 78-104. https://doi.org/10.1002/pits.22761

Jeong-Won, L. (1996). Linguistic coding deficit hypothesis in foreign language learning. *ENGLISH TEACHING (영어교육)*, *51*(3), 51-79.

Jüni, P., Altman, D. G., & Egger, M. (2001). Assessing the quality of controlled clinical trials. *Bmj*, *323*(7303), 42-46. https://doi.org/10.1136/bmj.323.7303.42

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement*, *20*(1), 141-151. https://doi.org/10.1177/001316446002000116

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive g and academic achievement g one and the same g? An exploration on the Woodcock–Johnson and Kaufman tests. *Intelligence*, *40*(2), 123-138. https://doi.org/10.1016/j.intell.2012.01.009

Kawaoka, N., Ohashi, K., Fukuhara, S., Miyachi, T., Asai, T., Imaeda, M., & Saitoh, S. (2022). Impact of school closures due to COVID-19 on children with neurodevelopmental disorders in Japan. *Journal of autism and developmental disorders*, *52*(5), 2149-2155. https://doi.org/10.1007/s10803-021-05119-0

Kee, C. E. (2021). The impact of COVID-19: Graduate students' emotional and psychological experiences. *Journal of human behavior in the social environment*, *31*(1-4), 476-488. https://doi.org/10.1080/10911359.2020.1855285

Kim, C., & Pekrun, R. (2014). Emotions and motivation in learning and performance. *Handbook of research on educational communications and technology*, 65-75. https://doi.org/10.1007/978-1-4614-3185-5_6

Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). New York: Guilford Press.

König, C., & Frey, A. (2022). The impact of COVID-19-related school closures on student achievement—A meta-analysis. *Educational Measurement: Issues and Practice*, *41*(1), 16-22. https://doi.org/10.1111/emip.12495

König, J., Jäger-Biela, D. J., & Glutsch, N. (2020). Adapting to online teaching during COVID-19 school closure: teacher education and teacher competence effects among early career teachers in Germany. *European journal of teacher education*, *43*(4), 608-622. https://doi.org/10.1080/02619768.2020.1809650

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and cognition*, *11*(2), 261-271. https://doi.org/10.1017/S1366728908003416

Kreutzer, M. A., Leonard, C., Flavell, J. H., & Hagen, J. W. (1975). An interview study of children's knowledge about memory. *Monographs of the society for research in child development*, 1-60. https://doi.org/10.2307/1165955

Kriegbaum, K., Becker, N., & Spinath, B. (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, *25*, 120-148. https://doi.org/10.1016/j.edurev.2018.10.001

Krueger, C., & Tian, L. (2004). A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. *Biological research for nursing*, *6*(2), 151-157. https://doi.org/10.1177/1099800404267682

Kuhfeld, M., Lewis, K., & Peltier, T. (2023). Reading achievement declines during the COVID-19 pandemic: evidence from 5 million US students in grades 3–8. *Reading and Writing*, *36*(2), 245-261. https://doi.org/10.1007/s11145-022-10345-8

Lado, R. (1957). Linguistics across cultures; applied linguistics for language teachers.

Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159174.

Lasagabaster, D. (2001). The effect of knowledge about the L1 on foreign language skills and grammar. *International Journal of Bilingual Education and Bilingualism*, *4*(5), 310-331. https://doi.org/10.1080/13670050108667735

Lassoued, Z., Alhendawi, M., & Bashitialshaaer, R. (2020). An exploratory study of the obstacles for achieving quality in distance learning during the COVID-19 pandemic. *Education sciences*, *10*(9), 232. https://doi.org/10.3390/educsci10090232

Lavis, L., & Mahy, C. E. (2021). "I'll remember everything no matter what!": The role of metacognitive abilities in the development of young children's prospective memory. *Journal of Experimental Child Psychology*, *207*, 105117. https://doi.org/10.1016/j.jecp.2021.105117

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.

Lee, M. (2014). *Achievement goals, emotions, and foreign language performance in German and Korean students* (Doctoral dissertation, Verlag nicht ermittelbar).

Leventhal, H., & Scherer, K. (1987). The relationship of emotion to cognition: A functional approach to a semantic controversy. *Cognition and emotion*, *1*(1), 3-28. https://doi.org/10.1080/02699938708408361

Li, C., & Wei, L. (2023). Anxiety, enjoyment, and boredom in language learning amongst junior secondary students in rural China: How do they contribute to L2 achievement?. *Studies in Second Language Acquisition*, *45*(1), 93-108. https://doi.org/10.1017/S0272263122000031

Lichtenfeld, S., Pekrun, R., Stupnisky, R. H., Reiss, K., & Murayama, K. (2012). Measuring students' emotions in the early years: the achievement emotions questionnaire-elementary school (AEQ-ES). *Learning and Individual differences*, *22*(2), 190-201. https://doi.org/10.1016/j.lindif.2011.04.009

Linnenbrink, E. A., & Pintrich, P. R. (2002). Achievement goal theory and affect: An asymmetrical bidirectional model. *Educational psychologist*, *37*(2), 69-78. https://doi.org/10.1207/S15326985EP3702_2

Lüdecke, D., Ben-Shachar, M. S., Patil, I., & Makowski, D. (2020). Extracting, computing and exploring the parameters of statistical models using R. *Journal of Open Source Software*, *5*(53), 2445. https://doi.org/10.21105/joss.02445

MacIntyre, P. D. (1995). How does anxiety affect second language learning? A reply to Sparks and Ganschow. *The modern language journal*, *79*(1), 90-99. https://doi.org/10.1111/j.1540-4781.1995.tb05418.x

Mangiavacchi, L., Piccoli, L., & Pieroni, L. (2021). Fathers matter: Intrahousehold responsibilities and children's wellbeing during the COVID-19 lockdown in Italy. *Economics & Human Biology*, *42*, 101016. https://doi.org/10.1016/j.ehb.2021.101016

Marchese, F., Grillo, A., Mantovani, M. P., Gargano, G., Limone, P., & Indrio, F. (2022). Specific learning disorders and special educational needs during covid-19 pandemic; pilot survey study performed in local district schools in italy. *Children*, *9*(6), 825. https://doi.org/10.3390/children9060825

Marinelli, C. V., Romani, C., Burani, C., McGowan, V. A., & Zoccolotti, P. (2016). Costs and benefits of orthographic inconsistency in reading: Evidence from a cross-linguistic comparison. *PLoS One*, *11*(6), e0157457. https://doi.org/10.1371/journal.pone.0157457

Maydeu-Olivares, A. (2017). Maximum likelihood estimation of structural equation models for continuous data: Standard errors and goodness of fit. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(3), 383-394. https://doi.org/10.1080/10705511.2016.1269606

McCombs, B. L. (2013). Self-regulated learning and academic achievement: A phenomenological view. In *Self-regulated learning and academic achievement* (pp. 63-117). Routledge.

McDonald, R.P. (1999). *Test Theory: A Unified Treatment* (1st ed.). Psychology Press. https://doi.org/10.4324/9781410601087

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological methods*, 23(3), 412. https://psycnet.apa.org/doi/10.1037/met0000144

Meherali, S., Punjani, N., Louie-Poon, S., Abdul Rahim, K., Das, J. K., Salam, R. A., & Lassi, Z. S. (2021). Mental health of children and adolescents amidst COVID-19 and past

pandemics: a rapid systematic review. *International journal of environmental research and public health*, *18*(7), 3432. https://doi.org/10.3390/ijerph18073432

Meyer, E., Abrami, P. C., Wade, C. A., Aslan, O., & Deault, L. (2010). Improving literacy and metacognition with electronic portfolios: Teaching and learning with ePEARL. *Computers & Education, 55*(1), 84-91. https://doi.org/10.1016/j.compedu.2009.12.005

Mirman, D. (2017). *Growth curve analysis and visualization using R*. Chapman and Hall/CRC. https://doi.org/10.1201/9781315373218

Moè, A. (2010). *La Motivazione*. Il mulino.

Muthén, L. K., & Muthén, B. O. (2017). *Version 8 Mplus user's guide*. Los Angeles, CA: Muthén & Muthén, Author.

Muthén, L. K., & Muthén, B. O. (2023). Mplus: Statistical analysis with latent variables.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory, 26*, 125-173. New York: Academic Press. https://doi.org/10.1016/S0079-7421(08)60053-5

Oga-Baldwin, W. Q., Nakata, Y., Parker, P., & Ryan, R. M. (2017). Motivating young language learners: A longitudinal model of self-determined motivation in elementary school foreign language classes. *Contemporary educational psychology*, *49*, 140-150. https://doi.org/10.1016/j.cedpsych.2017.01.010

Onofrei, S. G., & Iancu, L. (2015). The role of new technology in teaching through proverbs in primary school. *Procedia-Social and Behavioral Sciences*, *203*, 130-133. https://doi.org/10.1016/j.sbspro.2015.08.270

Orcan, F. (2020). Parametric or non-parametric: Skewness to test normality for mean comparison. *International Journal of Assessment Tools in Education*, *7*(2), 255-265.

Orsini, A., Pezzuti, L., & Picone, L. (2012). WISC–IV: Contributo alla taratura Italiana. [WISC–IV Italian Edition]. Florence: Giunti O. S.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic reviews*, *5*, 1-10. https://doi.org/10.1186/s13643-016-0384-4

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

Palladino, P. & Cornoldi, C. (2023). I disturbi specifici del linguaggio, le difficoltà di apprendimento della lingua straniera e il caso degli studenti bilingui o con DSA. In C. Cornoldi, I disturbi dell'apprendimento, Il mulino, Bologna.

Palladino, P., & Cornoldi, C. (2004). Working memory performance of Italian students with foreign language learning difficulties. *Learning and individual differences*, *14*(3), 137-151. https://doi.org/10.1016/j.lindif.2004.01.001

Palladino, P., & Ferrari, M. (2008). Phonological sensitivity and memory in children with a foreign language learning difficulty. *Memory*, *16*(6), 604-625. https://doi.org/10.1080/09658210802083072

Palladino, P., Bellagamba, I., Ferrari, M., & Cornoldi, C. (2013). Italian children with dyslexia are also poor in reading English words, but accurate in reading English pseudowords. *Dyslexia*, *19*(3), 165-177. https://doi.org/10.1002/dys.1456

Palladino, P., Botto, M., Bellagamba, I., Ferrari, M., Cornoldi, C. (2019). *English is fun!*. Vol. 1. Trento: Erickson.

Palladino, P., Cismondo, D., Ferrari, M., Ballagamba, I., & Cornoldi, C. (2016). L2 spelling errors in Italian children with dyslexia. *Dyslexia*, *22*(2), 158-172. https://doi.org/10.1002/dys.1522

Palladino, P., Trotta, E., Bonvino, A., Carlucci, L., Cottini, M. (under review).How do you feel during English class? Emotions and Metacognition in primary school children learning English as a second language. *Metacognition and Learning.*

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in psychology*, *8*, 422. https://doi.org/10.3389/fpsyg.2017.00422

Panagouli, E., Stavridou, A., Savvidi, C., Kourti, A., Psaltopoulou, T., Sergentanis, T. N., & Tsitsika, A. (2021). School performance among children and adolescents during COVID-19 pandemic: a systematic review. *Children*, *8*(12), 1134. https://doi.org/10.3390/children8121134

Panchal, U., Salazar de Pablo, G., Franco, M., Moreno, C., Parellada, M., Arango, C., & Fusar-Poli, P. (2023). The impact of COVID-19 lockdown on child and adolescent mental health: systematic review. *European child & adolescent psychiatry*, *32*(7), 1151-1177. https://doi.org/10.1007/s00787-021-01856-w

Pedone, R., Semerari, A., Riccardi, I., Procacci, M., Nicolò, G., & Carcione, A. (2017). Development of a self-report measure of metacognition: The Metacognition Self-Assessment

Scale (MSAS). Instrument description and factor structure. *Clinical Neuropsychiatry*, *14*(3), 185-194.

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development: Developing motivation and motivating development* (pp. 143–163). Elsevier Science. https://doi.org/10.1016/S0166-4115(00)80010-2

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*, *18*, 315-341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary educational psychology*, *36*(1), 36-48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Peterson, J., Welch, V., Losos, M., & Tugwell, P. J. O. O. H. R. I. (2011). The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. *Ottawa: Ottawa Hospital Research Institute*, *2*(1), 1-12.

Pfefferbaum, B. (2021). Challenges for child mental health raised by school closure and home confinement during the COVID-19 pandemic. *Current psychiatry reports*, *23*(10), 65. https://doi.org/10.1007/s11920-021-01279-z

Picca, M., Manzoni, P., Milani, G. P., Mantovani, S., Cravidi, C., Mariani, D., ... & Agostoni, C. (2021). Distance learning, technological devices, lifestyle and behavior of children and their family during the COVID-19 lockdown in Lombardy: a survey. *Italian Journal of Pediatrics*, *47*, 1-6. https://doi.org/10.1186/s13052-021-01156-8

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of educational Psychology*, *95*(4), 667. https://doi.org/10.1037/0022-0663.95.4.667

Pokhrel, S., & Chhetri, R. (2021). A literature review on impact of COVID-19 pandemic on teaching and learning. *Higher education for the future*, *8*(1), 133-141. https://doi.org/10.1177/2347631120983481

Prior, A., Goldina, A., Shany, M., Geva, E., & Katzir, T. (2014). Lexical inference in L2: Predictive roles of vocabulary knowledge and reading skill beyond reading

comprehension. *Reading and Writing*, *27*, 1467-1484. https://doi.org/10.1007/s11145-014-9501-8

Putwain, D. W., Sander, P., & Larkin, D. (2013). Using the 2× 2 framework of achievement goals to predict achievement emotions and academic performance. *Learning and Individual Differences*, *25*, 80-84. https://doi.org/10.1016/j.lindif.2013.01.006

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian journal of educational research*, *45*(3), 269-286. https://doi.org/10.1080/00313830120074206

Quílez-Robres, A., González-Andrade, A., Ortega, Z., & Santiago-Ramajo, S. (2021). Intelligence quotient, short-term memory and study habits as academic achievement predictors of elementary school: A follow-up study. *Studies in Educational Evaluation*, *70*, 101020. https://doi.org/10.1016/j.stueduc.2021.101020

R Core Team (2022). R: A Language and environment for statistical computing. (Version 4.1) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from CRAN snapshot 2023-04-07).

R Core Team (2024). *R: A Language and environment for statistical computing*. (Version 4.4) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from CRAN snapshot 2024-08-07).

Raccanello, D., Brondino, M., Moè, A., Stupnisky, R., & Lichtenfeld, S. (2019). Enjoyment, boredom, anxiety in elementary schools in two domains: Relations with achievement. *The Journal of Experimental Education*, *87*(3), 449-469. https://doi.org/10.1080/00220973.2018.1448747

Ranieri, M., Gaggioli, C., & Borges, M. K. (2020). La didattica alla prova del Covid-19 in Italia: uno studio sulla Scuola Primaria. *Praxis educativa*, *15*. https://doi.org/10.5212/praxeduc.v.15.16307.079

Rankin, T. (2017). (Working) memory and L2 acquisition and processing. *Second Language Research*, *33*(3), 389-399. https://doi.org/10.1177/0267658316645387

Raoofi, S., Chan, S. H., Mukundan, J., & Rashid, S. M. (2014). Metacognition and second/foreign language learning. *English Language Teaching*, *7*(1), 36-49. http://dx.doi.org/10.5539/elt.v7n1p36

Rapporto INVALSI (2022). Istituto Nazionale Per La Valutazione Del Sistema Educativo Di Istruzione E Di Formazione. Available from: https://invalsi-

areaprove.cineca.it/docs/2022/Rilevazioni_Nazionali/Rapporto/Rapporto_Prove_INVALSI_2022.pdf

Rapporto INVALSI (2023). Istituto Nazionale Per La Valutazione Del Sistema Educativo Di Istruzione E Di Formazione. Available from: https://invalsi-areaprove.cineca.it/docs/2023/Rilevazioni_Nazionali/Rapporto/Rapporto%20Prove%20INVALSI%202023.pdf

Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research. [R package]. Retrieved from https://cran.r-project.org/package=psych.

Roebers, C. M. (2002). Confidence judgments in children's and adult's event recall and suggestibility. *Developmental psychology*, *38*(6), 1052. https://psycnet.apa.org/doi/10.1037/0012-1649.38.6.1052

Roebers, C. M., & Spiess, M. (2017). The development of metacognitive monitoring and control in second graders: A short-term longitudinal study. *Journal of cognition and development*, *18*(1), 110-128. https://doi.org/10.1080/15248372.2016.1157079

Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: An explorative study on their interrelations and their role for school achievement. *Metacognition and Learning*, *7*, 151-173. https://doi.org/10.1007/s11409-012-9089-9

Roebers, C. M., Schmid, C., & Roderer, T. (2009). Metacognitive monitoring and control processes involved in primary school children's test performance. *British Journal of Educational Psychology, 79*(4), 749-767. https://doi.org/10.1348/978185409X429842

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118-137. https://doi.org/10.1016/j.intell.2015.09.002

Rubini, V., & Rossi, M. A. (1982). Analysis of the psychometric characteristics of the Primary Mental Abilities Tests (Level 4-sup-6, 1962 revision) applied to a group of Italian subjects. *Bollettino di Psicologia Applicata*, *161*, 87-105.

Saulle, R., De Sario, M., Bena, A., Capra, P., Culasso, M., Davoli, M., ... & Minozzi, S. (2022). School closures and mental health, wellbeing and health behaviours among children and adolescents during the second COVID-19 wave: a systematic review of the literature. *Epidemiologia e prevenzione*, *46*(5-6), 1-20. https://dx.doi.org/10.19191/EP22.5-6.A542.089

Scarpellini, F., Segre, G., Cartabia, M., Zanetti, M., Campi, R., Clavenna, A., & Bonati, M. (2021). Distance learning in Italian primary and middle school children during the COVID-19 pandemic: a national survey. *BMC public health*, *21*(1), 1035. https://doi.org/10.1186/s12889-021-11026-x

Scheer, D., & Laubenstein, D. (2021). The impact of COVID-19 on mental health: psychosocial conditions of students with and without special educational needs. *Social Sciences*, *10*(11), 405. https://doi.org/10.3390/socsci10110405

Schneider, W. J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. *Contemporary intellectual assessment: Theories, tests, and issues*, 73-163.

Schneider, W., & Niklas, F. (2017). Intelligence and verbal short-term memory/working memory: Their interrelationships from childhood to young adulthood and their impact on academic achievement. *Journal of Intelligence*, *5*(2), 26. https://doi.org/10.3390/jintelligence5020026

Schneider, W., & Schneider, W. (2015). The development of metamemory. *Memory development from early childhood through emerging adulthood*, 255-308. https://doi.org/10.1007/978-3-319-09611-7_9

Schraw, G., & Dennison, R. S. (1994). Metacognitive awareness inventory. *Contemporary Educational Psychology*. https://psycnet.apa.org/doi/10.1037/t21885-000

Schult, J., Mahler, N., Fauth, B., & Lindner, M. A. (2022, April). Long-term consequences of repeated school closures during the COVID-19 pandemic for reading and mathematics competencies. In *Frontiers in Education* (Vol. 13, p. 867316). Frontiers Media SA. https://doi.org/10.3389/feduc.2022.867316

Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools*, *22*(2), 208-223. https://doi.org/10.1002/1520-6807(198504)22:2<208::AID-PITS2310220215>3.0.CO;2-7

Sector E. Covid 19–A glance of national coping strategies on high-stakes examinations and assessments [Internet]. Working Document. Paris: Unesco. https://en. unesco. org/sites/default/files …; 2020 [cited 2023 Dec 20]. Available from: https://dubito.s3.eu-central-1.amazonaws.com/dubito-v2/2020/5/19/1592576185808-unesco_review_of_high-stakes_exams_and_assessments_during_covid-19_en.pdf

Segre, G., Campi, R., Scarpellini, F., Clavenna, A., Zanetti, M., Cartabia, M., & Bonati, M. (2021). Interviewing children: the impact of the COVID-19 quarantine on children's

perceived psychological distress and changes in routine. *BMC pediatrics*, *21*(1), 231. https://doi.org/10.1186/s12887-021-02704-1

Shao, K., Pekrun, R., & Nicholson, L. J. (2019). Emotions in classroom language learning: What can we learn from achievement emotion research?. *System*, *86*, 102121. https://doi.org/10.1016/j.system.2019.102121

Siedlecka, M., Paulewicz, B., & Wierzchoń, M. (2016). But I was so sure! Metacognitive judgments are less accurate given prospectively than retrospectively. *Frontiers in psychology*, *7*, 218. https://doi.org/10.3389/fpsyg.2016.00218

Smith, L., Fowler, S., Gabriel, F., & Brinkman, S. (2023). Measuring Metacognition in Children and Adolescents.

Snijders, T. A. B., & Bosker, R. J. (Eds.). (2012). Multilevel analysis: An introduction to basic and advanced multilevel modeling (2nd ed.). London, UK: Sage.

Sohn, Y. W., & Doane, S. M. (2003). Roles of working memory capacity and long-term working memory skill in complex task performance. *Memory & cognition*, *31*, 458-466. https://doi.org/10.3758/BF03194403

Sparks, R. L. (2019). Why reading is a challenge for US L2 learners: The impact of cognitive, ecological, and psychological factors in L2 comprehension. *Foreign Language Annals*, *52*(4), 727-743. https://doi.org/10.1111/flan.12432

Sparks, R. L., & Ganschow, L. (1993). The impact of native language learning problems on foreign language learning: Case study illustrations of the linguistic coding deficit hypothesis. *The Modern Language Journal*, *77*(1), 58-74. https://doi.org/10.2307/329559

Sparks, R. L., & Ganschow, L. (1995). A strong inference approach to causal factors in foreign language learning: A response to MacIntyre. *The Modern Language Journal*, *79*(2), 235-244. https://doi.org/10.1111/j.1540-4781.1995.tb05436.x

Sparks, R. L., & Patton, J. (2013). Relationship of L1 skills and L2 aptitude to L2 anxiety on the Foreign Language Classroom Anxiety Scale. *Language Learning*, *63*(4), 870-895. https://doi.org/10.1111/lang.12025

Sparks, R. L., Patton, J. O. N., Ganschow, L., & Humbach, N. (2009). Long-term relationships among early first language skills, second language aptitude, second language affect, and later second language proficiency. *Applied Psycholinguistics*, *30*(4), 725-755. https://doi.org/10.1017/S0142716409990099

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2012). Do L1 reading achievement and L1 print exposure contribute to the prediction of L2 proficiency?. *Language Learning*, *62*(2), 473-505. https://doi.org/10.1111/j.1467-9922.2012.00694.x

Sparks, R., & Ganschow, L. (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics*, *21*, 90-111. https://doi.org/10.1017/S026719050100006X

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., ... & Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *bmj*, *355*. https://doi.org/10.1136/bmj.i4919

Stewart, J. H. (2005). Foreign language study in elementary schools: Benefits and implications for achievement in reading and math. *Early Childhood Education Journal*, *33*(1), 11-16. https://doi.org/10.1007/s10643-005-0015-5

Sulla, F., Ragni, B., D'Angelo, M., Rollo, D. (2022) Teachers' emotions, technostress, and burnout in distance learning during the COVID-19 pandemic. In: Proceedings of the Third Workshop of Technology Enhanced Learning Environments for Blended Education–The Italian e-Learning Conference. p. 1–13. Available from: https://www.researchgate.net/publication/365172879_Teachers%27_emotions_technostress_and_burnout_in_distance_learning_during_the_COVID-19_pandemic

Swanson, H. L., & Alloway, T. P. (2012). Working memory, learning, and academic achievement. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol. 1. Theories, constructs, and critical issues* (pp. 327–366). American Psychological Association. https://doi.org/10.1037/13273-012

Szpunar, G., Cannoni, E., & Di Norcia, A. (2021). La didattica a distanza durante il lockdown in Italia: il punto di vista delle famiglie. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, (23), 137-155. https://doi.org/10.7358/ecps-2021-023-szpu

Tan, J., Mao, J., Jiang, Y., & Gao, M. (2021). The influence of academic emotions on learning effects: a systematic review. *International journal of environmental research and public health*, *18*(18), 9678. https://doi.org/10.3390/ijerph18189678

Taouki, I., Lallier, M., & Soto, D. (2022). The role of metacognition in monitoring performance and regulating learning in early readers. *Metacognition and Learning*, *17*(3), 921-948. https://doi.org/10.1007/s11409-022-09292-0

Taylor, J. G., & Fragopanagos, N. F. (2005). The interaction of attention and emotion. *Neural networks*, *18*(4), 353-369. https://doi.org/10.1016/j.neunet.2005.03.005

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *41*(2), 363-387. https://doi.org/10.1017/S0272263118000311

Termine, C., Dui, L. G., Borzaga, L., Galli, V., Lipari, R., Vergani, M., ... & Ferrante, S. (2021). Investigating the effects of COVID-19 lockdown on Italian children and adolescents with and without neurodevelopmental disorders: a cross-sectional study. *Current Psychology*, 1-17. https://doi.org/10.1007/s12144-021-02321-2

Termine, C., Galli, V., Dui, L. G., Berlusconi, V., Taras, R., Vergani, M., ... & Cavanna, A. E. (2022). Self-reported impact of the COVID-19 pandemic and lockdown on young patients with tic disorders: findings from a case–control study. *Neurological Sciences*, *43*(6), 3497-3501. https://doi.org/10.1007/s10072-022-05997-x

The jamovi project (2023). jamovi. (Version 2.4) [Computer Software]. Retrieved from https://www.jamovi.org.

The jamovi project (2024). *jamovi*. (Version 2.6) [Computer Software]. Retrieved from https://www.jamovi.org.

Thorell, L. B., Fuermaier, A. B., Christiansen, H., Steinmayr, R., Baeyens, D., de la Peña, A. G., ... & Skoglund, C. (2022). Distance learning during the COVID-19 pandemic for children with ADHD and/or ASD: A European multi-center study examining the role of executive function deficits and age. *Child and Adolescent Psychiatry and Mental Health*, *16*(1), 101. https://doi.org/10.1186/s13034-022-00540-4

Thorell, L. B., Skoglund, C., de la Peña, A. G., Baeyens, D., Fuermaier, A. B., Groom, M. J., ... & Christiansen, H. (2021). Parental experiences of homeschooling during the COVID-19 pandemic: Differences between seven European countries and between children with and without mental health conditions. *European child & adolescent psychiatry*, 1-13. https://doi.org/10.1007/s00787-020-01706-1

Thorn, W., & Vincent-Lancrin, S. (2021). *Schooling During a Pandemic The Experience and Outcomes of Schoolchildren During the First Round of COVID-19 Lockdowns: The Experience and Outcomes of Schoolchildren During the First Round of COVID-19 Lockdowns*. OECD Publishing. Available from: https://www.oecd-ilibrary.org/education/schooling-during-a-pandemic_1c78681e-en

Trotta, E., Cottini, M., Cannito, L., Palladino, P. (under review [A]). Difficoltà nel vocabolario L1 e apprendimento della lingua inglese L2: un'indagine sulle emozioni legate al contesto di apprendimento alla scuola primaria. *Ricerche di Psicologia, 2024, Vol. 47 (2).* ISSN 0391-6081, ISSNe 1972-5620.

Trotta, E., Serio, G., Bonvino, A., Favini, A., Cottini, M., Palladino, P. (under review [B]). English Metacognitive Questionnaire (EMQ): Analisi delle proprietà psicometriche dello strumento nella scuola primaria (EMQ–PS). Giornale Italiano di Psicologia.

Trotta, E., Serio, G., Monacis, L., Carlucci, L., Marinelli, C. V., Petito, A., ... & Palladino, P. (2024). The effects of the COVID-19 pandemic on Italian primary school children's learning: A systematic review through a psycho-social lens. *Plos one*, *19*(6), e0303991. https://doi.org/10.1371/journal.pone.0303991

Tucker, L. R., & MacCallum, R. C. (1997). Exploratory factor analysis. *Unpublished manuscript, Ohio State University, Columbus*, 1-459.

Tzohar-Rozen, M., & Kramarski, B. (2014). Metacognition, motivation and emotions: Contribution of self-regulated learning to solving mathematical problems. *Global Education Review*, *1*(4).

Viner, R., Russell, S., Saulle, R., Croker, H., Stansfield, C., Packer, J., ... & Minozzi, S. (2022). School closures during social lockdown and mental health, health behaviors, and well-being among children and adolescents during the first COVID-19 wave: a systematic review. *JAMA pediatrics*, *176*(4), 400-409. 10.1001/jamapediatrics.2021.5840

Vulchanova, M., Foyn, C. H., Nilsen, R. A., & Sigmundsson, H. (2014). Links between phonological memory, first language competence and second language competence in 10-year-old children. *Learning and Individual Differences*, *35*, 87-95. https://doi.org/10.1016/j.lindif.2014.07.016

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *The Journal of Educational Research*, *84*(1), 30-43. https://doi.org/10.1080/00220671.1990.10885988

Wang, N. (2024). How does basic psychological needs satisfaction contribute to EFL learners' achievement and positive emotions? The mediating role of L2 self-concept. *System*, *123*, 103340. https://doi.org/10.1016/j.system.2024.103340

Warrens, M. J. (2015). Some relationships between Cronbach's alpha and the Spearman-Brown formula. *Journal of Classification*, *32*, 127-137. https://doi.org/10.1007/s00357-015-9168-0

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of black psychology*, *44*(3), 219-246. https://doi.org/10.1177/0095798418771807

Whitebread, D., Coltman, P., Pasternak, D. P., Sangster, C., Grau, V., Bingham, S., ... & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and learning, 4*, 63-85. https://doi.org/10.1007/s11409-008-9033-1

Wickham, H. (2011). ggplot2. *Wiley interdisciplinary reviews: computational statistics*, *3*(2), 180-185. https://doi.org/10.1002/wics.147

Wickrama, K., Lee, T. K., O'Neal, C. W., & Lorenz, F. (2021). *Higher-order growth curves and mixture modeling with Mplus: A practical guide*. Routledge. https://doi.org/10.4324/9781003158769

Williams, J. P., & Atkins, J. G. (2009). The role of metacognition in teaching reading comprehension to primary students. In *Handbook of metacognition in education* (pp. 26-43). Routledge.

Wood, L., Egger, M., Gluud, L. L., Schulz, K. F., Jüni, P., Altman, D. G., ... & Sterne, J. A. (2008). Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *bmj*, *336*(7644), 601-605. https://doi.org/10.1136/bmj.39465.451748.AD

World Health Organization. Considerations for school-related public health measures in the context of COVID-19: annex to considerations in adjusting public health and social measures in the context of COVID-19, 14 September 2020 [Internet]. World Health Organization; 2020 [cited 2023 Dec 20]. Available from: https://apps.who.int/iris/bitstream/handle/10665/334294/WHO-2019-nCoV-Adjusting_PH_measures-Schools-2020.2-eng.pdf

Wu, C., Jing, B., Gong, X., Mou, Y., & Li, J. (2021). Student's learning strategies and academic emotions: their influence on learning satisfaction during the COVID-19 pandemic. *Frontiers in psychology*, *12*, 717683. https://doi.org/10.3389/fpsyg.2021.717683

Wu, R., & Yu, Z. (2022). Exploring the effects of achievement emotions on online learning outcomes: A systematic review. *Frontiers in psychology*, *13*, 977931. https://doi.org/10.3389/fpsyg.2022.977931

Yusuf, B. N., & Ahmad, J. (2020). Are we prepared enough? A case study of challenges in online learning in a private higher learning institution during the Covid-19 outbreaks. *Advances in Social Sciences Research Journal*, *7*(5), 205-212. 10.14738/assrj.75.8211

Zaboski II, B. A., Kranzler, J. H., & Gage, N. A. (2018). Meta-analysis of the relationship between academic achievement and broad abilities of the Cattell-Horn-Carroll theory. *Journal of school psychology*, *71*, 42-56. https://doi.org/10.1016/j.jsp.2018.10.001

Zaccoletti, S., Camacho, A., Correia, N., Aguiar, C., Mason, L., Alves, R. A., & Daniel, J. R. (2020). Parents' perceptions of student academic motivation during the COVID-19 lockdown: A cross-country comparison. *Frontiers in psychology*, *11*, 592670. https://doi.org/10.3389/fpsyg.2020.592670

Zhang, W., Wang, Y., Yang, L., & Wang, C. (2020). Suspending classes without stopping learning: China's education emergency management policy in the COVID-19 outbreak. *Journal of Risk and financial management*, *13*(3), 55. https://doi.org/10.3390/jrfm13030055

Zimmerman, B. J., & Schunk, D. H. (2004). Self-regulating intellectual processes and outcomes: A social cognitive perspective. In *Motivation, emotion, and cognition* (pp. 337-364). Routledge.