Open University of Cyprus

Faculty of Pure & Applied Sciences

Master's Programme of Study
Cognitive Systems

Postgraduate (Master's) Dissertation

Cognitive System Design for Motivation and Self-Regulation - A Proposed Theoretical Framework for Intelligent Tutoring Systems

Andreas Tsiridis

Supervisor
Dr Maria Sofologi

December 2023

Faculty of Pure & Applied Sciences

Master's Programme of Study
Cognitive Systems

Postgraduate (Master's) Dissertation

Cognitive System Design for Motivation and Self-Regulation - A Proposed Theoretical Framework for Intelligent Tutoring Systems

Andreas Tsiridis

Supervisor
Dr Maria Sofologi

The present Postgraduate (Master's) Dissertation was submitted in partial fulfilment of the requirements for the postgraduate degree in Cognitive Systems

Faculty of Pure & Applied Sciences

December 2023

BLANK PAGE

Abstract

The development of Intelligent Tutoring Systems has been a sought-after and rapidly growing field for research, especially now, with the emergence of Generative Artificial Intelligence. This study aspires to contribute to the topic of cognitive systems for education by proposing a Cognitive Architecture which encapsulates three contemporary psychological theories and constructs for learning, motivation, and self-regulation, namely the Zone of Proximal Development, Self-Determination Theory and Self-Regulated Learning by providing a theoretical blueprint for an Intelligent Tutoring Systems for children of 8 to12 years of age. The study employs surveys and experimental designs to preliminary tap into correlations between constructs of the three theories of contemporary approaches. By extracting data using instruments and cognitive tasks to the relevant population and their parents and teachers, the researcher attempts to identify any associations between items and factors and converge on a minimal set of variables and predictors, which in turn may lead to an efficient computational design model for a cognitive assistant that will employ optimal strategies for learning. However, the results suggest that more complex experimental designs may be needed to tap into the nuances of self-regulation and motivation. Finally, the study attempts to converge findings from the literature and offer a well-informed summarisation to psychologists, cognitive scientists, software architects and developers for future designs.

Acknowledgements

This master's degree has been a fulfilling experience, with thought-provoking courses, meaningful research, and tremendous personal growth along the way. A journey concluded with this thesis and, hopefully, continued after it. I am deeply grateful to all those who selflessly took time from their busy lives to participate in my research study. This project could not have succeeded without their input.

It would be unfair to thank the academic personnel of the Open University of Cyprus, my teachers, with a mere generic *"thank you"* note, so I will thank them one by one.
Therefore, I would like to thank Prof. Loizos Michael, Dr Ilianna Kollia, Prof. Dimitris Kalles, Dr Isidoros Perrikos, and Dr Panagiota Theodoni for all the valuable knowledge they offered to me. Also, special thanks are due to my teachers and members of the evaluation committee, Prof. Marios Avraamides and Dr Konstantinos Tsagkaridis.
Last, but *definitely* not least, I would like to thank my supervisor, Dr Maria Sofologi, for her guidance, support, and patience –especially for the latter as the date approached the submission deadline.

Most importantly, I wish to thank my daughter, Electra. Pursuing this degree demanded compromises in quality time spent together, with me trying to be a proper student at the expense of being a proper dad. Yet despite the sacrifices, she never stopped inspiring me and has always been my greatest supporter. I cannot fully express my graditude for her encouragement through every late night and early morning.
It is to *you*, my beloved Electra, that I proudly dedicate this thesis. You are a constant reminder that knowledge should be used to make the world a little brighter.

Table of contents

Chapter 1
Introduction

1.1 General Introduction

The concept of utilizing technological tools to support learning is not new. It stems back to Skinner's introduction of his teaching machines in the 1950s when he said that they would make teachers obsolete and the growth of cognitive-behavioural techniques in psychology (Skinner, 1954). Despite the extremely woeful, nihilistic, or hopeful –depending on the viewer's viewpoint– nature of this assertion, it stimulated the creation of early instructional software in the 1960s and the 1970s, when several AI researchers started focusing on education, creating automated teaching systems and other tools that, according to early cognitive scientists and theorists like Roger Schank and Seymour Papert, might change the educational system (Sawyer, 2022).

In the past twenty years, there has been a significant shift in how technology has developed and been interwoven into daily life. With the advancement of technology, we can now process more data more quickly than before; the advancements in the field of AI open opportunities offer new opportunities for a variety of fields, including adaptive learning (Singh, 2023).

Technology-enhanced learning refers to the use of systems that incorporate technology, enabling students to acquire knowledge and skills, with the guidance of instructors tutors, learning tools and technological resources (Gros, 2016). These systems have gained significance during the pandemic as they assisted educators, in re-evaluating and improving their course designs to provide more meaningful learning experiences for their students (Pappas & Giannakos, 2021).

Learning can be positively impacted by utilising the findings and 'effects' of cognitive psychology research, especially in computer-based adaptive systems (Sinatra, 2018). However, according to (Sawyer, 2022), the potential of computers in educational settings has not yet been fully exploited; up until recently, educational software was developed on *instructionalist principles,* with the computer filling roles that were previously handled by instructors — the program acts as an expert authority and conveys knowledge to the learner.

According to the same source, learning sciences research suggests that the computer should take on a more facilitative role, helping students have the kinds of experiences that lead to deep learning, which involves critical thinking and the assimilation of new concepts to pre-existing knowledge (Filius et al., 2018).

Due to the exponential rise in computing power and developments in machine learning, AI has established itself as a common companion in contemporary society, including education (Singh, 2023). As such, when it comes to an educational setting, AI should consider both the sought-over educational outcomes but also young students' *well-being;* design considerations cannot exist in a vacuum, since different choices may lead to a different outcome when the AI is put to use; therefore developers must be sensible of the human, psychological, and ethical considerations when creating an AI product, to ensure the qualities necessary for bringing long-term effects of AI to use on users' welfare (De Vreede et al., 2021). Transferring best practices applied in educational settings to AI system design could be beneficial for learners. For example, a *'Student-centered'* learning environment where student responsibility and participation are prioritised above course material or tutoring activities, enhances self-efficacy and autonomy, leading to improved levels of motivation (Smit et al., 2014). Therefore, it could be beneficial to keep this in mind when designing AI systems for education.

In this vein, human curiosity, inquisitive nature and eagerness to learn are opportunities that may be exploited (Loewenstein, 1994), promoting interest, and sought-after coherence in information (Ryan & Weinstein, 2009) as they guide their students in learning, keeping them in the best possible learning form, which seems to be important for both academic achievement and the process of acquiring knowledge. This may be achieved in a variety of methods depending on the learner's motivation, cognitive capacity, and interaction with their surroundings, classmates, and teachers, as well as by employing technological means (Sawyer, 2022).

AI-assisted systems may also be used to tackle practical issues, such as overcrowded classrooms. As low-income, minority and disadvantaged students tend to experience poorer outcomes in large classes compared to their peers (Bosworth, 2011; Schanzenbach, 2014), China has taken a turn towards using AI adaptive systems to tackle this issue, as junior secondary classes have an average of 52 pupils; Wang et al. (2020) report that using an adaptive tutoring system significantly improved pupils' performance over the ones who only attended a class. The results were also significant for two groups of students attending small

classes of 3. The findings from these two studies suggest that the benefits of using an Adaptive AI Learning system cannot be solely attributed to a decrease in class size, as students utilising the system outperformed the students who received both whole-class and small-group instruction.

1.2 Current situation on Intelligent Tutoring Systems

The development of Intelligent Tutoring Systems (ITS) is nothing but new; it has been a sought-after and rapidly growing field for research, especially now with the emergence of generative AI. According to Lambert and Stevens (2023), machine learning algorithms are used in ITS systems to analyse data on students' learning styles, strengths, weaknesses, and progress. This data analysis enables customised material and feedback to address students' knowledge gaps and improve understanding. These systems may constantly adjust to students' development, keeping them challenged and interested. Personalised learning systems can detect and forecast areas where students fail, allowing instructors to provide crucial interventions and help when necessary.

In this chapter, an attempt will be made to outline the current state of affairs regarding some prominent ITSs.

1.2.1 Early and legacy ITS systems

a. Metatutor

MetaTutor is an intelligent tutoring system (ITS) that uses hypermedia to teach complex STEM subjects, such as the human circulatory system. It was developed by Azevedo et al. (2012) and a team of interdisciplinary researchers from the University of Memphis, McGill University, Illinois Institute of Technology, North Carolina State University, and the University of Central Florida over the past ten years.

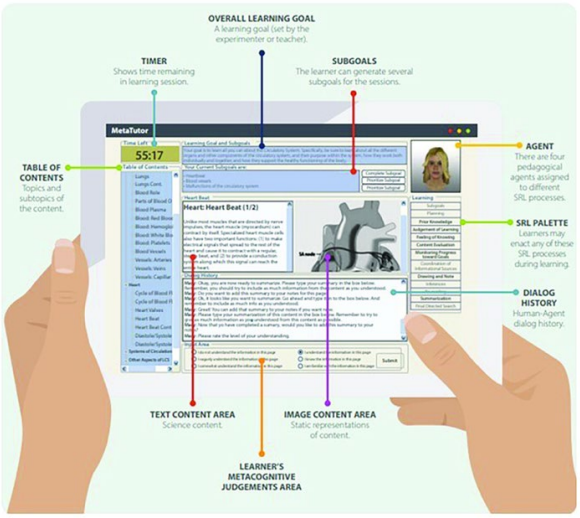
MetaTutor follows the theoretical, conceptual, and methodological principles of *Self-Regulated Learning* (SRL - i.e. the process in which learners actively engage and maintain their thoughts, emotions, and actions in a structured manner to achieve learning objectives), using advanced learning technologies. These principles are supported by several researchers such as Azevedo et al. (2017), Winne and Hadwin (1998), Zimmerman and Schunk (2011), and Greene (2017). Metatutor is truly *multimodal*, in the sense that it employs sensors to assess in real-time Cognitive, Affective, Motivational and *Metacognitive* (CAMM) 'signatures' and use these data to devise optimal learning strategies.

Metacognition refers to the ability of an individual to reflect, and understand their current state of knowledge, providing a basis for self-regulation, supervision, and evaluation, and it is a predictor of academic success (Conway-Smith & West, 2022; Teng & Yang, 2023; Wang et al., 2015; Winne, 2014).

Metatutor collects data consisting of facial expressions denoting emotion, physiological sensors, eye trackers, log files, and screen recordings of student-system interactions. The system also includes several *avatars* i.e. virtual persons acting as *'Pedagogical Agents'* (PAs), with different personas, each one specialised in different cognitive strategies (see Image 1); learners can interact with the agents via text inputs employing Natural Language Processing (Azevedo et al., 2022).

During the process of learning via the Metatutor, the prospective learning outcome is divided into goals and subgoals, whereas learners can select the SRL process they wish to enact. Also, they are able to return metacognitive feedback (i.e. a self-reporting evaluation regarding the level of their learning) to the system (Azevedo et al., 2022; Azevedo et al., 2019).

Empirical studies suggest that the self-regulated approach offered by Metatutor has a positive impact on learners. Cloude et al. (2021) by analysing the data on performance outcomes, self-reported affect and temporal traces regarding user interaction collected by the system were able to identify a deactivation of self-reported negative emotions like hopelessness, boredom, and sadness, as users engaged in cognitive strategies such as summarising, note taking and inference making. As for the role of PAs, Dever et al. (2022) points out that students who receive guidance from agents in an Intelligent Tutoring System (ITS) tend to employ a range of SRL strategies, across various tactics leading to greater improvements, in their learning; while students who tackle concepts within an ITS without assistance utilise the SRL technique more frequently but achieve lower learning gains. These findings suggest that employing a variety of different SRL methods instead of repeating a single method consistently, leads to better learning outcomes.



Agent Profile	Agent Name	Agent Role and Corresponding Cognitive and Metacognitive SRL Strategies	Example Cognitive and Metacognitive SRL Prompts
	Gavin the Guide	Orients and directs learners on navigating the learning environment and its various features.	"You will have 90 minutes to learn about the circulatory system. Your overall learning goal for this session is to learn all you can about the circulatory system."
	Mary the Monitor	Content Evaluation (CE)	"Do you think this page and image are related to your current subgoal of learning about Purposes of the circulatory system?"
		Feeling of Knowing (FOK)	"How strongly do you feel you already know the content from this page?"
		Judgment of Learning (JOL)	"How well do you think you comprehend the material on this page?"
		Monitoring Progress towards Goals (MPTG)	"Do you think you have finished the subgoal of learning about Blood components?"
	Pam the Planner	Planning (PLAN)	"Ok, let's set up some subgoals then."
		Prior Knowledge Activation (PKA)	"Alright, now we are going to start learning about Purposes of the circulatory system. Before we begin, tell me everything you already know about this topic..."
	Sam the Strategizer	Taking Notes (TN)	"It sometimes helps learners to draw what they are reading about. You can draw using the notepad on the desk now."
		Summarizing (SUMM)	"Please type your summary in the box below. Remember, you should try to include as much information from the content as you understood."

Image 1: Interface and the four Pedagogical Agents of Metatutor (Azevedo et al., 2022)

It should be mentioned though, that simpler and less multimodal ITS systems may have a good impact on learning. A recent study by Abdelshiheed et al. (2023) measured the impact of metacognition and motivation on transferring knowledge across two ITSs, the one teaching deductive logic using a default backward and an alternative backward chaining strategy, and the other teaching probability, using only backward chaining. More specifically, the study involved Discrete Mathematics students, who after being trained on solving problems in the two ITSs, were divided into three groups, based on their metacognitive skills measured by the ability to switch to the most efficient strategy each time. Their knowledge was classified either as *conditional*, either *procedural* or *rote*, with the students possessing the conditional knowledge to be able to select the appropriate strategy each time; the ones with the procedural knowledge having some understanding of the problem-solving strategies but without the comprehensive knowledge of when to switch between them; lastly, rote students used only the default strategy. The first group would switch early on whenever strategy was necessary to solve a problem, the second group would switch later when they couldn't easily solve it, whereas rote learners would stick to the default strategy. Based on their online traces while using the ITSs, they were divided into two *motivational* levels, high and low. The preliminary study showed that only the highly motivated students who possessed conditional knowledge were able to transfer their knowledge across the two ITSs. Subsequently, two consecutive experiments involving the groups with procedural and rote knowledge were conducted. The researchers altered the logic ITS to provide metacognitive support in the form of 'nudges', i.e. of written prompts in the first experiment,

and a combination of 'nudges' and worked examples in the second. The results indicated that possessing high levels of motivation had a significant benefit on the ability to transfer knowledge only for the most knowledgeable group. However, when the other two received support, they were able to catch up with their knowledgeable counterparts, with rote learners receiving the most benefit when they received nudges and examples, regardless of their levels of motivation. These results suggest that a combination of pre-existing knowledge, cognitive flexibility and motivation may promote learning.

b. nStudy

The software platform nStudy (Winne & Hadwin, 2013) was an early ITS developed by Philip Winne and fellow researchers at Simon Fraser University. Like Metatutor which came later, it was aimed at the modelling and support of metacognitive self-regulatory skills for university students (Winne, 2013). Key system components provided adaptable support including prompting planning tactics before studying content, tracking study strategies used during learning, evaluating note quality after readings, and reflective assessments aimed at rendering covert thought processes into more visible skills amenable to guidance (Bannert, 2009). A key component of nStudy was the pioneering notion of time-stamped trace-based feedback – by comprehensively capturing detailed temporal indicators of all reader's behaviours across actions like highlighting, annotations, access sequences, and response latencies logged within the digital course materials, the system could leverage this tracing data to provide feedback describing productive patterns as well as diagnose procedural lapses (Winne, 2014).

While nStudy was easily accessible as it was a plugin for browsers available in plugin marketplaces, and the initial testing periods were promising, as they were able to show gains on some targeted outcomes like note-taking quality and course satisfaction compared to more conventional digital learning platforms; despite this, the adoption of nStudy ultimately remained limited, and its development was abandoned. Nevertheless, nStudy conceptually exemplified a trailblazing ITS trait, the ability to leverage data mining approaches to reveal metacognitive and self-regulation processes through analyses of trace indicators; until today, it continues influencing current work on learner modelling and optimal tactics supporting the development of self-directed learning skills crucial for lifelong academic efficacy (Azevedo & Gašević, 2019).
There is limited empirical evidence to support nStudy efficiency in promoting improved learning outcomes for its users; however, its data collection approach remains influential.

Building on its principles, Marzouk (2016) proposed a theoretical framework incorporating nStudy data into the Interactive > Constructive > Active > Passive (ICAP) framework proposed by Chi and Wylie (2014). ICAP distinguishes many forms of observable behaviour, as well as the underlying knowledge-change processes and learning outcomes for each learning mode. Learning results are expected to vary from minimum, to shallow, and up to the deepest (for examples, see Table 1). According to the ICAP hypothesis, interactive behaviours improve learning the most, followed by constructive behaviours, active behaviours, and finally passive behaviours.

	PASSIVE Receiving	ACTIVE Manipulating	CONSTRUCTIVE Generating	INTERACTIVE Dialoguing
LISTENING to a lecture	Listening without doing anything else but oriented toward instruction	Repeating or rehearsing; Copying solution steps; Taking verbatim notes	Reflecting out-loud; Drawing concept maps; Asking questions	Defending and arguing a position in dyads or small group
READING a text	Reading entire text passages silently/aloud without doing anything else	Underlining or highlighting; Summarizing by copy-and-delete	Self-explaining; Integrating across texts; Taking notes in one's own words	Asking and answering comprehension questions with a partner
WATCHING a video	Watching the video without doing anything else	Manipulating the tape by pausing, playing, fast-forward, rewind	Explaining concepts in the video; Comparing and contrasting to prior knowledge or other materials	Debating with a peer about the justifications; Discussing similarities & differences

Table 1: Examples of activities classified by the ICAP framework (Chi & Wylie, 2014)

ICAP has seen empirical validation. For example, Menekse et al. (2013) tested its presumptions with students aged 19-20, by comparing four learning activity conditions utilising introductory materials of science topics in an experimental design with random assignment. As predicted by the ICAP theory, student gain scores grew gradually from passive to active to constructive and interactive settings. The results of the pairwise

comparisons indicated significant differences between all engagement modes. The mean difference was greatest between the interactive and passive modes (MD = 26.29) followed by the differences between interactive and active (MD = 17.76), constructive and passive (MD = 16.46), interactive and constructive (MD = 9.83), active and passive (MD = 8.53), and finally between constructive and active (MD = 7.93). The author provides evidence to conclude that there are graded differences in learning outcomes aligned to the level of engagement, with students demonstrating greater gains the more interactively they participate, at a rate of 8-10% across each mode.

With the data provided by nStudy, it could be assumed that an ITS system with access to a database of tagged domains could devise well-informed personalised learning approaches ranging from interactive to constructive for learners and enhance their knowledge and experience (Marzouk, 2016). It could also provide an element of social/constructive collaboration between peers, as active, interactive, and constructive modes often involve peer participation.

1.2.2 Factors that may hinder the adoption of Intelligent Tutoring Systems

Metatutor and nStudy were both successful but in different ways. The former was –and still is– able to produce measurable results for students. The second exemplified the usability of big data analytics to tap into the efficacy of ITSs. However, neither of them was broadly used by the public; Metatutor is still active, yet not available to the wider public, and nStudy which was available through browser plugin stores has been discontinued. Only a few ITSs have recently found their way into widespread public use, with most of them being confined to research settings, posing a question: *Why do most of the ITSs remain in the lab?*

There are several potential factors contributing to the limited adoption of certain intelligent tutoring systems, such as nStudy or MetaTutor, despite the presence of empirical evidence demonstrating their efficacy in promoting learning outcomes. The development of ITSs necessitates significant initial investments in AI, programming, content development, and related areas (Vanlehn, 2011). Researchers frequently face challenges in securing long-term funding to support the ongoing maintenance, updates, and widespread dissemination of ITSs beyond their initial trial phases (Baker, 2016). Scaling complex AI-based systems across schools and curricula poses significant challenges, like the knowledge of the mechanisms that underly learning; the need for adapting to teaching new skills adapted to today's needs; transforming interaction data into meaningful supporting strategies and more, highlighted

by Woolf (2009). Certain educators perceive ITS as a potential threat to their role and are hesitant to embrace its implementation (Heffernan & Heffernan, 2014). ITSs are frequently developed independently from educational curricula, standards, and assessments (Dede et al., 2005 2005). According to Walkington and Bernacki (2020), when systems are not designed with user-friendliness in mind for students and teachers, their utilisation becomes restricted. There is concern among scholars, such as Roll and Wylie (2016), regarding the potential negative impact of adaptive systems, such as ITS, on the development of independent learning skills. Although certain studies have demonstrated positive learning outcomes, the lack of extensive efficacy research has resulted in schools being cautious about adopting (Steenbergen-Hu & Cooper, 2013). According to a report issued by the U.S. Department of Education and its Office of Educational Technology, investors and publishers have often expressed uncertainty regarding the profitability and sustainability of the subject in question, showing some reluctance to fund such projects (Bienkowski et al., 2014 2012). Addressing these obstacles has the potential to facilitate the integration of intelligent systems from laboratory settings to practical educational environments.

Another serious constraint that restricted the widespread usage of ITSs is, according to Kurni et al. (2023), the need for *accountability* regarding the educational outcomes (e.g. validity and precision of assessments) confines their usage to formative rather than summative assessments. Formative assessments are a pivotal objective within the realm of intelligent tutoring systems (ITSs). According to the researchers, the fundamental objective of assessments lies in enhancing instructional methodologies and fostering educational advancements for students. One of the primary objectives of formative assessment is to enhance student learning by positioning the learner as an essential, resourceful, and self-reflective participant within the educational community. In formative assessment-based classrooms, it is common to find a pedagogical approach that emphasises individualised instruction and real-world practice. Kurni et al. (2023) suggest that the absence of standardisation and the comparatively less rigorous approach to formative assessment in contrast to summative assessment represent a fundamental flaw within this paradigm. Consequently, there is a potential for a decline in the quality of both the evaluation resources and the subsequent results. Hence, Intelligent Tutoring Systems (ITSs) are predominantly employed in formative assessments to enhance student learning outcomes and optimise instructional practices.

Also, the level of knowledge of teachers regarding technology and AI may be a contributing factor to the adoption of such tools. Based on the Technological, Pedagogical, and Content

Knowledge framework (TPACK), Celik (2023) developed a scale to assess the relationship between teachers' technological knowledge of AI and their attitudes towards AI-based instruction; The results indicated that technological knowledge positively predicts ethical assessments enabling teachers to evaluate decisions made by AI. However, the results of this study point out that such knowledge may be not enough for the integration of AI-based educational systems, whereas technological knowledge becomes meaningful when it is combined with pedagogical knowledge, forming a new construct, technological pedagogical knowledge.

One of the benefits associated with the integration of AI in the field of education is its capacity to effectively discern the areas of weakness and strength among pupils, subsequently offering prompt and relevant feedback. Furthermore, AI has the potential to alleviate the burden on educators by automating administrative duties and assessment processes. Nevertheless, it is crucial to bear in mind that artificial intelligence AI is incapable of supplanting the significant role of the educator in the educational journey; rather, it should function as an adjunctive and reinforcing instrument for teachers (Firdaus et al., 2023, p. 11).

While intelligent tutoring systems can provide adaptive instruction and data-driven insights, some teachers have expressed valid concerns about implementing these tools. Studies have found that teachers worry about how accurately the systems model struggling learners (Ostrow et al., 2015; San Pedro et al., 2015). Even with training, not all reluctance is alleviated (Baker, 2016). Some are hesitant about AI access to student data and how transparently it is used (Bodily et al., 2018). More recently, expert reports show teachers fear being replaced by AI completely, perceiving it as an intruder in key instructional roles and as hindering their mentoring relationships with students (García-Peñalvo & Reimann, 2016). A qualitative analysis by Kim and Kim (2022) investigated how teachers perceive and approach the use of an AI-based system called AISS (AI-supported writing) in STEM education. Teachers identified the advantages of using the AISS. Most expressed views regarding its ability to serve as a knowledgeable model providing well-structured and high-quality examples of scientific writing. The AISS was also praised for its capacity to offer feedback and suggestions tailored to each student's specific lexical, grammatical, and logical needs in their writing. It played a role by assisting students in enhancing their arguments effectively by incorporating evidence improving their creative thinking and problem-solving skills and facilitating the organization of ideas within their written work. Moreover, teachers made comparisons between the AISS and other educational technologies they had utilised

previously. They noted that the AISS appeared advanced in terms of customization and emulating human tutor support. Many believed that the AISS would be most suitable for high school students who already possess a solid foundation in scientific writing. These students could benefit significantly from self-directed learning approaches with a focus on evidence-based claims. However, there were concerns raised about relying on AI-powered systems like the AISS. Some teachers expressed worries that it might diminish the role of teachers, to assistants or supervisors by taking over crucial instructional tasks.

Overcoming perceptions of AI as a threat rather than a teaching aid relies on evidence affirming teachers' continued centrality in AI-enhanced classrooms. Addressing privacy and transparency concerns also appears crucial to acceptance. With judicious implementation aligning with teacher needs, intelligent systems have the potential to *augment* rather than *automate* quality education.

1.3 Generative AI & ITSs

The use of Large Language Models (LLMs) such as OpenAI's GPT, seem to have the potential to enhance the efficacy of the learning process by facilitating personalised learning experiences that cater to the unique skills of both learners and educators. AI-based learning tools, such as Khanmigo and Duolingo Max, may provide personalised assistance, promote critical thinking, and recommend relevant resources, therefore aiding students in acquiring the information and skills desired by employers (Luckin & Holmes, 2016).

Chen et al. (2023), by using qualitative research methods, describe interesting ways students interacted with OpenAI's ChatGPT; students were able to design tactics that would allow them to take advantage of ChatGPT's benefits while simultaneously developing techniques to deal with its flaws. The students used additional tools to verify the accuracy of the outputs generated by ChatGPT, committed time to the process of prompt engineering (i.e. providing the chatbot with the appropriate prompt to guide the expected outcomes) and sought guidance from the instructor whenever they encountered difficulties. It has been pointed out in previous work that while tools such as ChatGPT may deliver solutions in a short amount of time, they should never be used as a substitute for student's ability to think critically and solve problems (Dwivedi et al., 2023), with some students considering that the outputs of ChatGPT needed to be able to survive verification before they could be brought to their discussion area (Chen et al., 2023). This was because they positioned the use of ChatGPT within the context of knowledge development. According to Knight and Littleton (2017),

knowledge building offers a relevant setting in which the claims of knowledge made by the students themselves or knowledgeable others have important repercussions in terms of social and communicative interactions. By extending this notion, it may be possible that while the students' utilisations of ChatGPT are not necessarily innovative, the manner in which they combined ChatGPT to serve various cognitive tasks in a social and discursive setting gives a novel perspective on the digital potential that GAI has in the field of education (Tlili et al., 2023). The usage of ChatGPT by students was a *personalised experience* to them and tightly related to a common objective of acquiring knowledge with their peers (Chen et al., 2023), while at the same time, seeking personalised assistance from the chatbot (Dwivedi et al., 2023).

1.3.1 Tutoring systems with Large Language Model applications

As the rise of GAI and the abrupt spreading of its implementations are quite recent, research is currently at the state of observing GAI as a social phenomenon, with limited published examining the implementation of chatbots into ITSs.

Lambert and Stevens (2023) describe how ChatGPT and other AI-powered apps provide unique personalised learning (PL) methodologies. Teachers are entrusted with meeting an ever-increasing range of pupils, which is a demanding challenge. PL has been demonstrated to improve student outcomes, particularly techniques that use adaptive technology, and the longer kids are exposed to personalised learning, the higher their performance development (Pane et al., 2016).

Khan World School shows how one organisation is making use of the chance to utilise ChatGPT for personalised learning. Sal Khan (2023) the founder and CEO of Khan Academy, revealed *Khanmigo*, (see Image 2) a chatbot based on GPT-4 LLM, at a recent TED talk. Khanmigo is currently only available to users residing in the US.

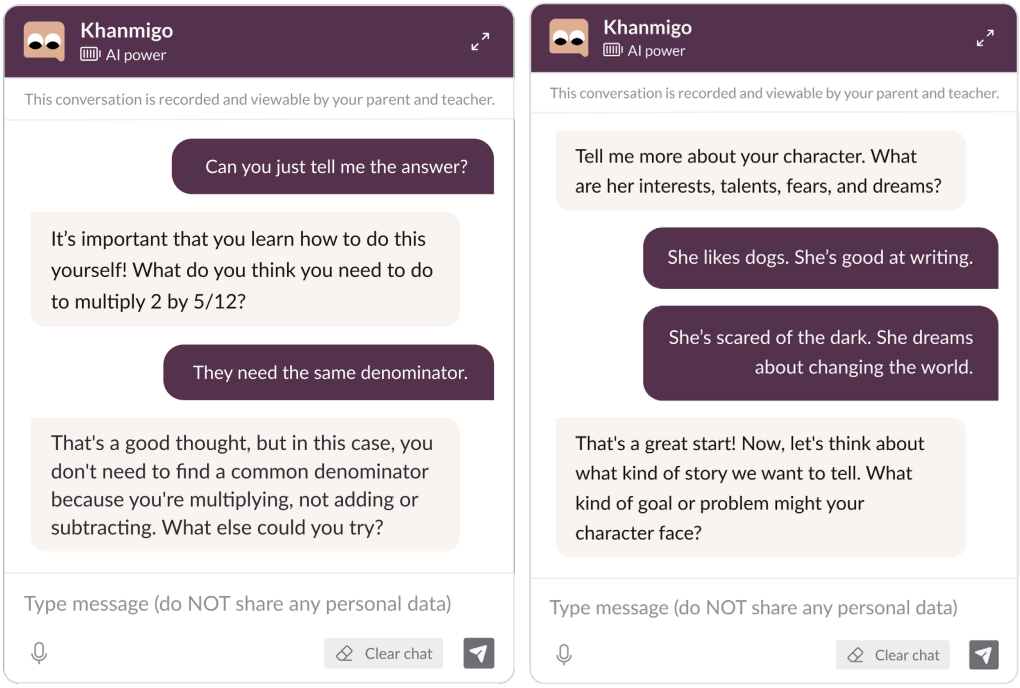


Image 2: The Khanmigo chatbox in action

Khan showed examples of how students utilised Khanmigo to aid them in math, computer programming, biology, and language arts. In the first case, a student was given a standard maths problem using distributive properties. When the student made a mistake, the tutor caught it, asked the student to clarify their thinking, and then reminded the student how to utilise the distributive principle. Essentially, the instructor did not provide a perfect answer but rather analysed what the student did incorrectly and offered advice for how to improve their work. In another case, a student inquired as to why they needed to learn about cell size. Khanmigo reacted by asking the pupil what they were interested in, to which the student replied, *"A professional athlete"* (Khan, 2023). The chatbot then emphasised how knowing cell size is important for understanding nutrition and how your body operates. In this case, the tutor recognised the context of the video the student was seeing and was able to ask the student a question to extract personal meaning. Another student utilised Khanmigo to acquire a character's viewpoint on a scene from F. Scott Fitzgerald's book *"The Great Gatsby"*. Using ChatGPT's capacity to take on a character, students may find literature, history, art, or politics to be much more intriguing and engaging (Khan, 2023). The Khanmigo tutor may also aid students with their writing by assisting them in developing an outline and providing comments on a first draught, much like a live writing coach. When providing feedback, the tutor might underline sections of a piece, inform the student that it does not support their assertion, and ask why. Again, the tutor serves as a real-time writing coach, scaffolding and

accelerating students' reading and writing skills. Furthermore, the tutor may be switched from student to teacher mode to provide instructors advice on how to teach a certain subject. Additionally, Khan (2023) said that they were actively looking at methods to employ generative AI to improve reading comprehension: A student may view a video and then, at times, click on a question, and the AI will quiz the student on the subject. The AI may highlight passages and ask inquiries such as, *"Why did the author choose that word? What was their goal? Does it support their case?"*.

The education technology firm Duolingo has leveraged artificial intelligence within its popular language learning application from the outset, accumulating over 500 million registered users of which 37 million actively utilise the app monthly spanning 95 courses across 38 languages. Recent partnerships with OpenAI starting in 2021 initialised the integration of advanced generative AI capabilities into the platform.

Specifically, Duolingo first adopted GPT-3 in 2021, enabling more eloquent explanations and examples responding to users' submitted answers. This year, the company incorporated GPT-4 features into a premium subscription extension called Duolingo Max (see Image 3). Two main offerings in Duolingo Max currently rely on generative AI - "Explain My Answer" and "Roleplay" (Kshetri, 2023).

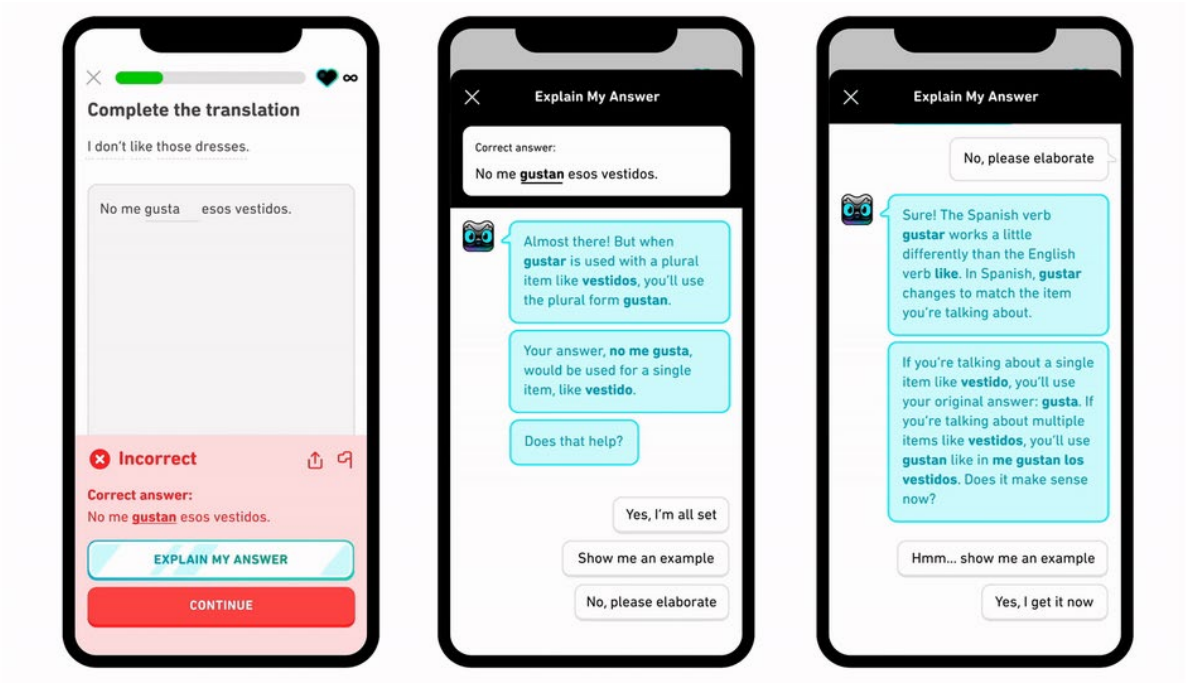


Image 3: Duolingo Max

Lambert and Stevens (2023) report several EduTech companies that are now integrating machine learning and AI into their learning platforms to accommodate personalised learning. Quizlet has been integrating AI technology to produce multiple-choice questions and example sentences for vocabulary acquisition with *Q-Chat*, which is based on ChatGPT technology, allowing a personalised learning coach to customise content especially to match the requirements of learners; *Knewton*, by Wiley company, a platform primarily targeted for college students; and finally *Century*, which provides a suite of personalised products for elementary schools through adults.

Century has drawn the attention of the author of this study as it is the only widely open system addressed specifically to elementary school children. The interaction with the platform is made by a rather minimalistic graphical user interface (see Image 4). The platform offers material for Math, English language, and also Verbal and Non-Verbal reasoning in the form of exercises chunked into *'Nuggets'.*

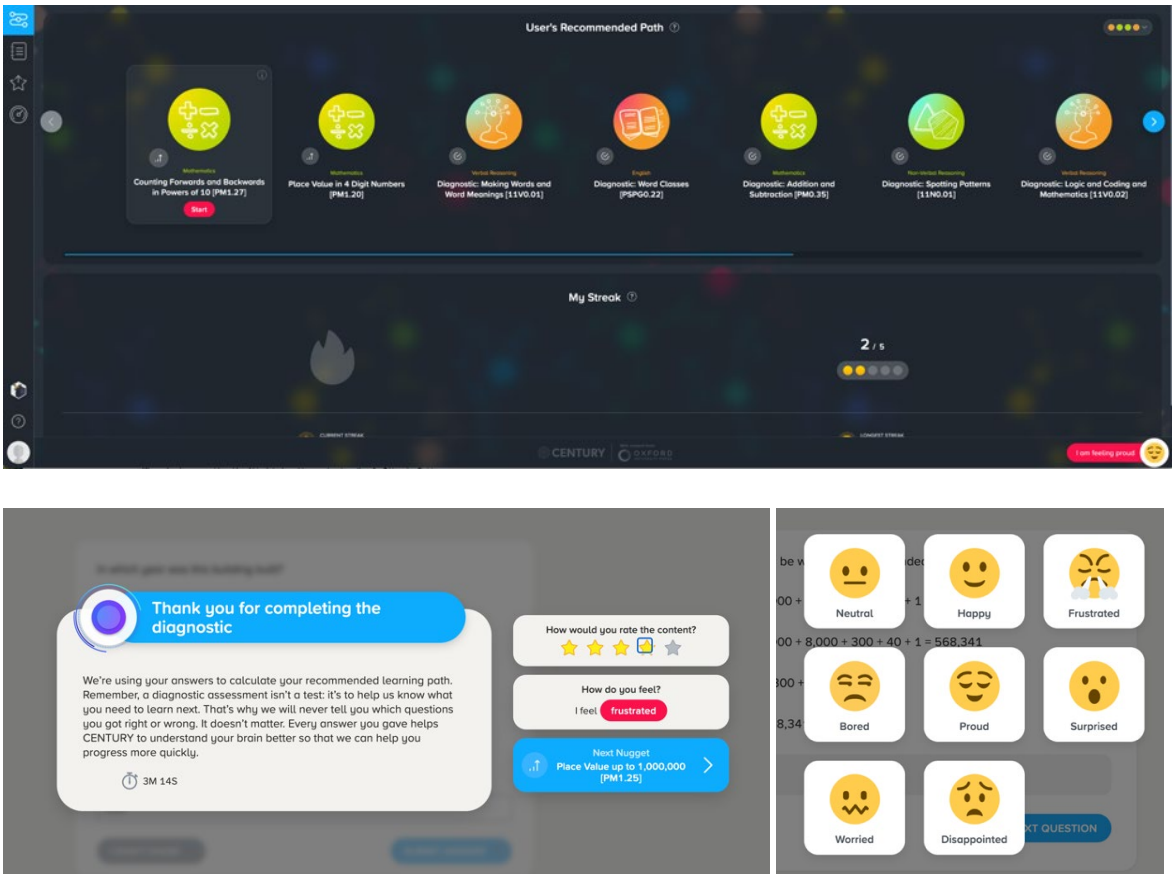


Image 4: On top, Century's dashboard, with offered courses.
In the middle, is the "thank you" screen after completion of the diagnostic test.

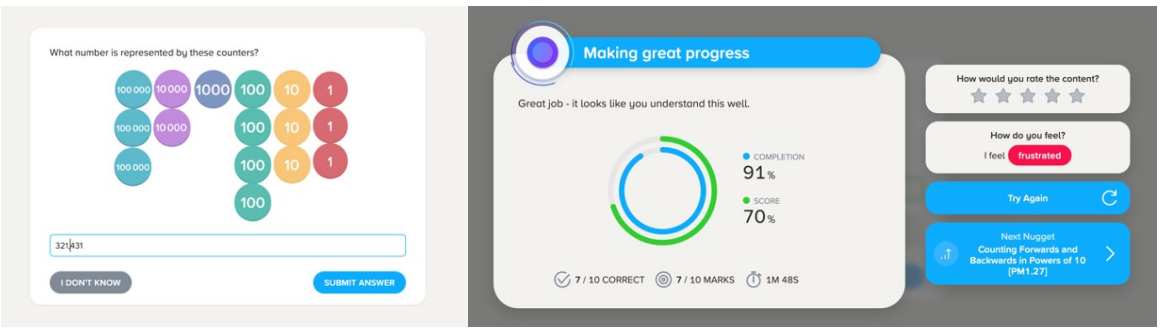


Image 4a: Left, the exercise interface. Right, the feedback interface.

A *'diagnostic'* test is initially offered to newcomers, which attempts to tap into their current knowledge level. Then, it adapts material based on the answers obtained. At the beginning of a nugget, a video or a slideshow is available to explain the concept and the requirements of the exercises that will follow. Also, at any time, users may provide feedback regarding their emotional states (see Image 4a). This input is also requested after completing a 'nugget'.

As the company has not yet disclosed any white papers regarding which psychological and pedagogical approaches have been endorsed, and judging solely from interaction with the platform, it could be assumed that Century has integrated some adaptive learning principles into its approach —hence the diagnostic test—, and that it does take into account the emotional states of the user during the learning procedure. However, it should be noted that Century does not include a chatbot, and confines itself to a structured, predefined set of exercises which are adapted and delivered to learners according to their performance.

Chapter 2
Theoretical framework for the proposed ITS

It is widely accepted that creating an educationally impactful ITS relies heavily on integrating insights from both artificial intelligence and cognitive psychological research (Woolf, 2009). When the learning processes and conversational capabilities modelled in AI tutors strongly align with well-established theories of human cognition, communication, motivation, and pedagogy, AI substantially enhances the naturality, adaptability, interactivity, and overall benefit of such systems for students (Graesser, 2016). By closely emulating scientifically validated frameworks of how students cognitively process subject material, store and retrieve knowledge, interact with instructors, and regulate their own learning, AI tutors have greater capability to provide more highly personalised and nuanced tutoring tailored to each individual learner's abilities, prior knowledge, interests, and learning styles (Shute & Zapata-Rivera, 2012, p. 9). Mimicking natural pedagogical discourse patterns additionally enables more fluid, intuitive, and human-like dialogue to occur between the intelligent tutor agent and the human learner during instructional interactions (Graesser, 2016). Such natural interactivity is key for learners to feel comfortable conversing with and confiding in the intelligent tutoring system.

Furthermore, the very process of implementing artificial intelligence systems founded on established learning science theories permits rigorous hypothesis testing and empirical refinement of those psychological theories, thereby mutually advancing both AI computing research and cognitive science research (Ritter et al., 2018; Ritter et al., 2007, p. 250). The interdisciplinary synthesis uncovers new insights into both effective personalised pedagogy and how to replicate elements of human learning in artificial agents. In summary, embedding psychological knowledge within the design of AI tutor systems confers critical advantages in terms of psychological authenticity, individually tailored adaptive instruction, natural and intuitive conversational interactivity, interpretability for students, and productive interdisciplinary contributions that cross-fertilise both fields. To maximally support and

enhance student learning, 21st century intelligent tutoring solutions must integrate and build upon the substantial prior research base developed in cognitive science illuminating the complexities and nuances of how human beings naturally communicate, teach, and learn.

According to Azevedo et al. (2019) computer assisted learning systems (CALSs) such as Intelligent Tutoring Systems (ITSs) have the potential to greatly impact education by helping students monitor and regulate cognitive, emotional, metacognitive, motivational and social processes. While this poses a challenge to existing frameworks and theories of self-regulation CALSs can now be equipped with technologies like natural language processing, machine vision, haptic devices, and physiological sensors. Azevedo et al. (2019) claim that These advancements have the potential to improve learners' self-regulation by enhancing CALSs used for research or teaching purposes. For example, integrating natural language processing, into CALSs can allow real-time collection of verbalizations that can be coded to assess judgments (e.g. determining the relevance of instructional material) and provide *scaffolding* when learners face specific metacognitive challenges. Despite the aspiration to incorporate cutting-edge technology into CALSs researchers should base their decisions on models, frameworks and theories of self-regulated learning that are supported by theoretical knowledge and empirical evidence.

2.1 Zone of Proximal Development

Lev Vygotsky, a renowned developmental psychologist, was the first person to publicly present the Zone of Proximal Development (ZPD). This concept captures an important discovery that serves as a cornerstone of social constructivist learning theories and sociocultural approaches to cognitive development (Scott & Palincsar, 2013). Vygotsky described a difference between the actual developmental level at which a child can independently carry out a task and problem-solve without assistance, which is referred to as the 'Zone of Actual Development', and what the child is capable of accomplishing when they are provided with guidance, support, or collaborate with more skilled partners, which is referred to as the 'Zone of Proximal Development' (Vygotsky, 1978). This theory captures this disparity at its core. The goal of *scaffolding* in teaching is to find a 'sweet spot' for optimum learning, which is represented by this assumed gap in accomplishment.

As part of the process of operationalising the key dimensions of ZPD, it is necessary to determine the lower threshold of unassisted ability towards a target outcome in conjunction with the upper ceiling of accomplishment manifested under a scaffolded partnership. This is

done to target instruction accordingly, although it is not an easy task, especially for complex, integrated pieces of knowledge and skills (Poehner & Infante, 2017). Zone of Proximal Development is a concept that outlines the region in which the cultivation of social support may promote the expansion of mastery beyond the existing level of competence via the use of graded encouragement and collaborative discussion. This process of assisted performance pulling instrumental competencies continually upward is central to sociocultural advancement (Shayer, 2003). As children participate in shared experiential learning with parents, carers, teachers, or mentors who provide adapted assistance, they gradually internalise understandings and strategies that were initially grasped interpersonally to propel skills forward at a faster rate. According to Miller (2011), Vygotsky proposed that community engagement and connections, which enable individuals to share partial knowledge while simultaneously using partners or cultural resources to fill in the gaps in their knowledge, remained significant drivers in the process of externalising and subsequently integrating socially co-constructed insights.

It is important to note that the Zone of Proximal Development is a fluid construct; it changes dynamically between domains and learners as their familiarity with the environment changes. Previous achievements constantly recreate the basis upon which guided experiences increase understanding (Shayer, 2003). Representing current independent capabilities is critical for curricula that are calibrated to student entry readiness (Miller, 2011). As a result, Vygotsky highlighted the significance of *continuous evaluation,* which reveals both rising capacities without assistance as well as increasing boundaries of potential with helped participation. This is an essential component of adaptive education, according to Scott and Palincsar (2013). According to Smagorinsky (2013) the ZPD is a concept that refers *to individualised pathways* that demarcate the areas in which adequately demanding collaborative interactions enhance new, emerging skills and promote motivated growth. Aligning curricula, mentoring interactions, and scaffolding for upcoming skills within each learner's fluctuating Zone remains Vygotsky's breakthrough, yielding remarkable results when implemented judiciously (Poehner & Infante, 2017; Shayer, 2003), even though it may be difficult to exactly pinpoint the Zone's boundaries, particularly across multidimensional knowledge.

The pioneering Zone of Proximal Development (ZPD) theory developed by Vygotsky, which highlights the difference between actual individual developmental levels and possible performance given social direction and participatory learning, continues to be very significant for current education and psychology (Shayer, 2003). The first thing that it does is

emphasise the need to precisely evaluate the students' existing independent competency in order to determine their readiness for growth. Additionally, it highlights the requirement of further customising proximal targets and scaffolding in order to fulfil current demands, which is a precursor of adaptive customisation (Tzuriel, 2000). In addition, Vygotsky's conceptualisation inherently acknowledges learning as being socially mediated and distributed across interdependent connections, rather than being an entirely individual endeavour. This recognition foreshadows more recent social constructivist learner-centred principles that aim to maximise engagement (Miller, 2011; Scott & Palincsar, 2013). Additionally, it recognises the huge diversity and domain variations in development that are necessary for the creation of flexible supports that are sensitive to the changing profiles of people, which is in line with the promise of current technology for personalisation. Vygotsky's emphasis on the cultivation of interpersonal mentoring relationships and graduated participation as a means of advancing understanding continues to provide direction for enhancing student development through the support of peers, even in the face of constrained resources or rigid curricula; as a result, the Zone of Proximal Development paradigm, when taken as a whole, makes a fundamental contribution to several long-lasting insights into adaptive, socially assisted education that calibrates collaborative activities to continuously expand developing skills (Shabani, 2016). Despite the nine decades since Lev Vygotsky conceived the ZPD his model of education and learning is still influential due to its relevance and practicality; ZPD produces a form of education that focuses on developing a student's thinking and personality rather than remembering and copying knowledge. Students go from their current level to their potential level with assistance and feedback When a skill is mastered, it becomes part of their skill and competencies, enhancing their *autonomy* (Billings & Walqui, 2018; Margolis, 2020; Vygotsky, 1978).

2.1.1 Zone of Proximal Development in Intelligent Tutoring Systems

The ZPD intrinsically provides a framework that fits the very concept of an adaptive ITS, as it reflects the ability of a tutor to identify the current knowledge level of a student and push forward by providing a *scaffolding* into new areas of knowledge to be conquered. Thus, the ZPD theoretical framework has been employed in numerous computerised adaptive educational systems, as it is intuitive in implementation; a system may use data from the user’s performance and choices and adapt difficulty and context to their learning pace, level, and personal preference; utility-based agents consider the environment, educational goals,

and a performance metric to decide how a goal might be accomplished then recommend educational material to the learner (Ferguson et al., 2022).

As long as a system follows this procedure for learning, then it may be just a matter of choosing the most appropriate and efficient algorithmic representation to be implemented in an ITS. Vainas et al. (2019) describe a challenge that most online learning platforms are facing by offering a predetermined sequence of tasks that students must complete in order: some students who are having difficulty grasping concepts may need more time and effort to do so before moving on to more advanced material, whereas other pupils, may get bored and uninterested, as they feel 'underserved' kept in a sluggish pace of learning. Consequently, some students will be under-challenged by information given in a non-adaptive way, while others will be over-challenged. Chounta et al. (2017) proposed a computational methodology to model the Zone of Proximal Development (ZPD) based on the predicted probability of correctness. With the help of a natural-language tutoring system, students go through high school physics problems the model attempts to predict the learner's ZPD by assessing several factors, such as the level of difficulty and their prior knowledge. They defined a "Grey Area" which is a region of uncertainty when the model is not able to tell whether it is within the student's ability to provide an accurate answer and assumes that whenever the system is found in this impasse, then the student is probably within their ZPD.

Vainas et al. (2019) introduced an adaptive learning engine called *'E-gostky'* that may determine the next exercise's difficulty level based on the student's performance on the previous one. As an example, it can bypass some tasks for a student who has shown competence in the subject matter, while omitting more challenging than usual tasks ("bonus exercises") for those who are still finding it difficult. Thus, E-gostky strives to ensure that the subsequent activity keeps children within their ZPD. To accommodate this purpose, Vainas et al. (2019) employed Dynamic Assessment (DA), another important idea in Vygotsky's theory, to this end. Lantolf and Poehner (2010) explain that in contrast to Static Assessments, DAs focus on the test taker's capacity to learn new material as they go, not on a predefined norm, which often suffers from bias and validity. In DA settings, a skill is first assessed, then taught, and then retested. This method provides the opportunity for the individual to acquire the information or skill that is being evaluated.

2.2 Self-Determination Theory

An influential theory of motivation in recent years, Self-Determination Theory (SDT) has found use in many domains, including medicine, business, and academia (Deci & Ryan, 2008). According to the theory every person has basic psychological needs, namely *Autonomy, Competence,* and *Relatedness*, that may be shaped by their environment, other people, and societal norms. Ryan and Deci (2017) offer the following definitions:

Autonomy is described as one's self-efficacy, bolstered by experiences of interest and worth, which entail a sense of initiative and control over one's activities. *Competence* is a feeling of mastery and accomplishment, that is fostered in well-organised environments using optimum challenges, constructive criticism, and growth opportunities. *Relatedness* describes a feeling of support and connection to others. These three needs are nurtured when one's actions or hindered by the level of *intrinsicality* of volition. Simply put, the more one acts based on their intrinsic motives, for the sake of their own satisfaction, the more these needs grow. The more one acts due to external pressure the more these needs are starved. To a great extent, well-being and happiness are outcomes of the degree to which autonomy, competence and relatedness are satisfied. Left unsatisfied, they lead to resentment, depression and other forms of psychopathology (Ryan & Deci, 2017).

Intrinsic and *extrinsic* motivation are the two most common forms in students. Situations that cater to the three fundamental psychological needs—happiness, fulfilment, and enjoyment—are conducive to the development of intrinsic motivation. *Extrinsic* motivation is not a homogenous construct. It may be either *controlled* (i.e. enforced by another person, or authority sometimes under threat of some punitive consequence), *introjected* (i.e. imposed by feelings of guilt or shame), *identified* (i.e. external but consciously valued e.g. to perform some action which one does not like, nevertheless acknowledges as beneficial) and finally *integrated* (i.e. assimilated into one's values, e.g. recycling). In Academic settings introjection frequently develops into a form of self-regulated regulation, taking the form of *"ego-involvement"* linked to success and playing a substantial role in an individual's sense of self-worth. In some cases, though, motivation may be extrinsic, in accordance with one's values and desires, leading to a form of healthy *internalisation.* In both cases, they are considered autonomous and, hence are beneficial. Contrastingly, less autonomous, and more controlled forms of motivation lead to a more fragile approach towards academic accomplishment (Bailey & Phillips, 2016; Ryan & Deci, 2000; Taylor et al., 2014). Autonomy has been observed as a key component of academic success, with significant outcomes for

children, adolescents, and adults, with autonomously driven students flourishing in a variety of educational contexts, especially when tutors encourage their autonomy. Furthermore, children may benefit from the cultivation of their autonomy not only in education but also in developmental terms (Reeve, 2002). Autonomy, along with Relatedness and Competence is one of the three pillars of the Self-Determination Theory (Deci & Ryan, 2008; Ryan & Deci, 2000), a macro-theory that supports that motivation leads to better short- and long-term outcomes regarding the achievement of goals and general well-being. Intrinsic motivation promotes perseverance, greater performance, and well-being (Reeve et al., 2008). It promotes improved information retention, fewer dropouts, and academic achievement in the domain of education (Ricard & Pelletier, 2016; Wentzel, 1998). It should be emphasised, nevertheless, that SDT implies a continuum rather than a bipole, as illustrated in Figure 1:

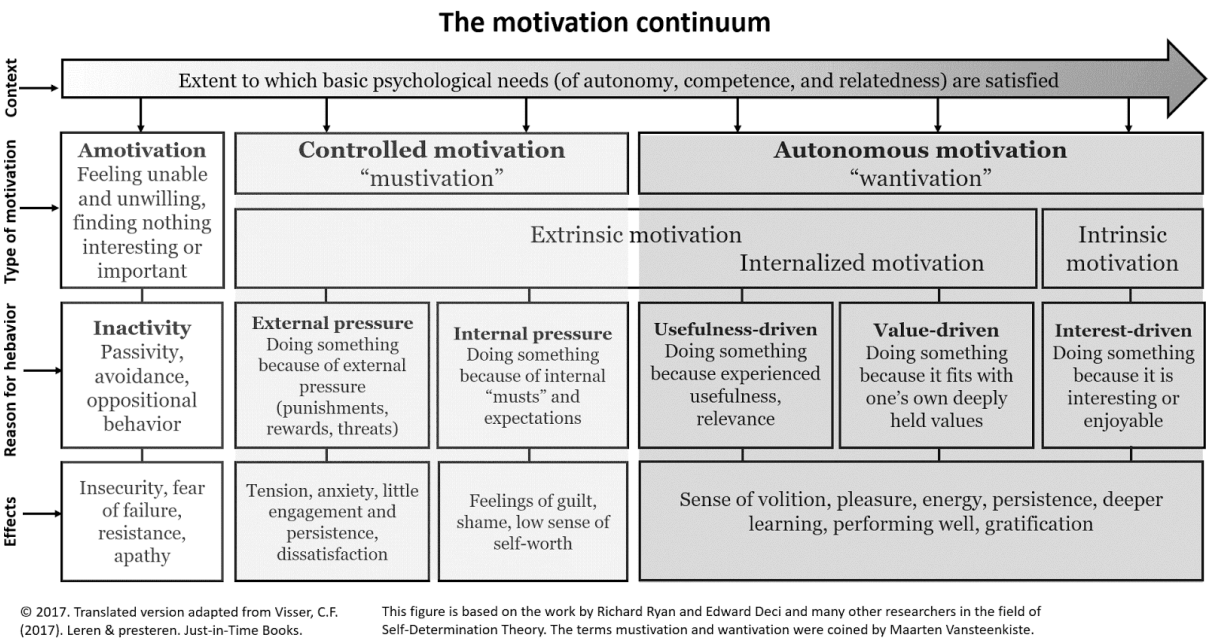


Figure 1: The motivation continuum (Visser, 2017).

2.2.1 Self-Determination Theory in ITS

The degree to which a participant puts effort into a cognitive task influences observable behaviour. The driving force behind this effort allocation is commonly referred to as motivation, and it is a significant impediment to correctly inferring individual characteristics from observations. E.g. a participant who performs poorly in a memory task may be unmotivated to complete the task rather than having limited Working Memory capacity; yet

motivation, despite its importance, is rarely modelled or accounted for in cognitive models (Yang & Stocco, 2023); while the utility of SDT is recognised in a variety of domains, a neglectable number of published works report SDT principles implemented in ITSs. There are, however some design principles for human-computer interaction (HCI), that are founded on satisfaction of the Autonomy, Competence and Relatedness; Peters et al. (2018) introduced the METUX (Motivation, Engagement & Thriving in User Experience) model, a framework for designing digital experiences that prioritise user motivation, engagement, and well-being. It is based on SDT, which posits that people are more likely to engage in sustained behaviour change when their psychological needs for autonomy, competence, and relatedness are met. According to Peters et al. (2018), METUX focuses on four *'spheres',* which are user experience, its outcomes, the psychological mediators for need satisfaction and finally the design of an HCI interface:

User Experience (UX) refers to the overall experience of a user while interacting with a product or service, e.g. as a website, app, or device. It comprises all aspects of the user's interaction, including the interface, functionality, usability, and aesthetics. *Outcomes* refer to the results or effects of a user's experience with a product or service. In the context of the METUX model, positive outcomes include motivation, engagement, and thriving, while negative outcomes might include frustration, disengagement, or even harm. *Mediators* are the psychological needs that are believed to mediate the relationship between user experience and outcomes. In the context of the METUX model, the mediators are the same as in SD: autonomy, competence, and relatedness, which are thought to be universal and fundamental to human motivation and well-being (Ryan & Deci, 2017). *Design* refers to the intentional creation of products, services, or experiences that meet the needs of users. In the context of the METUX model, design is focused on creating digital experiences that support users' psychological needs for autonomy, competence, and relatedness, thereby promoting positive outcomes such as motivation, engagement, and thriving. The model provides a framework for designers to evaluate and improve their designs, based on these criteria.

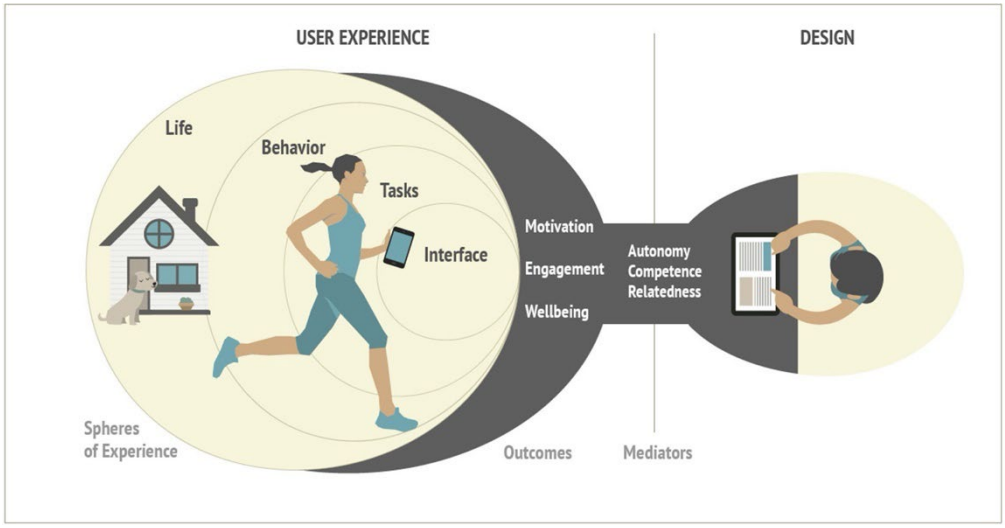


Figure 2 - METUX design: Autonomy, Competence and Relatedness act as mediators of positive user experience outcomes such as Motivation, Engagement and Wellbeing, thus constituting measurable parameters for HCI design aligned with SDT (Peters et al., 2018).

These spheres are divided into several sub-elements as described by Peters et al. (2018):

Adoption focuses on the decision-making experience between becoming aware of a new technology and acquiring it. The prominent psychological needs related to this element are autonomy and competence. Autonomy refers to the extent to which technology adoption is autonomously motivated, meaning that the user feels a sense of choice and volition in the decision to adopt the technology while competence refers to the extent to which the user expects to be competent at using the technology *Interface,* which focuses on the controls, navigation, information display, and aesthetics of the technology, components that may promote or hinder the overall need satisfaction of a user. *Task* focuses on the specific behaviours or activities that the technology is designed to support, and to which extent they promote their psychological needs; depending t the user's ability to perform a task and feel competent, to perform it on their own and feel autonomous and to associate with others who are also engaging in the task, promoting relatedness. *Behaviour* refers to the overarching activity that a task is intended to support. The difference between this element and the tasks element is important because some technology might support need-satisfying interaction at the interface and task levels, but may still not necessarily impact need-satisfaction regarding the behaviour it's designed to support (Burnell et al., 2023; Peters et al., 2018). For this purpose, any well-established and relevant to what it is to be measured metric should be employed.

Life refers to the link between technology and overall well-being. The SDT literature indicates that psychological need satisfaction increases mental and physical health. However, momentary need satisfaction relating to the use of a technology may not be sufficient to affect measurable improvements to individual flourishing, which can be assessed using numerous relevant instruments, e.g. the BPNS, or Basic Psychological Need Satisfaction scale (Burnell et al., 2023, Chen, 2015 #934). Finally, *Society* is the largest in scope and is the only one to step beyond the user experience. Societal well-being may be affected by using a technology both directly and indirectly. Within this element, ethical issues regarding the impact of an economic and environmental nature may become relevant.

Despite Peters et al. (2018) offering a solid, measurable framework of SDT-driven directives of METUX for efficient HCI, up to date no ITSs are reported to have adopted them, so its efficiency has not been sufficiently tested. However, in a recent study existing social network platforms like TikTok, Facebook and the well-known educational platforms Blackboard and Moodle have been used to evaluate some of the metrics proposed, i.e. Technology-based Experience of Need Satisfaction (TENS) -life, -behaviour, -task and -interface variants, with acceptable psychometric properties, with an $r > .75$ and CFIs between .90 - .93, except for TENS-life at .87. These metrics were reported after some modifications by Burnell et al. (2023). However, Chiu (2021) and Chiu and Chai (2020) report that several SDT principles have been incorporated into some small-scale ITS implementations of their own, but they do not report having used a solid methodology such as the one proposed in the METUX framework, which in terms of ITS design remains inadequately tested.

2.3 Executive functions

Executive Functions (EF) refer to a set of higher cognitive processes to engage, direct, or coordinate other lower processes, often in the service of goals (Miyake et al., 2000). However, it has also been described more and more reductively in recent years as a collection of distinct but connected component processes engaged in goal-directed cognition and behaviour; Working Memory, task switching, and Inhibition Control of prepotent thoughts and reactions are frequently listed as some of the most important executive functions (Best & Miller, 2010; Diamond, 2013). Executive functioning deficiencies may lead to difficulties in setting and maintaining objectives as well as the capacity to block out distractions may suffer. EFs are crucial components for successful Self-Regulation. There are several frameworks for investigating SR in childhood and adolescence. EF is sometimes referred to as Effortful Control (EC) and are two similar –and, to a significant extent,

overlapping– that have gained interest in several sectors of child development research (Rueda et al., 2005; Zhou et al., 2012). Some key elements of the EF are described, and their role in learning.

c. Working memory

To give a realistic model of primary memory, Baddeley and Hitch (1974) introduced the *Working Memory* model of human memory. Instead of viewing main memory as a single, cohesive unit, Working Memory divides it into distinct, yet interconnected parts, namely the *visuospatial sketchpad*, *the phonological loop,* and the *central executive,* which serves as a coordinator. This model was further improved by Baddeley and his colleagues by including a fourth component, the *episodic buffer,* and has now come to represent the prevailing theory in the study of Working Memory.

Three primary elements comprised Baddeley & Hitch's initial model. The phonological loop and the visuospatial sketchpad are the two major systems, to process linguistic and visual information, respectively. They are coordinated by the central executive, which codes new information, while it updates and/or replaces old information, combines data from several sources into coherent episodes, switches between tasks or retrieval techniques, blocks or suppresses dominant or automatic reactions and directs selective attention. Twenty-five years later, a fourth component was added, the episodic buffer (Baddeley, 2000). This component consists of a system with a limited storage capacity that offers short-term storage of data kept in a multimodal code that can combine data from the subsidiary systems and long-term memory into a single episodic representation. Presumably, the primary method of retrieval from the buffer is conscious awareness. The main way the updated model varies from the previous one is by placing more emphasis on the information integration procedures than on the separation of the subsystems. By doing so, it offers a better framework for addressing the more challenging facets of Working Memory executive control. According to Cowan (2014) planning, understanding, thinking, and problem-solving are all facilitated by Working Memory with a critical role in learning.

d. Response inhibition

Response inhibition, which ranks third in importance to Working Memory (WM) and attentional shifting, is essential to executive functioning (Miyake et al., 2000). Response inhibition is the capacity to stop a current response or to postpone the start of a response that is insufficient given the demands of the environment. Early arithmetic and reading

abilities in children as well as social functioning are all positively correlated with a functional response inhibition capacity. The *Stop Signal* task is one of the most widely used research paradigms for examining response inhibition (Zhao et al., 2018). This task is used in cognitive psychology research to assess response inhibition and impulse control abilities and disabilities in conditions like Attention Deficit and hyperactivity Disorder or ADHD for short (Schachar et al., 2007; Verbruggen & Logan, 2008) and the function of cognitive abilities involving the fronto-basal ganglia neural circuitry implicated in impulse regulation and higher order motor control and the inferior frontal cortex, which activate to implement 'braking' mechanisms suppressing actions (Aron et al., 2014; Zandbelt & Vink, 2010).

The representation of a Stop Signal task (Logan et al., 1984; Matzke et al., 2018; Schachar et al., 2007), which was used in this study involves a sequence of stimuli which are presented to a participant, to which they are instructed to respond when a stimulus is offered unless a stop signal is presented, in which case they need to suspend their reaction (e.g. by pressing a key when a green light is presented while pressing nothing if a red *stop* signal is presented shortly thereafter). In such a task, the possibility of initiating a response on a stop trial serves as the measure of Inhibition Control. Failure to withhold the response manifested by reacting after the stop signal is presented means that the response is erroneously produced and that inhibition has failed (Littman & Takács, 2017).

Inhibition tasks tap into key cognitive control capabilities essential for the broader self-regulatory faculty described as effortful control (Rothbart & Rueda, 2005; Rueda et al., 2005). Specifically, being able to withhold attention toward distracting stimuli in favour of focusing on goal-relevant information relies critically on inhibiting the processing of competing perceptual inputs (Diamond, 2013). Likewise, effortful governance over behaviour depends on the ability to stop reactive impulses or habitual responses when necessary to align actions with internalised standards or future objectives (Nigg, 2017). Even the act of regulating emotional expressions draws on Inhibition Control to override initial temptations to respond reflexively to provocation or cravings (Carlson & Wang, 2007). From a developmental point of view, expanding inhibitory capacity facilitates compliance with caregiver demands and social expectations during childhood by curbing rash reactivity (Riggs et al., 2010). In essence, effortful control reflects the integration of executive inhibitory mechanisms and motivational drives to fluidly modulate cognition, behaviour and emotion (Luna et al., 2010). Given the fact that inhibition provides the fundamental 'braking' mechanism for stabilizing these processes in line with intentions and situational norms,

quantitative markers of inhibitory efficiency derived from performance on paradigms like the stop-signal reaction time task offer direct biomarkers for effortful control competence.

2.4 Executive Functions and Learning

As being intrinsically motivated may not always be achievable, students must cope with negative emotions and distracting factors which are associated with *extrinsic* factors that often drive goal-oriented behaviours. As such, *Executive Functions* (EFs) become equally important. EFs describe the internal or transactional mechanisms that allow a person to guide goal-directed actions through time and across changing circumstances during which negative emotions such as frustration, boredom and loss of interest may emerge (Karoly, 1993). According to the Dictionary of the American Psychological Association (2020), EFs refer to a set of higher-level cognitive functions which include planning, decision-making, problem-solving, action sequencing, task assignment and organization, persistent goal pursuit, repression of conflicting impulses, flexibility in goal selection, and goal-conflict resolution (Meltzer, 2018). These frequently require the use of language, discretion, idea development, abstraction, and logic. They are commonly linked to the prefrontal cortex and neuronal networks that incorporate the prefrontal cortex (Stuss, 2011). Human cognition and performance depend on executive function, which individuals utilise to exert control over their thoughts and actions, particularly when attempting to do something that conflicts with inherent habits, inclinations, and desires (Rothbart & Rueda, 2005; Rueda et al., 2005; Simonds et al., 2007).

Executive functions (EFs) are essential self-regulatory skills that regulate basic or domain-specific cognitive processes (e.g., language, attention, sensory input, motor output) for goal-oriented problem-solving and behaviour. EFs are the controlling, supervisory functions of self-regulatory abilities that organise and guide cognitive activity, emotional reaction, and overt behaviour. EFs include several higher cognitive processes that govern behaviour, emotion, and cognition (Stern et al., 2016). Executive functions are needed to guide attention intentionally for the achievement of goals. These top-down brain functions develop early and underlie interweaved and interacting EFs including reasoning, emotional self-regulation, abstract thinking, problem-solving, planning-programming, and organisation essential to enable goal-directed behaviour and to adapt to novel situations and challenges (Malloy-Diniz et al., 2017), as well as social relationships and everyday tasks.

According to Miyake et al. (2000) the key EFs are Working Memory, Inhibition Control (IC) and cognitive flexibility; three substantial, different yet related brain functions. WM is the ability to store, update, and cognitively manipulate information. WM is essential for goal-directed planning, behaviour, and attention. Inhibition is the capacity to think before acting, also necessary for focus, as it controls urges and cuts out distractions, while cognitive flexibility is the ability to rearrange knowledge and redirect thoughts and actions to adapt to novel situations.

EFs appear as early as in the first year of infancy, rapidly developing between three and six years and continue to develop through adolescence and adulthood (Reznick et al., 2010) and it is crucial to assess them as early as possible (Huizinga et al., 2006) as during this extended period of time the prefrontal cortex matures undergoing o long developmental process (Diamond, 2013). This maturation process in this cortex, which is rapid in the early years and decelerates to a steady pace during adolescence, leads to an improvement in schoolchildren's EF. The first EF to develop is inhibition, followed by planning and problem-solving, while WM develops during 7-9 years of age. Thus, middle-aged children tend to have better WM and attention, quicker cognitive processes, and are becoming more adept at developing more complex cognitive strategies. From 3 to 5 years old, information processing develops fast, with 9- and 10-year-olds showing considerable advances (Margari et al., 2016), impacting, amongst others, academic accomplishment and emotion management (Sofologi et al., 2023). As executive functions remain critically tied to disciplined conduct, academic achievement, and psychosocial health throughout childhood and adolescence (Diamond, 2013), early appraisal enables promptly addressing deficiencies through cognitive or behavioural interventions, ideally minimizing adverse downstream impacts.

Working memory (WM) is described as the capacity to store and modify information at the same time WM is essential for making sense of everything that happens throughout time since it needs to remember what occurred before and link it to what occurs later. Thus, it is an ability required to understand written or spoken language, whether a phrase, a paragraph, or something lengthier. Similarly, practising math involves WM since it aids in problem-solving by converting instructions into action plans or integrating new knowledge into activities (Baddeley, 2014; Baddeley & Hitch, 1974). Some research has focused on particular parts of Working Memory important in early academic abilities, such as verbal and visual-spatial short-term memory (Bull et al., 2008). The first is responsible for storing and processing verbal information, while the second is responsible for storing and processing visual and spatial structures (Baddeley, 2014).

A large body of research shows that several executive function (EF) cognitive components, such as Working Memory but also Inhibition Control which are measured in this study, are associated with early academic abilities in preschool children (Montoya et al., 2019). Experiments have also shown that a deficiency in WM during the preschool years leads to issues with reading comprehension in elementary school, with children with WM deficits struggling with phonological awareness and letter-word recognition tasks (Alloway et al., 2004; Nation et al., 1999). According to Montoya et al. (2019) WM is an accurate predictor of nearly all early academic abilities, while all numeracy skills and receptive vocabulary were predicted by visual-spatial short-term memory. As such, it can be assumed that executive functions, including Inhibition Control and Working Memory, are key proponents of learning.

2.5 Learning in traditional classrooms

Traditional classroom learning has been the prominent paradigm since the conception of modern school. Learning in classrooms may promote learning as per the Vygotskian principles, as classrooms provide structured content where the ZPDs of different students interact, providing the social interaction necessary for learning (Chaiklin, 2003). Moreover, according to the SDT principles, classrooms may promote children's well-being, given that they enhance autonomy, competence, and relatedness, both in social and educational terms (Ryan & Deci, 2017). However, they may also have some negative issues associated with the latter. Some strong arguments regard the very concept of traditional classrooms as inherently problematic. Below, some of these arguments are presented in brief:

2.5.1 The "one-size-fits-all" approach.

The standardised, *"one-size-fits-all"* approach in traditional classrooms is a frequently cited reason why they often fail to meet all students' needs; standard classrooms are naturally inclined towards this approach, presenting the same content and pace for all students regardless of individual learning needs, abilities, interests, etc. leading to a failure to adapt to students' variability (Tomlinson, 2015). Teachers naturally struggle to provide differentiated instruction for the variety of learners in each class, especially when bound to standardised curricula and class schedules; more often than not, a tutor can't facilitate learning in a personalised matter, given the 1 teacher per 25 children ratio which is the mainstream in a classroom of an urban elementary school with things becoming even get harder for teachers who are teaching multiple classrooms in high schools (Dixon et al., 2014). Consequently, students are more likely to disengage when content lacks relevance to personal contexts,

cultural perspectives, prior experience, or real-world value. Different subgroups see curriculum priorities through distinct lenses not accounted for (Parsons et al., 2017).

The assumption that providing the same instruction, content, assignments, pace, and assessments for everyone in a class is fundamentally flawed for several reasons:

Firstly, students arrive with dramatic variability in their background knowledge, academic readiness, prerequisite skill levels, interests, and optimum learning modalities. A student struggling with reading comprehension, or a student good in math will not be well served receiving the same instruction as classmates. The uniform approach disregards the vast diversity in capacities and needs of learners (Santangelo & Tomlinson, 2012),

Secondly, enforcing rigid standards prevents customising pedagogy, materials, feedback or rate of progress due to learner differences. Students who quickly grasp concepts get held back when they require acceleration, while struggling students, due to a lack of personalised support, are unable to keep up and are left behind (Tomlinson, 2017).

In summary, *"one-size-fits-all"* education superficially treats surface similarities while neglecting deeply rooted developmental, experiential and motivational individual differences (Miyake & Friedman, 2012) requiring a more personalised alternative approach attuned to each learner's evolving Zone of Proximal development (Vygotsky, 1978).

2.5.2 Passive learning

The prevalent lecture-based classroom model has come under scrutiny for perpetuating passive learning that fails to adequately challenge or engage students, with teachers lecturing and students listening; this does not actively engage students or develop critical thinking skills (Chi & Wylie, 2014). Passivity may also reinforce surface learning study tactics focused on memorization rather than a deep foundational understanding (Baeten et al., 2010). Direct instruction lectures involve largely one-way transmission of information from teacher to students. While it may seem efficient for offering the content, this approach reduces opportunities for active participation, dialogue, or deeper cognitive engagement (Chi, 2009). Students end up playing a receptive role focused on auditory and note-taking tasks rather than interactive discussion, debate, collaborative discovery or problem-solving seen in active learning (Freeman et al., 2014). According to the same source, results differ dramatically when active learning is employed; it seems to be a factor that increases examination performance and decreases failure rates compared to traditional lecturing for STEM-related courses. Freeman et al. (2014) identified a performance increase of .47

standard deviations on student examination when students engaged in active learning practices, with students in traditional lecture-based classes being 1.5 times more likely to fail than students in classes with active learning. Moreover, active learning is effective across all class sizes, but the greatest effects are observed in small classes. The results support active learning as the preferred teaching practice in regular classrooms and raise questions about the continued use of traditional lecturing as a control in research studies.

While lectures have merits in specific applications, over-reliance in standard classrooms limits generative thought, restricted to lower-level cognitive processes. Promoting critical thinking and self-directed engagement implies balancing transmission models with increased learner empowerment (Kahl & Venette, 2010).

2.5.3 Standardised testing and learning

Classrooms often focus instruction around high-stakes, standardised tests. This encourages teaching *to the test* and obstructs a deeper, more enriching approach to instruction and learning that could lead students to develop deeper learning competencies (Au, 2011). With schools judged based on aggregate test performance, preparing students for state exams often becomes teaching's de facto –if not explicit– aim (Barnes et al., 2000). Consequently, the fixed scope of knowledge and skills assessed constrain classroom priorities and activities to focus intensively on drilling students for tested content rather than broader conceptual development or thinking skills (Firestone et al., 1998). Beyond test-taking proficiencies, learning how to craft evidence-based arguments, analyse diverse perspectives, solve ill-defined problems or self-monitor understanding tend to receive diminished emphasis when instructional time is narrowed to a *"teaching to the test"* approach (Popham, 2001).

While standardised assessment data offers accountability benefits, experts argue wider competency-based evaluation better captures the range of higher-order learning goals like critical thought, metacognition, collaborating productively, persistently overcoming obstacles, and communicating clearly (Rotherham & Willingham, 2010). Thus refocusing instruction, curriculum, and testing on developing such multilayered expertise could help shift classrooms from punitive *exam factories* toward *incubators* cultivating creative, self-directed lifelong learners (Sahlberg, 2010). More balanced assessment frameworks may likewise restore teaching's emphasis on igniting individual passion and potential rather than uniform test results.

2.5.4 Rigid, industrialised model

The regimented organization of traditional classrooms - with lectures delivered to rows of quiet students timed to bells - echoes factory production lines, oriented towards efficiency and control rather than meaningful engagement; reflecting an outdated industrial model rather than how people best learn (Robinson & Aronica, 2018; Tyack & Tobin, 2016) and rather than reflecting contemporary insights into effective learning (Labaree, 2011; Reich et al., 2020). This standardised industrial-era model engraves conformity, obedience and *rote knowledge* acquisition; it stifles creativity, active participation, passion for discovery and development of social-emotional skills (Jennings & Greenberg, 2009).

Current learning sciences recognise knowledge as socially constructed through guided participation in communities, not merely transmitted from expert to novice (Salomon, 1997). Students engage in higher-order critical thinking and problem-solving via collaborative dialogue and peer learning, which struggle to thrive within rigidly controlled hierarchies that fixate on metrics of content delivery over meaning-making (Dole, 2017).

Contrastingly, high-performing educational systems credit success to learner-centred policies fostering class discussion, project-based learning and passion-driven inquiry while balancing structure with autonomy support tailored to individuals and teams (Darling-Hammond et al., 2020).

2.5.5 Poor feedback

Classroom setting may also hinder effective and productive feedback, in numerous ways. Teachers, bound to specific curriculums, often overcumbered by the workload of delivering learning content to a large number of students end up struggling to follow up with their progress, and fail to deliver productive feedback to their students; they often resort to delayed, vague, or solely quantitative feedback (i.e. grades), which has been the mainstream in classrooms for many decades and has been measured to be less effective than personalised, verbal feedback, impeding both motivation and performance of learners (Lefevre & Cox, 2016; Tsirides, 2022). However, feedback plays a substantial role in learning. Formative feedback that directly addresses learners' just-completed work with targeted, corrective guidance in real-time consistently improves learning outcomes (Marwan et al., 2022).

These are just a few of the most prominent deficiencies that arise in classrooms, which most of the time end up promoting rote learning, memorisation, superficial learning, inadequate

cultivation of critical thinking, and poorly motivated students. Proponents of personalised learning argue new approaches like personalised instruction, project-based learning, adaptive technology, and student-driven learning address these inherent limitations. However, others maintain the classroom endures because it provides structure, socialization, and a shared community of learning when executed effectively. There are merits to both viewpoints in the ongoing debate over the effectiveness of the classroom model, however, it is widely accepted that the shortcomings of the traditional classroom models need to be tackled.

2.6 Evidence for superior results from 1-to-1 tutoring

Bloom (1984) identified the enhanced efficacy of 1-to-1 tutoring in his seminal paper known as “The 2-sigma problem”. Bloom describes the work of two doctoral students Anania (1982, 1983), and (Burke, 1980) who compared student learning in a conventional 30-person with one tutor whole class, used as a control, to two alternative methods: mastery learning and individual 1-to-1 tutoring. In mastery learning classrooms of around 30 students, formative tests were used for feedback and corrective work before parallel versions were given to check for mastery. In tutoring conditions, each student worked consistently with an individual tutor using the same formative test and corrective procedure approach. Through randomised experiments in grades 4, 5, and 8 in probability and cartography, lasting only 11 class periods over 3 weeks, they found substantial differences in final achievement between conditions. Tutored students averaged 2 standard deviations higher than conventional classes, with over 98% exceeding typical control students. Mastery learning averaged one SD higher, with 84% exceeding typical performance. The variation in scores also decreased dramatically in the enhanced conditions. Positive impacts were also found on student time on task and attitudes. Aptitude-achievement correlations dropped from +.60 in conventional classes to +.35 in mastery and +.25 for tutoring, showing prior measures became less predictive. The key conclusion was that the average tutored student performed at a level reached by only the top 2% of those in traditional instruction. The author argues that this demonstrates that most students can potentially achieve much higher levels of learning if optimal conditions like 1-to-1 tutoring are offered to them. Therefore, the "2 sigma problem" poses the challenge of finding feasible and affordable education models that can enable the level of learning gains seen from individualised tutoring, but on a larger scale.

2.6.1 Human tutors vs computer tutors

There are several attitudes regarding how well computer tutors perform compared to human counterparts, and which factors contribute to these performance discrepancies or similarities, which fluctuate across domains. Vanlehn (2011) examined a comprehensive set of prominent theories leading to the popular belief that a human tutor would most probably outperform a computer tutor when they both deliver the same content, leading to a key question for ITS developers: *"What are human tutors doing that computer tutors are not, and why are they more effective?".*

In order to customise their tutoring to suit each student's needs human tutors often try to understand the specific areas where a student lacks competence or has misunderstandings. However, it is important to note that this hypothesis has not been validated. Human tutors typically do not understand which knowledge components their students have yet to master. They are seldom aware of any misconceptions, false beliefs or faulty skills that the students may have (Chi et al., 2004; Putnam, 1987). They rarely ask questions that could help identify misconceptions held by individual students (McArthur et al., 1990; Putnam, 1987). They tend to modify their behaviours towards learners only when they identify errors or misconceptions, failing to do so when they identify mastery (Sleeman et al., 1989). Finally, staying for prolonged periods of time with the same students does not seem to improve their effectiveness, despite the intuitive assumption that it could offer them a better understanding of their student's strengths, weaknesses and preferences (Siler, 2004).

Another popular misconception is that human tutors are more successful than computer tutors because they can choose tasks that meet each student's needs. However, research indicates that human tutors typically follow a predetermined curriculum script, adjusting the pace of instruction based on their assessment of students' understanding of the material (Chi et al., 2008). Another hypothesis suggests that human tutors employ strategies, like Socratic irony, reciprocal teaching, and the inquiry method. However, studies have revealed that these strategies are rarely utilised in tutoring sessions conducted by humans suggesting that it is not these tutorial techniques alone that explain why human tutors outperform their computer counterparts (Ohlsson et al., 2007; Woolf et al., 2008).

Another theory is that human tutoring allows for mixed-initiative dialogues in which the student can ask questions or change the subject, an option that –supposedly– an ITS cannot offer. However, analyses of human tutorial dialogues have revealed that, while students take

the initiative more frequently than in classroom settings, the frequency remains low (Chi et al., 2001; Graesser et al., 1995).

Human instructors understand the subject better than machine tutors since they can explore related ideas and provide more detailed explanations for complex issues, that students may find counterintuitive and hard to grasp. These conversations, however, are uncommon when teaching cognitive skills (Merrill et al., 1995). When teaching less procedural subjects, human tutors frequently provide deeper explanations; yet, cutting off such explanations does not affect learning gains (Chi et al., 2001; Graesser et al., 1995). Human tutors may not always demonstrate their deeper and broader expertise during tutoring, and when they do, it does not appear to result in significantly bigger learning gains. As a result, while human tutors have a larger and deeper understanding, they may not always deliver the same level of explanation (Vanlehn, 2011).

The effectiveness of human tutoring might be linked to increased student motivation (Cordova & Lepper, 1996) although it is unclear how these strategies impact learning; the common perception is that praising is a motivational approach often associated with higher interest and better learning outcomes while computer-generated praise may have limited or negative effects (Vanlehn, 2011). However, the influence of tutors' praise on students is complex and could potentially hinder learning progress (Boyer et al., 2008; Henderlong & Lepper, 2002). Additionally, some tutors provide feedback for incorrect answers, which may impede learning but boost students' belief in their abilities (Lajoie et al., 1993).

Another theory is that human tutors assist students by monitoring and correcting their reasoning. If the student appears to be making progress, the tutor does not intervene; but, if the student becomes stuck or makes a mistake, the tutor can assist the student in resolving the lack of knowledge and getting back on track. Contrastingly students who use a computer may build up a long line of reasoning that leads to a false response, and then struggle to uncover the mistakes in their reasoning and fix their knowledge (Vanlehn, 2011). An immediate intervention provided by human coaching makes it much easier for pupils to identify and correct faults in their thinking (Merrill et al., 1992). Similarly, human tutors may provide "scaffolding" to their tutees, i.e. assist them during the reasoning process by asking mediating questions or 'hints' that help them think and provide explanations on their own, without giving them an answer straight away (Chi et al., 2001; Graesser et al., 1995; Merrill et al., 1995). According to (Vanlehn, 2011) the argument regarding immediate feedback and

scaffolding may provide a credible explanation for why human tutoring is more effective than computerised tutoring, unlike all the previous.

Wang et al. (2023) conducted a systematic literature review examining research on the applications and effectiveness of intelligent tutoring systems (ITS) in real educational contexts using social experimentation methods from 2011–2022. The authors conducted a comprehensive search across four databases, which yielded 40 relevant studies that met the inclusion criteria of using social experiments to evaluate ITS in real settings with sufficient sample sizes and intervention durations.

A social experiment approach allows causal claims about an intervention's effectiveness by implementing it in an authentic educational environment while controlling for confounding factors (Riecken et al., 1974, pp. 5-9). The reviewed studies spanned K–12, higher education, and adult learning contexts, with secondary and postsecondary education being the most common. ITSs were applied in diverse subjects, especially math, languages, and science. The majority of experiments (which accounts for over 60%) occurred in the United States, revealing a geographical imbalance in ITS research globally (Nye, 2015). The scope of ITS functionalities focused on tutoring, personalization, assessment, conversation, and games. Cognitive Tutors were the most widely studied system.

In terms of methodology, the reviewed experiments utilised mixes of quasi-experiments, RCTs, natural experiments, and longitudinal designs with sample sizes ranging from 100–3000+ students over 8 weeks to 5+ years. The most common benchmark was comparing ITS to "business-as-usual" instruction without ITS. Learning performance was the predominant outcome measure, with 90% of studies examining impacts on achievement.

Results were mixed, with 62.5% of experiments finding positive effects of ITSs on learning, 37% showing no differences, and 12.5% demonstrating negative effects relative to comparison conditions. Limited evidence existed for impacts on other outcome variables like help-seeking, engagement, and teachers' perceptions. The authors synthesise the challenges of social experiments with ITSs, including student attrition, individual differences, technological constraints, methodological complexities, maintaining fidelity, and teacher adaptation needs (Wang et al., 2023). Overall, the review reveals that ITSs can positively influence learning, but their effectiveness depends on contextual and implementation factors. The authors advocate for more rigorous, globally diverse ITS research that attends to student processes and uses social experiments to make causal claims.

Another systematic review by Al-aqbi (2019) examining 45 studies and a total of 12,105 students using ITSs vs controls who only attended traditional teaching approaches, showed a clear advantage for the use of the former, in a variety of fields, such as math, physics, physiology, computer science and more. According to this study, this edge may be attributed to the ability of ITSs to satisfy the requirements of students who may feel discouraged and unable to achieve their educational goals in conventional education systems, by offering a personalised learning environment in which students may create more learning approaches than standard teaching methods.

2.7 Education & AI: new tools for tackling the problems of classroom teaching

One-on-one tutoring, feedback, progress monitoring and cooperative learning may offer significant benefits over traditional classroom teaching, in terms of achievement (Dietrichson et al., 2017), however in many cases it may be available to people with higher socio-economic status (SES). As it is not feasible to generalise the one-to-one tutoring to all students, it is crucial to come up with interventions that offer some form of personalisation to more students, regardless of their SES, with AI-powered ITS being a means of *democratisation* of quality personalised education (Kucirkova & Leaton Gray, 2023), even when only limited funds are available (Muranga et al., 2023). Recent findings comparing learning outcomes with the use of ITSs vs learning in traditional settings are promisingly in favour of the former (Akyuz, 2020; Kulik & Fletcher, 2016; Wang et al., 2020).

With the rise of Generative AI (GAI), researchers are now exploring ways to use artificial intelligence to improve teaching and learning, by incorporating GAI chatbots into ITSs, thus taking advantage of *'the best of both worlds'*; the extremely natural, sophisticated conversational ability and original, real-time content generation, with the ITSs ability to model each student's knowledge and give personalised feedback and lessons based on constant assessments (Sharples, 2023). Work is ongoing into automatically generating questions for these AI tutors - methods based on understanding language structure look most useful so far (Ferster, 2022; Sharples, 2023).

These developments are quite recent and the lack of empirical evidence regarding the impact of GAI-enhanced ITSs does not allow any conclusive assumption; however, the abilities of chatbots and their fast-paced improvement are promising. By streamlining them, both generative chatbots and ITSs have the potential to improve education in the future.

However, there are still significant ethical and technical obstacles to overcome, such as ensuring that the systems interact organically and that learning is promoted fairly for every student (Chauncey & McKenna, 2023).

Chauncey and McKenna (2023) claim that qualitative evaluations show great promise for the usability of the now widespread ChatGPT 3.5, as expert humans rated the text outputs of the system in terms of linguistic quality and accuracy of information to be in accordance with the English Language & Arts (ELA) standard used in the schools of the state of New York.

Recently, GPT-4 which is the latest iteration of OpenAI's LLM, reported an above-human average performance on a variety of standardised academic measures (see Figure 3).

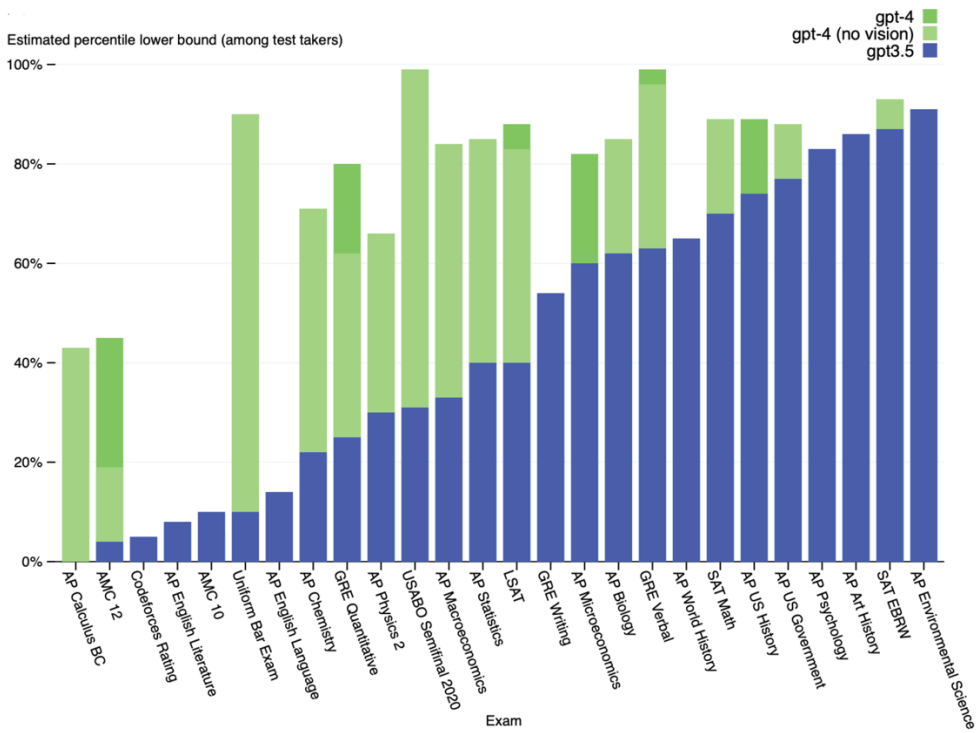


Figure 3: GPT performance has been reported as of human level on numerous standardised academic performance tests (OpenAi, 2023)

However, according to the same technical report "it still is not fully reliable (it "hallucinates" facts and makes reasoning errors). Great care should be taken when using language model outputs, particularly in high-stakes contexts" (OpenAi, 2023, p. 10)

To facilitate a functional, ethical and responsible incorporation of AI into the educational system, (Gibson et al., 2023) proposes a three-level learning model that synthesises and

integrates current learning theories. It uses artificial intelligence (AI) to improve educational methods. A causal learning process is included in the model, which describes how learning takes place at the micro, meso, and macro levels. It is influenced by developmental psychology, computational biology, instructional design, cognitive science, complexity theory, and sociocultural theory. The micro level is characterised as an individual's core and minimum process of learning, advancing through four levels with four connecting dynamics. The meso level focuses on teamwork and knowledge communities for collaborative learning. It reinterprets micro-level mechanisms from Piagetian and Kauffman concepts into terms from collaborative learning that are more readily recognised. The macro level takes into account cultural historical action as well as cultural change.
The results from these studies may suggest that developing AI systems to enhance teaching and learning, may require increased collaboration between educators and AI specialists. When human teaching is unavailable, technology may make it possible to create cognitive assistants for learning that would relieve tutors of laborious tasks so they can concentrate on coaching and mentoring students. It may also make it possible to give students learning opportunities options in more relaxed settings, outside of classrooms. AI may offer the potential to revolutionise the field of education. Yet, research is still in its infancy. It will take work to figure out how students learn best as we attempt to create artificial mentors and tutors.

The implementation of the three theories presented in this study in an ITS may be a computational challenge, and one may wonder why not using a less computationally sophisticated approach such as a Feedback Loop combined with a simple ZPD-based algorithm; feedback in natural language has been documented as having promising results on learning environments, even when teachers are not physically present when using advanced parsers of Natural Language Processing (NLP) techniques (Troussas et al., 2023). Self-Determination Theory (SDT) may be a framework that better explains the interaction and change in the content of the goal, motivation in the pursuit of that goal, and the self-regulatory processes through which the goal is pursued. It focuses on the degree to which behaviours are enacted with a sense of volition, as opposed to feeling controlled by external actions or internal compulsions. More specifically, Day et al. (2022, pp. 3-4) suggest that the SDT characteristic associated with higher efficacy in Self-Regulation interventions is competence, especially when encouragement and feedback are provided. Challenge, as a component of competence, also contributes to the strength of the effect. According to the same source, autonomy is also present in effective interventions, but its specific impact on

SR development is not well-established; however, there seems to be less evidence to support the SR effects of relatedness, indicating that more research is needed for successful intervention results. Nevertheless, it should be stressed that there have been some modelling attempts of other motivational theories into ACT-R cognitive architecture, such as the Expected Value of Control (EVC) theory (Yang & Stocco, 2023). The EVC theory of motivation and the SDT theory of motivation share some similarities and differences. Both theories acknowledge the significance of motivation, in influencing behaviour (Ünlü, 2023). However, the EVC theory primarily focuses on how expectancy, value and cost impact motivation (Knee & Browne, 2023), whereas SDT highlights the importance of autonomy, competence and relatedness, in fostering self-determined motivation (Mercader-Rubio et al., 2022). According to the EVC theory individuals feel motivated to participate in activities that they anticipate will lead to desired results and hold value. SDT, on the other hand, puts the need for satisfying the basic psychological needs for autonomy, competence, and relatedness are satisfied at the epicentre of motivation. Moreover, EVC theory emphasizes the role of cost-benefit equilibrium, suggesting that individuals balance the effort needed and the resources required to engage –or not– in an activity (Ünlü, 2023). SDT does not explicitly consider the concept of cost in its framework, differing in its focus and conceptualisation of the underlying factors that drive human behaviour. Freund and Lohbeck (2021), by taking advantage of the fact that SDT is considered a continuum (which, in turn, is reflected by the Relative Autonomy Index that can be expressed in a linear scale), propose an *Item Response Theory (IRT)* based model which assesses a learner's attitude congruency to the SDT four aspects. The basic concept of IRT models (or *latent trait models*) is that there is an underlying property, such as a skill, knowledge, or attitude, that is reflected in each answer to the items of a test or survey that can be modelled as a function of the responder's attitude and the item properties. IRT models are much more complex and computationally demanding, however, given the increase in computational power and data richness, it may be beneficial to incorporate them in contemporary Cognitive Architectures like ACT-R.

2.8 Aspiration & research strategy

This prospective study aspires to examine and propose a theoretical framework for the creation of Intelligent, AI-driven, adaptive tutoring systems (e.g. cognitive assistants, smart educational material aggregators, smart reminders) using the aforementioned constructs and theories (i.e. Zone of Proximal Development, Self-Determination Theory, and Self-Regulation) in order to motivate middle childhood students of age 8-12, by promoting

autonomy, enhancing competence and relatedness and by strengthening Self-Regulation, while keeping them into optimal learning states, as described by ZPD). No ITS using machine learning techniques supported by the theoretical frameworks proposed in this study could be identified as of the writing of these lines in the literature. As a result, by aiming to offer some fundamental design principles for a cognitive assistant that are supported by some widely accepted theories that have gained footing in motivation and learning, this work contributes to the field of ITS-facilitated learning.

Up to this date, only a handful of studies have attempted to incorporate some of the aforementioned theories into ITSs (see: Wei et al. (2021), Vainas et al. (2019), Pratama et al. (2016), Ballard and Butler (2011)), with even more limited literature for ITS for elementary school children. This study aspires to contribute to the topic of cognitive systems for education by approaching their design concept using an integrative, unitary model which encapsulates all three contemporary psychological theories and constructs and to propose some basic design principles of intelligent tutoring systems. However, this study will confine itself to the theoretical psychological framework and propose strategies and design guidelines, leaving aspirational space for computer engineers and programmers to design the software for ITS, using new or already available resources. This study will use a dual research strategy in an attempt to devise a theoretical outline for cognitive assistants for educational purposes by synthesizing well-known means of computerised or mobile-based paradigms that seem to perform well in helping 3rd to 6th-grade elementary students learn.

On the one hand, it will try to find convergence points in the literature on the described theories to propose a model that is computationally plausible, efficient, effective, programmer- and user-friendly. On the other hand, the study will attempt a hands-on approach by employing surveys and experimental designs to preliminary tap into similarities and collinearities in the proposed SDT and EF instruments and to the relevant population, i.e. middle childhood pupils of age 8-12 and their parents or teachers; by doing so, the researcher will attempt to pinpoint the strongest associations between items and factors and converge on a minimal set of variables and predictors which in turn may lead to an efficient computational design model. The decision to include self- and evaluator-reporting means along with performance-based cognitive task attempts to investigate the underlying cognitive processes that children employ while using CALSs on the principles employed by Azevedo et al. (2017) and Winne (2014) by employing self-reports assessing learners' self-perceptions of strategy usage, metacognition, motivation, and emotions.

2.9 An overview of the cognitive architecture for the proposed model

To create a theoretical model of an ITS based on a Cognitive Architecture that incorporates (a) the basic modules that the prominent Cognitive Architectures like ACT-R (Laird, 2019; Laird, 2022; Laird et al., 2017; Laird et al., 1987) and SOAR include, while (b) adding some dedicated modules that utilise the macro-theory of *Self-Determination Theory* and its application in the cultivation of *intrinsic motivation* (b) and some dedicated modules that utilise the theory of effortful control to enhance *self-regulation* and (c) use Vygotskyan principles of Zone *of Proximal Development* and *Scaffolding* to promote learning in an optimal matter for middle children (age 8-11).

The ITS proposed should no use wirings or sensors that monitor e.g. skin responses, as that should require specialised hardware, but may use sensors like cameras or microphones and also lingual inputs from the mouse and keyboard, and use widely known NLP techniques like sentiment analysis to tap into the affective, motivational and learning states of the users while during usage.

The objective is to develop a theoretical model for an ITS that draws upon current knowledge of human psychology, using a biologically inspired cognitive architecture, while attempting to maintain computational plausibility, by minimizing the model to an extent that impacts its potential real-life performance the least. Through the process of emulating the cognitive processes of the human mind, the cognitive architecture may be able to enhance the natural and efficient behaviour of a digital tutor. The integration of many cognitive science ideas serves as an initial foundation for the development of an assistant that exhibits enhanced learning and reasoning capabilities like those of a human being.

Cognitive architectures encompass theoretical frameworks that aim to replicate the cognitive processes and information-processing mechanisms observed in the human mind. Throughout the years, numerous designs have been developed, each grounded in distinct theories of cognition. However, the primary objective of these endeavours is to replicate human cognitive processes in a manner that can provide guidance to artificial intelligence and machine learning systems (Vernon, 2014). Currently, two prominent frameworks in the field of cognitive research are ACT-R and SOAR. The SOAR framework is predicated on the notion that the human brain employs elementary *if-then* production rules to determine appropriate responses to stimuli. Furthermore, it serves as a representation of Working Memory, wherein individuals engage in the active storage and manipulation of information

(Laird, 2019; Laird et al., 1987). On the other hand, the ACT-R framework decomposes cognition into distinct modules that collaborate synergistically, analogous to the functional segregation observed in various brain regions responsible for different cognitive processes (Anderson, 2007; Ritter et al., 2018). The architectural framework under discussion in this study is informed by both the SOAR and ACT-R models, however, in its organisation, it is most similar to the latter; this strategy is commonly referred to as a *hybrid* approach (Kotseruba & Tsotsos, 2020). Additionally, several aspects of the design were derived from established ideas of memory and attention; i.e. the Multi-store model of memory (Atkinson & Shiffrin, 1968), the model of Working Memory (Baddeley, 2000; Baddeley, 2014; Baddeley & Hitch, 1974), and Semantic and Episodic memory theory (Tulving, 1985; Tulving & Markowitsch, 1997). The notion that distinct stores exist for short-term and long-term memory is a prevalent concept in cognitive architecture design, as it allows for a clean-cut allocation of resources and categorization of data. Also, the model uses an attention module that is inspired by prevalent attention theories and models (Jonides, 1981; Posner, 1980; Treisman, 1969; Yantis et al., 2002). ACT-R architecture serves as a highly commendable basis for the development of intelligent tutoring systems (ITSs) due to its exceptional ability to meticulously examine cognitive processes at a fine-grained level (Anderson, 2007). This characteristic aligns with the meticulousness often observed in laboratory studies. Furthermore, ACT-R can seamlessly integrate these distinct cognitive components into a comprehensive model, thereby enabling the modelling of intricate and multidimensional cognitive phenomena (Dimov et al., 2019; Laird, 2022). Due to this characteristic, it establishes a plausible connection between fundamental cognitive psychology and the field of education as it was specifically designed to serve as a versatile framework that can be applied in various settings, including education (Ritter et al., 2018; Ritter et al., 2007). In recent years, it has been extensively utilised as the basis for a new wave of ITSs. These strategies have demonstrated efficacy in enhancing student learning outcomes, hence substantiating their value in the realm of education (Anderson & Gluck, 2013).

2.10 System organisation

The Cognitive Architecture proposed in this study is organised into three main parts, which in turn are subdivided into more specialised modules, responsible for handling information. A visualised description can be seen in Figure 4.

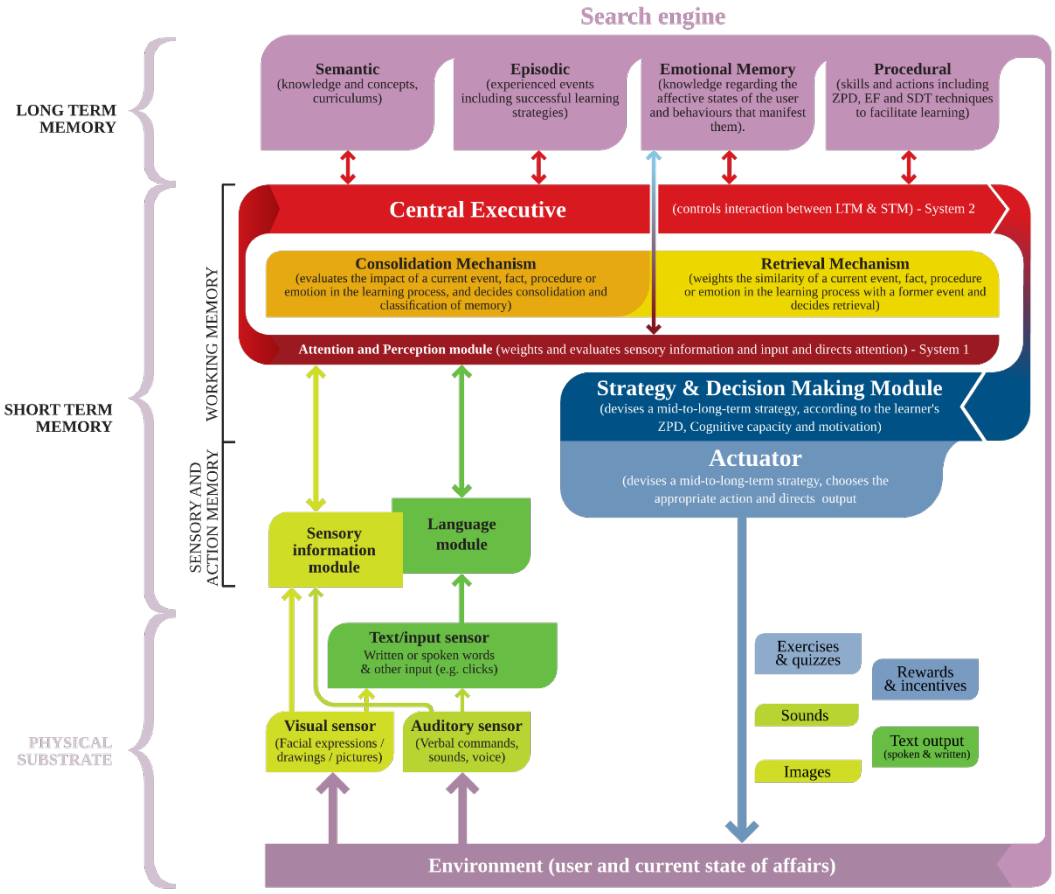


Figure 4: A visualised description of the proposed cognitive architecture

2.10.1 The Physical Substrate

The Physical Substrate represents the basic sensory interface between the system and the external environment including the user (Vernon, 2014). This physical linkage relies on sensors for input perception along with actuators for executing system decisions via observable output channels.

The included sensors encompass the full range of human sensory modalities, extracting visual, auditory, haptic and textual stimuli from the scene. Camera input registers settings and user appearance, microphones capture verbalizations, mice and touchpads capture motions and haptic inputs, and text input devices log typed entries. The proposed system avoids other sensors that could also relay signals from emotion detectors or neuro-imaging devices used in earlier systems like MetaTutor (Azevedo et al., 2022), as per this aspect, it favours simplicity, availability, plausibility and non-intrusiveness over perceptual richness.

These gathered data streams are intended to flow continuously amid user interaction, encoding rich details on identity traits, knowledge states, emotional expressions, comprehension markers, or misconception cues observable across channels (Anderson, 2007; Paas et al., 2003). Videos of facial movements and vocal fluctuations demonstrate user reactions. Environmental data (e.g. temporal or locational information) may also hold valuable educational context, as they may be indicators of a learner's habits.

Output channels subsequently relay customised feedback tailored to contextual needs, current states, and interaction history. Modalities span verbal dialogue, graphical illustrations, query exercises, incentive cues and performance summaries with explanatory rationales (Graesser et al., 2018; Graesser, 2016; Graesser et al., 2005).

This collective input/output cycle fuels an active sensorimotor loop enabling tight interactivity aligned with the broad scope of human communication and expression modalities It grounds system representation within the external learning environment.

2.10.2 Short-Term Memory

The Short-Term Memory (STM) is located on top of the Physical Substrate. This section is in charge of handling sensory data in a specialised module called Sensory & Action Memory (SAM) for retaining and assessing information during and shortly after the user interaction, as well as a specialised Language Module for performing natural language processing actions on lingual input (spoken, written, or typed). It also has another sub-section called Working Memory (WM) that stores the present state. The Working Memory (WM) contains an Attention and Perception Module (APM) that directs the corresponding modules and resources to where they are most needed, as well as a Central Executive (CE) that controls the interaction between the Working Memory and the LTM, including memory consolidation and retrieval (Vernon et al., 2007).

a. Sensory and Action Memory

The Sensory and Action Memory (SAM) stores input obtained from the outside environment via sensors for a brief amount of time, long enough to identify and forward to higher-level modules of the Working Memory. This module does just basic data processing, such as encoding and categorization. A higher-level module in the WM does the actual filtering of the information (which corresponds to human perception and attention) later in the system.

Language and Sensory Information Modules

The SAM is made up of two modules: the Sensory Information Module (SIM), which stores visual, aural, and input data from sensors, and the Language Module (LaM), which specialises in language analysis. This kind of specialisation is required since the system is intended to deal with linguistic data. This is also true for the human brain, which has language-processing and production modules such as the Broca and Wernicke regions (Eysenck & Keane, 2015). The SAM and LaM modules collaborate to extract data from sensors, which is subsequently passed on to the WM's Attention and Perception Module. As recent developments in LLMs are exponential, this module may be utilised in novel and efficient ways to understand human language in an extremely fine-grained manner (OpenAi, 2023).

The Actuator

The Actuator is similar to the Motor Cortex of the ACT-R (Anderson, 2007; Laird, 2022) is the module in charge of making the ultimate decision and carrying out the orders given by the Central Executive and the Strategy and Decision-Making Modules of the WM. The Actuator will begin delivering output depending on the most successful available method. The Actuator, like the SAM, only keeps data for as long as it is required to generate the output for the user before discarding it. Sounds, voice, written material, graphics, quizzes, exercises, incentives, and prizes may all be used as output. Again, LLMs and GAI in general offer opportunities for unprecedently refined, customised outputs (Porsdam Mann et al., 2023).

b. Working memory

Working Memory (WM) is a part of the brain that is in charge of keeping and processing information (Anderson, 2007) provided by the SAM and LaM, as well as pushing information towards them. The Central Executive (CE), the Action and Perception Module (APM), and the Strategy and Decision-Making Module (SDMM) are all parts of it.

The Central Executive

The Central Executive (CE) is a critical component that regulates the relationship between Long-Term Memory (LTM) and WM, consolidates, or decomposes information and determines whether a memory will be retained in LTM and in which module, or discarded altogether. The CE also selects which memories will be recovered from the LTM and which will be stored by using a consolidation and retrieval procedure. These two processes analyse the effect and resemblance of a current event, fact, procedure, or emotion in the learning

process and determine whether memory should be consolidated into the LTM or retrieved from it. The CE is also in charge of tagging LTM memories (Vernon et al., 2007). These tags may be *symbolic,* i.e. qualitative (for example, the memory "Athens" is tagged with "geography", "capital", "Greece", and so on) or *subsymbolic* i.e. qualitative (for example, a rule may have a numerical rating of 1 indicating success and 0 indicating failure, with values ranging from .01 to .99 in between). In comparison to the function of a human brain, the CE can be classified as a *System 2* module (Conway-Smith & West, 2022), which means it is slower and specialises in higher-level operations that demand complex thinking.

The Attention and Perception Module

The Attention and Perception Module (APM) is in charge of assessing and analysing sensory data, as well as directing attention to the most relevant input from Sensory Memory and Long-Term Memory (LTM) (Ritter et al., 2018). With a Consolidation Mechanism, the APM collaborates with the CE to assess the influence of a current event, information, procedure, or emotion on the learning process and determine whether it will be categorised and maintained in the LTM. In addition, the APM works with the Central Executive for Retrieval. It detects and retrieves the resemblance of a current event to a previous event, information, process, or emotion stored in the LTM, resulting in a comparable set of strategies, choices, and behaviours. The APM monitors environmental stimuli and employs the mechanism to compare them to existing data from the LTM by reading their tags and allocating resources where they are most required (Lieto, 2021). The tags may either be qualitative or quantitative. For example, if the visual sensor sees a map of France, the APM will compare the data from the Sensory Module to what is accessible in the LTM, discover anything comparable, and read its tags. If the tags are "France" and "Map," the memory will be retrieved, and resources will be directed to semantic content with the same tags. Apart from similarities, it may also evaluate significance. If, for example, the map is shown in the context of computing the surface area, the relevance is minimal, the memory is not recovered, and the module accesses the LTM for more relevant data.

The APM is a System 1 module (Conway-Smith & West, 2022), which implies it is primarily in charge of quick action and decision-making. This module has direct connectivity (a shortcut) to the Emotional memories in LTM to mimic the immediate actions that may be prompted by emotionally strong stimuli or memories.

Strategy and Decision-Making Module

This module communicates with the Central Executive and generates an action strategy by categorising possible strategies from the LTM (i.e., Procedural memory information in the form of rules). The Strategy and Decision-Making Module (SDMM) categorises regulations based on their evaluations as more or less effective, depending on their result. It then sends the data to the Actuator, which retrieves or generates output for the user. If the strategy is successful, it receives a positive rating, and the tag is modified and put in the LTM's Procedural memory with a new higher value. The SDMM is the module in which the ZPD, EF and SDT strategies are executed, by retrieving metrics from the LTM memory modules and fitting them into current situations and goals, thus devising a short-term micro-strategy each time it is needed (Anderson, 2007).

The proposed architecture comprised of the modules described draws inspiration from the existing ACT-R architecture to a large degree; however, it proposes a set of specialised modules, that may help the classification of data, as they may provide a clear picture for assessment and improvement. However, it is not intended to become a rigid hardware or software design, but rather outline basic functionalities for a ITS system focused on learning with specific purposes and features.

2.10.3 Long-Term Memory

The Long-Term Memory (LTM) has four distinct sub-modules: the *Semantic Memory* module, which keeps concepts and knowledge; the *Episodic Memory* module, stores past events; the *Emotional Memory* module, keeps representation of emotions connected to events and knowledge, and also conditions the system to respond quickly to strong emotional input; the *Procedural Memory* module, saves skills and actions in the form of chunks of rules. These sub-modules may hold user-related data but also may have access to a centralised database holding information from the interaction of the systems with other users. The system may also have filtered (i.e. moderated) access to general knowledge databases, and search engines.

a. The Semantic Memory module

Factual information important to the role of the tutoring system is kept in this module, such as the full curriculum of the user's classes (e.g. science, , literature, history, geography, math etc.) but also educational material such as exercises, documents etc. Some of these info may be injected to the system before launching it, or retrieved from databases, upon determining

user information, such as their age and class. Data may be tagged by subject to accommodate easy retrieval and storage. Old and new knowledge from curriculums may also be tagged as "known" or "not known" to the user using a numerical value ranging from 0 (="completely unknown") to 1 (="mastered").

b. The Episodic memory module

Past events and user experiences are saved here, also classified with tags and ratings (e.g., -1 = "absolutely negative" to 1= "absolutely positive"), with all the in between values. This module is vital for maintaining instances of success and failure, which are important to keep track of the user's performance. During the user's learning process, all matching sets of production rules that fired and led to some outcomes are sorted here, rated with any value from 0 to 1 corresponding to successful or unsuccessful, and the impact on the user's emotional and motivational states.

c. The Emotional memory module

As the relation between emotion and learning is strong, with the former acting as an covariant for the latter; this applies not only to humans, but also to machines (Kowalczuk & Czubenko, 2016; Minsky, 1988). Also, emotion may describe motivation. As intrinsically motivated actions, may bring more joy, whereas externally motivated actions driven by pressure or obligation, may cause frustration, this memory module may also rate user emotions that reflect their motives, by recognising words and phrases like "I have to", "I must", "I want" and "I like" (Deci & Ryan, 2008). This module may also keep track of the user's *self-regulation* during learning (Azevedo et al., 2017; Rothbart & Rueda, 2005), which can be measured by finding patterns in usage, tasks commitment, self-reporting etc.

d. The Procedural memory module

The system uses a *hybrid* approach; the system may start with a basic set of predefined rules, based on the proposed theoretical frameworks for self-regulation and motivation, while striving to keep the learner into an optimal learning zone. However, as the system 'grows' along the user, an *emergent* (Vernon, 2014) approach is embraced which helps it adapt to their knowledge and learning profile. As such, this module keeps the initial set of rules, which are updated or replaced as the system adapts and learns. This may be facilitated by a reinforced learning approach to ensure that the system will self-improve, given the time. The Central Executive and the Strategy and Decision-Making Module are the sub-systems of WM

responsible for the identification and the classification of the used strategies, by ranking and storing them in this module.

Chapter 3
Experimental Study

3.1 Rationale

Even though this study is mainly theoretical, the researcher attempts to assess the relationship between the constructs to be employed in the proposed Cognitive Architecture; the rationale of this decision is to keep the proposed model for the prediction of the student's motivation and cognitive abilities, it might be useful to know whether there are any interconnections or predictive relationship, resulting in a more well-informed decision model. Simply put, if a close (or even a predictive) relationship between e.g. extrinsic motivation and low WM capacity is to be found, then the model might assume that children with are more highly motivated may need more material to improve their WM. Furthermore, it might be useful to employ a gamified version of a cognitive test in order to assess and reassess the EF of children, or even ask questions regarding their attitudes towards the material given to them.

3.1.1 Assessment of Inhibition Control and Working Memory

Stop Signal tasks are a type of behavioural assessment used to evaluate Inhibition Control and impulsivity in children (Durston et al., 2002). These tasks require the child to respond, or “go”, following one signal but inhibit their response, or “Stop” signal, to a different signal. For example, the child may be asked to press a button when they see a green circle appear on a screen but not press the button when a red circle appears (Simpson & Riggs, 2006). Researchers can measure factors like the child's reaction time to press on go trials and the number of commission errors they make by incorrectly pressing on “Stop” trials (Schachar et al., 2007).

Stop Signal tasks are well-suited for studying Inhibition Control in early childhood because they are simple and engaging, making them appropriate for young children as well as children with developmental disorders like ADHD (Byrd, Loe, Pribram, & Casey, 2015). (Zhou et al., 2012; Zhou et al., 2022). Versions of the Stop Signal task have been created using stimuli like shapes, letters, cartoons, and sounds (Wiebe et al., 2011). Child-friendly

stimuli and game-like features help maintain young children's attention and motivation to complete what can be hundreds of boring, repetitive trials (Howard et al., 2014).

Research using Stop Signal tasks has revealed important developmental trends in Inhibition Control. In typically developing children, commission errors decrease gradually over the preschool and early school years, suggesting Inhibition Control strengthens over this period (Simpson & Riggs, 2006; Tottenham et al., 2011). However, children with disorders like ADHD that involve inhibitory deficits commit more errors on Stop Signal tasks than typical children, starting as early as age 4 (Berlin & Bohlin, 2002; Rubia et al., 2007). Such findings demonstrate the sensitivity of these tasks for detecting atypical development of cognitive control.

The neural systems supporting Inhibition Control on Stop Signal tasks also show protracted development over childhood. Functional magnetic resonance imaging (fMRI) studies reveal that, compared to adults, young children show weaker activation in brain regions linked to Inhibition Control like the dorsolateral prefrontal cortex and basal ganglia (B. Casey et al., 2005; Durston et al., 2002). The strength of activation in these regions, along with functional connectivity between them, increases over childhood and adolescence, paralleling observed behavioural improvements in control (B. J. Casey et al., 2005)

Stop Signal performance can be affected by experimental manipulations targeting cognitive processes like Working Memory and cognitive flexibility as well as motivational/emotional processes like reward and threat (Hutchinson et al., 2008; Pauli-Pott et al., 2014; Schloß et al., 2021). For example, introducing Working Memory demands by having children remember cue rules while performing the task increases commission errors among both typical and ADHD groups, but more dramatically for the clinical group (Jarrold et al., 2023). Such findings elucidate how taxing cognitive resources essential for top-down control can worsen behavioural regulation problems. Meanwhile, offering rewards for correct responses or having children perform under conditions of social evaluation improves ‘go’ trial reaction times and reduces “Stop” errors; these motivational enhancements likely involve activation of the brain’s dopamine-rich reward system and its connections to prefrontal control regions (Padmanabhan et al., 2011).

Taken together, research using developmentally appropriate Stop Signal tasks provides insights into the typical and atypical development of cognitive and motivational processes supporting children’s behavioural regulation. The simple, game-like nature of these tasks allows them to be used to study Inhibition Control and impulsivity starting early in

childhood. Findings from Stop Signal studies have informed models of self-control development and disorders involving control deficits like ADHD (Diamond, 2013; Nigg, 2001). When applying Stop Signal measures some concerns should be kept in mind. According to Müller et al. (2012) there is a potential for more variability in impulsivity among younger children compared to older children.

While the Stop Signal paradigm remains popular for evaluating behavioural inhibition, Simpson and Riggs (2006) delineated salient limitations in relying solely on such tasks to index effortful control capacities. They note performance integrates multiple cognitive processes beyond just inhibition per se. Accurately responding on "Stop" trials requires sustaining vigilant attention, retaining the task objective in Working Memory, and efficiently executing responses on go trials. Poor performance could therefore stem from non-inhibitory failures. Additionally, the gradual developmental improvements observed partly reflect structural maturation of the prefrontal regions implicated in top-down control, rather than improved competence in effortful regulation specifically. Furthermore, children's verbal competency provides a confounding factor, as privately using self-directed speech or rules can enhance inhibitory success in ways unrelated to motivational self-control. The impact of concurrent Working Memory loads also remains equivocal - while some secondary demands like emotional stimuli may diminish Inhibition Control, others like external incentives can sometimes boost stopping performance (Jarrold et al., 2023). The tasks also lack specificity in terms of isolating purely behavioural inhibition from affective forms. Similar overarching neural circuitry seems to support response inhibition broadly across both motivational and non-affective situations. Therefore, performance gains may not necessarily indicate improved regulation of negative emotions in real-world contexts. Finally, Simpson and Riggs (2006) argued developmental improvements likely integrate external experiential exposures and inherent biological maturation, rather than purely Inhibition Control gains alone. In summary, while unquestionably valuable, Stop Signal tasks require complementary measures to clarify the precise cognitive and motivational components contributing to developmental enhancements in wilful, effortful self-governance over thoughts, actions, and emotions. Despite some issues and confounding elements with Stop Signal tasks, their capacity to probe component processes like Working Memory and motivation that affect control, seem to be able to advance the understanding of children's emerging ability to regulate their thoughts, emotions, and actions.

Corsi Block-Tapping Test (CBTT) is a cognitive task designed to examine Working Memory. In its traditional form devised by Corsi (1972) nine wooden blocks are presented to a child in a

random arrangement on a wooden board. The researcher is pointing at a sequence of blocks, and the participants are required to replicate the sequence the researcher types in either the same order (forward) or the opposite order (backwards). Forward CBTT tests mainly the visuospatial sketchpad capacity, whereas tapping the sequence backward tests more aspects of Working Memory (Kessels et al., 2008; Kessels et al., 2000). After a correct response, the sequence increases by one block. If the participant fails to give the correct sequence for two consecutive times, then the test terminates, and the last successful trial is recorded as an integer number which indicates the span of the relevant Executive Function.

The CBTT comes also in several computerised versions; even though normative data on the electronic implementation are absent in current literature, on several occasions computerised versions of CBTT have performed as well as the traditional, with similar reliability scores, .77 for the former and .82 for the latter (Siddi et al., 2020). Additionally, Brunetti et al. (2014) reported that the computerised versions of CBTT may be more consistent in several ways; the eCorsi differs from standard CBTT in that it has more control over the Inter-Stimulus presentation times. With manual tapping, the examiner has a particularly difficult time controlling the temporal precision, which might (inadvertently) be slower or quicker depending on a variety of conditions. Furthermore, the examiner might alter the tapping finger, the amplitude of hand and arm motions, and the posture of the limb during interlapping periods. Most studies do not even mention how the examiner directed the participants' attention. The same source reports that neuropsychologists often comment that when giving especially lengthy sequences, they are obliged to slow down the tempo of block tapping in order to retain the sequence. Moreover, Brunetti et al. (2014) claims that the use of eCorsi can significantly reduce many sources of inter-trial and inter-test variability, both between subjects and between examiners, while retaining an extreme plasticity to be adapted to different research and clinical purposes through customization of timing intervals and its various task modes. Another clear benefit is that all trials are automatically logged and hence error checked.

3.2 Hypotheses

The purpose of this study is to devise a model of prediction of effortful control, based on two student-related traits using a minimalistic set of questions.

First Hypothesis: Higher scores of *Executive Functioning* and *relative autonomy* will predict higher *Inhibition Control* for elementary students, indicating a positive relationship between the predictors and the outcome.

Second Hypothesis: Higher scores of *Executive Functioning* and *relative autonomy* will predict higher *Working Memory* capacity for elementary students, indicating a positive relationship between the predictors and the outcome.

Third Hypothesis: There will be a significant positive correlation between *Inhibition Control* scores and *Working Memory* capacity.

3.3 Method

3.3.1 Design

This is a correlational study, conducted using a linear regression model with two predictors and one outcome, performed for each hypothesis. As the first and second hypotheses involve two predictors, Multiple Regression will be used, whereas, for the third, a Simple Linear Regression will be used with variables of at least interval or scale.

3.3.2 Participants

The sample were Greek-speaking elementary school students (n=71) between 8-12 years old (Age Mean(SD)=10.2 (1.2)) with their parent or teacher (n=71). Participants and their parents/teachers were recruited via an email campaign which was circulated via online school discussion groups between 20.2.2003 to 10.4.2023. 91 children participants and their accompanying adults attempted the study, however only 76 of them completed it and returned their results, so the N>68 criterion for R^2=0.15 with power=0.85 was marginally fulfilled (see Appendix 4). The sample comprised of 40 boys (Age Mean(SD)=10.2(1.2)) and 31 girls (Age Mean(SD)=10.4(1.2)). From the 76 coupled participants, 5 were screened out of the sample, as their inputs suggest that they did not follow the exams. The sample is considered stratified, as the study attempted to explore student-specific attitudes and traits. Parents and teachers offered demographic data for themselves and their children/students and, completed the CHEXI-GR translated in Greek, and then gave their place to the younger participants for the SRQ-A-GR translated in Greek; followingly, they attempted the 2 cognitive tasks, the Backwards Corsi Blocks Tapping Test and the Stop Signal task.

After the completion of the test, the participants were given a unique code to be used in the case of a withdrawal request. After completion of the online procedure for all the participants, data were downloaded, screened, and analysed.

3.3.3 Materials

Two questionnaires and two cognitive tests were utilised, and administered in sequence via the Psytoolkit online platform (Stoet, 2010, 2017) and only on desktop or laptop computers, excluding tablets or smartphone devices which typically have smaller screen sizes and different, touch-screen-based interfaces.

a. Instruments

The Children's Executive Function Inventory (CHEXI) devised by Thorell and Nyberg (2008) serves as a valuable measurement tool for assessing core cognitive self-control capacities in children that continue maturing throughout the school-age years. Specifically, Executive Functions such as inhibiting impulses, retaining information in mind, and flexibly adjusting actions comprise top-down mental processes that substantially shape voluntary regulation of behaviour and emotion as children navigate environments with expanding autonomy (Garon et al., 2008).

While cognitive control faculties have biologically embedded underpinnings, environmental exposures and learning experiences also critically influence developmental trajectories. Assessing progress regularly can identify atypical delays or may signal a need for support. As laboratory tests remain impractical for widespread screening, CHEXI offers an efficient proxy, using adult observer ratings on typical Inhibition Control, Working Memory, and mental flexibility displays in everyday settings. Analyses substantiate concordance between inventory ratings and direct measures, supporting validity. Thorell & Nyberg's (2008) CHEXI questionnaire supplies an accessible tool for this crucially informative screening via adult reports on practical Executive Function employment.

Numerous aspects of behaviour rely on efficient Executive Functions (EFs), which include higher-order cognitive skills such as Working Memory, Inhibition Control, and planning. To concentrate primarily on Executive Functioning, Thorell and Nyberg (2008) devised the Childhood Executive Functioning Inventory (CHEXI). Their study examined the concordance between Executive Function (EF) laboratory measures of inhibition and Working Memory, and rating instruments measuring ADHD symptoms and early academic skills in children. 113 Parents and 89-105 teachers provided ratings of children's EF difficulties, ADHD

symptoms (hyperactivity/impulsivity and inattention), and academic skills (language and mathematics). Correlational analyses were conducted to assess the degree of agreement between EF test performance and EF ratings across informants. Results showed several moderate positive correlations between EF lab measures and EF ratings by parents and teachers. For example, parent ratings of Working Memory difficulties correlated .33 ($p < .001$) with children's performance on Working Memory tests. Teacher Working Memory ratings also correlated significantly with Working Memory tests ($r = .29$, $p < .01$ with parents; $r = .39$, $p < .001$ with teachers). Similar correlations were found between inhibition test scores and parent/teacher ratings of inhibition difficulties (r values ranging from .28 to .35). EF lab measures also correlated as expected with parent and teacher ratings of ADHD symptoms and academic skills deficits. Based on the results, Thorell and Nyberg (2008) concluded that EF lab-based tests and CHEXI, designed to assess related EF constructs show convergence across informants.

This instrument has shown interesting properties when it comes to collecting more general behavioural information. However, it is not meant to be a replacement for laboratory EF measurements; Toplak et al. (2013) suggest that performance-based and ratings-based measures of Executive Function do not represent the same degree of analysis, underlying process, or neural substrate. It is based on teacher or parent ratings; the issue of reporting bias should be considered. As such, in conjunction with the self-reporting and evaluator reporting instruments, two cognitive tasks were employed directly by the young participants; the first task is a simple Stop Signal task (Criaud & Boulinguez, 2013; Verbruggen & Logan, 2008) to examine the relationship between response inhibition, Executive Function and effortful control whereas the second is a Backwards Corsi Block-Tapping task. For the complete instrument, please see Appendix 3.

For the purpose of this study, the Greek version of CHEXI (Thorell & Nyberg, 2008) was administered. It is a 25-item, 5-scale Likert-based instrument (1=Definitely not true, 2=Not true, 3=Partially true, 4=true, 5=Definitely true) and with statements of the type “When asked to do several things, he/she only remembers the first or last”, “Has clear difficulties doing things he/she finds boring”, “Has difficulty stopping an activity immediately upon being told to do so. For example, he/she needs to jump a couple of extra times or play on the computer a little bit longer after being asked to stop” etc. The instruments' 24 items are clustered into 4 subscales, with the two first representing Working Memory and Planning; the last two representing Regulation and Inhibition; the first two add up to the Working Memory while the last two add up to Inhibition. The CHEXI-GR reported an excellent

internal consistency (Cronbach's α=.95) which is in accordance with previous measurements of α=.93 (Sofologi et al., 2022).

The second instrument was the Greek version of Academic Self-Regulation Questionnaire (SRQ-A-GR, Standard Version), devised by Grolnick and Ryan (1987), which is a 4-scale Likert-based instrument (4=Very true, 3=Sort of true, 2=Not very true, 1=Not at all true) which consists of 4 sections, each with a question like "Why do I try to answer hard questions in class?" followed by statements like "Because I want the other students to think I'm smart", "Because I feel ashamed of myself when I don't try", "Because I feel ashamed of myself when I don't try" which students are asked to report the degree of truth, in the 4-level scale provided. The instrument's 32 items are clustered in four factors, reporting the type of motivation, i.e. External, Introjected, Identified or Intrinsic. The instruments provide scores for each factor, along with an accumulative Relative Autonomy Index (RAI) which is calculated as RAI=(2*Intrinsic + Identified - Introjected – 2*External). The SRQ-A_GR was translated according to the International Test Commission (ITC) standards (www.intestcom.org, accessed on 12 March 2023). A back-translation procedure was also implemented in order to eradicate any inconsistencies that would disrupt the accuracy of the results. The SQR-A-GR reported a good internal consistency (Cronbach α=.79) which is well within acceptable boundaries (Field, 2018).

Cognitive tests

The two cognitive tasks were administered again via the Psytoolkit platform (Stoet, 2010, 2017). Stimuli were presented in an 800 X 600 pixel rectangular on the computer screen of the user.

b. Implementation of the Stop Signal task in this study

The "Stop Signal" task initiates with instructions: A white circle (150px diameter, 10px thick) with a white fixation cross appears on centred on the screen. After 2 seconds, the circle will display either a left or right-pointing green arrow, randomised in direction across trials. For green arrow "go" trials, participants should press the "Z" key on their keyboard for left arrows or the "M" key for right arrows, as quickly and accurately as possible before the 500ms response timeout. However, on a subset of "Stop" signal trials occurring randomly, 150ms after the green arrow appears, the fixation circle will turn red. When participants see the red circle "Stop" signal, they must inhibit their key press response within 500ms before the trial ends. There is a 2 sec rest interval between trials.

The first block serves as “go” familiarization of up to 12 “go” only trials, ending after either 6 consecutive correct “go” responses or a maximum of 12 trials, ensuring that the participants had familiarised themselves. The second block consists of 50 baselines, “go” only trials ending after 20 correct consecutive “go” trials or after a maximum of 50. Then comes a familiarisation block with the “Stop” task, with 12 trials, of which 8 “go” trials and 4 randomly sequenced, intermixed “Stop” signal trials. Finally, the critical test block delivers 30 “go” trials and 10 randomly sequenced, intermixed “Stop” signal trials, demanding participants' response inhibition when the sudden red circle appears 150ms into the initially cued go response. Response speed and accuracy are emphasised for go trials, scored correct if the correct arrow key is pressed in under the 500ms timeout. Correct responses within the 150ms interval before the red “Stop” signal circle trials are also considered correct.

c. Implementation of the Corsi Blocks Backward Task in this study

The reverse Corsi blocks task displays 9 purple blocks in random locations within an 800x600 pixel black area on the centre of the screen. At the start of each trial, signified by an audible cue, a sequence of blocks briefly flashes yellow in succession. Once the sequence ends, the participant must click the blocks in the reverse order from which they flashed. The blocks are 100px squares that flash yellow for 600ms with a 600ms interval between flashes in the sequence. Participants have up to 20 seconds after the audible tone to respond by clicking the blocks in reverse order before the trial times out.

The participants received written instructions on the screen, and by pressing the space button, they were given a 3-2-1 countdown, before a training block of trial presents a sequence of just 2 blocks. The participants must respond by clicking the squares in reverse order. The next training trial is a sequence of 3 blocks.

After the training concluded, the actual measured trials commenced, starting with a sequence of 2 blocks and moving onwards to 3, and so on, each time a participant responds correctly by clicking the squares in the reverse order. For each correct response, the next sequence increases by one additional block up to a maximum possible sequence of 9 blocks. Participants complete trials at increasingly longer span lengths until they fail to correctly reverse 2 successive span lengths. The task ends after 2 consecutive failures, and the participant's reverse Corsi Block Working Memory span score is determined by the longest sequence length they successfully responded to.

3.4 Procedure

The participants (children and parents/teachers) were invited to the experiment via an email campaign, including a briefing form which informed them of the procedure and link to the online experimental platform. Consent was given via the online form. All children and their accompanying adults were informed that they were free to withdraw from the evaluation process at any time. The children were asked to be examined individually in a quiet room and to complete the study without the adult intervening. Practice trials were offered to ensure comprehension of every task. Due to the specific type of the current research, demographic data such as age, gender, and educational level were included. Since these are considered personal data, the European Union law that has existed since May 28, 2018, was applied. According to European law, the use of sensitive personal data is allowed solely for research reasons. The study's protocol followed the principles outlined in the Helsinki Declaration.

3.5 Results

3.5.1 Descriptive statistics

The sample consisted of 71 elementary school students, attending 3rd, 4th, 5th and 6th grade and between the age of 8 to 12 (Age Mean(SD)=10.2(1.15). Of them, 40 where boys (Age Mean(SD)=10.23(1.28) and 31 where girls (Age Mean(SD)=10.2(1.07). For details, see the following table. The reporting adults were 44 women and 27 men. They were either the parents of the child (N=50) or their teacher (N=21).

d. Table 2: Sociodemographic Characteristics of Participants

Children

	Age		Class			
	mean	*SD*	*3rd*	*4th*	*5th*	*6th*
Boys (n=40)	10.23	1.28	4	15	11	10
Girls (n=31)	10.2	1.07	5	8	9	9
Total (n=71)	10.2(1.15)		9	23	20	10

Reporting adults

	Relationship		Level of Education			
	Parent	*Teacher*	*Basic*	*Post Basic*	*BSc*	*Postgrad*
Women (n=44)	24	20	3	18	8	15
Men (n=27)	26	1	4	7	10	6
Total (n=71)	50	21	9	23	20	10

3.5.2 Exploratory Data Analysis

After extraction of the demographic data, frequency tables were created (see table above). The parent/teachers' reported scores from the CHEXI-GR, the child participants' scores from the SRQ-A-GR and the cognitive tasks were analysed and mean scores were calculated.

Literature suggests that correlational studies do not require a normal distribution in the sense that methods examining causal relationship do (Field, 2018); however, in this study the sample was relatively small, and as such, the sample was examined for normal distributions to have a better understanding of the data but also for the sake of data exploration. The input data was examined via observation of distribution plots and normality test, and by the examination of standardised values ($Z_{values.}$).

The normality tests reveal statistically significant deviations from normality. The Kolmogorov – Smirnov test revealed non-normality of distribution for all the variables, except for the CHEXI-GR, whereas the Shapiro Wilk reported non-normality for all of the variables.

e. Table 3: Normality tests

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	p
RAI	.124	71	.009	.900	71	<.001
CHEXI-GR	.079	71	.200	.969	71	.077
Correct "Stop" responses	.162	71	<.001	.939	71	.002
Corsi Back. Span	.166	71	<.001	.942	71	.003

The box plots revealed some potential extreme values for the RAI scores and the Corsi Backward Span. However, the graphical representation for small sample sizes (n<200) may be inaccurate, and for a better evaluation of the data, Standardised Z-values were calculated for the scores, the predicted scores, and the residuals.

The examination of the standardised scores (Z_{scores}) and standardised residuals ($Z_{residuals}$) for all variables, revealed no values beyond the acceptable boundaries of ±3 for Z_{scores} and the ±3.29 range for the $Z_{residuals}$ (Field, 2018), therefore, no outliers were identified.

The scores of the variables were normally distributed for the most part. However, some normality issues were observed on the RAI score histogram, as negative kurtosis was observed on the distribution bell. This assumption was confirmed by examining the $Z_{kurtosis}$ value (Kurtosis / Std.Error = 6.74> ±2.58). The normality tests for the standardised predicted value ($Z_{predicted}$) yielded non-significant results (K-S: $D(71)=.076$ with $p=.200>\alpha=.05$, S-W $D(71)=.990$ $p=.833>\alpha=.05$) but contrastingly, revealed significance for the $Z_{residual}$ scores (K-S: $D(71)=.107$ with $p=.042<\alpha=.05$, S-W: $D(71)=.961$ $p=.027<\alpha=.05$).

For the first hypothesis testing, the Durwin-Watson test yielded acceptable results (*D-W*=1.565) and the Variance Inflation Factors (VIF) revealed normal results (*CHEXI:* VIF=1.012<10, *RAI:* VIF=1.012<10). No multicollinearity issues were observed as the Variance Inflation Factors (VIF) revealed normal results (*CHEXI:* VIF=1.012<10, *RAI:* VIF=1.012<10). No heteroscedasticity issues were observed, as both the Breusch-Pagan test and the Koenker test revealed no significant results (B-P: χ^2 (1, *N*=71) = .022, p=.99 > α=.05, K: χ^2 (1, *N*=72) = .028, p=.99 > α=.05).

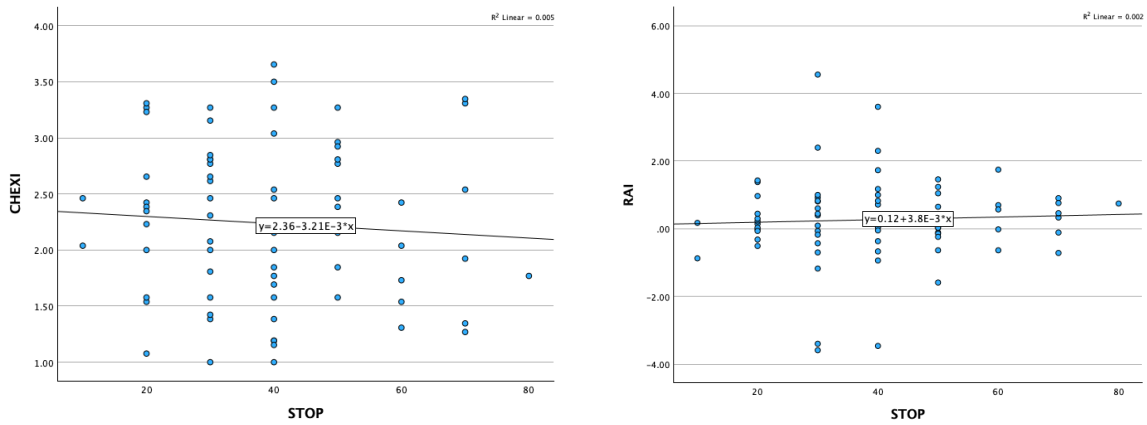
For the second hypothesis testing, the Durwin-Watson test yielded acceptable results (*D-W*=1.732), suggesting no autocorrelation issues (1<*DW*<3); the Variance Inflation Factors (VIF) revealed normal results (*CHEXI:* VIF=1.012<10, *RAI:* VIF=1.012<10). No multicollinearity issues were observed as the Variance Inflation Factors (VIF) revealed normal results (*CHEXI:* VIF=1.012<10, *RAI:* VIF=1.012<10). No heteroscedasticity issues were observed, as both the Breusch-Pagan test and the Koenker test revealed no significant results (B-P: χ^2 (1, *N*=71) = .198, p=.91 > α=.05, K: χ^2 (1, *N*=72) = .211, p=.90 > α=.05).

For the third hypothesis testing, the Durwin-Watson test yielded acceptable results (*D-W*=1.778), suggesting no autocorrelation issues (1<*DW*<3). No heteroscedasticity issues were observed, as both the Breusch-Pagan test and the Koenker test revealed no significant results (B-P: χ^2 (1, *N*=71) = .148, p=.700 > α=.05, K: χ^2 (1, *N*=72) = .153, p=.70 > α=.05).

Despite that, the score normality tests yielded some significant results, in correlational models, the normality of distribution is not a critical assumption. If the homoscedasticity criterion is met and the variables are at least of scale or interval the assumption for the multiple regression is and therefore the study can proceed without any transformation over the dataset.

A correlational design was employed to examine the first hypothesis; to determine whether *Executive Functioning,* measured by the CHEXI score, and *relative autonomy* in academic settings, measured by the RAI score predict higher *Inhibition Control* for elementary students, measured by "Stop" Signal error rate scores.

Two scatterplots (1,2) for each predictor were examined for goodness of fit, which visualised loosely scattered datapoints around low-sloped fit lines. The low slopes are also noticed at the regression equations (y_{chexi} =2.36-3.21E-3x_{stop}, y_{RAI}=0.12+3.8E-3*x_{stop}).

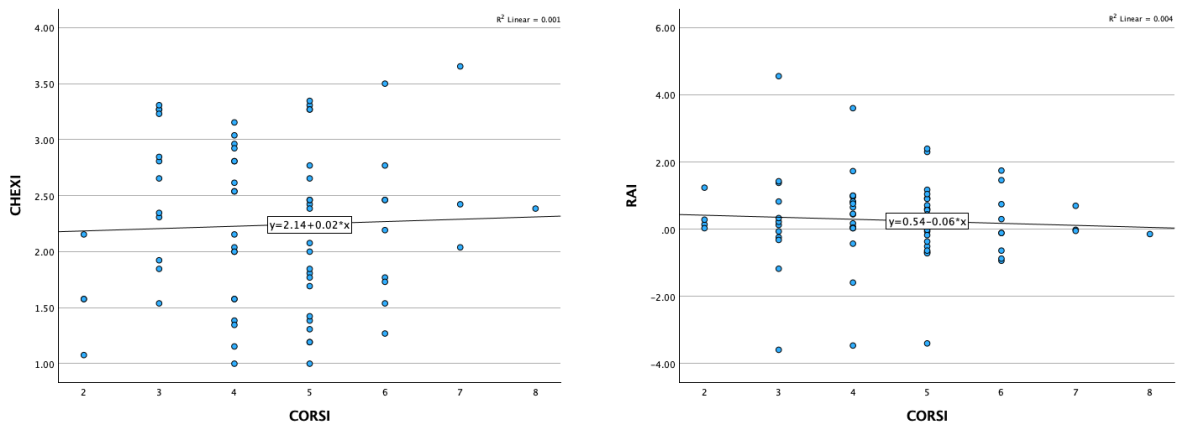


Scatterplots 1,2: Goodness of fit for CHEXI / Stop Signal Inhibitions and RAI / Stop Signal Inhibitions

Data were assessed by Multiple Regression using the Enter Method. The regression equation produced a minimal effect size (R^2 = .007, R^2_{adj} *=-.022)* indicating that Executive Function, as measured by CHEXI score and relative autonomy, as measured by RAI score, were not significant predictors of the Inhibition Control, measured by error rate ($F(2,68) = 2.44$, $p = .784 > \alpha = .025$, one-tailed). No significant predictors were identified.

A correlational design was employed to examine the second hypothesis; to determine whether if *Executive Functioning,* measured by the CHEXI score and *relative autonomy* in academic settings, measured by the RAI score predict higher *Working Memory* capacity for elementary students, measured by Corsi Backward Span.

Two scatterplots (3,4) for each predictor were examined for goodness of fit, which visualised loosely scattered datapoints around low-sloped fit lines. The low slopes are also noticed at the regression equations (y_{chexi} =2.14+0.02*x_{corsi}, y_{RAI}=0.54-0.06*x_{corsi}).



Data were assessed by Multiple Regression using the Enter Method. The regression equation produced a minimal effect size (R^2 = .068, R^2_{adj} =-.025) indicating that Executive Function, as measured by the CHEXI score and relative autonomy, as measured by the RAI score, were not significant predictors of Working Memory capacity measured by Corsi Backward Span (F(2,68) = .162, p = .853 > α = .025, one-tailed). No significant predictors were identified.

A correlational design was used to examine the predictive relationship between Inhibition Control and Working Memory, via a Simple Linear Regression using the Enter Method. The regression equation produced a very small effect size (R^2 = .09), indicating that Inhibition Control was a weak, though significant predictor of Working Memory (F(1,69) = 8.123, p = .006 < α = .050, two-tailed). There was a significant positive relationship between Inhibition Control ability as measured by correct inhibitions at the Stop Signal Task and Working Memory capacity as measured by the Corsi Blocks Backward Task (t = 2.850, df = 70, p = .001 < α = .050) with Working Memory capacity increasing with Inhibition Control ability. The model predicts that one unit change in Working Memory capacity would result in an increase in Inhibition Control ability by .026 units.

3.6 Discussion on the experimental study

The purpose of this study was to examine whether two instruments, namely the parent/teacher report-based CHEXI-GR and the self-reporting SRQ-A-GR can predict two components of Executive Function (EF), namely (a) the Inhibition Control and (b) the Working Memory (WM) of school children of 8-12, and the association between Inhibition Control and WM. The study was conducted using two distinct Multiple Linear Regression analyses. The first used two instruments as predictors and the number of correct inhibitions on a stop signal task as outcomes, while the second used the two instruments as predictors and the correct size of a sequence on a Corsi blocks reverse task. The association between inhibition and WM was examined by a Simple Linear Regression analysis.

The statistical analysis did not confirm the two first hypotheses. This may indicate that the predictors and the outcomes are independent, and therefore no significant predictive value can be attributed to any of the factors. The results suggest that the predictive model has minimal practical implications over the devised ITS model, pointing out that either another predictive model is needed, based on another instrument(s), cognitive task and/or

experimental design, either that they all need to be assessed without assuming that they may coexist linearly, or that the specific design had serious design flaws that affected the results.

The first hypothesis is rejected, as *Executive Functioning* and *relative autonomy* did not seem to predict *Inhibition Control* of elementary school students. The second hypothesis is also rejected, as *Executive Functioning* and *relative autonomy* do not seem to predict higher *Working Memory* capacity for elementary students.

These findings are in accordance with previous research findings which cannot conclusively associate reported measures with actual performance. Georgiou and Zhang (2023) investigated how different aspects of EF relate to reading and math ability using a hierarchical regression approach; the researchers used the Working Memory Test Battery and the Inquisit Tower Task for assessing the WM, and the Colour Word Interference Test and DKEFS Trail Making Test for EFs. The study specifically focused on four EF components; planning, inhibition, shifting and Working Memory. Only Working Memory showed an association with reading and math performance based on EF measures. The study also considered ratings of EF using the CHEXI, but these ratings were not associated when reading and math abilities when other factors were taken into account. Surprisingly, though, in the same study, a significant correlation *was* found, between the results of the Backwards Digit Span test, which is in some respects fairly similar to the Corsi; with the crucial differentiation that the latter assesses the visuospatial memory, whereas the former also involves the phonological loop. Varying degrees of intercorrelation between behavioural-rated measures and EF tasks are also reported for other instruments; Tamm and Peugh (2019) examined correlations between various Executive Function and cognitive control measures in 244 children. Measures included parent/teacher ratings of behaviours related to self-regulation (e.g. Behavioural Regulation, Inhibition Control), direct assessments of Executive Function skills (e.g. Head-Toes-Knees-Shoulders Revised (HTKS), Trails Switching), and temperament factors reflecting attentional and Inhibition Control (e.g. Attentional Control, Suppression). Several moderate intercorrelations emerged between the ratings and direct measures associated with similar constructs. For example, parent-rated Behavioural Regulation correlated -.53 ($p < .01$) with Inhibitory Self-Control and -.45 ($p < .01$) with overall Executive Functioning. The HTKS Executive Function test also correlated with parent-rated Attentional Control (.29, $p < .01$) and the temperament factor of Concurrent Self-Control (.31, $p < .01$). However, some measures displayed divergence across constructs. For example, the temperament factor of Suppression showed non-significant or negative

correlations with direct tests of Executive Functioning and Inhibition Control. Additionally, several measures were associated with age, indicating developmental changes in Executive Functions over childhood. Older age correlated with improvements in attentional control and some direct tests but decreases in parent-reported behavioural regulation. In summary, the varying degree of intercorrelations provides evidence for both convergence and divergence between ratings and direct tests aiming to measure children's development of Executive Function and self-regulation abilities. The complex pattern indicates measurement methods may assess both overlapping and distinct subsurface processes. The literature review results suggest that EF is not a unitary concept but rather of system of constantly interacting subsystems; as their performance has a direct impact on the manifestation of behaviours, building a predictive model based on EF measures only or reported behaviour measures only may be inadequate.

Another study by Camerota et al. (2018) on a large sample (n=846) of preschool children of 4 and 5 years of age, examined correlations between parent ratings on the CHEXI Working Memory (WM) scale (mean = 2.08, SD = .67), the CHEXI inhibition scale (mean = 2.68, SD = .71), and a direct performance-based measure of Executive Function of children (EF Touch task; mean = .65, SD = .14). The CHEXI scales showed a strong positive correlation (r = .76), providing evidence for convergent validity. However, the CHEXI WM and inhibition scales showed weak negative correlations with the EF Touch task performance (r = -.10 and r =-.05 respectively). This divergence between ratings and the direct EF measure highlights differences between assessment methods which suggests that multiple tools and measurement methods may provide a more comprehensive picture of children's Executive Function. However, this study addresses a different population, within a different developmental stage; therefore, it is mentioned only as a reference that the CHEXI may have some predictive capacity for younger children.
A post-hoc analysis for the current study reported similar descriptive measures for the CHEXI WM (mean = 2.05, SD = .67), and similar measures for CHEXI EF (mean = 2.47, SD = .83). In terms of the direction of correlation, the regression analysis also revealed a non-significant negative correlation for the CHEXI WM / Corsi correlation (r = -.05, p=.97 > α=0.05) and a non-significant negative correlation for the CHEXI inhibition (r = -.04, p=.10 > α=0.05). This might be an indication that given a larger sample size, the results might have been significant.

There is extremely limited data available to compare the findings of this study to earlier findings regarding the relationship between type of motivation and Executive Function;

during the literature research, only one study attempted to correlate SQR-A with EF (Sosic-Vasic et al., 2015). According to this study which was conducted on 208 children of a diverse range of ages (5-14 years), intrinsic motivation negatively predicted differences in Executive Function error rates across congruent and incongruent trial conditions (B=-.21 to -.40), indicating enhanced Executive Functioning for intrinsically motivated children. Conversely, external regulation positively predicted Executive Function error rate differences (B =.36 to .40), linking this negative motivational style to poorer cognitive self-regulation. Also, teachers' provision of autonomy support itself was associated with reduced Executive Function error rates in students (B =-.014 to -.037). This connects higher autonomy support to better child performance on tasks demanding executive control. Additional findings showed that girls exhibited both elevated intrinsic motivation and lower Executive Function error rates than boys. Junior high school students similarly displayed faster reaction times and lower error rates than primary school children. However, it can be observed that the correlations are weak, despite being significant; moreover, for some measurements, the significance level was at α=.10 which suggests that the results were marginal. Sosic-Vasic et al. (2015) comment that EF may be intercorrelated with the locus of motivation.

A post-hoc analysis for the current study, analysing similar variables i.e. type of motivation and inhibition performance, revealed a non-significant, weak negative correlation between the intrinsic motivation of children and inhibition performance (B = -.116, p = .500 > a =.05) and a non-significant weak positive correlation between the externally controlled motivation of children (B = .107, p = .539 > a =.05). Despite the fact that we cannot directly compare non-significant and significant results, a comment should be made; the measure of inhibition performance in the current study is the number of correct trials, therefore the trend in this study is reversed compared to Sosic-Vasic et al. (2015). Therefore, the assumption that a larger sample size might have yielded significant results may be highly speculative in this case. The results are extremely marginal; therefore, no conclusive comments can be made.

As the results from the first two hypotheses testing were not significant, they are not suited for conclusive remarks; however, in conjunction with previous studies, they suggest that the parent/teacher reported assessment of cognitive function using the CHEXI-GR may not be an accurate indicator of the actual Executive Function of children, at least as per the visuospatial Working Memory and the Inhibition Control. Therefore, using such instruments in an ITS to assess the relative cognitive domains solely by reports, may be counterproductive. The same applies to the SQR-A-GR; as the results from this study and previous literature suggest, the type of motivation may only account for extremely small

variation in Inhibition Control and WM. Again, this may be only taken as an indication, as the current research on the matter is extremely limited.

The *third hypothesis* was confirmed, as a *positive correlation* between *WM* and *EF* was identified in this study. No other study was found regarding the relationship between WM and EF measured by the tasks employed and to the population examined in this study. However, the results revealed only a small effect size. Previous literature assesses the correlation between WM and EF as two closely interconnected whole constructs using numerous tasks; no study was found for the relationship between Visuospatial Memory and Inhibition Control *in particular*. For example, McCabe (2011) measured an extremely high correlation between EF and WM, at r = .97, based on a series of measurements by batteries of numerous instruments, such as the Computation Span task, a Letter Rotation Span task, a Reading Span task and a Match Span task for WM; and the Wisconsin Card Sorting Test, a Verbal Fluency task, a Mental Arithmetic task and a Mental Control task for EF. The Corsi Blocks task was not employed, in either of its forms, backwards or forward. The two tasks are also used as measures for clinical purposes; indicatively, a study by De Jong et al. (2009) pinpointed the effect of Attention Deficit Hyperactivity Disorder on the performance of WM and EF using a a Corsi Blocks forward task and a Stop Sign task, but no correlation between the two is reported. Therefore, this study contributes to the field by associating to some small extent, Inhibition Control and Working Memory capacity. Future studies may need to replicate the results using larger sample sizes & controlled conditions.

The lack of empirical studies reporting results from one or a small number of task-based measures, especially in high-stake situations like clinical assessment or for educational purposes, may be explained by the notion of *Task Impurity* (Miyake et al., 2000) which states that no *single* task can be exclusively associated with Executive Function. Therefore, to assess EF functions, it seems to be essential to incorporate batteries of tasks. This may explain why the results of this study were marginal in terms of correlation coefficients, as they only can explain a fraction of variance.

3.6.1 Limitations of the experimental study

Each participant attended the study from their home, on their own space and computer each one at a different time of the day. This may provide some arguments for ecological naturality; however, this cannot screen out numerous confounding factors, like noise, distractions, interruptions, the participants receiving help from others, not adhering to the

instructions and more. Also, the researcher could not check who was the person who participated, as complying with the screening process was at the discretion of the adult who was overseeing the process.

It should also be noted that high attrition was observed. 91 participants attempted the study, but only 73 completed it; moreover, by observing the raw data it can be derived that some of the children did not complete the cognitive tests, abandoning the study during the process or in between the two cognitive tasks. Also, by observing the results from the keyboard inputs as they were recorded from the experimental platform, it may be derived that some child participants may have "gamed" the process by intentionally doing the opposite of what was asked. Others may have failed to understand that they should have repeated the Corsi Blocks input in the reverse order. All these may be indicative of design flaws in terms of human-computer interaction. Some researchers use more game-oriented interfaces e.g. the Inhibition Control test is designed as a children's computer game, e.g. the young participant is requested to try and catch fish as they appear on the screen by pushing a button in a well-timed manner or refrain from it when a shark shows up (Howard & Okely, 2015). The lack of this kind of implementation may be even more crucial for the present study, as the proposed ITS model is intended for young children. An effort was made to gamify the process as much as possible for both the CBBT and the Stop Signal tasks, however, the results may be far from what children consider a game nowadays.

The attrition observed during the experimental design may also indicate that the process may have been toiling for the youngest children or the ones with lower self-regulating capabilities; it might have been useful to examine the incomplete raw data of the children who dropped out while filling in the SQR-GR-A or during the cognitive task, however, they were not available due to some restrictions of the experimental platform.

Also, the sample size may have been small; even though the N>68 criterion –indicated by the *a-priori* analysis was fulfilled (see Appendix 4) larger samples may be required, allowing for a stricter screening process.

Another possible design flaw may have been the rather broad age range of the young participants. Ranging from 8 to 12 years of age, the participant group may have been quite heterogenous, as different ages may also represent different developmental states of EFs (Maldonado et al., 2020). Boys (n=40) were also more than girls (n=31), thus the developmental disparity may also have been confounded by gender differences, which is a common finding in current literature (Schirmbeck et al., 2020).

Finally, the translation procedure of the SRQ-A-GR from the English original has some flaws. The instrument was translated and back-translated, and checked for its internal consistency, reporting an acceptable result. However, the procedure lacks a Confirmatory Factor Analysis to measure the performance of variables in terms of construct representation.

Chapter 4
General Discussion

The purpose of this study was to propose a Cognitive Architecture for an adaptive Intelligent Tutoring System for elementary school students. By employing principles of empirically well-established psychological theories for learning and motivation, the proposed model is an attempt to transfer knowledge from the domain of human-to-human instruction to computer-assisted instruction. The recent developments in AI, and especially the rise of Generative AI that the world is witnessing as these lines are being written, brings new prospect into the field of cognitive systems, in an unprecedented manner, making it hard for any researcher in this field to keep up with current trends. The field of Computer Assisted Learning is expected to experience (if not *already* experiencing) a large-scale transformation after the advent of Large Language Models (LLMs) and Generative AI (GAI) and their prospective incorporation into new ITSs. New studies are published each day at an unprecedented rate, bringing new applications for these models. The sought-after field of ITS is also expected to grow with new systems exploiting the current and emergent abilities that these systems bring. At the same time, as the performance of LLMs is improving, there are mixed responses, fluctuating from awe and enthusiasm to reluctance and suspicion for the use of AI in educational settings, including an existential fear of educators, considerations of explainability, accountability issues, and ethical considerations. This may be the reason why despite this exponential growth in the performance and penetration of LLMs and GAI, only a few widely available ITS platforms adopted these new technologies; two EdTech companies, Khan Academy and Duolingo have incorporated chatbots in their learning platforms. Therefore, by delving into the current literature on the domain, a gap was identified; currently, there are no Intelligent Tutoring Systems, neither theoretical nor applied, informed by the principles of the Zone of Proximal Development (ZPD) Ryan and Deci (2000, 2017) Self-Determination Theory (SDT), while attempting to assess their Executive Function abilities of users to assist their learning elementary school children by adapting to their knowledge level, enhance their motivation and promote self-regulated learning. In that sense, this study contributes to the field by bringing together existing knowledge to provide a theoretical outline of a Cognitive Architecture and to inspire future ITS designs. This contribution is not made in terms of mathematical modelling or

programming. This effort is theoretical, with a focus on psychological theories of learning, motivation, and cognition, taking into account contemporary approaches to cognitive design and cognitive science. The implementation of the cognitive architecture into real software is left to the programmers who might find inspiration in this work. The researcher attempted to delve into the literature, extract and synthesise knowledge that may be incorporated into this proposed model and create a novel approach, building, however, over previous, long-lasting knowledge on the subject. To provide a model that may be able to include future developments, rather than an implementation based on today's available tools. It is an effort to compose general, long-lasting knowledge that may be relevant regardless of the turn things may take in AI and Cognitive Science; it aspires to provide general guidelines, or so to say a general cognitive *blueprint* that will make the work of future developers easier by guiding them into what *concepts* they need to take into account, rather which *tools* they need to use. In brief, this study is *concept*-oriented work, rather than implementation-oriented. Given the speed at which things develop in terms of AI implementations, it could render an attempt for the latter obsolete in a matter of months. In the same vein, the researcher adopted some of the most contemporary, well-researched and well-established, theories of learning and motivation.

As such, the design is based on an ACT-R, one of the most lasting and predominately used models of cognitive architecture (Anderson, 2007; Laird, 2019; Laird, 2022). As ACT-R relies more on a *humanesque,* biologically-inspired framework, mimicking the components and the functions of a human brain, it may facilitate the incorporation of theories of cognition, motivation and learning more intuitively; moreover, ACT-R is considered an innately hybrid approach, more able to develop emergent features, i.e. characteristics and capabilities that were not imported during the initial development of the system, but rather incorporated through an interaction with the environment (Ritter et al., 2018). Furthermore, recent research has focused on the development of theoretical models that can accommodate some features of human cognition, and has seen some implementations in ITS systems (Dimov et al., 2019; Ritter et al., 2007). Moreover, the ACT-R architecture has been developed into a complete software nowadays, leaping a theoretical framework to a development platform for cognitive scientists and developers, available for download at the website of Carnegie Mellon University (see http://act-r.psy.cmu.edu/). Additionally, ACT-R adapts a *symbolic* and *sub-symbolic* representation of knowledge i.e. the classification of knowledge into explicit pieces of meaning and accompanying implicit set of rules and instructions on how to handle existing and synthesise new information. For example, the subsymbolic system uses

mathematical equations to represent memory retrieval. These equations, amongst others, define the possibility and speed of memory retrieval or memory consolidation (Lieto, 2021, pp. 16-17). This feature of the ACT-R cognitive architecture creates opportunities for extremely fine-tuned representations, up to the point of predicting the neural activity of users (Dimov et al., 2019).

The Zone of Proximal Development (ZPD) is a conceptual framework almost a century old, even though it came into worldwide prominence by the late 70s to the early 80s. Vygotskian principles are widely acknowledged to provide an excellent theoretical and practical guide for children's development, teaching approaches and principles for all sorts of knowledge, student support, curricula tutoring, and more. This study begins by attempting to find similarities in ways both human tutors and Intelligent Tutoring Systems interact with learners and suggests that ZPD and its key concept 'scaffolding' provide an excellent frame of reference to model the latter in human manners (Ballard & Butler, 2011; Chounta et al., 2017; Ferguson et al., 2022; Murray & Arroyo, 2002; Vainas et al., 2019). Surprisingly enough, the conceptual frameworks of all the software designs for adaptive and personalised learning the researcher came across were inherently Vygotskian; in the sense that any system of this sort will attempt to offer the appropriate material that will advance the knowledge of a learner optimally for their knowledge advancement and their motivation, at a pace that best suits their learning ability.

It could be argued that ZPD has become even more relevant and up to date, as the internet has made, and it is still making more knowledge available every day; and an ITS may have access to vast amounts of educational material to choose what is deemed appropriate not only to the learners' abilities but also to the learner's preference. Also, the increasing computational power of interconnected home computer systems and handheld devices with which users interact many times daily may provide rich ZPD-related data for a system to use.

The incorporation of ZPD into an ITS has the advantage of being intuitive in modelling for a computational system. It involves iterative circles of evaluation, delivery, and re-evaluation, that may provide a safeguard for the learner, in the sense that the users of such a system will remain in a state of learning, even when the emotional, motivational, and cognitive aspects of their learning experience cannot be effectively measured.

However, the benefits of tapping into a learner's experience of learning may significantly enhance it; learning is a complicated, interwoven process that involves many levels of cognition. For that purpose, the proposed model should be able to evaluate the behavioural

and cognitive traits of its users and be able to develop personalised cognitive profiles for each one of them. Therefore, it encompasses an interface to accommodate sets of non-intrusive sensors (e.g. cameras, microphones, mice, keyboards etc.), being able to extract real-time data from the perceived environment including the users' affective states, their motives and learning habits, devise personalised patterns of data that can be used to develop personalised and adaptive strategies to promote optimal learning (Ballard & Butler, 2011).

This feature was inspired by two seminal ITSs who employed Self-Regulation Learning (SRL) strategies, nStudy and Metatutor (Cloude et al., 2022). The first paved the way for data mining of students' behavioural traces, and it took the *'leap of faith'* as it became available to wide audiences as a plugin for browsers; the latter, was one of the first –if not *the* first– to be authentically multimodal, having access to the learners' physiological responses by analysing Cognitive, Affective, Motivational and Metacognitive (CAMM) 'signatures' captured through sensors. Moreover, it delivered educational support using Pedagogical Assistants (PAs), i.e. *avatars*, each one specialised in a learning strategy (e.g. planning, notetaking, progress monitoring etc.). However, Metatutor did not reach the public. Its use was confined to several universities that participated in its development and the offered curriculum was restricted to the study of the human circulatory system and later it expanded to STEM. Additionally, Metatutor was aimed at young adults only.

Metatutor inspired the current study the most, for its attempt to tap into the emotional and cognitive and motivational states of learners; especially the notion of Metacognition and its critical connection to the ability of students to regulate their *own* learning may implicitly assert some SDT principles into the learning process; indeed, Azevedo et al. (2022, p. 15) commending on Cloude et al. (2022) findings that suggest that not all students motivation was positively affected by the PAs suggestion, explained that

"(…)performance-approach learners benefited from PA support, mastery-approach learners did not. Self-determination Theory (SDT) can help to explain these distinct patterns and in particular why some learners may have less positive reactions to PAs. According to SDT, when an individual's (…) need for competence or autonomy is impeded, the individual is likely to react negatively to regulation efforts, whereas when these needs are met, they are likely to react positively, given that the regulation is more internalised".

This quote may imply a direct connection between SDT and EF coexisting a *continuum;* in other words, it may be the case that when intrinsic motivation ends, the need for self-regulation begins. This study builds on this very notion, by inserting a specialised sub-

module into the Strategy and Decision-Making Module of the Cognitive Architecture. Thus, if the system identifies a failure in the intrinsic motivation of the child, it will revert to an SRL strategy to maintain learning, and it will fall back when the child engages in a learning activity intrinsically. Or, if the system identifies that a child is approaching a module extrinsically motivated because it does not enjoy it, it may use SRL strategies to enhance the child's competence by offering a more solid *scaffolding* or enhance the child's relatedness by proposing a shared activity with a knowledgeable peer. If the strategy is successful, it may enhance the child's autonomy which in turn may transcend them into a more intrinsically motivated state (Deci & Ryan, 2008; Ryan & Deci, 2017).

Following the learner on a continuum of intrinsically driven actions and actions that need self-regulation and employing the most appropriate strategies each time, may lead to improved learning outcomes, given that the learner is kept into their ZPD. Day et al. (2022, pp. 3-4) identified competence as the SDT component that is associated with self-regulation interventions the most, given that feedback and support are provided. In other words, successful self-regulation that helps a learner resolve a challenge enhances competence. Extending this notion one can assume that self-regulation might be a form of competence in an overarching form. This interplay between SDT and SR is intriguing and has inspired some recent theoretical approaches (e.g. Champ et al. (2023); Morsink et al. (2021)) that attempt to use SDT to explain some behaviours in ADHD where the absence of effective SR is salient. However, no empirical evidence exists to support or reject this.

This study attempted to measure EF (Working Memory and Inhibition Control) and Motivation by employing questionnaires and cognitive tasks. Metatutor and nStudy on the other hand, measured CAMM and behavioural traces of students while performing real studying. Having some initial input of a learner's EF and Motivational states in a domain may give some indications, as tasks have inherent value that interacts with a person's values and desires (Meece, 2023). Therefore, measurements during a performed action relevant to a domain may have more meaning. To propose meaningful measurements of motivation. Touré-Tillery and Fishbach (2014) make a distinction between *outcome-focused motivation* and *process-focused motivation.* The former refers to actions performed with a goal in mind and with a *'let's get it done'* attitude; the latter refers to actions that offer enjoyment during the time that are performed. According to this classification, outcome-focused motivation utilises constructs directed toward successfully attaining a target result, while process-focused motivation emphasizes personal perceptions, feelings, and reactions related to

engaging in the behaviour inherently. The former is concerned with effectively reaching the destination; the latter with fulfilling qualities within the 'journey' itself.

The aspiration for creating an ITS for personalised learning derives from the "2-sigma problem" (Bloom, 1984) which is nested in the inability of a human teacher to attend to the individual differences and learning abilities of each student individually especially in large classrooms, in conjunction with the inability of an average family to provide tutors for each student. However, the 2-sigma problem may stand, as long as the continued research on the subject yields significant results in favour of 1-to-1 tutoring. The benefit of personal tutoring seems to be well established; moreover, as Vanlehn (2011) implies that the common belief for the superiority of a human tutor over a computer tutor, is not validated by empirical research, providing additional motivation for the current study. An attempt to aggregate some of the main problems that hinder the learning experience of students in classrooms, to pinpoint the issues, the proposed (and possibly *any* proposed) ITS should address. The list of problems with traditional classroom learning provided gives a non-exhaustive overview, however, there may be more –or less– of them in some classroom settings. A team of researchers attempting to design an ITS may need to also identify what classrooms offer in terms of positive learning outcomes and the well-being of young students and try to transfer the knowledge. It should also be stressed again that the purpose of the proposed ITS is not by any means to replace a human teacher or tutor, but rather to assist them in settings where they are not available. An ITS may provide an interactive and intuitive means of studying at home. Also, some concerns regarding the acceptance of an ITS as a tool for learning are outlined in this study, trying to alleviate –to some extent– the inherent bias in favour of ITS in this study. Children will be the end users of the proposed system; therefore, asserting the concerns of those responsible for them, parents and teachers included, regarding its efficacy, validity, precision, and transparency is something that should be kept in mind.

This study attempts to outline some basic measurable principles of design that may promote motivation. The proposed Cognitive Architecture encompasses several theories for motivation and self-regulation; it is supposed to underpin an ITS software design, and for that purpose, it may be useful to apply design principles that comply with them. These include the Motivation, Engagement & Thriving in User Experience (METUX) model described by Peters et al. (2018) and Burnell et al. (2023) the provided scales to evaluate whether a technological tool is in accordance with SDT principles can promote the user's well-being by enhancing their competence, relatedness and autonomy. However, a point of

concern is the limited literature regarding the METUX applicability principles for HCI targeted to children. Another framework that designers and developers may need to have in mind is the ICAP (Chi & Wylie, 2014). There seems to be a significant number of studies that suggest interactive learning activities offer superior learning outcomes, followed by constructive and active, whereas passive learning has the least positive effect.

This study also employed an experimental, to examine whether EF and relative autonomy can predict Inhibition Control in elementary school students; also, whether Executive Functioning and relative autonomy can predict higher Working Memory capacity for elementary students; and finally, if a positive correlation exists between WM and EF. The first two hypotheses were not confirmed, and the third was confirmed, however yielding a very correlation. Detailed results for the experimental design are given in the relevant section of this study. However, an assumption based on Miyake and Friedman (2012) is that the maximum a study can expect when employing a minimalistic set of instruments and tasks to assess EFs and Motivation is to find a relationship between instruments and tasks, between instruments, or between tasks; to assess the construct as a whole, there may be a need for more sophisticated designs, such as batteries of tests and also advanced Natural Language Processing techniques offered by employing LLMs and extract meaning by qualitatively assess each user's input on an individualised manner.

This study again resorts to Azevedo et al. (2022) and Winne Winne (2014), in conjunction with Miyake and Friedman (2012); EFs *are* personal differences and as such any generalization should only be made for reference, which in turn will self-adjust by data gathered from each user individually, in order to create individualised experiences, outputs and outcomes. Otherwise, an ITS that remains fixed on generalised ideas, may fall into the trap of replicating the 'one size fits all' problem of traditional learning.

4.1 Limitations and suggestions for future studies

This study has several limitations that need to be addressed. The limitations of the experimental design have already been discussed in detail in the relevant section, so in this, they will be omitted.

The first limitation is that it is mostly theoretical. The current experience regarding ITSs and the failure of some successful models to scale up and become available to the broad public and widespread adoption, suggests that launching an ITS, especially for children is something that is *'more easily said than done'.* Developing such systems requires exhaustive

research, prototyping, experimentation, and funding. As such, the framework proposed here has an inherent limitation related to its theoretical nature. The study did not attempt any feasibility check nor any prototyping or even modelling for the proposed architecture and relied mostly on published work and existing models. The next step towards the realisation of this architecture may be to work on the mathematical formalisation of the modules. Then, the development of a prototype, first by creating a functional implementation on a specific domain of knowledge (e.g. math, geography, language), and after thorough iterations testing and modification to expand it to more domains. As more tools are being developed and launched, there may be many opportunities for actual implementation in the future.

Another problem that this study came across is the relatively limited amount of research for the population the proposed system is targeted. Therefore, many of the methods, metrics and tools that are described need to be redesigned and revalidated for school children between 9 and 12 years old. An effort was made to adopt general principles that might also apply to children, however, this choice was made by informed intuitions and based on the general knowledge of the researcher rather than grounded on evidence.

Additionally, this model did not adequately address ethical issues and considerations such as the privacy of data. An attempt was made to pinpoint possible caveats by enumerating some well-known, overarching concerns regarding the use of ITSs for children. However, this issue is of uttermost importance and needs to be addressed not only from a psychological but also from a legal standpoint.

4.2 Conclusive remarks

This study is a theoretical attempt to devise a Cognitive Architecture for an Intelligent Tutoring System for children, informed by the contemporary psychological constructs of the Self-Determination Theory of Motivation, Self-Regulation, and the Zone of Proximal Development theory of learning. It aspires to offer a blueprint for future designs, and by applying some basic principles, to promote the advantages of personalised learning for elementary school students. It suggests that an ITS should probably refrain from extracting knowledge from metrics and questionnaires, and rather assess the data from natural interaction with the user, using its sensors and advanced tools provided by the Large Language Models. The results from this study indicate that it is possible to benefit by inserting into the system some well-established knowledge for motivation, self-regulation, and learning.

References

Abdelshiheed, M., Barnes, T., & Chi, M. (2023). How and When: The Impact of Metacognitive Knowledge Instruction and Motivation on Transfer Across Intelligent Tutoring Systems. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-023-00371-0

Akyuz, Y. (2020). Effects of Intelligent Tutoring Systems (ITS) on Personalized Learning (PL). *Creative Education*, *11*(06), 953-978. https://doi.org/10.4236/ce.2020.116069

Al-aqbi, T. Q. A. F., Yousif Saleh, Basma Jumaa Al-juaifari, Nadia Muwafaq Abdulhassan, Lamees. (2019). The Effect of the Intelligent Tutoring Systems on the Education. *Journal of Intelligence Research*(27), 698-683. https://doi.org/10.36302/jir.v0i27.32

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A.-M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, *87*(2), 85-106. https://doi.org/10.1016/j.jecp.2003.10.002

American Psychological Association. (2020). Executive Functions. In *APA Dictionary of Psychology*. Retrieved October 19 2022, from https://dictionary.apa.org/executive-functions

Anania, J. (1982). *The effects of quality of instruction on the cognitive and affective learning of students*. The University of Chicago.

Anania, J. (1983). The influence of instructional conditions on student learning and achievement. *Evaluation in Education*, *7*(1), 1-92. https://doi.org/10.1016/0191-765x(83)90002-2

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.

Anderson, J. R., & Gluck, K. A. (2013). What role do cognitive architectures play in intelligent tutoring systems? In *Cognition and instruction* (pp. 227-261). Psychology Press. https://doi.org/10.1184/R1/6619124.v1

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends in Cognitive Sciences*, *18*(4), 177-185. https://doi.org/10.1016/j.tics.2013.12.003

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Elsevier.

Au, W. (2011). Teaching under the new Taylorism: high-stakes testing and the standardization of the 21stcentury curriculum. *Journal of Curriculum Studies*, *43*(1), 25-45. https://doi.org/10.1080/00220272.2010.521261

Azevedo, R., Bouchet, F., Duffy, M., Harley, J., Taub, M., Trevors, G., Cloude, E., Dever, D., Wiedbusch, M., Wortha, F., & Cerezo, R. (2022). Lessons Learned and Future Directions of MetaTutor: Leveraging Multichannel Data to Scaffold Self-Regulated Learning With an Intelligent Tutoring System. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.813632

Azevedo, R., Bouchet, F., Harley, J., Feyzi Behnagh, R., Trevors, G., Duffy, M., Taub, M., Pacampara, N., Agnew, L., Griscom, S., Mudrick, N., Stead, V., & Yang, W. (2012). *MetaTutor: An Intelligent Multi-Agent Tutoring System Designed to Detect, Track, Model, and Foster Self-Regulated Learning*. https://doi.org/10.13140/RG.2.1.1334.6640

Azevedo, R., & Gašević, D. (2019). Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in Human Behavior*, *96*, 207-210. https://doi.org/10.1016/j.chb.2019.03.025

Azevedo, R., Mudrick, N. V., Taub, M., & Bradbury, A. E. (2019). Self-Regulation in Computer-Assisted Learning Systems. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge Handbook of Cognition and Education* (pp. 587-618). Cambridge University Press. https://doi.org/10.1017/9781108235631.024

Azevedo, R., Taub, M., & Mudrick, N. V. (2017). Understanding and reasoning about real-time cognitive, affective, and metacognitive processes to foster self-regulation with advanced learning technologies. In *Handbook of self-regulation of learning and performance* (pp. 254-270). Routledge.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417-423. https://doi.org/10.1016/s1364-6613(00)01538-2

Baddeley, A. D. (2014). *Essentials of human memory* (Classic edition. ed.). Psychology Press, Taylor & Francis Group.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Elsevier.

Baeten, M., Kyndt, E., Struyven, K., & Dochy, F. (2010). Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review*, *5*(3), 243-260. https://doi.org/10.1016/j.edurev.2010.06.001

Bailey, T. H., & Phillips, L. J. (2016). The influence of motivation and adaptation on students' subjective well-being, meaning in life and academic performance. *Higher Education Research & Development*, *35*(2), 201-216. https://doi.org/10.1080/07294360.2015.1087474

Baker, R. S. (2016). Stupid Tutoring Systems, Intelligent Humans. *International Journal of Artificial Intelligence in Education*, *26*(2), 600-614. https://doi.org/10.1007/s40593-016-0105-0

Ballard, J., & Butler, P. (2011). Personalised Learning: Developing a Vygotskian Framework for E-learning [Article]. *International Journal of Technology, Knowledge & Society*, *7*(2), 21-36. https://doi.org/10.18848/1832-3669/CGP/v07i02/56198

Bannert, M. (2009). Promoting Self-Regulated Learning Through Prompts. *Zeitschrift für Pädagogische Psychologie*, *23*(2), 139-145. https://doi.org/10.1024/1010-0652.23.2.139

Barnes, M., Clarke, D., & Stephens, M. (2000). Assessment: The engine of systemic curricular reform? *Journal of Curriculum Studies*, *32*(5), 623-650. https://doi.org/10.1080/00220270050116923

Berlin, L., & Bohlin, G. (2002). Response Inhibition, Hyperactivity, and Conduct Problems Among Preschool Children. *Journal of Clinical Child & Adolescent Psychology*, *31*(2), 242-251. https://doi.org/10.1207/S15374424JCCP3102_09

Best, J. R., & Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Development*, *81*(6), 1641-1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bienkowski, M., Feng, M., & Means, B. (2014). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief*.

Billings, E., & Walqui, A. (2018). The zone of proximal development: An affirmative perspective in teaching ELLs/MLLs. *Retrieved August*, *6*.

Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, *13*(6), 4-16. https://doi.org/10.3102/0013189x013006004

Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). *Open learner models and learning analytics dashboards* Proceedings of the 8th International Conference on Learning Analytics and Knowledge,

Bosworth, R. (2011). Class size, class composition, and the distribution of student achievement. *Education Economics*, *22*(2), 141-165. https://doi.org/10.1080/09645292.2011.568698

Boyer, K. E., Phillips, R., Wallis, M., Vouk, M., & Lester, J. (2008). Balancing Cognitive and Motivational Scaffolding in Tutorial Dialogue. In (pp. 239-249). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-69132-7_28

Brunetti, R., Del Gatto, C., & Delogu, F. (2014). eCorsi: implementation and testing of the Corsi block-tapping task for digital tablets. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00939

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

Burke, A. J. (1980). *Students' potential for learning contrasted under tutorial and group approaches to instruction*. The University of Chicago.

Burnell, R., Peters, D., Ryan, R. M., & Calvo, R. A. (2023). Technology evaluations are associated with psychological need satisfaction across different spheres of experience: an application of the METUX scales. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1092288

Camerota, M., Willoughby, M. T., Kuhn, L. J., & Blair, C. B. (2018). The Childhood Executive Functioning Inventory (CHEXI): Factor structure, measurement invariance, and correlates in US preschoolers. *Child Neuropsychology*, *24*(3), 322-337. https://doi.org/10.1080/09297049.2016.1247795

Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development*, *22*(4), 489-510. https://doi.org/10.1016/j.cogdev.2007.08.002

Casey, B., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*(3), 104-110. https://doi.org/10.1016/j.tics.2005.01.011

Casey, B. J., Galvan, A., & Hare, T. A. (2005). Changes in cerebral functional organization during cognitive development. *Current Opinion in Neurobiology*, *15*(2), 239-244. https://doi.org/10.1016/j.conb.2005.03.012

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, 107468. https://doi.org/10.1016/j.chb.2022.107468

Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. *Vygotsky's educational theory in cultural context*, *1*(2), 39-64.

Champ, R. E., Adamou, M., & Tolchard, B. (2023). Seeking connection, autonomy, and emotional feedback: A self-determination theory of self-regulation in attention-deficit hyperactivity disorder. *Psychological Review*, *130*(3), 569-603. https://doi.org/10.1037/rev0000398

Chauncey, S. A., & McKenna, H. P. (2023). A framework and exemplars for ethical and responsible use of AI Chatbot technology to support teaching and learning. *Computers and Education: Artificial Intelligence*, *5*, 100182. https://doi.org/10.1016/j.caeai.2023.100182

Chen, B., Zhu, X., & Díaz del Castillo H, F. (2023). Integrating generative AI in knowledge building. *Computers and Education: Artificial Intelligence*, *5*, 100184. https://doi.org/10.1016/j.caeai.2023.100184

Chi, M. T. H. (2009). Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science*, *1*(1), 73-105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T. H., Roy, M., & Hausmann, R. G. M. (2008). Observing Tutorial Dialogues Collaboratively: Insights About Human Tutoring Effectiveness From Vicarious Learning. *Cognitive Science*, *32*(2), 301-341. https://doi.org/10.1080/03640210701863396

Chi, M. T. H., Siler, S. A., & Jeong, H. (2004). Can Tutors Monitor Students' Understanding Accurately? *Cognition and Instruction*, *22*(3), 363-387. https://doi.org/10.1207/s1532690xci2203_4

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471-533. https://doi.org/10.1207/s15516709cog2504_1

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219-243. https://doi.org/10.1080/00461520.2014.965823

Chiu, T. K. F. (2021). Digital support for student engagement in blended learning based on self-determination theory. *Computers in Human Behavior*, *124*, 106909. https://doi.org/10.1016/j.chb.2021.106909

Chiu, T. K. F., & Chai, C.-S. (2020). Sustainable Curriculum Planning for Artificial Intelligence Education: A Self-Determination Theory Perspective. *Sustainability*, *12*(14), 5568. https://doi.org/10.3390/su12145568

Chounta, I.-A., Albacete, P., Jordan, P., Katz, S., & McLaren, B. (2017). *The "Grey Area": A Computational Approach to Model the Zone of Proximal Development*. https://doi.org/10.1007/978-3-319-66610-5_1

Cloude, E. B., Azevedo, R., Winne, P. H., Biswas, G., & Jang, E. E. (2022). System design for using multimodal trace data in modeling self-regulated learning [Original Research]. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.928632

Cloude, E. B., Wortha, F., Dever, D. A., & Azevedo, R. (2021). Negative emotional dynamics shape cognition and performance with MetaTutor: Toward building affect-aware systems. 2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII),

Conway-Smith, B., & West, R. (2022). *System-1 and System-2 realized within the Common Model of Cognition*.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, *88*(4), 715-730. https://doi.org/10.1037/0022-0663.88.4.715

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain.

Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, *26*(2), 197-223. https://doi.org/10.1007/s10648-013-9246-y

Criaud, M., & Boulinguez, P. (2013). Have we been asking the right questions when assessing response inhibition in go/no-go tasks with fMRI? A meta-analysis and critical review. *Neuroscience & Biobehavioral Reviews*, *37*(1), 11-23. https://doi.org/10.1016/j.neubiorev.2012.11.003

Darling-Hammond, L., Schachner, A., & Edgerton, A. K. (2020). Restarting and Reinventing School: Learning in the Time of COVID and Beyond. *Learning Policy Institute*.

Day, N., Paas, F., Kervin, L., & Howard, S. J. (2022). A Systematic Scoping Review of Pre-School Self-Regulation Interventions from a Self-Determination Theory Perspective. *International Journal of Environmental Research and Public Health*, *19*(4), 2454. https://doi.org/10.3390/ijerph19042454

De Jong, C. G. W., Van De Voorde, S., Roeyers, H., Raymaekers, R., Oosterlaan, J., & Sergeant, J. A. (2009). How Distinctive are ADHD and RD? Results of a Double Dissociation Study. *Journal of Abnormal Child Psychology*, *37*(7), 1007-1017. https://doi.org/10.1007/s10802-009-9328-y

De Vreede, T., Raghavan, M., & De Vreede, G.-J. (2021). Design Foundations for AI Assisted Decision Making: A Self Determination Theory Approach.

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne*, *49*(3), 182.

Dede, C., Honan, J. P., Peters, L., & Harvard University. Graduate School of, E. (2005). *Scaling up success : lessons learned from technology-based educational improvement* (1st ed.). Jossey-Bass. http://www.loc.gov/catdir/toc/ecip0422/2004019590.html

Dever, D. A., Sonnenfeld, N. A., Wiedbusch, M. D., & Azevedo, R. (2022). Pedagogical Agent Support and Its Relationship to Learners' Self-regulated Learning Strategy Use with an Intelligent Tutoring System. In (pp. 332-343). Springer International Publishing. https://doi.org/10.1007/978-3-031-11644-5_27

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Dietrichson, J., Bøg, M., Filges, T., & Klint Jørgensen, A.-M. (2017). Academic Interventions for Elementary and Middle School Students With Low Socioeconomic Status: A Systematic Review and Meta-Analysis. *Review of Educational Research*, *87*(2), 243-282. https://doi.org/10.3102/0034654316687036

Dimov, C., Khader, P. H., Marewski, J. N., & Pachur, T. (2019). How to model the neurocognitive dynamics of decision making: A methodological primer with ACT-R. *Behavior Research Methods*, *52*(2), 857-880. https://doi.org/10.3758/s13428-019-01286-2

Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated Instruction, Professional Development, and Teacher Efficacy. *Journal for the Education of the Gifted*, *37*(2), 111-127. https://doi.org/10.1177/0162353214529042

Dole, S. F. (2017). Creating cultures of thinking: the 8 forces we must master to truly transform our schools. *Interdisciplinary Journal of Problem-Based Learning*, *11*(2), 13.

Durston, S., Thomas, K. M., Worden, M. S., Yang, Y., & Casey, B. J. (2002). The Effect of Preceding Context on Inhibition: An Event-Related fMRI Study. *NeuroImage*, *16*(2), 449-453. https://doi.org/10.1006/nimg.2002.1074

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., . . . Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Ferguson, C., van den Broek, E. L., & van Oostendorp, H. (2022). AI-Induced guidance: Preserving the optimal Zone of Proximal Development. *Computers and Education: Artificial Intelligence*, *3*, 100089. https://doi.org/10.1016/j.caeai.2022.100089

Ferster, B. (2022). Intelligent Tutoring Systems. In *Intelligent Tutoring Systems*. https://doi.org/10.4324/9781138609877-ree6-1

Field, A. (2018). *Discovering statistics using IBM SPSS statistics 5th ed*. Sage.

Filius, R. M., de Kleijn, R. A. M., Uijl, S. G., Prins, F. J., van Rijen, H. V. M., & Grobbee, D. E. (2018). Strengthening dialogic peer feedback aiming for deep learning in SPOCs. *Computers & Education*, *125*, 86-100. https://doi.org/10.1016/j.compedu.2018.06.004

Firdaus, M. F., Wibawa, J. N., & Rahman, F. F. (2023). Utilization of GPT-4 to Improve Education Quality Through Personalized Learning for Generation Z in Indonesia. *IT for Society*, *8*(1).

Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-Based Assessment and Instructional Change: The Effects of Testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, *20*(2). https://doi.org/10.3102/01623737020002095

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410-8415. https://doi.org/10.1073/pnas.1319030111

Freund, P. A., & Lohbeck, A. (2021). Modeling Self-Determination Theory Motivation Data by Using Unfolding IRT. *European Journal of Psychological Assessment*, *37*(5), 388-396. https://doi.org/10.1027/1015-5759/a000629

García-Peñalvo, F. J., & Reimann, D. (2016). An overview of the most relevant literature on coding and computational thinking with emphasis on the relevant issues for teachers.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31-60. https://doi.org/10.1037/0033-2909.134.1.31

Georgiou, G. K., & Zhang, L. (2023). Do Performance-Based Measures and Behavioral Ratings of Executive Functioning Complement Each Other in Predicting Reading and Mathematics in Chinese? *Behavioral Sciences*, *13*(10), 823. https://www.mdpi.com/2076-328X/13/10/823

https://mdpi-res.com/d_attachment/behavsci/behavsci-13-00823/article_deploy/behavsci-13-00823-v2.pdf?version=1696838642

Gibson, D., Kovanovic, V., Ifenthaler, D., Dexter, S., & Feng, S. (2023). Learning theories for artificial intelligence promoting learning processes. *British Journal of Educational Technology*, *54*(5), 1125-1146. https://doi.org/10.1111/bjet.13341

Graesser, A., Hu, X., & Sottilare, R. (2018). Intelligent Tutoring Systems. In *International Handbook of the Learning Sciences*. Routledge. https://doi.org/10.4324/9781315617572-24

Graesser, A. C. (2016). Conversations with AutoTutor Help Students Learn. *International Journal of Artificial Intelligence in Education*, *26*(1), 124-132. https://doi.org/10.1007/s40593-015-0086-4

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An Intelligent Tutoring System With Mixed-Initiative Dialogue. *IEEE Transactions on Education*, *48*(4), 612-618. https://doi.org/10.1109/te.2005.856149

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, *9*(6), 495-522. https://doi.org/10.1002/acp.2350090604

Greene, J. A. (2017). Handbook of self-regulation of learning and performance.

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology*, *52*(5), 890-898. https://doi.org/10.1037/0022-3514.52.5.890

Gros, B. (2016). The design of smart educational environments. *Smart Learning Environments*, *3*(1). https://doi.org/10.1186/s40561-016-0039-x

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments Ecosystem: Building a Platform that Brings Scientists and Teachers Together for Minimally Invasive Research on Human Learning and Teaching. *International Journal of Artificial Intelligence in Education*, *24*(4), 470-497. https://doi.org/10.1007/s40593-014-0024-x

Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: a review and synthesis. *Psychol Bull*, *128*(5), 774-795. https://doi.org/10.1037/0033-2909.128.5.774

Howard, S. J., Johnson, J., & Pascual-Leone, J. (2014). Clarifying inhibitory control: Diversity and development of attentional inhibition. *Cognitive Development*, *31*, 1-21. https://doi.org/10.1016/j.cogdev.2014.03.001

Howard, S. J., & Okely, A. D. (2015). Catching Fish and Avoiding Sharks. *Journal of Psychoeducational Assessment*, *33*(6), 585-596. https://doi.org/10.1177/0734282914562933

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017-2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Hutchinson, A. D., Mathias, J. L., & Banich, M. T. (2008). Corpus callosum morphology in children and adolescents with attention deficit hyperactivity disorder: A meta-analytic review. *Neuropsychology*, *22*(3), 341-349. https://doi.org/10.1037/0894-4105.22.3.341

Jarrold, C., Li, K., & Wang, T. (2023). A novel approach to measuring the developmental interactions between working memory and inhibitory control in young children. *Cognitive Development*, *67*, 101362. https://doi.org/10.1016/j.cogdev.2023.101362

Jennings, P. A., & Greenberg, M. T. (2009). The Prosocial Classroom: Teacher Social and Emotional Competence in Relation to Student and Classroom Outcomes. *Review of Educational Research*, *79*(1), 491-525. https://doi.org/10.3102/0034654308325693

Jonides, J. (1981). *Attention and performance IX*. Erlbaum Hillsdale.

Kahl, D. H., & Venette, S. (2010). To Lecture or Let Go: A Comparative Analysis of Student Speech Outlines from Teacher-Centered and Learner-Centered Classrooms. *Communication Teacher*, *24*(3), 178-186. https://doi.org/10.1080/17404622.2010.490232

Karoly, P. (1993). Mechanisms of Self-Regulation: A Systems View. *Annual Review of Psychology*, *44*(1), 23-52. https://doi.org/10.1146/annurev.ps.44.020193.000323

Kessels, R. P. C., van den Berg, E., Ruis, C., & Brands, A. M. A. (2008). The Backward Span of the Corsi Block-Tapping Task and Its Association With the WAIS-III Digit Span. *Assessment*, *15*(4), 426-434. https://doi.org/10.1177/1073191108315611

Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi Block-Tapping Task: Standardization and Normative Data. *Applied Neuropsychology*, *7*(4), 252-258. https://doi.org/10.1207/S15324826AN0704_8

Khan, S. (2023). *How AI Could Save (Not Destroy) Education*. https://www.youtube.com/watch?v=hJP5GqnTrNo

Kim, N. J., & Kim, M. K. (2022). Teacher's Perceptions of Using an Artificial Intelligence-Based Educational Tool for Scientific Writing. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.755914

Knee, C. R., & Browne, L. (2023). 160C7Relationships Motivation Theory. In R. M. Ryan (Ed.), *The Oxford Handbook of Self-Determination Theory* (pp. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197600047.013.8

Knight, S., & Littleton, K. (2017). Socialising Epistemic Cognition. *Educational Research Review*, *21*, 17-32. https://doi.org/10.1016/j.edurev.2017.02.003

Kotseruba, I., & Tsotsos, J. K. (2020). 40 years of cognitive architectures: core cognitive abilities and practical applications. *Artificial Intelligence Review*, *53*(1), 17-94. https://doi.org/10.1007/s10462-018-9646-y

Kowalczuk, Z., & Czubenko, M. (2016). Computational Approaches to Modeling Artificial Emotion – An Overview of the Proposed Solutions [Review]. *Frontiers in Robotics and AI*, *3*. https://doi.org/10.3389/frobt.2016.00021 (Frontiers in Robotics and AI)

Kshetri, N. (2023). The Economics of Generative Artificial Intelligence in the Academic Industry. *Computer*, *56*(8), 77-83. https://doi.org/10.1109/MC.2023.3278089

Kucirkova, N., & Leaton Gray, S. (2023). Beyond Personalization: Embracing Democratic Learning Within Artificially Intelligent Systems. *Educational Theory*, *73*(4), 469-489. https://doi.org/10.1111/edth.12590

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring Systems:A Meta-Analytic Review. *Review of Educational Research*, *86*(1), 42-78. https://doi.org/10.3102/0034654315581420

Kurni, M., Mohammed, M. S., & Srinivasa, K. G. (2023). Intelligent Tutoring Systems. In M. Kurni, M. S. Mohammed, & S. K G (Eds.), *A Beginner's Guide to Introduce Artificial Intelligence in Teaching and Learning* (pp. 29-44). Springer International Publishing. https://doi.org/10.1007/978-3-031-32653-0_2

Labaree, D. F. (2011). *Someone has to fail: The zero-sum game of public schooling*. Harvard University Press.

Laird, J. E. (2019). *The cognitive architecture*. MIT press.

Laird, J. E. (2022). An Analysis and Comparison of ACT-R and Soar. *ArXiv*, *abs/2201.09305*.

Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. *AI Magazine*, *38*(4), 13-26. https://doi.org/10.1609/aimag.v38i4.2744

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. *Artificial Intelligence*, *33*(1), 1-64. https://doi.org/10.1016/0004-3702(87)90050-6

Lajoie, S. P., Derry, S. J., & Lajoie, S. P. (1993). *Computers as cognitive tools*. L. Erlbaum Associates. https://locate.coventry.ac.uk/openurl/COV/COV_VU1?u.ignore_date_coverage=true&rft.mms_id=996921562602011

https://www.taylorfrancis.com/books/9781136475474

Lambert, J., & Stevens, M. (2023). ChatGPT and Generative AI Technology: A Mixed Bag of Concerns and New Opportunities. *Computers in the Schools*, 1-25. https://doi.org/10.1080/07380569.2023.2256710

Lantolf, J. P., & Poehner, M. E. (2010). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research*, *15*(1), 11-33. https://doi.org/10.1177/1362168810383328

Lefevre, D., & Cox, B. (2016). Delayed instructional feedback may be more effective, but is this contrary to learners' preferences? *British Journal of Educational Technology*, *48*(6), 1357-1367. https://doi.org/10.1111/bjet.12495

Lieto, A. (2021). *Cognitive design for artificial minds*. Routledge.

Littman, R., & Takács, Á. (2017). Do all inhibitions act alike? A study of go/no-go and stop-signal paradigms. *PLoS One*, *12*(10), e0186774. https://doi.org/10.1371/journal.pone.0186774

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75-98. https://doi.org/10.1037/0033-2909.116.1.75

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: a model and a method. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(2), 276.

Luckin, R., & Holmes, W. (2016). Intelligence Unleashed: An argument for AI in Education.

Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the Development of Cognitive Control through Adolescence? *Brain and Cognition*, *72*(1), 101-113. https://doi.org/10.1016/j.bandc.2009.08.005

Maldonado, T., Orr, J. M., Goen, J. R. M., & Bernard, J. A. (2020). Age Differences in the Subcomponents of Executive Functioning. *The Journals of Gerontology: Series B*, *75*(6), e31-e55. https://doi.org/10.1093/geronb/gbaa005

Malloy-Diniz, L. F., Miranda, D. M., & Grassi-Oliveira, R. (2017). Editorial: Executive Functions in Psychiatric Disorders. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01461

Margari, L., Craig, F., Margari, F., Legrottaglie, A., Palumbi, R., & De Giambattista, C. (2016). A review of executive function deficits in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Neuropsychiatric Disease and Treatment*. https://doi.org/10.2147/ndt.S104620

Margolis, A. A. (2020). Zone of Proximal Development, Scaffolding and Teaching Practice. *Cultural-Historical Psychology*, *16*(3), 15-26. https://doi.org/10.17759/chp.2020160303

Marwan, S., Akram, B., Barnes, T., & Price, T. W. (2022). Adaptive Immediate Feedback for Block-Based Programming: Design and Evaluation. *IEEE Transactions on Learning Technologies*, *15*(3), 406-420. https://doi.org/10.1109/TLT.2022.3180984

Marzouk, Z. (2016). Generating Learning Analytics to Improve Learners' Metacognitive Skills Using nStudy Trace Data and the ICAP Framework. International Conference on Learning Analytics & Knowledge, Edinburgh.

Matzke, D., Verbruggen, F., & Logan, G. (2018). The stop-signal paradigm. *Stevens´handbook of experimental psychology and cognitive neuroscience*, *5*, 383-427.

McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring Techniques in Algebra. *Cognition and Instruction*, *7*(3), 197-244. https://doi.org/10.1207/s1532690xci0703_2

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition*, *39*(3), 462-476. https://doi.org/10.3758/s13421-010-0035-2

Meece, J. L. (2023). The role of motivation in self-regulated learning. In *Self-regulation of learning and performance* (pp. 25-44). Routledge.

Meltzer, L. (2018). *Executive function in education: From theory to practice*.

Menekse, M., Stump, G. S., Krause, S., & Chi, M. T. H. (2013). Differentiated Overt Learning Activities for Effective Instruction in Engineering Classrooms. *Journal of Engineering Education*, *102*(3), 346-374. https://doi.org/10.1002/jee.20021

Mercader-Rubio, I., Ángel, N. G., Ruiz, N. F. O., & Carrión-Martínez, J. J. (2022). Emotional intelligence as a predictor of identified regulation, introjected regulation, and external regulation in athletes. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.1003596

Merrill, D. C., Reiser, B. J., Merrill, S. K., & Landes, S. (1995). Tutoring: Guided Learning by Doing. *Cognition and Instruction*, *13*(3), 315-372. https://doi.org/10.1207/s1532690xci1303_1

Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective Tutoring Techniques: A Comparison of Human Tutors and Intelligent Tutoring Systems. *Journal of the Learning Sciences*, *2*(3), 277-305. https://doi.org/10.1207/s15327809jls0203_2

Miller, R. (2011). *Vygotsky in perspective*. Cambridge university press.

Minsky, M. (1988). *Society of mind*. Simon and Schuster.

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions. *Current Directions in Psychological Science*, *21*(1), 8-14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Montoya, M. F., Susperreguy, M. I., Dinarte, L., Morrison, F. J., San Martín, E., Rojas-Barahona, C. A., & Förster, C. E. (2019). Executive function in Chilean preschool children: Do short-term memory, working memory, and response inhibition contribute differentially to early academic skills? *Early Childhood Research Quarterly*, *46*, 187-200. https://doi.org/10.1016/j.ecresq.2018.02.009

Morsink, S., Oord, S., Antrop, I., Danckaerts, M., & Scheres, A. (2021). Studying Motivation in ADHD: The Role of Internal Motives and the Relevance of Self Determination Theory. *Journal of Attention Disorders*, *26*. https://doi.org/10.1177/10870547211050948

Müller, U., Kerns, K. A., & Konkin, K. (2012). Test–Retest Reliability and Practice Effects of Executive Function Tasks in Preschool Children. *The Clinical Neuropsychologist*, *26*(2), 271-287. https://doi.org/10.1080/13854046.2011.645558

Muranga, K., Muse, I. S., Köroğlu, E. N., & Yildirim, Y. (2023). Artificial Intelligence and Underfunded Education. *London Journal of Social Sciences*(6), 56-68. https://doi.org/10.31039/ljss.2023.6.105

Murray, T., & Arroyo, I. (2002). *Toward Measuring and Maintaining the Zone of Proximal Development in Adaptive Instructional Systems*.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working Memory Deficits in Poor Comprehenders Reflect Underlying Language Impairments. *Journal of Experimental Child Psychology*, *73*(2), 139-158. https://doi.org/10.1006/jecp.1999.2498

Nigg, J. T. (2001). Is ADHD a disinhibitory disorder? *Psychological Bulletin*, *127*(5), 571-598. https://doi.org/10.1037/0033-2909.127.5.571

Nigg, J. T. (2017). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *58*(4), 361-383. https://doi.org/10.1111/jcpp.12675

Nye, B. D. (2015). Intelligent Tutoring Systems by and for the Developing World: A Review of Trends and Approaches for Educational Technology in a Global Context. *International Journal of Artificial Intelligence in Education*, *25*(2), 177-203. https://doi.org/10.1007/s40593-014-0028-6

Ohlsson, S., Di Eugenio, B., Chow, B., Fossati, D., Lu, X., & Kershaw, T. (2007). *Beyond the code-and-count analysis of tutoring dialogues*.

OpenAi. (2023). GPT-4 Technical Report. *arXiv pre-print server*. https://doi.org/arxiv:2303.08774

Ostrow, K., Donnelly, C., Adjei, S., & Heffernan, N. (2015). *Improving Student Modeling Through Partial Credit and Problem Difficulty* Proceedings of the Second (2015) ACM Conference on Learning @ Scale,

Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist - EDUC PSYCHOL*, *38*, 63-71. https://doi.org/10.1207/S15326985EP3801_8

Padmanabhan, A., Geier, C. F., Ordaz, S. J., Teslovich, T., & Luna, B. (2011). Developmental changes in brain function underlying the influence of reward processing on inhibitory control. *Developmental Cognitive Neuroscience*, *1*(4), 517-529. https://doi.org/10.1016/j.dcn.2011.06.004

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2016). Effectiveness of Cognitive Tutor Algebra I at Scale. *Educational Evaluation and Policy Analysis*, *36*(2), 127-144. https://doi.org/10.3102/0162373713507480

Pappas, I. O., & Giannakos, M. N. (2021). Rethinking Learning Design in IT Education During a Pandemic. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.652856

Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2017). Teachers' Instructional Adaptations: A Research Synthesis. *Review of Educational Research*, *88*(2), 205-242. https://doi.org/10.3102/0034654317743198

Pauli-Pott, U., Dalir, S., Mingebach, T., Roller, A., & Becker, K. (2014). Attention deficit/hyperactivity and comorbid symptoms in preschoolers: Differences between subgroups in neuropsychological basic deficits. *Child Neuropsychology*, *20*(2), 230-244. https://doi.org/10.1080/09297049.2013.778236

Peters, D., Calvo, R. A., & Ryan, R. M. (2018). Designing for Motivation, Engagement and Wellbeing in Digital Experience. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00797

Poehner, M. E., & Infante, P. (2017). Mediated Development: A Vygotskian Approach to Transforming Second Language Learner Abilities. *TESOL Quarterly*, *51*(2), 332-357. https://doi.org/10.1002/tesq.308

Popham, W. (2001). Teaching to the Test? *Educational Leadership*, *58*, 16-20.

Porsdam Mann, S., Earp, B. D., Møller, N., Vynn, S., & Savulescu, J. (2023). AUTOGEN: A Personalized Large Language Model for Academic Enhancement—Ethics and Proof of Principle. *The American Journal of Bioethics*, 1-14. https://doi.org/10.1080/15265161.2023.2233356

Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3-25. https://doi.org/10.1080/00335558008248231

Pratama, M., Lu, J., Lughofer, E., Zhang, G., & Anavatti, S. (2016). Scaffolding type-2 classifier for incremental learning under concept drifts. *Neurocomputing*, *191*, 304-329. https://doi.org/10.1016/j.neucom.2016.01.049

Putnam, R. T. (1987). Structuring and Adjusting Content for Students: A Study of Live and Simulated Tutoring of Addition. *American Educational Research Journal*, *24*(1), 13-48. https://doi.org/10.3102/00028312024001013

Reeve, J. (2002). Self-determination theory applied to educational settings. *Handbook of self-determination research*, *2*, 183-204.

Reeve, J., Ryan, R., Deci, E., & Jang, H. (2008). Understanding and promoting autonomous self-regulation: A self-determination theory perspective. *Motivation and self-regulated learning: Theory, research, and applications*, 223-244.

Reich, J., Buttimer, C. J., Coleman, D., Colwell, R. D., Faruqi, F., & Larke, L. R. (2020). *What's Lost, What's Left, What's Next: Lessons Learned from the Lived Experiences of Teachers during the 2020 Novel Coronavirus Pandemic*. Center for Open Science. https://dx.doi.org/10.35542/osf.io/8exp9

Reznick, J. S., Morrow, J. D., Goldman, B. D., & Snyder, J. (2010). The Onset of Working Memory in Infants. *Infancy*, *6*(1), 145-154. https://doi.org/10.1207/s15327078in0601_7

Ricard, N. C., & Pelletier, L. G. (2016). Dropping out of high school: The role of parent and teacher self-determination support, reciprocal friendships and academic motivation. *Contemporary Educational Psychology*, *44-45*, 32-40. https://doi.org/10.1016/j.cedpsych.2015.12.003

Riecken, H. W., Boruch, R. F., Campbell, D. T., Caplan, N., Glennan, T. K., Pratt, J. W., Rees, A., & Williams, W. (1974). The Purposes of Social Experimentation. *Educational Researcher*, *3*(11). https://doi.org/10.3102/0013189x003011005

Riggs, N. R., Spruijt-Metz, D., Sakuma, K.-L., Chou, C.-P., & Pentz, M. A. (2010). Executive Cognitive Function and Food Intake in Children. *Journal of Nutrition Education and Behavior*, *42*(6), 398-403. https://doi.org/10.1016/j.jneb.2009.11.003

Ritter, F. E., Tehranchi, F., & Oury, J. D. (2018). ACT-R: A cognitive architecture for modeling cognition. *WIREs Cognitive Science*, *10*(3). https://doi.org/10.1002/wcs.1488

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive Tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review*, *14*(2), 249-255. https://doi.org/10.3758/bf03194060

Robinson, K., & Aronica, L. (2018). *You, your child, and school : navigate your way to the best education*. Viking,.

Roll, I., & Wylie, R. (2016). Evolution and Revolution in Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*, *26*(2), 582-599. https://doi.org/10.1007/s40593-016-0110-3

Rothbart, M. K., & Rueda, M. R. (2005). The Development of Effortful Control. In *Developing individuality in the human brain: A tribute to Michael I. Posner.* (pp. 167-188). https://doi.org/10.1037/11108-009

Rotherham, A. J., & Willingham, D. T. (2010). 21st-century" skills. *American educator*, *17*(1), 17-20.

Rubia, K., Smith, A. B., Taylor, E., & Brammer, M. (2007). Linear age-correlated functional development of right inferior fronto-striato-cerebellar networks during response inhibition and anterior cingulate during error-related processes. *Human Brain Mapping*, *28*(11), 1163-1177. https://doi.org/10.1002/hbm.20347

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The Development of Executive Attention: Contributions to the Emergence of Self-Regulation. *Developmental Neuropsychology*, *28*(2), 573-594. https://doi.org/10.1207/s15326942dn2802_2

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54-67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory : basic psychological needs in motivation, development, and wellness*. Guilford Press.

Ryan, R. M., & Weinstein, N. (2009). Undermining quality teaching and learning. *Theory and Research in Education*, *7*(2), 224-233. https://doi.org/10.1177/1477878509104327

Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change*, *11*(1), 45-61. https://doi.org/10.1007/s10833-008-9098-2

Salomon, G. (1997). *Distributed cognitions: Psychological and educational considerations*. Cambridge University Press.

San Pedro, M. O. Z., Baker, R. S., Heffernan, N. T., & Ocumpaugh, J. L. (2015). *Exploring college major choice and middle school student behavior, affect and learning* Proceedings of the Fifth International Conference on Learning Analytics And Knowledge,

Santangelo, T., & Tomlinson, C. A. (2012). Teacher Educators' Perceptions and Use of Differentiated Instruction Practices: An Exploratory Investigation. *Action in Teacher Education*, *34*(4), 309-327. https://doi.org/10.1080/01626620.2012.717032

Sawyer, R. K. (2022). *The Cambridge handbook of the learning sciences* (Third Edition. ed.). Cambridge University Press,.

Schachar, R., Logan, G. D., Robaey, P., Chen, S., Ickowicz, A., & Barr, C. (2007). Restraint and Cancellation: Multiple Inhibition Deficits in Attention Deficit Hyperactivity Disorder. *Journal of Abnormal Child Psychology*, *35*(2), 229-238. https://doi.org/10.1007/s10802-006-9075-2

Schanzenbach, D. W. (2014). Does class size matter? *Boulder, CO: National Education Policy Center.* . http://nepc.colorado.edu/publication/does-class-size-matter.

Schirmbeck, K., Rao, N., & Maehler, C. (2020). Similarities and differences across countries in the development of executive functions in children: A systematic review. *Infant and Child Development*, *29*(1). https://doi.org/10.1002/icd.2164

Schloß, S., Derz, F., Schurek, P., Cosan, A. S., Becker, K., & Pauli-Pott, U. (2021). Reward-Related Dysfunctions in Children Developing Attention Deficit Hyperactivity Disorder—Roles of Oppositional and Callous-Unemotional Symptoms. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.738368

Scott, S., & Palincsar, A. (2013). *Sociocultural theory*.

Shabani, K. (2016). Applications of Vygotsky's sociocultural approach for teachers' professional development. *Cogent Education*, *3*(1), 1252177. https://doi.org/10.1080/2331186x.2016.1252177

Sharples, M. (2023). Towards social generative AI for education: theory, practices and ethics. arXiv:2306.10063. https://ui.adsabs.harvard.edu/abs/2023arXiv230610063S https://arxiv.org/pdf/2306.10063.pdf

Shayer, M. (2003). Not just Piaget; not just Vygotsky, and certainly not Vygotsky as alternative to Piaget. *Learning and Instruction*, *13*(5), 465-485. https://doi.org/10.1016/s0959-4752(03)00092-6

Shute, V. J., & Zapata-Rivera, D. (2012). Adaptive Educational Systems. In *Adaptive Technologies for Training and Education* (pp. 7-27). https://doi.org/10.1017/cbo9781139049580.004

Siddi, S., Preti, A., Lara, E., Brébion, G., Vila, R., Iglesias, M., Cuevas-Esteban, J., López-Carrilero, R., Butjosa, A., & Haro, J. M. (2020). Comparison of the touch-screen and traditional versions of the Corsi block-tapping test in patients with psychosis and healthy controls. *BMC Psychiatry*, *20*(1). https://doi.org/10.1186/s12888-020-02716-8

Siler, S. A. (2004). *Does tutors' use of their knowledge of their students enhance one-to-one tutoring?* University of Pittsburgh].

Simonds, J., Kieras, J. E., Rueda, M. R., & Rothbart, M. K. (2007). Effortful control, executive attention, and emotional regulation in 7–10-year-old children. *Cognitive Development*, *22*(4), 474-488. https://doi.org/10.1016/j.cogdev.2007.08.009

Simpson, A., & Riggs, K. J. (2006). Conditions under which children experience inhibitory difficulty with a " button-press" go/no-go task. *Journal of Experimental Child Psychology*, *94*(1), 18-26. https://doi.org/10.1016/j.jecp.2005.10.003

Sinatra, A. M. (2018). Leveraging Cognitive Psychology Principles to Enhance Adaptive Instruction. In *Augmented Cognition: Intelligent Technologies* (pp. 69-77). https://doi.org/10.1007/978-3-319-91470-1_7

Singh, S. V. (2023). *The Prospects for Advancing Adaptive Learning Technology through AI Methods* 2023 Future of Educational Innovation-Workshop Series Data in Action, https://ieeexplore.ieee.org/document/10105111/

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review*, *24*, 86-97.

Sleeman, D., Kelly, A. E., Martinak, R., Ward, R. D., & Moore, J. L. (1989). Studies of Diagnosis and Remediation with High School Algebra Students. *Cognitive Science*, *13*(4), 551-568. https://doi.org/10.1207/s15516709cog1304_3

Smagorinsky, P. (2013). What does Vygotsky provide for the 21st-century language arts teacher? *Language Arts*, *90*(3), 192-204.

Smit, K., De Brabander, C. J., & Martens, R. L. (2014). Student-centred and teacher-centred learning environment in pre-vocational secondary education: Psychological needs, and motivation. *Scandinavian Journal of Educational Research*, *58*(6), 695-712. https://doi.org/10.1080/00313831.2013.821090

Sofologi, M., Kamari, A., Kougioumtzis, G., Tatsiopoulou, P., Bonti, E., Papadimitropoulou, P., & Sarris, D. (2023). Emphasizing the Developmental Trajectory of Executive Functions in Preschool, School Children, Adolescence, and Its Relations With School Performance. In *Developmental Language Disorders in Childhood and Adolescence* (pp. 224-238). https://doi.org/10.4018/979-8-3693-0644-4.ch017

Sofologi, M., Papantoniou, G., Avgita, T., Lyraki, A., Thomaidou, C., Zaragas, H., Ntritsos, G., Varsamis, P., Staikopoulos, K., Kougioumtzis, G., Papantoniou, A., & Moraitou, D. (2022). The Gifted Rating Scales-Preschool/Kindergarten Form (GRS-P): A Preliminary Examination of Their Psychometric Properties in Two Greek Samples. *Diagnostics*, *12*(11), 2809. https://doi.org/10.3390/diagnostics12112809

Sosic-Vasic, Z., Keis, O., Lau, M., Spitzer, M., & Streb, J. (2015). The impact of motivation and teachers' autonomy support on children's executive functions. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00146

Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, *105*(4), 970-987. https://doi.org/10.1037/a0032447

Stern, A., Pollak, Y., Bonne, O., Malik, E., & Maeir, A. (2016). The Relationship Between Executive Functions and Quality of Life in Adults With ADHD. *Journal of Attention Disorders*, *21*(4), 323-330. https://doi.org/10.1177/1087054713504133

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096-1104. https://doi.org/10.3758/brm.42.4.1096

Stoet, G. (2017). PsyToolkit. *Teaching of Psychology*, *44*(1), 24-31. https://doi.org/10.1177/0098628316677643

Stuss, D. T. (2011). Functions of the Frontal Lobes: Relation to Executive Functions. *Journal of the International Neuropsychological Society*, *17*(5), 759-765. https://doi.org/10.1017/S1355617711000695

Tamm, L., & Peugh, J. (2019). Concordance of teacher-rated and performance-based measures of executive functioning in preschoolers. *Child Neuropsychology*, *25*(3), 410-424. https://doi.org/10.1080/09297049.2018.1484085

Taylor, G., Jungert, T., Mageau, G. A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: the unique role of intrinsic motivation. *Contemporary Educational Psychology*, *39*(4), 342-358. https://doi.org/10.1016/j.cedpsych.2014.08.002

Teng, M. F., & Yang, Z. (2023). Metacognition, motivation, self-efficacy belief, and English learning achievement in online learning: Longitudinal mediation modeling approach. *Innovation in Language Learning and Teaching*, *17*(4), 778-794. https://doi.org/10.1080/17501229.2022.2144327

Thorell, L. B., & Nyberg, L. (2008). The Childhood Executive Functioning Inventory (CHEXI): A New Rating Instrument for Parents and Teachers. *Developmental Neuropsychology*, *33*(4), 536-552. https://doi.org/10.1080/87565640802101516

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1). https://doi.org/10.1186/s40561-023-00237-x

Tomlinson, C. A. (2015). Teaching for Excellence in Academically Diverse Classrooms. *Society*, *52*(3), 203-209. https://doi.org/10.1007/s12115-015-9888-0

Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms* (3rd Edition. ed.). ASCD.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*(2), 131-143. https://doi.org/10.1111/jcpp.12001

Tottenham, N., Hare, T. A., & Casey, B. J. (2011). Behavioral Assessment of Emotion Discrimination, Emotion Regulation, and Cognitive Control in Childhood, Adolescence, and Adulthood. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00039

Touré-Tillery, M., & Fishbach, A. (2014). How to Measure Motivation: A Guide for the Experimental Social Psychologist. *Social and Personality Psychology Compass*, *8*(7), 328-341. https://doi.org/10.1111/spc3.12110

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*(3), 282-299. https://doi.org/10.1037/h0027242

Troussas, C., Papakostas, C., Krouska, A., Mylonas, P., & Sgouropoulou, C. (2023). Personalized Feedback Enhanced by Natural Language Processing in Intelligent Tutoring Systems. In C. Frasson, P. Mylonas, & C. Troussas, *Augmented Intelligence and Intelligent Tutoring Systems* Cham.

Tsirides, A. (2022). Exploring (de) motivating factors in school years. The role of students, schools, and families from a Self-Determination Theory perspective. A thematic analysis. *Dialogues in Clinical Neuroscience & Mental Health*, *5*(1), 18-34. https://doi.org/10.26386/obrela.v5i1.218

Tulving, E. (1985). How many memory systems are there? *American psychologist*, *40*(4), 385.

Tulving, E., & Markowitsch, H. J. (1997). Memory beyond the hippocampus. *Current Opinion in Neurobiology*, *7*(2), 209-216. https://doi.org/10.1016/s0959-4388(97)80009-8

Tyack, D., & Tobin, W. (2016). The "Grammar" of Schooling: Why Has it Been so Hard to Change? *American Educational Research Journal*, *31*(3), 453-479. https://doi.org/10.3102/00028312031003453

Tzuriel, D. (2000). Dynamic Assessment of Young Children: Educational and Intervention Perspectives. *Educational Psychology Review*, *12*(4), 385-435. https://doi.org/10.1023/a:1009032414088

Ünlü, A. (2023). Qualitative motivation with sets and relations. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.993660

Vainas, O., Bar-Ilan, O., David, Y. B., Gilad-Bachrach, R., Lukin, G., Ronen, M., Shillo, R., & Sitton, D. (2019). E-Gotsky: Sequencing Content using the Zone of Proximal Development. *ArXiv*, *abs/1904.12268*.

Vanlehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, *46*(4), 197-221. https://doi.org/10.1080/00461520.2011.611369

Verbruggen, F., & Logan, G. D. (2008). Automatic and controlled response inhibition: Associative learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology: General*, *137*(4), 649-672. https://doi.org/10.1037/a0013170

Vernon, D. (2014). *Artificial cognitive systems : a primer*. The MIT Press.

Vernon, D., Metta, G., & Sandini, G. (2007). A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents. *IEEE Transactions on Evolutionary Computation*, *11*(2), 151-180. https://doi.org/10.1109/tevc.2006.890274

Visser, C. (2017). *The motivation continuum: self-determination theory in one picture.* http://www.progressfocused.com/2017/12/the-motivation-continuum-self.html

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes. L. S. Vygotsky*. Harvard U Press.

Walkington, C., & Bernacki, M. L. (2020). Appraising research on personalized learning: Definitions, theoretical alignment, advancements, and future directions. *Journal of Research on Technology in Education*, *52*(3), 235-252. https://doi.org/10.1080/15391523.2020.1747757

Wang, H., Tlili, A., Huang, R., Cai, Z., Li, M., Cheng, Z., Yang, D., Li, M., Zhu, X., & Fei, C. (2023). Examining the applications of intelligent tutoring systems in real educational contexts: A systematic literature review from the social experiment perspective. *Education and Information Technologies*, *28*(7), 9113-9148. https://doi.org/10.1007/s10639-022-11555-x

Wang, M., Geneva, H., & Walberg, H. (2015). What Influences Learning? A Content Analysis of Review Literature. *The Journal of Educational Research*, *84*, 30-43. https://doi.org/10.1080/00220671.1990.10885988

Wang, S., Christensen, C., Cui, W., Tong, R., Yarnall, L., Shear, L., & Feng, M. (2020). When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction. *Interactive Learning Environments*, *31*(2), 793-803. https://doi.org/10.1080/10494820.2020.1808794

Wei, X., Sun, S., Wu, D., & Zhou, L. (2021). Personalized Online Learning Resource Recommendation Based on Artificial Intelligence and Educational Psychology [Methods]. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.767837

Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of educational psychology*, *90*(2), 202.

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, *108*(3), 436-452. https://doi.org/10.1016/j.jecp.2010.08.008

Winne, P., & Hadwin, A. (1998). Studying as self-regulated learning. In (pp. 291-318). Routledge.

Winne, P. H. (2013). Self-regulated learning viewed from models of information processing. *Self-regulated learning and academic achievement*, 145-178.

Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning*, *9*(2), 229-237. https://doi.org/10.1007/s11409-014-9113-3

Winne, P. H., & Hadwin, A. F. (2013). nStudy: Tracing and Supporting Self-Regulated Learning in the Internet. In (pp. 293-308). Springer New York. https://doi.org/10.1007/978-1-4419-5546-3_20

Woolf, B., Aimeur, E., Nkambou, R., & Lajoie, S. (2008). *Intelligent Tutoring Systems: 9th International Conference on Intelligent Tutoring Systems, ITS 2008, Montreal, Canada, June 23-27, 2008, Proceedings* (Vol. 5091). Springer.

Woolf, B. P. (2009). *Building intelligent interactive tutors : student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann Publishers/Elsevier. Table of contents only http://www.loc.gov/catdir/toc/ecip0822/2008026963.html

Publisher description http://www.loc.gov/catdir/enhancements/fy0901/2008026963-d.html

Yang, Y. C., & Stocco, A. (2023). Allocating Mental Effort in Cognitive Tasks: A Model of Motivation in the ACT-R Cognitive Architecture. *Topics in Cognitive Science*. https://doi.org/10.1111/tops.12711

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience*, *5*(10), 995-1002. https://doi.org/10.1038/nn921

Zandbelt, B. B., & Vink, M. (2010). On the Role of the Striatum in Response Inhibition. *PLoS One*, *5*(11), e13848. https://doi.org/10.1371/journal.pone.0013848

Zhao, X., Chen, L., & Maes, J. H. R. (2018). Training and transfer effects of response inhibition training in children and adults. *Developmental Science*, *21*(1), e12511. https://doi.org/10.1111/desc.12511

Zhou, Q., Chen, S. H., & Main, A. (2012). Commonalities and Differences in the Research on Children's Effortful Control and Executive Function: A Call for an Integrated Model of Self-Regulation. *Child Development Perspectives*, *6*(2), 112-121. https://doi.org/10.1111/j.1750-8606.2011.00176.x

Zhou, X., Planalp, E. M., Heinrich, L., Pletcher, C., Dipiero, M., Alexander, A. L., Litovsky, R. Y., & Dean, D. C. (2022). Inhibitory Control in Children 4–10 Years of Age: Evidence From Functional Near-Infrared Spectroscopy Task-Based Observations. *Frontiers in Human Neuroscience*, *15*. https://doi.org/10.3389/fnhum.2021.798358

Zimmerman, B. J., & Schunk, D. H. (2011). *Handbook of self-regulation of learning and performance*. Routledge/Taylor & Francis Group.

Appendices

Appendix 1: Online forms

Briefing form

Σχετικά με αυτή την έρευνα

OPEN UNIVERSITY OF CYPRUS
www.ouc.ac.cy

20 YEARS DISTANCE EDUCATION

Αγαπητέ γονέα,
Η παρούσα έρευνα πραγματοποιείται στα πλαίσια διπλωματικής εργασίας του φοιτητή Ανδρέα Τσιρίδη για το μεταπτυχιακό με τίτλο "MSc Cognitive Systems" που προσφέρεται από τη Σχολή Θετικών & Επιστημών του Ανοικτού Πανεπιστημίου Κύπρου (ΑΠΚΥ) και της Σχολής Ψυχολογίας του Πανεπιστημίου Κύπρου. Σκοπός της έρευνας είναι να εντοπιστούν συσχετισμοί μεταξύ του κινήτρου για μελέτη των σχολικών μαθημάτων, της αυτορύθμισης και των γνωστικών λειτουργιών, ειδικώς της εκτελεστικής λειτουργίας και της εργαζόμενης μνήμης. Ο εκτιμώμενος συνολικός χρόνος για τη συμμετοχή είναι 35 λεπτά. Στην έρευνα καλούνται να συμμετάσχουν παιδιά από την τετάρτη μέχρι την έκτη δημοτικού καθώς και οι δάσκαλοι ή οι γονείς τους. Για κάθε παιδί αντιστοιχεί ένας γονέας ή δάσκαλος. Η έρευνα είναι ανώνυμη και πραγματοποιείται σύμφωνα με τον κανονισμούς της ΕΕ (GDPR) ώστε να προστατευθούν τα προσωπικά και ερευνητικά δεδομένα που ενδεχομένως να χρησιμοποιηθούν, καθώς και η ψυχική και σωματική ευημερία των συμμετεχόντων. Προκειμένου να διασφαλιστεί η ποιότητα και η εγκυρότητα της έρευνας, παρακαλώ τηρήστε τις οδηγίες που θα σας δοθούν. Οι ερωτήσεις καθώς και τα γνωστικά τεστ που θα πραγματοποιηθούν είναι σχεδιασμένα έχουν σταθμιστεί και έχουν δοκιμαστεί ώστε να μην προξενήσουν κάποια δυσφορία ή συναισθηματική ένταση στους συμμετέχοντες. Εάν παρ' όλα αυτά σε οποιαδήποτε στιγμή της έρευνας οποιοσδήποτε από τους συμμετέχοντες θελήσει να κάνει διάλειμμα ή διακόψει εντελώς τη διαδικασία, μπορεί να το κάνει. Ωστόσο, κατά τη διάρκεια του γνωστικού τεστ (του τρίτου μέρους) επειδή παίζει ρόλο ο χρόνος απόκρισης είναι απαραίτητο να μην διακοπεί, εκτός αν υπάρξει απόλυτη ανάγκη.

Συνίσταται να βάλετε τον φυλλομετρητή (browser) σε λειτουργία πλήρους οθόνης

Στοιχεία Επικοινωνίας

Πληροφορίες για αυτή την έρευνα :

Για οποιαδήποτε ερώτηση παρακαλώ επικοινωνήστε με τον φοιτητή Ανδρέα Τσιρίδη στο andreas.tsiridis@st.ouc.ac.cy ή την επιβλέπουσα καθηγήτρια Δρ Μαρίας Σοφολόγη στο maria.sofologi@ouc.ac.cy

Ηλεκτρονικό ταχυδρομείο : andreas.tsiridis@st.ouc.ac.cy

Σημαντικές τεχνικές απαιτήσεις για τον υπολογιστή σας

Φαίνεται να χρησιμοποιείτε τον ακόλουθο περιηγητή (αριθμός έκδοσης στις παρενθέσεις): Chrome (119)
Ο περιηγητής σας υποστηρίζει τις απαιτήσεις αυτής της έρευνας.

Για αυτή την έρευνα χρειάζετε να έχετε ένα φυσικό πληκτρολόγιο.

Για αυτή την έρευνα ο υπολογιστής σας θα πρέπει να έχει ήχο και το μεγάφωνο πρέπει να είναι ενεργό.Ελέγξτε τον ήχο σας τώρα.

Ελέγξτε το ΑΡΙΣΤΕΡΟ μεγάφωνο/ακουστικό | Ελέγξτε το ΔΕΞΙΟ μεγάφωνο/ακουστικό

Επιβεβαιώστε ότι θέλετε να κάνετε αυτή την έρευνα

Παρακαλώ επιβεβαιώστε ότι θέλετε να συμμετάσχετε σε αυτήν την έρευνα. Οι πληροφορίες σας (συμπεριλαμβανομένης της IP του υπολογιστή) θα αποθηκευτούν και ενδέχεται να χρησιμοποιηθούν για έρευνα.

☐ Έλαβα γνώση και δίνω τη συγκατάθεσή μου για τη δική μου συμμετοχή και τη συμμετοχή του παιδιού στην έρευνα, καθώς έχω ενημερώσει τον γονέα ΄/ κηδεμόνα του και έχω λάβει τη συγκατάθεσή του. Αν είμαι ο γονέας, δηλώνω πως έχω δώσει τη συγκατάθεσή μου.

Πατήστε αυτό το κουμπί για να ξεκινήσετε την έρευνα

Important data protection information

When you start, this survey will store your answers, your internet address, and browser information on the PsyToolkit server.The responsibility for this survey rests entirely with the researcher(s) listed above. Click here if you do not want to participate now.

Demographics form

Γράψτε εδώ την τάξη του παιδιού του παιδιού

- ◯ Τρίτη Δημοτικού
- ◯ Τετάρτη Δημοτικού
- ◯ Πέμπτη Δημοτικού
- ◯ Έκτη Δημοτικού

Ποιο είναι το βιολογικό φύλλο του παιδιού;

- ◯ Αγόρι
- ◯ Κορίτσι

Ηλικία παιδιού

Πόσων ετών είναι το παιδί; [ ]

Σχέση με το παιδί

- ◯ Γονέας/κηδεμόνας
- ◯ Δάσκαλος/-α

Βιολογικό φύλο γονέα/κηδεμόνας ή δάσκαλου/-ας

- ◯ Άνδρας
- ◯ Γυναίκα

Ηλικία γονέα/κηδεμόνα ή δάσκαλου/-ας

Πόσων ετών είστε; [ ]

Πατήστε το κουμπί για να συνεχίσετε

Appendix 2: Instruments /

SRQ-A (English- standard version)

WHY I DO THINGS

Name: _____ Age: _____

Grade: _____ ( ) Boy or Girl ( ) Teacher: _____

A. Why do I do my homework?

1. Because I want the teacher to think I'm a good student.
   Very True | Sort of true | Not very true | Not at all true
2. Because I'll get in trouble if I don't.
   Very True | Sort of true | Not very true | Not at all true
3. Because it's fun.
   Very True | Sort of true | Not very true | Not at all true
4. Because I will feel bad about myself if I don't do it.
   Very True | Sort of true | Not very true | Not at all true
5. Because I want to understand the subject.
   Very True | Sort of true | Not very true | Not at all true
6. Because that's what I'm supposed to do.
   Very True | Sort of true | Not very true | Not at all true
7. Because I enjoy doing my homework.
   Very True | Sort of true | Not very true | Not at all true
8. Because it's important to me to do my homework.
   Very True | Sort of true | Not very true | Not at all true

B. Why do I work on my classwork?

9. So that the teacher won't yell at me.

   Very True | Sort of true | Not very true | Not at all true

10. Because I want the teacher to think I'm a good student.

    Very True | Sort of true | Not very true | Not at all true

11. Because I want to learn new things.

    Very True | Sort of true | Not very true | Not at all true

12. Because I'll be ashamed of myself if it didn't get done.

    Very True | Sort of true | Not very true | Not at all true

13. Because it's fun.

    Very True | Sort of true | Not very true | Not at all true

14. Because that's the rule.

    Very True | Sort of true | Not very true | Not at all true

15. Because I enjoy doing my classwork.

    Very True | Sort of true | Not very true | Not at all true

16. Because it's important to me to work on my classwork.

    Very True | Sort of true | Not very true | Not at all true

C. Why do I try to answer hard questions in class?

17. Because I want the other students to think I'm smart.

    Very True | Sort of true | Not very true | Not at all true

18. Because I feel ashamed of myself when I don't try.

    Very True | Sort of true | Not very true | Not at all true

19. Because I enjoy answering hard questions.

    Very True | Sort of true | Not very true | Not at all true

20. Because that's what I'm supposed to do.

    Very True | Sort of true | Not very true | Not at all true

21. To find out if I'm right or wrong.

    Very True | Sort of true | Not very true | Not at all true

22. Because it's fun to answer hard questions.

    Very True | Sort of true | Not very true | Not at all true

23. Because it's important to me to try to answer hard questions in class.
    Very True | Sort of true | Not very true | Not at all true

24. Because I want the teacher to say nice things about me.

D. Why do I try to do well in school?

25. Because that's what I'm supposed to do.
    Very True | Sort of true | Not very true | Not at all true

26. So my teachers will think I'm a good student.
    Very True | Sort of true | Not very true | Not at all true

27. Because I enjoy doing my school work well.
    Very True | Sort of true | Not very true | Not at all true

28. Because I will get in trouble if I don't do well.
    Very True | Sort of true | Not very true | Not at all true

29. Because I'll feel really bad about myself if I don't do well.
    Very True | Sort of true | Not very true | Not at all true

30. Because it's important to me to try to do well in school.
    Very True | Sort of true | Not very true | Not at all true

31. Because I will feel really proud of myself if I do well.
    Very True | Sort of true | Not very true | Not at all true

32. Because I might get a reward if I do well.
    Very True | Sort of true | Not very true | Not at all true

Scoring the SRQ-A (standard version). First, you calculate the subscale score for each of the four subscales by averaging the items that make up that subscale. Very true is scored 4; Sort of true is scored 3; Not very true is scored 2; and Not at all true is scored 1. The four subscales are: external regulation, introjected regulation, identified regulation, and intrinsic motivation. Listed below are the item numbers associated with each of the four subscales.

External Regulation: 2, 6, 9, 14, 20, 24, 25, 28, 32

Introjected Regulation: 1, 4, 10, 12, 17, 18, 26, 29, 31

Identified Regulation: 5, 8, 11, 16, 21, 23, 30

Intrinsic Motivation: 3, 7, 13, 15, 19, 22, 27

You can use the individual subscale scores in your analyses, and you can also use the Relative Autonomy Index (RAI). To form the RAI for this scale, use the following formula to combine the subscale scores:

2 X Intrinsic + Identified - Introjected - 2 X External

Childhood Executive Functioning Inventory (CHEXI)
for parents and teachers

Below, you will find a number of statements. Please read each statement carefully and thereafter indicate how well that statement is true for the child. You indicate your response by circling one of the numbers (from 1 to 5) after each statement.

Definitely not true 1	Not true 2	Partially true 3	True 4	Definitely true 5

1. Has difficulty remembering lengthy instructions	1	2	3	4	5
2. Seldom seems to be able to motivate him---/herself to do something that he/she doesn't want to do	1	2	3	4	5
3. Has difficulty remembering what he/she is doing, in the middle of an activity	1	2	3	4	5
4. Has difficulty following through on less appealing tasks unless he/she is promised some type of reward for doing so	1	2	3	4	5
5. Has a tendency to do things without first thinking about what could happen	1	2	3	4	5
6. When asked to do several things, he/she only remembers the first or last	1	2	3	4	5
7. Has difficulty coming up with a different way of solving a problem when he/she gets stuck	1	2	3	4	5
8. When something needs to be done, he/she is often distracted by something more appealing	1	2	3	4	5

9. Easily forgets what he/she is asked to fetch	1	2	3	4	5
10. Gets overly excited when something special is going to happen (e.g., going on a field trip, going to a party)	1	2	3	4	5
11. Has clear difficulties doing things he/she finds boring	1	2	3	4	5
12. Has difficulty planning for an activity (e.g., remembering to bring everything necessary for a field trip or things needed for school)	1	2	3	4	5
13. Has difficulty holding back his/her activity despite being told to do so	1	2	3	4	5
14. Has difficulty carrying out activities that require several steps (e.g., for younger children, getting completely dressed without reminders; for older children, doing all homework independently)	1	2	3	4	5
15. In order to be able to concentrate, he/she must find the task appealing	1	2	3	4	5
16. Has difficulty refraining from smiling or laughing in situations where it is inappropriate	1	2	3	4	5
17. Has difficulty telling a story about something that has happened so that others may easily understand	1	2	3	4	5
18. Has difficulty stopping an activity immediately upon being told to do so. For example, he/she needs to jump a couple of extra times or play on the computer a little bit longer after being asked to stop	1	2	3	4	5
19. Has difficulty understanding verbal instructions unless he/she is also shown *how* to do something	1	2	3	4	5
20. Has difficulty with tasks or activities that involve several steps	1	2	3	4	5

21. Has difficulty thinking ahead or learning from experience	1	2	3	4	5
22. Acts in a wilder way compared to other children in a group (e.g., at a birthday party or during a group activity)	1	2	3	4	5
23. Has difficulty doing things that require mental effort, such as counting backwards	1	2	3	4	5
24. Has difficulty keeping things in mind while he/she is doing something else	1	2	3	4	5

SCORING

Fill in the total score for the respective subscales in the boxes below, and the total score for the two factors WORKING MEMORY and INHIBITION. For an example of ADHD and control group means and SDs, as well as cut off scores, see Catale, Meulemans, & Thorell (in press)[1].

Subscale 1: Working memory
Total score for items: 1, 3, 6, 7, 9, 19, 21, 23, 24 = _____

Subscale 2: Planning
Total score for items: 12, 14, 17, 20 =_____

Working Memory - Total score =_____

[1] Catale, C., Meulemans, T., & Thorell, L. B. (in press). The Childhood Executive Function Inventory (CHEXI): Confirmatory Factorial analyses and cross---cultural clinical validity in a sample of 8–11 years old Children. Journal of Attention Disorders, doi: 10.1177/1087054712470971

Subscale 3: Regulation

Total score for items: 2, 4, 8, 11, 15 = _____

Subscale 4: Inhibition

Total score for items: 5, 10, 13, 16, 18, 22 =_____

Inhibition Memory - Total score = _____

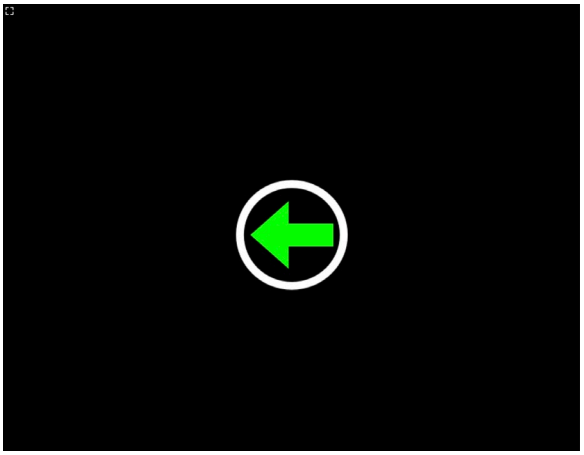
The CHEXI includes four different subscales tapping Working Memory, planning, regulation and inhibition. However, factor analysis in children in kindergarten was only able to identify two factors referred to as **WORKING MEMORY** (Working Memory and planning subscales) and **INHIBITION** (regulation and inhibition subscales).

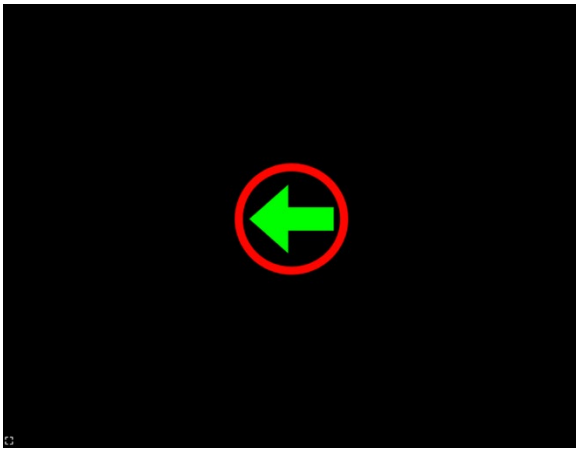
Appendix 3: Cognitive Tasks (screenshots)

The experimental platform is available here

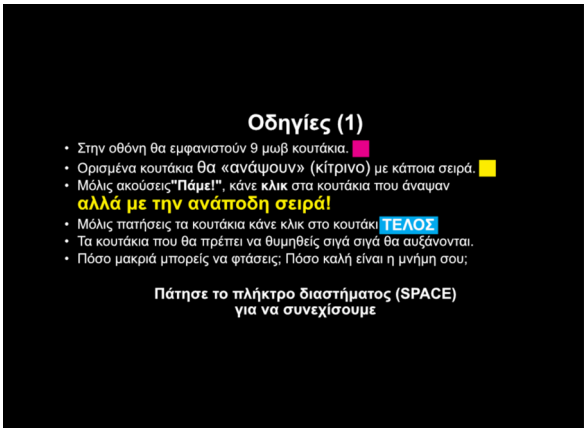
https://www.psytoolkit.org/c/3.4.4/survey?s=EEdNr

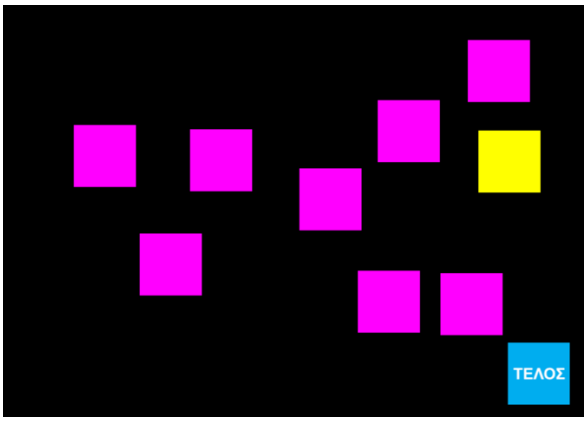
Stop Signal

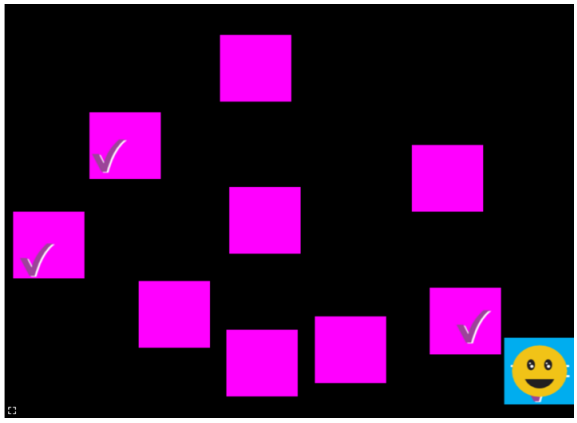




Corsi Blocks Backwards Task







Appendix 4: a-priori power analysis

Exact - Linear multiple regression: Random model

Options:	Exact distribution				
Analysis:	A priori: Compute required sample size				
Input:	Tail(s)	=	One		
	H1 ρ^2	=	0,15		
	H0 ρ^2	=	0		
	α err prob	=	0,05		
	Power (1-β err prob)			=	0,85
	Number of predictors			=	2
Output:	Lower critical R^2	=	0,0880557		
	Upper critical R^2	=	0,0880557		
	Total sample size	=	68		
	Actual power	=	0,8502427		

(^_^)