UNIVERSITY OF PAVIA – IUSS SCHOOL FOR ADVANCED STUDIES PAVIA

Department of Brain and Behavioral Sciences (DBBS)

MSc in Psychology, Neuroscience and Human Sciences

A Systematic Review and Meta-Analysis of Temperament and Learning Abilities

Supervisors:

Prof. Sara Mascheretti

Prof. Valentina Lampis

Thesis written by

Nadin Stereva

Academic year 2023/2024

Index

Index 1
Abstract 3
Introduction & Background 4
Temperament 4
Learning 15
The relationship between temperament and learning: the state of the art. 19
Hypotheses and aims 22
Method 24
Search methods for identification of studies 25
Inclusion/Exclusion Criteria 25
Types of Studies 27
Types of participants 27
Language 27
Date of Publication 27
Temperament and Academic Measures 28
Selection Process 28
Data Collection 29
Statistical analysis 29
Results 30
Descriptive statistics 30
Temperamental Factors 32
Hyperactivation & Distractibility 32
Goal Orientation & Regulation 34
Negative Tendency 35
Positive Tendency & Social Skills 36
Behavioral Inhibition 37
Data extraction and management 37
Meta-analysis results: Math 38
The relationship between Hyperactivation & Distractibility and Math 38
The relationship between Goal Orientation & Regulation and Math 41
The relationship between Negative Tendency and Math 44
The relationship between Positive Tendency & Social Skills and Math 46
The relationship between Behavioral Inhibition and Math 48
Meta-analysis results: Reading 50
The relationship between Hyperactivation & Distractibility and Reading 50
The relationship between Goal Orientation & Regulation and Reading 52
The relationship between Negative Tendency and Reading 54
The relationship between Positive Tendency & Social Skills and Reading 56

The relationship between Behavioral Inhibition and Reading 58
Discussion **61**
Hyperactivation & Distractibility 61
Goal Orientation & Regulation 64
Negative Tendency 65
Positive Tendency & Social Skills 66
Behavioral Inhibition 67
Limitations 67
Sources of Heterogeneity 68
Temperamental Dimensions Definitions 68
Reporter Bias 69
Standardized Test Scores 70
Age 71
Conclusion **73**
Bibliography **75**

Abstract

This systematic review and meta-analysis aimed to synthesize the different findings of previous empirical research concerning the relationship between temperament and learning abilities (math and reading) in school-aged children. We included 19 published studies between 1990 and the present day across six countries with a cumulative sample of 9,847 children across 28 temperamental dimensions collected using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) method. Five temperamental macro-dimensions (factors) were created and a random-effects model was fitted for the data of each factor looking for the central tendency of the relationship between that specific temperamental profile and either math or reading. The main findings of this meta-analysis indicated positive correlations between Goal Orientation & Regulation profile and both reading and math, and between Positive Tendency & Social Skills profile and reading and math. The findings for Hyperactivation & Distractibility and Negative Tendency factors affirmed negative correlations with both reading and math. Our final temperamental factor, Behavioral Inhibition, did not result in a significant correlation.

Introduction & Background

Temperament is a concept which can be traced back to ancient Greek philosophers, such as Hippocrates and Galen, who postulated people's behavior to be caused by varying humors. Temperament has come far from its Mesopotamian and Greek roots in humorism. Moving to modern times, the first predominant publication on child temperament came from Thomas and colleagues and was based on the New York Longitudinal Study (Thomas, 1963). This study led to further reports concerning temperamental development throughout childhood, identifying nine basic temperamental variables and providing questionnaires for parents of infants and both parents and teachers of children aged 3 to 7 years (Thomas & Chess, 1977). This work constituted a great influence on the paradigm shift from an essentially unidirectional, environmental approach of child development, to one recognizing the active role of the child in the developmental process. In further research the emphasis on how temperament affects different areas of development led to the recognition that these differences in children's capacity to regulate their emotions, attention and reactions to stimuli impact a child's adjustment and performance in school (Keogh, 2003).

This review aims to identify the relationship between the temperament of a child and their academic performance and learning skills, specifically in the areas of math and reading, as presented in scientific literature to this point.

Temperament

Temperament is a prevalent interest in the area of psychology and scientific research. One of the main reasons why temperament is a main interest of psychologists, educators, and parents alike is because there is theoretical work that suggests that not only does temperament develop early in a child but additionally, it remains rather stable throughout that individual's lifetime. The early development of temperament, in broad, refers to the emergence and

stability of individual differences in emotional, motor, and attentional reactivity, as well as self-regulation, in infancy and early childhood (Calkins & Degnen, 2005). Temperament is considered to be biologically based and observable from a young age, with roots in genetic and neurobiological processes. Temperamental traits begin to emerge in infancy, being first observable in a baby's reactions to sensory stimuli and new experiences. Some children may be more easily soothed while displaying a tendency for positive affect, while others may react strongly to unfamiliar situations or stimuli, indicating differences in reactivity and self-regulation (Thomas & Chess, 1977). Despite recent genome-wide association studies depicting that temperament is strongly influenced by more than 700 genes (Cloninger et al., 2019) other studies have additionally illustrated how environmental factors such as parenting style, cultural context, and early experiences also play a significant role in shaping and moderating temperamental traits. For example, sensitive parenting can boost the development of effortful control in children, leading to better self-regulation; on the contrary, children with tendencies toward frustration, impulsivity and low in effortful control are more vulnerable to the adverse effects of negative parenting. The same studies however, show that this is a bidirectional effect (Kiff et al., 2011).

In addition to theoretical support for the early development of temperament and its multiple channels, there is also work observing the stability of temperament, which refers to the degree to which temperamental traits are maintained across different developmental stages, situations, and contexts. There are two main perspectives to consider when looking at the stability of temperament across ages. One is what is referred to as Homotypic stability which examines the consistency of similar behaviors across time. As maturation or normative development may impact the stability of temperament over time, this may explain why homotypic stability may be more visible after puberty (Putman et al. 2008). Kagan (1969) argues that as a result of the rapid and widespread developmental changes in the early years

of life the stability of temperamental traits among children will be heterotypic. Heterotypic stability refers to the consistency of the underlying psychological attribute that may have different behavioral manifestations at different ages. A study by Putman, Rothbart and Gartstein (2008) investigated the stability of temperamental traits measured by some of the most used tool to measure temperament, i.e., the Infant Behavior Questionnaire Revised (IBQ-R) (Gartstein & Rothbart, 2003), Early Childhood Behavior Questionnaire (ECBQ) (Putnam et al., 2006) and the Children's Behavior Questionnaire (CBQ) (Rothbart et al., 2001). This study found both considerable homotypic and heterotypic stability. Homotypic stability for Surgency and Negative Affect was seen across all time points, while Effortful Control and Self-Regulation were seen to be stable across adjacent time periods. Additionally, this study determined that other temperamental dimensions such as Activity, Impulsivity and Intensity strongly supported the stability of Surgency, while Sadness and Frustration played major roles in the stability of Negative Affect. Heterotypic stability was also found as high levels of infant Surgency were seen to be predictive of toddler Effortful Control, attention Shifting and Low-Intensity Pleasure (Putnam at al., 2008). However, as stability in temperament can have different denotations, researchers propose at least four types of stability in longitudinal research (Caspi & Shiner, 2006; De Fruyt et al., 2006; Putnam et al, 2008). The first is ipsative which examines the degree to which the prevalence, priority and order of temperamental traits of an individual are preserved across time. The second is structural which defines the amount of continuity in the associations among temperamental traits across time. The next two are the most commonly expanded on from my research when specifically looking at temperament rather than personality. Mean-level stability observes whether the average levels of a temperamental traits remain constant over time while rank-order stability (relative order) examines the consistency of rank order of individuals on a trait. A study by Josefsson et al. (2013) examined the developmental patterns

of temperament through a longitudinal study of a large group of Finnish men and women (20 - 45 years olds). The examination of the mean-level stability depicted clear-cut, qualitative patterns for the temperamental traits of Novelty Seeking, Harm Avoidance, Reward Dependence and Persistence. The finding illustrated a decrease in Novelty Seeking and a slight increase in Persistence. A study by Dyson et al. (2015) observed rank-order stability on 447 children in laboratory settings using the Laboratory Temperament Assessment Battery (Lab-TAB). They looked for the temperamental dimensions of Positive Affect/Interest, Sociability, Dysphoria, Fear/Inhibition, and Impulsivity vs. Constraint. The result depicted moderate, yet significant rank-order stability from the age of 3 to 6 and significant heterotypic associations of Sociability at age 3 and Positive Affect/Interest at age 6, as well as Impulsivity vs. Constraint at age 3 and Fear/Inhibition at age 6. Despite the lack of agreement on the definition of stability in temperament and the differing perspectives on the consideration of developmental change, the majority of researchers agree that there are patterns and consistency of temperament over time.

Despite its vast allure there is yet to be a predominant theoretical approach and a specified denotation of its dimensions. Some prevalent theoretical approaches on temperament have been put forth by Thomas and Chess (1977), Rothbart and Derryberry (1981), Kagan (1984), and Cloninger (1993).

As aforementioned, Thomas, Chess and colleagues (1963) play a vital role in the initial outlooks on child temperament. Their discussions often focused on how a child's temperament affected their experiences in school, at home and with other children. They adopted the nine temperamental traits put forth by Dr. Herbert Birch, which were; Activity, Adaptability, Distractibility, Initial Reaction, Intensity of Emotions, Mood, Persistence and Attention Span, Regularity and Sensitivity. All of these temperamental traits were rated for

children on a continuum in which the presumed best place to be was in the middle, not too low or high. Activity level describes a child's physical energy, if high it would denote that the child may have difficulty sitting still in class and may be hyperactive at home and in social situations, whilst a low score indicates a child who prefers sedentary activities. Adaptability refers to the time it takes a child to adjust to a new environment or situation, a child with low adaptability is resilient to conform and adapt to a new circumstance or routine. Distractibility, in its two extremes, indicates a child that is highly probable to become distracted by external stimuli or one that maintains great attention to the task at hand despite the presence of interrupters. What was then measured as Initial Reaction is now most commonly known as approach or withdrawal, depicting how a child responds to novelty (environments or people). Intensity of emotions is relatively, straight forward in that a child with high intensity of emotions will exhibit extreme reactions when experiencing either negative or positive emotions, whilst a child with a low score may not display notable emotions. A child's tendency toward a generally positive or negative demeanor was measured as Mood. Persistence and Attention Span is almost a reverse measure of Distractibility, it looks at a child's ability to maintain focus and the duration of time for which a child can attend to a specific task. Children with low scores in this variable often became frustrated after losing interest in any given activity. Regularity indicates how predictable and routine a child's biological responses are; a child with a high score will have a stable routine of when to eat, sleep, etc. Regularity is often referred to as rhythmicity. Finally, Sensitivity measures a child sensory threshold, a child with high sensitivity will be preoccupied or disturbed by noises, lights, smells and even textures while a child with a low sensitivity will not pay attention to these factors and maintain focus on a task. These temperamental dimensions were used to create three profiles of children; easy, difficult and slow-to-warm (Thomas and Chess, 1977). What was referred to as 'easy' children were relatively calm, highly adaptable, stable in their

routines and tended toward positive moods and emotions. ‘Difficult’ children on the other hand were highly active, emotional (leaning toward negative emotions and frustration), irritable and had unstable eating and sleeping patterns. The final profile of ‘slow-to-warm’ children was that they scored low in activity, were relatively shy and were slow to adapt to new situations and people, but with time they did. However, the limitation of this approach is that not all children fit into these categories.

Thomas and Chess propose that temperament is present from birth and represents a set of biological individual differences in behavior that are evident early in life. They argued that these temperamental traits are innate, meaning they are not learned or acquired and that these core temperamental traits are relatively stable, and provide the foundation for later personality development. Temperamental traits are seen as the precursors to personality suggesting that a child who is highly adaptable and sociable may develop into an outgoing and socially skilled adult (Thomas and Chess, 1977). They recognized that while temperament has a biological basis, its manifestation can be significantly influenced by environmental factors. While temperament is considered moderately stable the manner in which its traits are expressed can change based on environmental influences, such as parenting styles, family dynamics, cultural context, and life experiences. In their discussion on environmental influences, Thomas and Chess explain the concept of "goodness of fit", which suggests that a child's development depends on how well the environment, particularly their caregivers, match or conflict with the child's temperamental traits. A good fit between a child's temperament and their environment can lead to positive developmental outcomes, whereas a poor fit can lead to difficulties in connections and maladaptive outcomes. Thomas and Chess (1977) highlighted several potential developmental outcomes based on the interplay between temperament and environment such as the adaptive (positive self-esteem or social competence) or maladaptive (anxiety or poor relationships) outcomes based on

"goodness of fit" (Hipson & Séguin, 2017). Similarly, they discuss how children with generally more difficult temperaments, such as those characterized by high Activity, low Adaptability, and Negative Mood, are more at risk for developing adjustment problems. Thomas and Chess developed an instrument to assess temperament, The New York Longitudinal Study (NYLS) Parent Interview was one of the first systematic efforts to study temperament. The NYLS parent Interview utilized parent interviews to assess the nine dimensions in infants and young children. Finally, the Carey Temperament Scales (CTS) were created by Dr. William B. Carey, the CTS provides age-specific questionnaires (Infant Temperament Questionnaire, Toddler Temperament Scale, etc.) to measure temperament from infancy through adolescence, assessing similar dimensions to those provided by Thomas and Chess (Carey, 1970).

Mary K. Rothbart is a prominent psychologist known for her influential work on the study of temperament, particularly in children. Her contributions have significantly advanced the understanding of how individual differences in temperament emerge, develop, and influence behavior. Rothbart explored the ways in which temperament influences developmental trajectories (Rothbart & Derryberry, 1981). She studied how different temperamental traits interact with environmental factors to shape personality development and behavior over time. Her research has shown that temperament is relatively stable. However, her studies have provided evidence for the fact that the development of temperament can be influenced by parenting, culture, and other environmental factors. The theoretical emphasis of Rothbart's work concerns the importance of integrating the biological, cognitive and emotional aspects of temperament to create a better understanding of a child's innate temperaments disposition and the role of life experiences as building blocks toward adult personality.

The perspective on temperament of Mary K. Rothbart depicts temperament as the individual differences of children that are present before the complete development of personality and higher cognitive functions. Building on the prior theories describing temperament as the dimensions of self-regulation and reactivity manifested in a child's emotions, activities, and attention, Rothbart and colleagues found three consistent broad factors that apply to children between the ages of 3 and 8 years old thus creating the Three-Factor Model of Temperament. Using scale-level factor analyses of the Children's Behavior Questionnaire the three emerging factors were *Surgency/Extraversion,* which was defined by the scales of Approach, High Intensity Pleasure, Activity Level, Impulsivity, and Shyness. Secondly *Negative Affectivity* was defined by scales of Discomfort, Fear, Anger/Frustration, Sadness and Soothability. Finally, the third dimension was *Effortful Control*, defined by Inhibitory Control, Attention Focusing, Low Intensity Pleasure, and Perceptual Sensitivity (Rothbart & Ahadi, 1994).

Rothbart views temperament as biologically based and observable from infancy, occurring in the first few months. She supports that temperament is influenced over time by heredity, maturation, and experience. Rothbart's research suggests that traits such as Fear, Irritability, Activity level, and Soothability are observable in infancy, while other traits like Effortful Control emerge in early childhood as cognitive and emotional systems develop. Rothbart's perspective on temperament stability emphasizes both continuity and change across development, acknowledging the influence of both maturation and environment. Parental behaviors, socialization, cultural influences, and life experiences can affect how temperament is expressed and modified throughout development. According to her beliefs, two children with equally high levels of Negative Affectivity develop differently as one may learn how to better emotional regulation from supportive parents, which in turn can lead to changes in how this temperamental trait is expressed. Rothbart emphasizes that temperament

plays a crucial role in shaping developmental outcomes; traits such as Effortful Control are linked to the development of social competence, emotional regulation, and moral behavior. Effortful Control is also a strong predictor of cognitive outcomes, such as academic achievement. Additionally, Rothbart's research indicates that certain temperamental traits can increase the risk for developing internalizing and externalizing disorders, especially when combined with environmental stressors or a lack of supportive relationships.

Finally, Mary K. Rothbart also played a crucial role in the development of a multitude of temperament measurement tools. Rothbart developed the Infant Behavior Questionnaire (IBQ) which is an instrument used to assess temperament in infants (from 3 to 12 months of age), allowing researchers to study temperament from a very early age (Rothbart, 1981). The original IBQ was developed in the early 1980s and was first reported in “Measurement of Temperament in Infancy” (Rothbart, 1981). This early form of the instrument assessed 6 infant temperamental domains (activity level, soothability, fear, distress to limitations, smiling and laughter, and duration of orienting). The items on the IBQ ask parents to rate the frequency of temperament-related behaviors observed over the past week. The IBQ has been revised and several new scales were added (IBQ-R; Gartstein & Rothbart, 2003). Additionally, there is now a Short (91 items; 14 scales) and Very Short (37 items; 3 broad scales) versions of the IBQ-R developed by Sam Putnam and colleagues in 2008. She developed a widely used parent-report questionnaire to assess temperament in children aged 3 to 7 years. The Children’s Behavior Questionnaire (CBQ) is a highly differentiated assessment of temperament in early to middle childhood. It is using this instrument that the three factors (Negative Affectivity, Surgency Extraversion, and Effortful Control) were reliably derived. The CBQ is widely used in developmental research and adapted for many of the needs of this field. Rothbart and Samuel Putnam developed a short and very short version of the instrument. Hedy Teglasi created a teacher-report version of the CBQ Short Form.

The work of Jerome Kagan on temperament primarily focuses on the dimension of Reactivity. Kagan presents two different temperamental types of children based on the manner in which they react when faced with unfamiliarity (Kagan, 1997), those with high and low reactivity. The 'High Reactive' children showed high levels of motor activity and distress when presented with novelty were deemed inhibited, while the 'Low Reactive', uninhibited children continued to be motorically relaxed and were emotionally unfazed by the same unfamiliar stimuli. If these differences in children's reactivity profiles remained when they were 4.5 years old this could indicate a risk for developing an anxiety or conduct disorder. His research suggests that temperamental biases, such as the tendency toward inhibition or un-inhibition, can be detected in infants at four months old, indicating that temperament has an early onset that is largely genetically influenced (Kagan & Snidman, 1999). High Reactivity infants exhibit strong physiological and behavioral reactions to unfamiliar stimuli. He suggests that this reactivity reflects a sensitive and easily aroused amygdala, a brain structure involved in processing emotions. Kagan's perspective on the stability of temperament focuses on the idea that early temperament is moderately stable over time but is affected by environmental influences such as parenting style, cultural context, and social experiences. For example, supportive parenting and positive social experiences can help highly reactive children become more adaptable and less inhibited over time.

Cloninger's model of temperament involves four dimensions of personality that reflect the heritable aspects of the automatic responses to perceptual stimuli. These four dimensions are referred to as temperamental factors as they are defined by the individual differences in associative learning in response to novelty, danger, punishment or reward (Cloninger, 1993). The first dimension is called Novelty Seeking, it alludes to a heritable

tendency to be compelled by and brought to take action when faced with novelty. Additionally, it describes a disposition for impulsive decision making, quick reaction when faced with the possibility of reward, quick loss of temper and a drive to avoid boredom and frustration. Second there is Harm Avoidance, which indicates a disposition for a pessimistic, anxious outlook, passive avoidant behaviors, fear of uncertainty and shyness. The third temperament factor, Reward Dependence, is viewed as the maintenance of ongoing behaviors, due to sentimental reasons, social attachment, and dependence on approval of others. The final temperamental dimension is Persistence, which describes an individual's tendency to persevere despite frustration, loss of motivation or fatigue. It was originally proposed as a continent of Reward Dependence but it was uncorrelated to sentimental and social motives. Cloninger supported that the stable nature of temperament along with the measure of these dimensions in children could be predictive of adolescent and adult behavior. These specific factors are denoted as temperament dimension and not merely aspect of personality as they are heritable, manifest in childhood, and are based in preconceptual or unconscious biases of learning (Cloninger, 1993). Cloninger views temperament as having a biological and genetic basis, with a child's temperamental traits becoming apparent early in life. The aforementioned four temperamental dimensions he created are thought to be linked to specific neurotransmitter systems (dopamine for Novelty Seeking, serotonin for Harm Avoidance). Cloninger developed several instruments used to assess temperament and character across different age groups. The first was the Temperament and Character Inventory (TCI), which is designed for adults and includes multiple scales corresponding to the four temperament dimensions and three character dimensions (Self-Directedness, Cooperativeness, and Self-Transcendence; Cloninger, 1994). Starting from these tests, a Junior Temperament and Character Inventory (JTCI) was created for children and adolescents from ages 7 to 17 years old (Luby et al., 1999). The Revised Temperament and Character

Inventory (TCI-R) was created which provides a more nuanced and detailed measurement of Cloninger's dimensions and can be used in both clinical and research settings (Gutierrez et al., 2015). Finally, a Short Temperament and Character Inventory (STCI) was developed for quick and efficient assessments.

Each of the aforementioned theories of temperament offer unique perspectives and frameworks for understanding temperament in children and its development across the lifespan. While they share some similarities, such as the recognition of temperament as being a biologically based component of personality, they differ in their conceptualization and definitions of temperament dimensions, the role of environmental influences, and their approach to measuring temperament.

Learning

A predominant portion of early life for almost all children is spent in school. This is why for parents, researchers, educators and policy makers, a child's academic competence is a topic of great attention. Understanding the developmental pathways, environmental influences and individual differences that lead to the variations in academic outcomes for children and adolescents may not only lead to new manners of helping these students improve academic performance, but provide ways of promoting better psychological and functional outcomes during the school years. While academic achievement refers to the competences in reading, writing, mathematics, sciences and the thinking skills that facilitate a student's success in school, learning refers to the acquisition of the knowledge and skills which permit any of this to happen. Learning abilities are present in infancy, and are crucial for adaptation. They range from habituation and responses to stimuli to much more complex learning capacities that evolve throughout one's lifespan. During development, learning abilities become more easily integrated across different cognitive tasks and modalities, allowing the

encoding of more complex information, and in larger amounts. This leads to an increase in knowledge acquisition which affects anything from our behavior, decision to our academic capabilities (Lafontaine et al. 2020). These learning abilities early on in development help children develop their mathematical and reading skills which continue to grow and become more complex with age given stimulated learning.

In this study we focused on the performances of school aged children in mathematics and reading skills, as they are two of the most similar and stable subjects across cultures. Mathematical skills have been shown to have a stable trajectory of development, along with a great susceptibility to influence by external factors. In 1992 Karen Wynn presented evidence suggesting that five-month-old babies react with surprise when numerical excitations are violated. This led to the proposition that even despite the acquisition of any numerical language, infants are capable of encoding numerical information. This began long debates on the development of mathematical skills, whether or not it is an innate skill and the origins of math skills. However, a general understanding of math development is that in infancy, between the ages of 0 to 12 months old, there is some basic numerical awareness, basic understanding of differences in quantities (recognizing a difference between two and three objects) and of changes in small numbers, a skill known as "subitizing" which is an innate ability (Van Aster & Shalev, 2007). During toddlerhood, between the ages of one- and three-years old, children start learning the sequence of number words and can often recite them in order around the age of two, which is known as "rote counting". At this stage, they begin to understand that numbers correspond to specific quantities and can recognize small numbers/sets of up to three objects (around two or three years old). Toddlers begin to grasp simple concepts of addition and subtraction when done with small groups of objects, for example adding or taking one away. In early childhood, between the ages of three and five, children understand cardinality, which is the fact that the last number when counting

represents the total value, at this point they can also usually count higher than ten. Around ages four or five, children start to perform basic addition and subtraction, often using physical objects, fingers, or drawings to help them visualize the operations. Around the age of five to six years old with children entering kindergarten, they can usually count to 100 and beyond, they recognize and write numbers and hold a functional understanding of values which allows for them to develop the first strategies to help with mental math along with visualization concepts like length, width etc. The development of math skills in the early years is a dynamic process that builds on children's innate abilities to recognize quantities and patterns. As children grow, their math skills evolve from basic numerical awareness and counting to more complex arithmetic, spatial reasoning, and problem-solving skills. Early math experiences are crucial as they form the foundation for future mathematical learning and cognitive development. Examining the correlation between a child's temperament and math skills is important as it can provide valuable insights into children's cognitive and emotional development and inform us on the functionality of educational strategies. It is particularly useful when trying to understand how children's individual differences affect their academic achievement and development.

The development of reading skills in the early years of life is a process that involves building a foundation for language, literacy, and cognitive skills. This development starts from infancy and continues through early childhood, forming the basis for later reading proficiency. In infancy, from 0 to 12 months old, children begin developing pre-reading skills through exposure to language. They learn to recognize phonemes specific to their own native language and begin to distinguish between different sounds within the language and sounds of different languages. Toddlers from the age of one to three years old rapidly expand their vocabulary and understanding of spoken language, they also begin to understand that pictures

in books represent real objects and can identify and name objects in pictures. Toddlers develop a (phonological) awareness of sounds in words, starting with recognizing rhymes. At this stage, children mimic reading behaviors, such as holding a book, turning pages, and pretending to read. Preschoolers begin recognizing letters, especially those in their own names, and learn that letters represent specific sounds. They also develop an understanding of the concept of print (i.e., the fact that reading happens from left to right and top to bottom). Around the ages of four to five years, children begin connecting letters and the corresponding sounds and start decoding simple, one syllable words. In kindergarten children begin to read simple books with familiar vocabulary and begin to develop sight-word recognition eliminating the need for decoding, which is the first step to automated reading. Fluency emerges as children read more confidently and smoothly, and comprehension skills improve as does the understanding of sentences and stories.

Examining children's mathematical and reading skills in relation to their temperament is important for understanding how personality traits influence academic achievement. It is important to look at math and reading as they are representative of two subjects based in different cognitive domains, and in assessing the influence of temperament we should find that it has subject-specific effects. Math and reading require different cognitive abilities. Mathematics requires problem-solving skills and logical reasoning, while reading demands comprehension and linguistic processing. Outside of just looking at math and reading in relation to temperament, examining children's math and reading skills is crucial because these foundational abilities significantly influence their academic success, cognitive development, and future opportunities. Math and reading are fundamental skills that underpin learning across all subjects, therefore early skills in these two domains predict all future academic success. Math and reading development are linked to critical cognitive skills such as memory, attention, problem-solving, and executive function and the more engagement a child has with

these two practices the better they become at logical thinking, pattern recognition, and the ability to follow sequences. Additionally, competence in math and reading can boost a child's confidence and self-esteem, bettering their social competences. Finally, well developed math and reading skills are beneficial for future learning and are crucial for success in higher education, many career paths and even easing everyday tasks.

The relationship between temperament and learning: the state of the art.

Temperament traits such as attention control or negative affect might impact mathematical and reading domains differently. For instance, a child who is easily frustrated tends to struggle more with math problems requiring persistence, while a child who is less focused may find reading comprehension more challenging (Martin et al., 1994). By looking at both mathematical and reading skills, researchers can understand how specific temperament traits influence distinct cognitive tasks.

From the moment a child begins schooling their unique temperament affects how they interact with others, respond to their environments and engage with learning tasks. As the child ages and transitions through the different grades their temperamental differences impact how they adapt to the new experiences, challenges and people. Considering a child's temperament is important when examining academic achievement as its effects many influencing factors like learning styles, classroom behavior, motivation, how children face academic distress, social interactions with peers and teachers (Keogh, 2003). Temperament traits, like effortful control, influence a child’s ability to concentrate, focus, and persist with tasks, which are crucial for academic success (Curtindale et al., 2007). Children with high effortful control develop better abilities to regulate their attention, leading to more effective learning. Another example could be a child’s level of activity (which is either considered to be one measure of temperament or as a part of the Surgency/Extraversion dimension) can

affect how children engage with learning tasks. Highly active children might struggle with tasks that require prolonged focus and quiet study, while children with lower activity levels may excel in such environments. Outside of merely the learning, a large part of the academic experience is based in classroom interactions. Children with better self-regulation are more likely to exhibit positive classroom behaviors such as following instructions, staying on task, and controlling impulses. These behaviors contribute to a more favorable learning environment and better academic outcomes. Opposingly, children who are highly reactive or have high negative affect might experience more frustration, anxiety, or sadness, which can interfere with their ability to learn and participate fully in classroom activities. Temperament also affects children's motivation, for example curiosity (part of Surgency) can drive a child's intrinsic motivation to learn, explore, and engage in learning. Children who are naturally curious and eager to learn are more likely to perform better academically. High levels of effortful control lead to more persistence, allowing children to tackle challenging tasks and persist through difficulties, which is critical for academic achievement. Temperament influences how children interact with peers, which can affect their learning success and their general experience in school settings. Children with positive social temperaments may find it easier to collaborate, seek aid, and engage in group learning, all of which can enhance their academic performance and it can also affect how children respond to peer influences, which can impact their attitudes towards school and learning. A child's temperament impacts their relationship with peers and teachers alike. Positive relationships with teachers can further motivate children leading to better academic outcomes. Children whose temperament aligns well with their teacher's style may develop stronger, more supportive relationships. For example, a child with high Surgency may thrive in an interactive, hands-on learning environment, while a child with high Effortful Control may do well with structured, independent work. Understanding temperament helps educators tailor their teaching methods

to meet the needs of different students. Finally, a child's temperament may dictate how they respond to stress or discourse. For instance, children with high negative affectivity are more sensitive to stress, which can negatively impact their academic performance, which is resilience (another aspect of effortful control) enabling children to recover from setbacks and persist in the face of academic challenges. Understanding these differences allows parents and educators to provide the appropriate support to help children cope with academic pressures.

In conclusion, it is important to study temperament not merely because it is a stable, early arising sign of a child's personality that can give us insight into how a child may act and react in social and academic environments but also because, as it is relatively stable, it can provide us with the opportunity to build profiles of children's specific needs. Looking at temperament along with math skill and reading ability can inform us to create tailored interventions. By assessing both reading and math skills, educators and parents can identify specific areas where a child's temperament might be a barrier to success. For instance, a child with low persistence might benefit from strategies that build perseverance to aid in math. Understanding the correlation between temperament and academic skills enables educators to tailor teaching methods and provide personalized learning for children with different needs. For instance, a child with a temperament that favors structured environments might excel in math with clear, step-by-step instructions, while a more flexible, creative approach might be better for another child. Temperament is closely tied to emotional regulation, which in turn affects learning. By examining how temperament influences both math and reading skills, educators can better integrate social and emotional learning (SEL) into academic curricula, helping children develop the emotional and cognitive skills necessary for success across all subjects. The observation of the interaction of temperament and academic performance could exhibit early predictors of risk factors. Some temperament traits, like low self-regulation or

high emotional reactivity, can predict general lower academic achievement which is can predict somatic health risks (Alatupa et al., 2010), depression (Chen, Rubin, & Li, 1995), future educational attainment (Marjoribanks, 2005), socioeconomic position (Guglielmi, 2008), and the risk of unemployment in adulthood (Caspi, Wright, Moffitt, & Silva, 1998).

Hypotheses and aims

Before even beginning our search for information and data, we set forth a simple question: what evidence is there in the existing research that a child's temperament affects their learning abilities and consequently school performances? As with any review, we continued from that point onward hoping to find enough data to be able to analyze any aspect of the relationship of interest between temperament and specific learning domains. We therefore created five distinct temperamental factors/profiles (Hyperactivation & Distractibility, Goal Orientation & Regulation, Negative Tendency, Positive Tendency & Social Skills, and Behavioral Inhibition) and we analyzed the correlation between each of these factors and academic learning.

Concerning our first factor, Hyperactivation & Distractibility, we expect a negative correlation between temperament and either of our academic domains. As children's score in Activity, Impulsivity, Distractibility, Low Task Orientation, Emotionality, Reactivity, Surgency and Intensity increase, both assessments of math and reading will decrease. From our delve into previous literature we cannot predict a specific difference between mathematical and reading ability based on the temperamental dimensions included in this factor.

We predict a positive correlation between our second factor, Goal Orientation & Regulation, and both math and reading. As Effortful Control, Attention, Persistence, Perseverance, Self-Regulation and Sensory Regulation increase, so will the score of academic

achievement. However, in this case we expect to see a more substantial effect on math as our previous research suggest that Persistence (Xiao & Sun, 2021) and Self-Regulation (Fauzi & Widjajanti, 2018) are an important factor that affects the development of math ability and the motivation from which math performance may improve over a student academic career.

Our third factor, Negative Tendency, consisting of Negative Affect, Negative reactivity and Anger, is expected to negatively correlate to both math and reading performance. Once more we cannot necessarily state that we expect a specific difference between mathematical and reading scores when it comes to their interaction with Factor 3 as negative emotions have been repeatedly shown to negatively affect all aspects of academic performance.

We predict that the fourth factor, Positive Tendency & Social Skills, will have a positive relationship with both math and reading as, opposingly to Factor 3, positive emotions and good social skills have been shown to have a beneficial effect on academic performance. We predict that as children's temperamental assessment scores increase in Adaptability/Agreeableness, Positive Affect, Positive Mood, Positive Reactivity, Mood and Resilience, Sociability, Affiliation and Soothability so will their math and reading scores.

Our final factor is called Behavioral Inhibition, this factor depicts the greatest difficulty in setting a hypothesis concerning a relationship exclusively with academic skills. The effects of Shyness, Social Inhibition and Inhibition on children's social experiences in school settings is more focused on in previous literature than on the relationship with their math and reading abilities. However, we believe these to be connected and therefore expect our fifth factor to have a negative correlation with math and reading scores.

Method

A systematic review of literature referring to temperamental dimensions and academic achievement was conducted. It focused on empirical studies of children of all school ages, starting from pre-kindergarten to grade 12. We used a search protocol based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2015). A summary of this process can be seen in our PRISMA chart (Table 1).

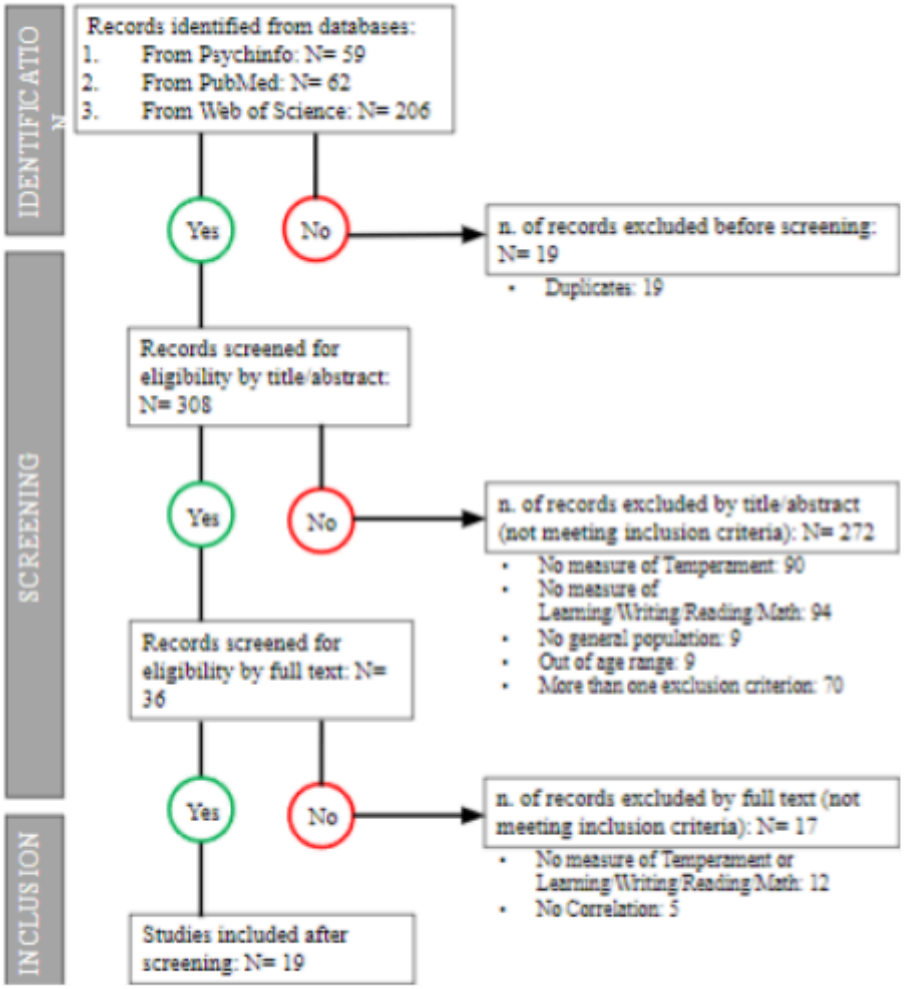

Table 1 | Flowchart of the identification and selection of studies

Search methods for identification of studies

The literature search was conducted in May 2024, across three databases; Web of Science, PsychINFO and PUBMED. The following keywords were used to conduct an advanced search with the combined the terms:

1. “temperament”, “temperamental dimension”, “temperamental trait” (all OR)
2. “academic achievement”, “academic performance”, “learning abilit”, “reading”, “reading skill”, “reading comprehension”, “literacy”, “math”, “numeracy”, “writing” (all OR)
3. “school age”, “school child”, “grade”, “kindergarten”, “preschool”, “prekindergarten”, “preK” and “adolescent” (all OR)

The rest of the criteria were specified through the side filters; peer reviewed, publication years, language and general or control participants. Additionally, we made sure to only look at studies that conducted correlation analyses rather than regressions as our primary interest was the relationship between temperament and learning. The results of these combinations of terms were depicted in either titles, abstracts, and/or subject terms. The results for the three databases consisted of 327 texts which were then reduced to 308 when duplicates from across the different databases were removed.

Inclusion/Exclusion Criteria

Before beginning to search for the texts from which the review would derive its data, inclusion and exclusion criteria were set. In particular, we considered the characteristics of studies, the participants of interest and the internal measures and topics of the articles themselves. Table 2 depicts all of the inclusion and exclusion criteria set in place.

Table 1 – Inclusion and Exclusion Criteria for Systematic Review

Inclusion Criteria	Exclusion Criteria
The document is accessible, in its full version. The document is from a peer-reviewed journal or published by a test publisher/editor The document provide result of original empirical data The document in in English Document must be published between 1987-Present	**Missing Document:** Text could not be found **Unpublished Document:** Document in unpublished, or in press (e.g. dissertations or theses are excluded) **Non-Reviewed Document:** Document is not published in a peer-reviewed journal (e.g. books, book chapters or conference proceedings, are excluded) **Non-Empirical Study:** Does not provide new, original data (e.g. meta-analyses, theoretical articles, literature reviews, letters or editorials, are excluded) **Language:** Article is written in a language other than English. **Old Publication:** Document has been published before 1987
Measure done with human participants The measure is administered to school aged participants (3 to 18 years of age). Studies with participants ages 3 and 18 or over are included if the participants are attending an academic institution. The sample is included if the age is not disclosed but rather the academic level/grade. Preschool and kindergarten aged children also included Participants reflect general population	**Non-Human Subjects:** Not administered on humans (e.g. animal studies) **Outside Age Range:** younger children or adults are excluded. **No General Population:** Measure administered only on individual of specific population (e.g. case studies or studies with no control group are excluded)
A measure aiming to evaluate temperament/ temperamental dimensions in normative sample A measure aiming to quantify/determine academic achievement or propensity for specific learning skill	**Does Not Measure Temperament:** The measure does not assess participants temperament (e.g. measures of personality and social behaviors are excluded) **Does Not Measure Academic Ability:** The measure does not assess participants abilities in math, reading or language.

Types of Studies

We focused solely on original empirical research in peer-reviewed journal publications. This review excluded dissertations, theses, books, meta-analyses and other potentially non peer reviewed articles. Additionally, articles that could not be found in their full form were clearly not included.

Types of participants

To be included in the present study, samples have to be composed by school-aged children representative of the general population. Some of the studies included were more specific, such as Blackson T. C. (1995) which looks at how temperament and IQ mediate the effects of family history of substance abuse on academic achievement. We included this study as there was a matched pair case-control group, and we only considered the data collected for the children of the control group. When specific socio-economic status (SES) groups were the primary focus of the study, data was included if the analysis controlled for variables.

Language

One of our main exclusion criteria was language of publication. The studies included in this review are only those published in English. In cases in which the academic measure of the studies was reading, the language of the assessment did not have to be English and we allowed for data of any language speaking population. We included multiple studies which observed reading in English and Finnish, and one study which recorded reading scores on multi-lingual participants.

Date of Publication

The collection of journal articles was limited to a specific set of years. Included articles were published between and present day. The reason for which we set the publication

limitation to any article postdating 1977 was because that was the year in which a theory of temperament was published for the first time. As mentioned in the introduction Thomas and Chess had published an article defining their nine temperamental dimensions and creating their Childhood Temperament Questionnaire in 1977.

Temperament and Academic Measures

As the review is examining whether there is a correlation between temperament and academic achievement, each article has to contain both a measure of at least one temperamental dimension and academic achievement or learning skills. The temperamental dimensions in all the articles included were present in one of the four theories of temperament presented in the introduction.

Selection Process

The abstracts were read for all 308 articles and 272 articles were excluded because they did not contain a measure of temperament or of academic achievement, a control group, participants within the defined age range and any combination of these reasons. The remaining 36 articles were read in full; however, 17 articles were excluded for either not having a correlation analysis but merely a regression, or for not including a viable measurement of temperament or academic performance in mathematic or reading abilities. The final number of articles used for this review was 19 from which a total of 226 effects size describing the relation between a temperamental dimension/trait and an academic assessment were taken.

Data Collection

Once all the articles were selected, we began to collect all our qualitative and quantitative data of interest. As aforementioned we had three academic categories; math, reading and other. The category that we created named other mainly included pre-reading measures in preschoolers from longitudinal studies, however, we put it in a separate category as to not include it and allow it to skew the meta-analysis. Across all the data collected we had measures for 32 temperamental dimensions which sometimes represent different names for temperamental traits with the same definition (such as Persistence and Perseverance) and sometimes represent reverse measures (such as Attention and Distractibility). In the majority of the articles, the authors had either created profiles or factors in which they had combined different temperamental traits that would often coincide in a child. We referenced our original theoretical literature along with the definitions set forth in the creation of five temperamental dimensions, i.e., Hyperactivation & Distractibility, Goal Orientation & Regulation, Negative Tendency, Positive Tendency & Social Skills, and Behavioral Inhibition.

Statistical analysis

Following the collection of all our data of interest and their separation into academic domains, a meta-analysis was conducted using the MAJOR module in Jamovi Version 2.3.26. We created different databases, one for each academic domain. For each factor a Correlation Coefficients meta-analysis was run including information about Pearson's r, sample size and the study label.

Results

The total of 19 studies which met the inclusion criteria, were comprised of 9847 children. The studies varied in design, with most utilizing either longitudinal or cross-sectional data. Each of the studies included at least one measure of temperament along with an assessment of either or both math and reading skills leading to 226 correlations.

Descriptive statistics

From the final studies included in this review, we have 9847 participants from studies across the United States of America, Asia, Canada and Europe. The percentage of participants per region included can be seen in Chart 1.

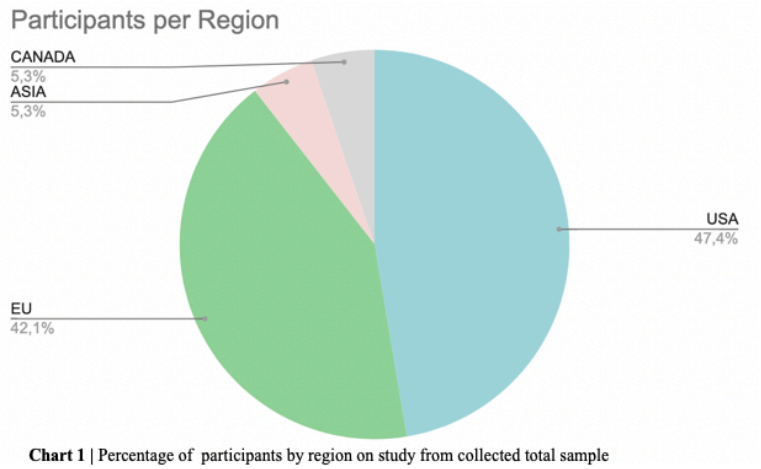

Chart 1 | Percentage of participants by region on study from collected total sample

As previously mentioned, we included original articles which reported both a measure of at least one temperamental dimension and learning skills. We included child temperament that was reported by different sources (parents, teachers, psychologists or self-report).

The proportion of each category of temperamental rater per each academic domain can be seen in Chart 2. For academic measures we had three categories; reading, math and other. The reading measures we included had participants ranging from 1st Grade until the 9th Grade. For math we have measures starting at the end of kindergarten until 9th grade. Our final category that will not be included in the analysis portion of this thesis was measures that were either written out as literacy, phonological awareness or etc.

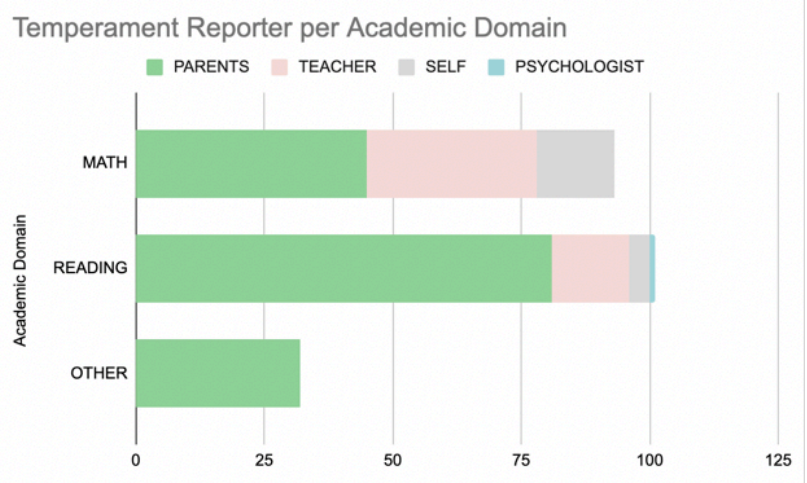

Chart 2 | Prevalence of specific temperament reporters per academic domain across studies included

In general, this category, which we named "other", measures pre-reading skill in preschoolers and pre-kindergarten children in longitudinal studies from which we are only taking the academic data measured by standardized reading tests once the participants develop their reading abilities.

For the analysis, the correlations were organized by the factor that their temperamental dimension belonged to. This meant that every single study was considered in every factor for which temperamental dimensions were available. Additionally, the analysis and therefore the results were divided into math and reading in correlation with each factor.

The analysis used the Fisher r-to-z transformed correlation coefficient as the outcome measure, as it yields a variable which is approximately normally distributed stabilizing the variance, and a random- effects model was fitted to the data.

Temperamental Factors

The five factors created from the original 32 temperamental dimensions were based on commonly used child profiles throughout the studies and theories applied for this review. All temperamental dimensions were divided into these five factors except for Threshold and Rhythmicity as they were only applied to the academic domain we labeled as 'other' which will not be used in the analysis for this thesis.

Hyperactivation & Distractibility

The first factor we created was the most commonly occurring throughout the literature we used as reference. It drew inspiration from the profile of the 'Difficult' child from Thomas and Chess and the 'high reactive' child by Kagan. This factor included eight of the original temperamental dimensions paired up and divided into four groups (Chart 3 depicts the distribution of data collected across the different dimensions within the factor labeled

Chart 3 | Frequency of temperamental measures per learning domain Hyperactivation & Distractibility

Hyperactivation & Distractibility across academic domains). The first pair is Activity and Impulsivity; there were seven that reported temperament through Activity and one through Impulsivity. The CBQ defines Activity as the level of gross motor activity including rate and extent of locomotion (Rothbart et al., 2001), while Impulsivity was only used in a single article by Svens-Liavåg et al. (2023) in which they measure impulsivity/activity through questions such as the following which was used as a reverse measure (e.g., "I sit calmly on my chair when somebody is reading or telling me something"). The second pair in this factor is Distractibility and Low Task Orientation which were placed together as Distractibility is the tendency to get distracted or break attention (Thomas and Chess, 1977) and Low Task Orientation was only used in one article in which it was measured with items such as the following (e.g., "Whatever the student is doing, nothing can distract him or her" [reversed item]) (Viljaranta et al., 2015). The third is Emotionality and Reactivity which were both only used by one article each Reactivity was defined as the latency, duration, and intensity of a child's reaction to external and internal stimuli (Auerbach et al, 2019) which Emotionality was defined as the tendency of a child toward react with great arousal/emotion either positive or negative. Finally, the last pair of our first factor was Surgency and Intensity, where Surgency was widely used and defined as a temperament dimension that reflects an individual's disposition toward positive affect, approach, sociability, high-intensity pleasure, reward seeking, and a high activity level (Rothbart, 2011). Intensity was only used in one article by Schraeder et al. (1990) but it was never defined by itself within this article, however, they did use the altered version of the IBQ, the Toddler Behavior Questionnaire which defines Intensity through high and low intensity pleasure; it is the amount of pleasure or enjoyment related to high or low stimulus intensity, rate, complexity, novelty, and incongruity.

Goal Orientation & Regulation

The second factor created was based on the second most common child profile described in our theoretical literature, the profile of the 'Easy' child from Thomas and Chess and the 'low reactive' child by Kagan. This factor consisted of three pairs of our collected temperamental dimensions. The first pair is based highly on the definition of Effortful Control by Mary K. Rothbart, it contains all measures collected of Effortful Control and of Attention, this was our second most measured pair and can be seen in Chart 4. The second pair included Persistence and Perseverance which was one of our initial combination because while reading the single article by Svens-Liavåg et al. (2023) that reported Perseverance in their tables when discussing the items used to get these measures they use the word 'persistence' interchangeably, the following was a revered item (e.g. "I give up easily when meeting a difficult task, like a difficult homework or a difficult game"). Our final pair for this factor was Self-Regulation and Sensory Regulation, the latter was only measured in one article by Auerbach at al. (2019) as they based their temperamental denotation on a study (Gouze et al., 2009) which suggests that Sensory Regulation is a distinct factor, but the majority of our theoretical background views it as a component of temperamental self-regulation.

Chart 4 | Frequency of temperamental dimensions per learning ability for Goal Orientation & Regulation

Negative Tendency

The third factor created consisted of two groups based primarily on the denotation put forth by Rothbart & Ahadi (1994) of Negative Affectivity. The first pair was Negative Affect and Negative Emotionality which was our most measured pair and consisted of these two terms which were interchangeable throughout our collected literature depending on which tools were being used. When the EATQ-R (Capaldi & Rothbart, 1992) or the CBQ (Rothbart et al., 2001) were used Negative Affect was being measured as both were based on the theories of Mary K. Rothbart, but when the TABC (Martin, 1994) was the primary tool Negative Emotionality was measured as this tool was based on the temperament theory of Thomas and Chess (1977). The second category in this factor merely consisted of the temperamental measure of Anger, it was only measured in one study by Valiente (2021) as the tool used was the TMCQ which divided negative affect into anger, frustration and fear and measured them separately as they were triggered by different stimuli.

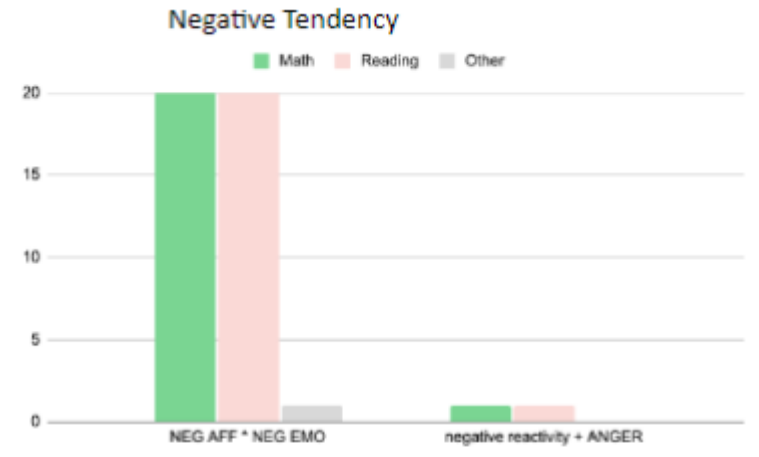

Chart 5 | Frequency of temperamental dimensions per learning ability for Negative Tendency

Positive Tendency & Social Skills

Positive Tendency & Social Skills was created as a reverse of Negative Tendency with some aspects influenced by Cloninger's description of Reward Dependence. This was a factor created to reflect the profile of a positive and socially motivated child, as through-out the studies collected this profile seem to differ from the rest and it wasn't necessarily as academically inclined as the children described by Goal Orientation & Regulation and differed greatly from all our other temperamental profiles. This factor was composed of ten temperamental dimensions divided into two groups, one more so implying a child's tendency toward positive emotions and reaction which included Adaptability/Agreeableness, Positive Affect, Positive Mood, Positive Reactivity, Mood and Resiliency. The second group within this factor defined a child's social inclinations whether that be focused toward their peers, parents or teachers, it included Sociability, Affiliation and Soothability.

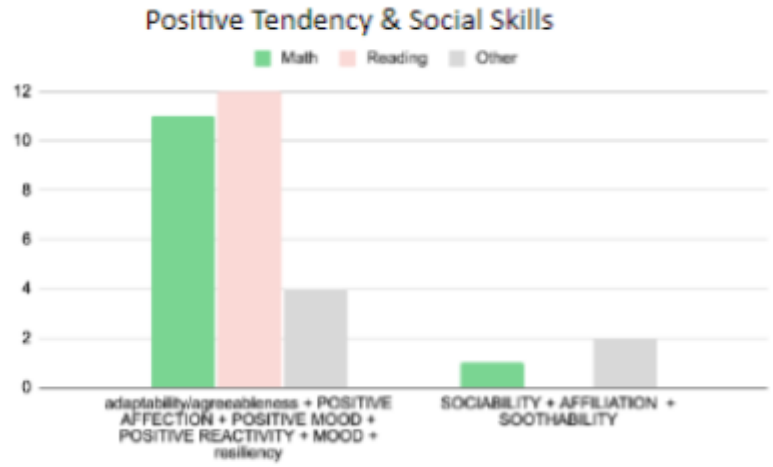

Chart 6 | Frequency of temperamental dimensions per learning domain for Positive Tendency & Social Skills

Behavioral Inhibition

The final factor was based on the 'slow-to-warm' child by Thomas and Chess, with some consideration of Cloninger's definition of Harm Avoidance children. This factor consists of three temperamental dimensions; Shyness, Social Inhibition and Inhibition. Shyness was measured with the TMCQ by Valiente et al. (2021) where it was defined as a slow or inhibited approach in situations involving novelty or uncertainty. Social Inhibition was also explored in one study by Newman and colleagues (1998) and was defined as a child who makes new friends easily, but is slow to warm up to new adults. Finally, Inhibition was used across the four studies in which it was most often observed as withdrawal, weariness and avoidance of novelty or uncertainty.

Chart 7 | Frequency of temperamental dimensions per learning domain for Behavioral Inhibition

Data extraction and management

Data collection from the 19 final studies was conducted using a pre-defined, comprehensive form which included the following variables: author(s), year of publication, study design, region/country of the study, age range of control/normative participants in the

studies and gender distribution among participants. We then summaries the temperamental information that could be extractor from the articles; name of measure/tool, specific temperamental dimension and what factor it belonged to, reporter of child's temperament, reliability of the measure, temperament mean, maximum and minimum and the age at which temperament was measured. We then focused on collecting the academic data of interest starting with the academic domain, learning test, test reliability, age at which learning test was administered and the mean score of the test. Finally, we collected all the data relevant to our correlation of interest; Pearson's r, Fisher effect size, whether there was significance or not (coded 1 if significant 0 if there was none) and the number and type of covariates. From that point onward once all the data was collected, it was divided into 3 sheets depending on the academic domain; math, reading and other. Descriptive statistics were added in a final sheet to provide us with a manner to visualize our library of information. Whatever data was missing from a specific study was reported as 999 to signify that this element was not reported in the original article.

Meta-analysis results: Math

The relationship between Hyperactivation & Distractibility and Math

The analysis of math scores in correlation with Hyperactivation & Distractibility included a total of 60 correlations from our 19 studies. The correlation coefficients observed across the included collected correlations, after applying the Fisher r-to-z transformation, ranged from -0.5493 to 0.1003, with the majority of estimates being negative (78%). The meta-analysis estimated an average Fisher r-to-z transformed correlation coefficient of -0.1409(95% CI: -0.1836 to -0.0982). The Random-Effects model accounts for variability both within and between studies, suggesting that the negative relationship observed is likely

consistent across different study populations and contexts. The Confidence Interval (CI) indicates that we can be 95% confident that the true average correlation is between -0.1836 and -0.0982, since this range does not include zero, it suggests that the negative relationship is statistically significant. The average outcome differed significantly from zero ($z = -6.4633$, $p < 0.0001$), the z-score represents how many standard deviations the average correlation is away from zero and a z-score of -6.4633 is large, indicating that the average correlation is significantly different from zero. Additionally, the p-value depicts the probability of observing such a strong result if there were no actual relationship (if the true correlation was zero) and the small p-value (< 0.0001) reaffirms that this result is statistically significant, stating that there is a strong likelihood that the observed negative correlation is not merely due to chance. Overall, the results concerning the correlation between our first temperamental factor and mathematical ability depicts a consistent negative relationship across the studies which is statistically significant. This relationship, although weak (average correlation of -0.1409), is reliably different from zero, indicating that as a child's score in Activity, Impulsivity, Distractibility, Low Task Orientation Emotionality, Reactivity, Surgency and Intensity increases, the math scores or mathematical ability tends to decrease slightly. The forest plot (Chart 8) illustrates the effect sizes from the studies, with the overall pooled effect size indicating a negative relationship.

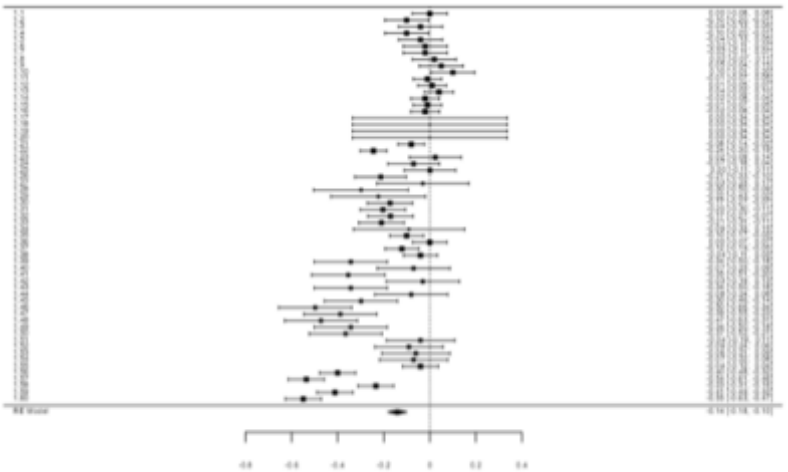

Chart 8 | A forest plot of the effect sizes for Hyperactivation & Distractibility and math

The results of the heterogeneity statistics for the meta-analysis provide insights into the variability and reliability of the findings across the studies included. According to the Q-test, the true outcomes appear to be heterogeneous ($Q(59) = 659.1325$, $p < 0.0001$, $tau2 = 0.0240$, $I2 = 90.9229\%$). The Q statistic illustrates whether the observed variation in effect sizes across studies is greater than what would be expected by chance. The value 659.1325 is the computed Q statistic, and the degrees of freedom (df) are 59 (number of studies minus one). Once again, the very small p-value indicates that the variation in effect sizes is statistically significant, this insinuates that the true effect sizes across the studies are heterogeneous (the difference is more than what would be expected by random sampling variability alone). Tau-Squared is the estimated variance of the true effect sizes (between-study variance). A Tau-Squared value of 0.0240 indicates there is some variability in the true effects across the studies, suggesting that not all studies are estimating the same underlying effect. Finally, the I-Squared quantifies the percentage of total variability in effect sizes that is caused by heterogeneity and not by chance. An I-Squared value of 90.9229% indicates that a very large proportion of the observed variance is due to actual differences between the studies rather than random error. The 95% prediction interval for the true outcomes is given by -0.4476 to 0.1658 which once more indicates that despite the majority of the studies/correlations included depicting a negative relationship some may illustrate a weak positive correlation. When looking to detect outliers and influence, the examination of the residuals revealed that none of the studies had a value larger than ± 3.3415 and hence there was no indication of outliers in the context of this meta-analysis model. However, according to Cook's distances, two studies (1.57; 1.60) could be considered to be overly influential. Cook's distance measures the influence of each study on the overall meta-analysis results. Two studies have Cook's distances which implies that these studies have a large

impact on the estimated average effect size, and that their removal might significantly change the results. The rank correlation test indicated funnel plot asymmetry (p = 0.0047) but not the regression test (p = 0.5968) (Chart 9).

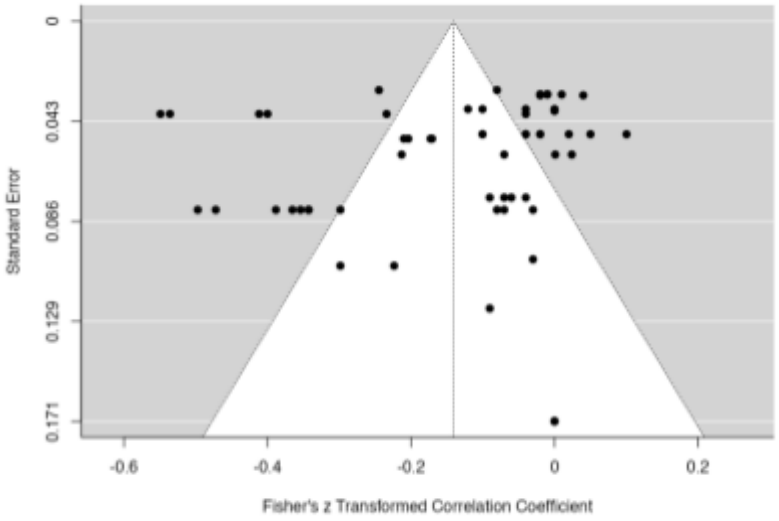

Chart 9 | Funnel Plot for Hyperactivation & Distractibility and Math Meta-Analysis

The Rank Correlation test evaluates whether there is asymmetry in the funnel plot, which could indicate publication bias. The p-value of 0.0047 suggests significant funnel plot asymmetry, implying potential publication bias. The regression test is another method for assessing funnel plot asymmetry. However, the non-significant p-value (0.5968) indicates that there is no strong evidence of funnel plot asymmetry. The inconsistency between the rank correlation test and the regression test may indicate that any potential bias is not strong.

The relationship between Goal Orientation & Regulation and Math

The analysis of math scores in correlation with Goal Orientation & Regulation was based on 52 correlations from our studies looking at measurements of Effortful Control, Attention, Persistence, Perseverance, Self-Regulation and Sensory Regulation. The observed

Fisher r-to-z transformed correlation coefficients ranged from -0.4973 to 0.6625, with the majority of estimates being positive (88%). The estimated average Fisher r-to-z transformed correlation coefficient based on the random-effects model was 0.1876 (95% CI: 0.1424 to 0.2327). Therefore, the average outcome differed significantly from zero ($z = 8.1424$, $p < 0.0001$). The CI indicates that we can be 95% confident that the true average correlation is between 0.1424 and 0.2327, since this range does not include zero, it suggests that the positive relationship is statistically significant. The z-score of 8.1424 is quite large, indicating that the average correlation is significantly different from zero. The very small p-value (< 0.0001) supports that this result is statistically significant, and that this positive correlation is not a result of chance. The results concerning the correlation between our second temperamental factor and mathematical ability illustrates a stable, positive, statistically significant relationship. This relationship is significantly different from zero depicting the positive effect that a child possessing Effortful Control, Attention, Persistence, Perseverance, Self-Regulation and Sensory Regulation would have on their mathematical ability. The forest plot (Chart 10) illustrates the effect sizes from the studies, with the overall pooled effect size indicating a positive relationship.

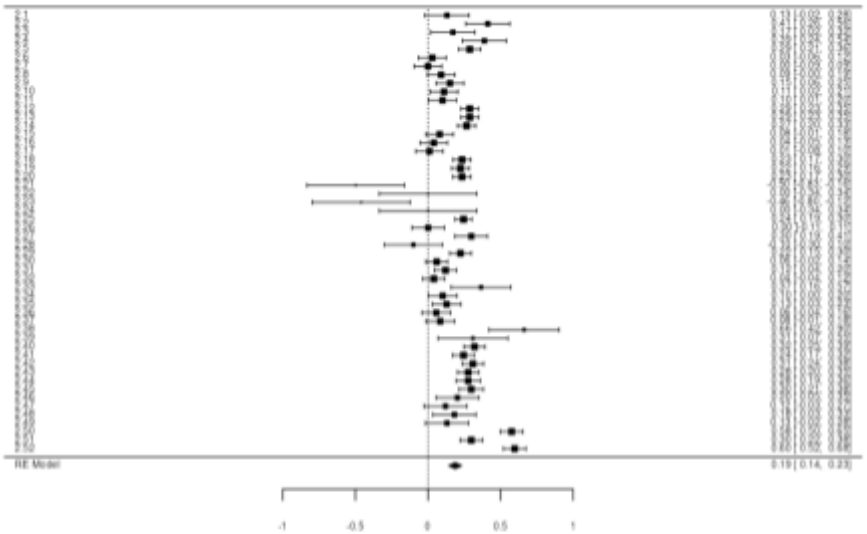

Chart 10 | A forest plot depicting the size effects of the analysis for Goal Orientation & Regulation and math

According to the results of the heterogeneity statistics the true outcomes appear to be heterogeneous (Q(51) = 473.6226, p < 0.0001, tau2 = 0.0234, I2 = 91.5331%). The Q statistic value of 473.6226 signifies a pretty strong heterogeneity seeing as the degrees of freedom is just 51. The small p-value insinuates that the variation in effect sizes is statistically significant. A Tau-Squared value of 0.0234 depicts some variability in the true effects across the studies. An I-Squared value of 91.5331% indicates that a very large proportion of the observed variance is not due to random error. The 95% prediction interval for the true outcomes is given by -0.1159 to 0.4910 which suggests that the majority of the studies depict a positive correlation and some illustrate a weak negative correlation. The residuals revealed that none of the studies had a value larger than ± 3.3015 and hence there was no indication of outliers. However, according to Cook's distances, several studies (2.21; 2.50; 2.52) could be considered to be overly influential. In this case the regression test indicated funnel plot asymmetry (p = 0.0074) but not the rank correlation test (p = 0.0825) (Chart 11).

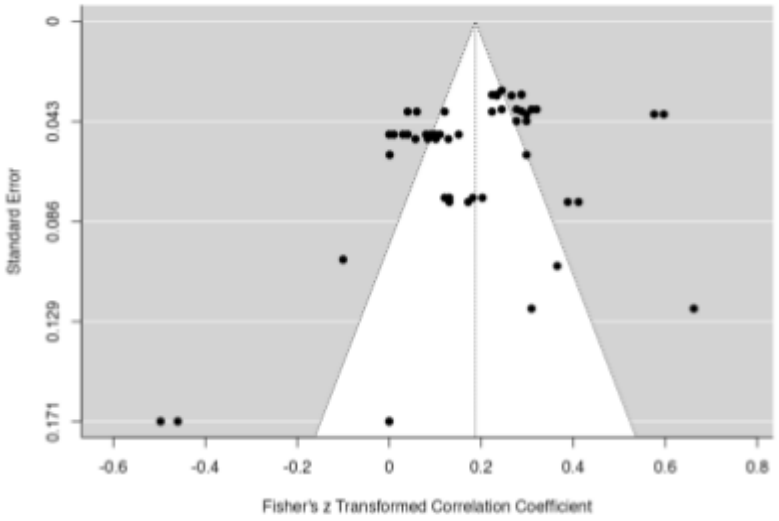

Chart 11 | Funnel plot for Goal Orientation & Regulation - Math Meta-Analysis

The relationship between Negative Tendency and Math

The meta-analysis of Negative Tendency correlations withs math scores was run with the 43 correlations across our 19 studies that measured children's Negative Affect, Negative reactivity and Anger. The observed Fisher r-to-z transformed correlation coefficients ranged from -0.3654 to 0.3575, with the majority of estimates being positive (93%). The estimated average Fisher r-to-z transformed correlation coefficient based on the random-effects model was -0.1066 (95% CI: -0.1431 to -0.0701). Therefore, the average outcome differed significantly from zero ($z = -5.7265$, $p < 0.0001$). The CI indicates that we can be 95% confident that the true average correlation is between -0.1431 to -0.0701, since this range does not include zero, it suggests that the negative relationship is statistically significant and consistent. The z-score of -5.7265 indicates that the average correlation is significantly different from zero. The very small p-value (< 0.0001) supports that this result is statistically significant, and that this negative correlation is not due to chance. The results concerning the correlation between our third temperamental factor and mathematical ability illustrates that a child with higher scores in the temperamental dimensions Negative Affect, Negative reactivity and Anger may be more likely to have lower math scores. The forest plot (Chart 12) illustrates the effect sizes from the studies, with the overall pooled effect size indicating a negative correlation.

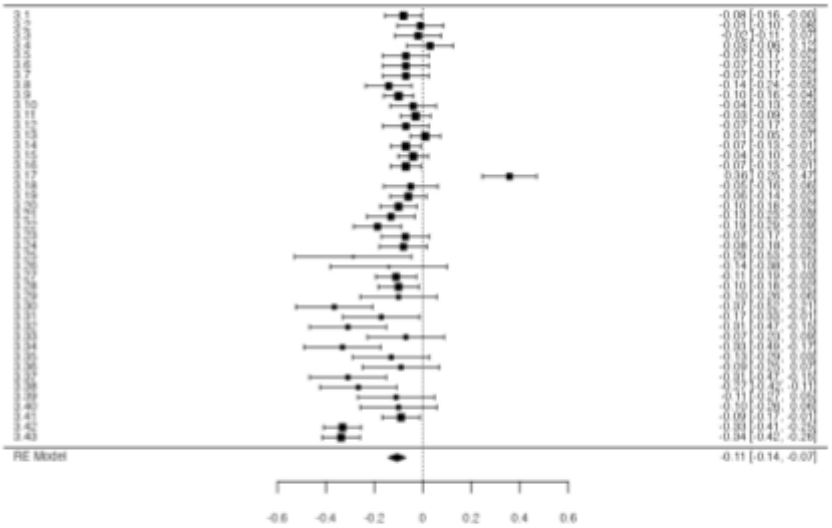

Chart 12 | A forest plot depicting the effects sizes of the meta-analysis of Negative Tendency and math

According to the results of the heterogeneity statistics the true outcomes appear to be heterogeneous (Q(42) = 217.8965, p < 0.0001, tau2 = 0.0116, I2 = 83.5457%). The Q statistic value of 217.8965 defines decent heterogeneity for a degree of freedom as small as 42. The small p-value suggests the variation in effect sizes is statistically significant. A Tau-Squared value of 0.0116 depicts some variability in the true effects across the studies. An I-Squared value of 83.5457% indicates that an adequate amount of the observed variance is not due to random error. The 95% prediction interval for the true outcomes is given by -0.3207 to 0.1074 which suggests that the majority of the studies lean toward a negative correlation but some have a positive relationship between the temperamental factors and math. The residuals revealed that one study had a value larger than ±3.2479 and may be a possible outlier in the context of this meta-analysis model. According to Cook's distances, several studies (3.17; 3.42; 3.43) could be overly influential to the results of this factor's analysis. Both the Rank and Regression tests indicated potential funnel plot asymmetry (p = 0.0041 and p = 0.0223) (Chart 13). Overall, the analysis of Factor 3 seems to be the most likely thus far to be affected by some publication bias or highly influential studies, shifting the overall effect of these temperamental dimensions on math achievement.

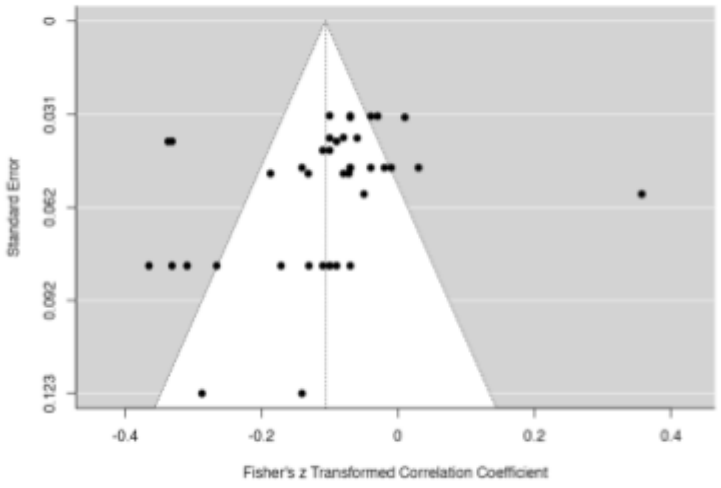

Chart 13 | Funnel plot for Negative Tendency - Math Meta-Analysis

The relationship between Positive Tendency & Social Skills and Math

The analysis of Positive Tendency & Social Skills with math scores consisted of 40 correlations from the total 19 studies that measured children's Adaptability/Agreeableness, Positive Affect, Positive Mood, Positive Reactivity, Mood and Resilience, Sociability, Affiliation and Soothability. The observed Fisher r-to-z transformed correlation coefficients ranged from -0.2554 to 0.8107, with the majority of estimates being positive (75%). The estimated average Fisher r-to-z transformed correlation coefficient based on the random-effects model was 0.1447 (95% CI: 0.0842 to 0.2052). Therefore, the average outcome differed significantly from zero ($z = 4.6847$, $p < 0.0001$). The CI indicates that we can be 95% confident that the true average correlation is between 0.0842 and 0.2052, since this range does not include zero, it suggests that the positive relationship is statistically significant and consistent. The z-score of 4.6847 indicates that the average correlation is significantly different from zero. The very small p-value (< 0.0001) supports that this result is statistically significant, and that this positive correlation is not due to chance. The forest plot (Chart 14) illustrates the effect sizes from the studies, with the overall pooled effect size indicating a positive relationship.

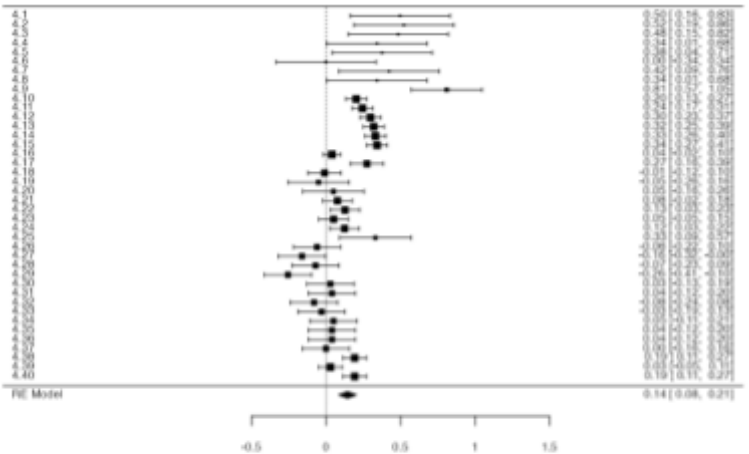

Chart 14 | A forest plot depicting the effects sizes of the meta-analysis of Positive Tendency & Social Skills

According to the results of the heterogeneity statistics the true outcomes appear to be heterogeneous (Q(39) = 282.7518, p < 0.0001, tau2 = 0.0304, I2 = 90.1420%) . The Q statistic value of 282.7518 suggests strong heterogeneity for a degree of freedom as small as 39. The small p-value suggests the variation in effect sizes is statistically significant. A Tau-Squared value of 0.0304 illustrates variability in the true effects across the studies. An I-Squared value of 90.1420% signifies that a large part of the observed variance is not due to random error. The 95% prediction interval for the true outcomes is given by -0.2021 to 0.4915 which suggests that the majority of the correlations are positive however, not all are. The studentized residuals revealed that one study (4.9) had a value larger than ±3.2272 and may be a possible outlier in the context of this model. According to Cook's distances, the same study (4.9) could be considered overly influential. Neither the Rank nor Regression tests indicated any potential funnel plot asymmetry (p = 0.8117 and p = 0.0701) (Chart 15).

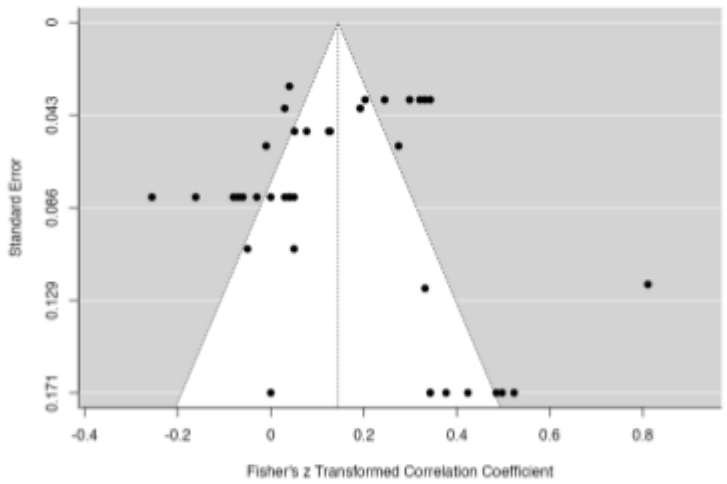

Chart 15 | Funnel plot for Positive Tendency & Social Skills - Math Meta-Analysis

Overall, the analysis of Positive Tendency & Social Skills indicated that the temperamental dimensions related to a tendency toward positive emotions and being socially adept generally is correlated with higher mathematical ability, however, there are instances in

which children who have a high score in these temperament dimensions may score lower in math.

The relationship between Behavioral Inhibition and Math

The final analysis for mathematical ability is looking at its correlation with our defined Behavioral Inhibition which consists of 23 correlations from the total19 studies that measured children's temperament scores in Shyness, Social Inhibition and general Inhibition. The observed Fisher r-to-z transformed correlation coefficients ranged from -0.2769 to 0.1923, with the majority of estimates being negative (52%). The estimated average Fisher r-to-z transformed correlation coefficient based on the random-effects model was -0.0263 (95% CI: -0.0847 to 0.0321). Therefore, the average outcome did not differ significantly from zero (z = -0.8827, p = 0.3774). The CI indicates that we can be 95% confident that the true average correlation is between -0.0847 and 0.0321, this range does include zero. This suggests that the relationship is not statistically significant. The z-score of -0.8827 indicates that the average correlation is not significantly different from zero. In this case the p-value also signifies non significance, suggesting that whatever correlation we may find between Factor 5 and math ability is either not large enough to signify any real effect or whatever effect there is may just be due to chance. The forest plot (Chart 16) illustrates how the effect sizes from the studies always surround zero and do not differ much. Although we can see a tendency toward a positive relationship, it was not significant.

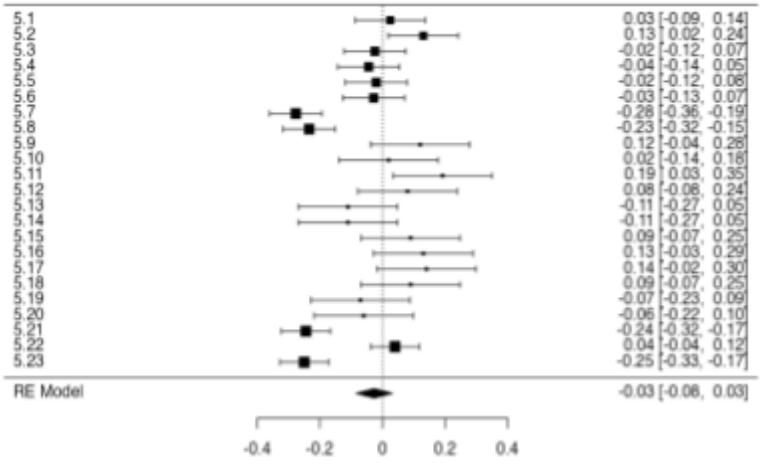

Chart 16 | A forest plot depicting the effects sizes of the meta-analysis of Behavioral Inhibition

According to the results of the heterogeneity statistics the true outcomes appear to be heterogeneous (Q(22) = 145.6970, p < 0.0001, tau2 = 0.0161, I2 = 82.8904%) . The Q statistic of 145.6970 is much larger than the degrees of freedom (22). This large value suggests that the variability in the effect sizes across the studies is much greater than what would be expected if the true effect sizes were homogeneous. The small p-value suggests the variation in effect sizes is statistically significant. A Tau-Squared value of 0.0161 depicts some variability in the true effects across the studies. An I-Squared value of 82.8904% illustrates that a relatively large part of the observed variance is not due to random error. The 95% prediction interval for the true outcomes is given by -0.2820 to 0.2294 which suggests that most of the correlations are negative but in general the distribution is quite balanced. The residuals revealed that none of the studies had a value larger than ±3.0654 and therefore there are no indications of the presence of an outlier in this model. According to Cook's distances, there are no studies that should be considered overly influential. Finally, both the Rank and the Regression tests indicated that there is potential funnel plot asymmetry (p = 0.0440 and p = 0.0007)(Chart 17).

From the analysis of all the created factors correlations with math ability and mathematical achievement, Behavioral Inhibition is the first to turn up as not significant indicating that Shyness, Social Inhibition and general Inhibition don't affect math performance from our exploration of previous literature.

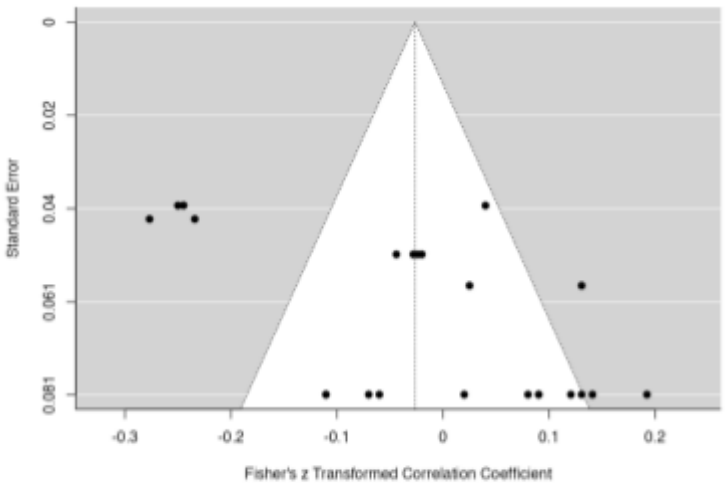

Chart 17 | Funnel plot of Behavioral Inhibition and Math Meta-Analysis

Meta-analysis results: Reading

The relationship between Hyperactivation & Distractibility and Reading

The meta-analyses run for reading followed the same approach as that which was done previously for math. The first analysis looked at the correlations of Hyperactivation & Distractibility with reading scores/ability through 28 correlations across our 19 studies. The observed Fisher r-to-z transformed correlation coefficients ranged from -0.4973 to 0.0400, with the majority of estimates being negative (93%). The estimated average Fisher r-to-z transformed correlation coefficient based on the random-effects model was -0.1407 (95% CI: -0.1927 to -0.0886). Therefore, the average outcome differed significantly from zero ($z = -5.2956$, $p < 0.0001$). The CI indicates that we can be 95% confident that the true average correlation is between -0.1927 to -0.0886, since this range does not include zero, it suggests that the negative relationship is statistically significant and consistent. The z-score of -5.2956 indicates that the average correlation is significantly different from zero. The very small p-value (< 0.0001) supports that this result is statistically significant, and that this negative correlation is not due to chance. These results depict a correlation between our first temperamental factor and reading ability which denotes that as the temperamental score of a child in regard to Activity, Impulsivity, Distractibility, Low Task Orientation Emotionality, Reactivity, Surgency and Intensity increase their performance in reading tasks decreases. The forest plot (Chart 18) illustrates the effect sizes from the studies, with the overall pooled effect size indicating a negative correlation.

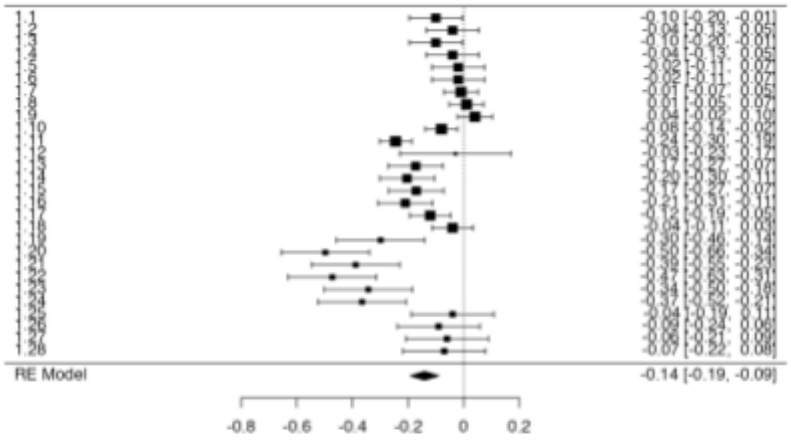

Chart 18 | A forest plot depicting the effects sizes of the meta-analysis of Hyperactivation & Distractibility and reading

According to the results of the heterogeneity statistics the true outcomes appear to be heterogeneous (Q(27) = 174.1866, p < 0.0001, tau2 = 0.0163, I2 = 87.7805%). The Q statistic value of 174.1866 defines decent heterogeneity for a degree of freedom as small as 27. The small p-value suggests the variation in effect sizes is statistically significant. A Tau-Squared value of 0.0163 depicts some variability in the true effects across the studies. An I-Squared value of 87.7805% indicates that a rather substantial amount of the observed variance is not due to random error. The 95% prediction interval for the true outcomes is given by -0.3964 to 0.1151 which suggests that the majority of the studies lean toward a negative correlation but some have a positive relationship between the temperamental factors and reading. The residuals revealed that none of the studies had a value larger than ±3.1237 and therefore there is no indication of an outlier in the context of this meta-analysis model. Cook's distances reinforce this as none of the studies should be considered as overly influential to the results of this factor's analysis. However, both the Rank test and the Regression test indicated potential funnel plot asymmetry (p = 0.0020 and p = 0.0011)(Chart 19).

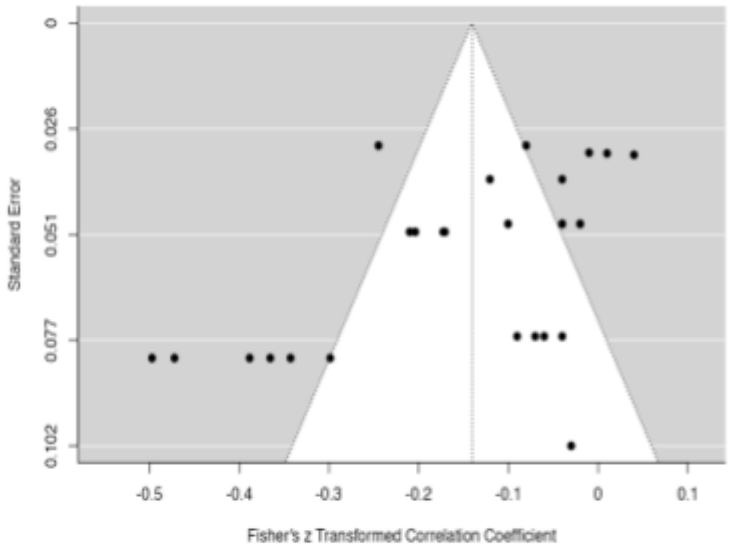

Chart 19 | Funnel plot for Hyperactivation & Distractibility - Reading Meta-Analysis

The relationship between Goal Orientation & Regulation and Reading

The analysis of reading assessments in correlation with Goal Orientation & Regulation was based on 26 correlations from our collected studies which measured children's Effortful Control, Attention, Persistence, Perseverance, Self-Regulation and Sensory Regulation. The observed Fisher r-to-z transformed correlation coefficients ranged from -0.1003 to 0.3884 with the majority of estimates being positive (92%). The estimated average Fisher r-to-z transformed correlation coefficient based on the random-effects model was 0.1621 (95% CI: 0.1201 to 0.2042). Therefore, the average outcome differed significantly from zero ($z = 7.5623$, $p < 0.0001$). The CI indicates that we can be 95% confident that the true average correlation is between 0.1201 and 0.2042, since this range does not include zero, it illustrates a statistically significant, positive correlation. The z-score of 7.5623 depicts an average correlation which significantly differs from zero. The very small p-value (< 0.0001) supports that this result is statistically significant, and that this positive relationship is not merely a result of chance. These results illustrate how as the temperamental scores increase, meaning that a child has a more developed Effortful Control, Attention, Persistence or Self-regulation the better they perform on reading assessments. The forest plot (Chart 20) illustrates the effect sizes from the correlations, with the overall positive relationship.

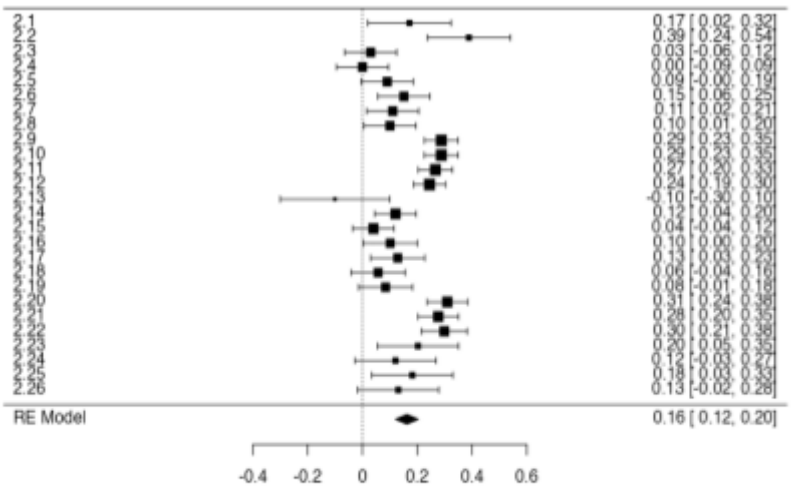

Chart 20 | A forest plot depicting the effects sizes of the meta-analysis of Goal Orientation & Regulation

The results of the heterogeneity statistics depict that the true outcomes appear to be heterogeneous (Q(25) = 137.8908, p < 0.0001, tau2 = 0.0092, I2 = 81.5769%). The Q statistic value of 137.8908 illustrates strong heterogeneity for a degree of freedom as small as 25. The small p-value suggests that the variation in effect sizes is statistically significant. A Tau-Squared value of 0.0092 depicts some variability in the true effects across the studies. An I-Squared value of 81.5769% indicates that a rather substantial amount of the observed variance is not due to random error. The combination of a high Q statistic, a very low p-value, a non-zero τ^2, and a high I^2 indicates significant heterogeneity among the studies included in the meta-analysis. The 95% prediction interval for the true outcomes is given by -0.0303 to 0.3545 which suggests that the majority of the studies lean toward a positive correlation but a few of the included have a positive negative between the temperamental dimensions and reading. Once again, the residuals revealed that none of the studies had a value larger than ±3.1019 and therefore there is no indication of an outlier. Cook's distances additionally depicted that none of the studies should be considered as overly influential. Finally, neither the Rank test nor the Regression test indicated potential funnel plot asymmetry (p = 0.1881 and p = 0.0881)(Chart 21). The overall results of this meta-analysis model illustrated how as the score for the temperamental dimensions in Goal Orientation & Regulation increased so did scores in reading, however this was accompanied by high scores of heterogeneity amongst the studies included in this model.

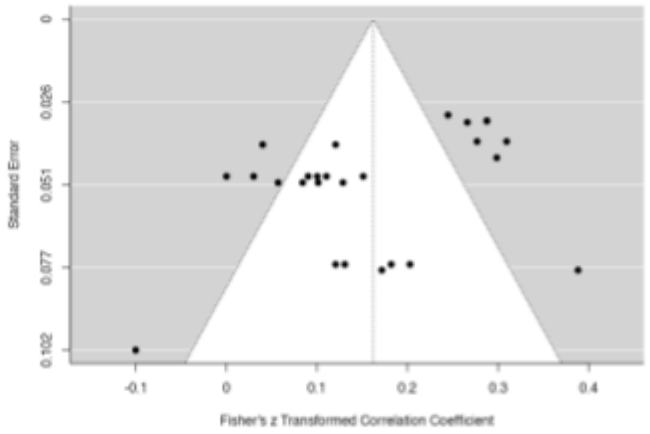

Chart 21 | Funnel plot for Goal Orientation & Regulation - Reading Meta-Analysis

The relationship between Negative Tendency and Reading

The meta-analysis of Negative Tendency correlations with reading consisted of 21 correlations across our 19 studies that measured children's Negative Affect, Negative reactivity and Anger. The observed Fisher r-to-z transformed correlation coefficients ranged from -0.3095 to 0.0300, with the majority of estimates being negative (90%). The estimated average Fisher r-to-z transformed correlation coefficient based on the random-effects model was -0.0812 (95% CI: -0.1107 to -0.0517). Therefore, the average outcome differed significantly from zero ($z = -5.3981$, $p < 0.0001$). This suggests that we can be 95% confident that the true average correlation is between 0.1107 and -0.0517, since this range does not include zero, it suggests that the negative relationship is statistically significant. However, the very small range between the upper and lower bound alludes to a small to moderate true effect. The z-score of -5.3981 indicates that the average correlation is significantly different from zero. The very small p-value (< 0.0001) supports that this result is statistically significant, and that this negative correlation is not a result of chance. These results suggest that as scores for Negative Tendency increase and a child displays more Negative Affect, Negative reactivity and Anger reading ability decreases, or at least the scores on reading assessments decrease. The forest plot (Chart 22) illustrates the effect sizes from the studies, with the overall pooled effect size indicating a negative correlation. As we can see from Chart 22 in this case the effect sizes did barely cross over into the realm of positive effects illustrating the consistency of this negative relationship.

According to the results of the heterogeneity statistics the true outcomes appear to be heterogeneous ($Q(20) = 39.4470$, $p = 0.0059$, $tau2 = 0.0021$, $I2 = 47.5698\%$). The Q statistic value of 39.4470 is still larger than the degrees of freedom (20) suggesting that there is some level of heterogeneity among the studies. However, it is not extremely large compared to the

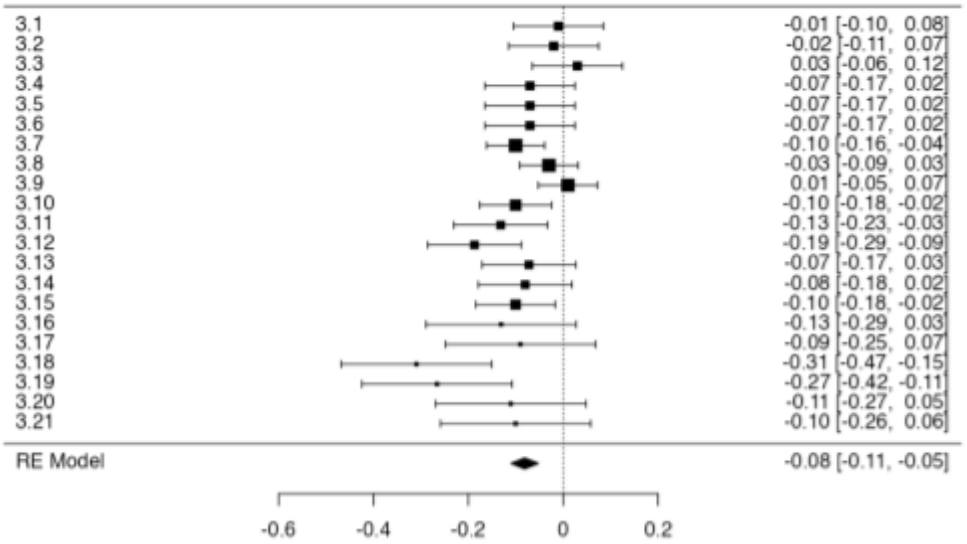

Chart 22 | A forest plot depicting the effects sizes of the meta-analysis of Negative Tendency

degrees of freedom, indicating that the heterogeneity is moderate. The p-value of 0.0059 is still below the threshold of 0.05 therefore the heterogeneity among the studies is statistically significant, meaning it is unlikely to be due to random chance alone. A Tau-Squared value is 0.0021, this small value indicates that there is some variability in the true effect sizes, but it is relatively modest. This suggests that while the effect sizes differ across studies, the magnitude of these differences is not very large. An I^2 value of approximately 47.6% implies that about half of the observed variability in effect sizes is due to actual differences between studies, while the other half may likely be due to sampling error. This level of heterogeneity is considered moderate. The 95% prediction interval for the true outcomes is given by -0.1755 to 0.0131 which suggests that more of the studies lean toward a negative correlation than those that have a positive relationship between Negative Tendency and reading. The examination of the residuals revealed that none of the studies had a value larger than ±3.0381 hence there being no indication of an outlier in this context. Cook's distances reinforce this as none of the studies should be considered as overly influential to the results of this factor's

analysis. However, both the Rank test and the Regression test indicated potential funnel plot asymmetry (p = 0.0210 and p = 0.0061) (Chart 23).

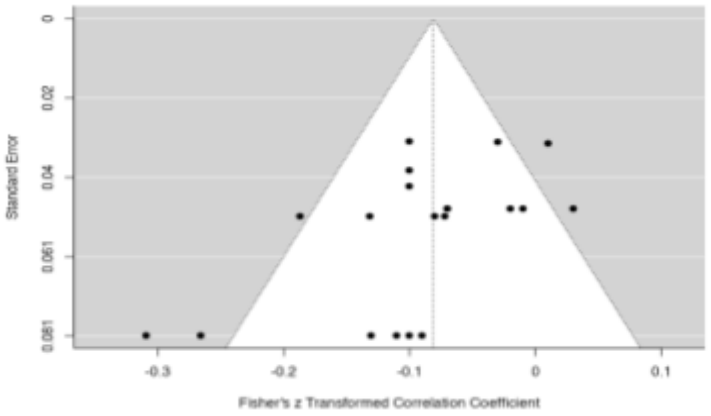

Chart 23 | Funnel plot for Negative Tendency - Reading Meta-Analysis

The relationship between Positive Tendency & Social Skills and Reading

The analysis of Positive Tendency & Social Skills with reading consisted of 15 correlations from the 19 studies that measured children's Adaptability/Agreeableness, Positive Affect, Positive Mood, Positive Reactivity, Mood and Resilience, Sociability, Affiliation and Soothability. The observed Fisher r-to-z transformed correlation coefficients ranged from -0.0802 to 0.8107, with the majority of estimates being positive (80%). The estimated average Fisher r-to-z transformed correlation coefficient based on the random-effects model was 0.1453 (95% CI: 0.0452 to 0.2454). Therefore, the average outcome differed significantly from zero (z = 2.8226, p = 0.0044). An average of 0.1453 indicates a small positive correlation across the studies. The fact that the entire interval is above zero indicates that the average effect size is positively correlated, and this positive effect is statistically reliable. The lower bound (0.0452) and the upper bound (0.2454) suggest that the true effect is likely small to moderate in magnitude. The z-score of 2.8226 indicates that the average correlation is significantly different from zero. A p-value of 0.0044 is well below the conventional threshold of 0.05, indicating that the average effect size is

significantly different from zero. The forest plot (Chart 24) illustrates the effect sizes from the studies, with the overall pooled effect size indicating a positive relationship.

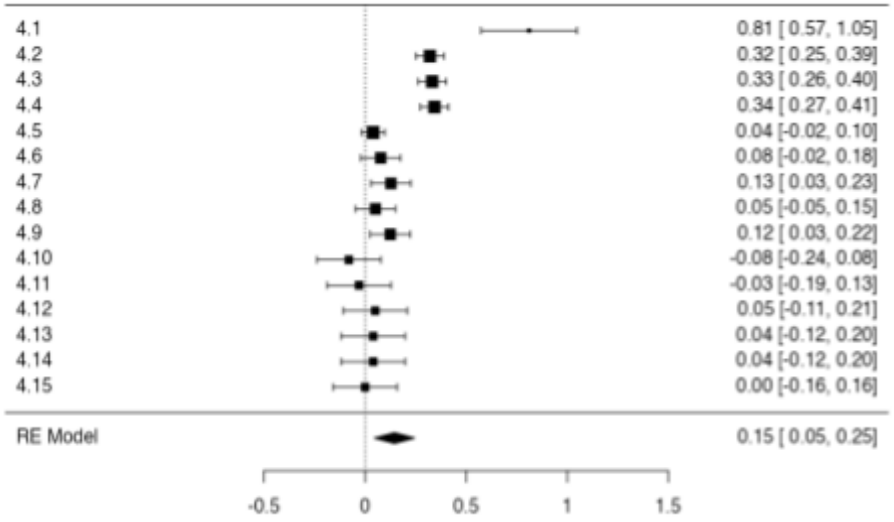

Chart 24 | A forest plot depicting the effects sizes of the meta-analysis of Positive Tendency & Social Skills

The results of the heterogeneity statistics the true outcomes appear to be heterogeneous (Q(14) = 146.6736, $p < 0.0001$, tau2 = 0.0348, I2 = 93.0800%). This Q value is much larger than the degrees of freedom (14), indicating substantial heterogeneity among the studies. The very low p-value suggests that this heterogeneity is statistically significant, meaning that the observed variability in effect sizes is unlikely due to random chance alone. A Tau-Squared value of 0.0348 suggests a moderate amount of variance in the true effect sizes. An I-Squared value of 93.08% indicates very high heterogeneity, meaning that most of the observed variability in effect sizes is due to differences between the studies. The 95% prediction interval for the true outcomes is given by -0.2340 to 0.5246, this wide interval suggests that while the average effect size is estimated to be positive, some studies could have a negative effect size, further indicating considerable uncertainty in predicting the outcomes. The residuals revealed that one study (4.1) had a value larger than ±2.9352 and may be a possible outlier in the context of this model. According to Cook's distances, the

same study (4.1) could be considered overly influential. Neither the Rank nor Regression tests indicated any potential funnel plot asymmetry (p = 0.4973 and p = 0.7230) (Chart 25), meaning there is no clear indication of publication bias in the studies analyzed. The results of the meta-analysis conducted for Positive Tendency & Social Skilss, concerning its effects of reading ability depict a positive correlation. As Adaptability/Agreeableness, Positive Affect, Positive Mood, Positive Reactivity, Mood and Resilience, Sociability, Affiliation and Soothability ratings on children increase so do their score on reading assessments, however these results seen slightly inconsistent and not necessarily stable even to conduct predictions.

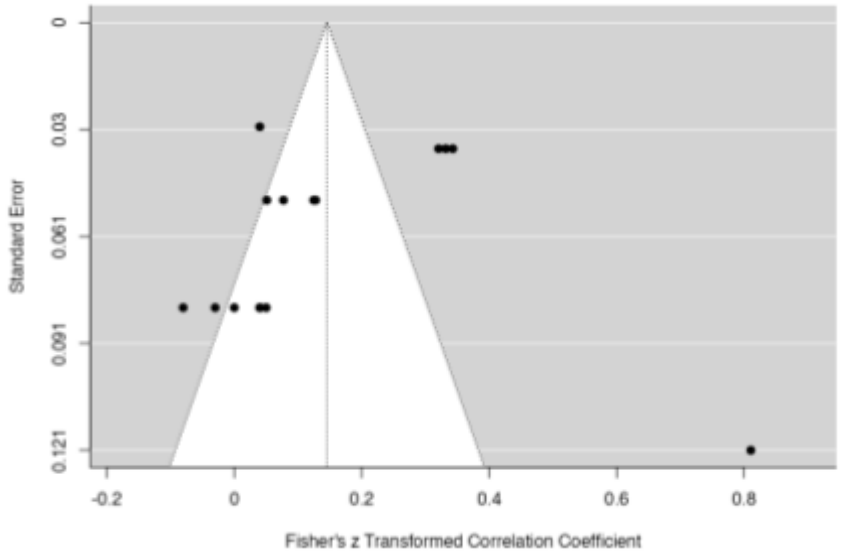

Chart 25 | Funnel plot for Positive Tendency & Social Skills - Reading Meta-Analysis

The relationship between Behavioral Inhibition and Reading

The final analysis for reading assessment/ability considers its correlation with our temperamental factor Behavioral Inhibition. This random-effects model considers 11 correlations from the total 19 studies, specifically those that measured children's Shyness, Social Inhibition and general Inhibition. The observed Fisher r-to-z transformed correlation coefficients ranged from -0.2342 to 0.1409, with the majority of estimates being negative (64%). The estimated average Fisher r-to-z transformed correlation coefficient based on the

random-effects model was -0.0137 (95% CI: -0.0808 to 0.0535). Therefore, the average outcome did not differ significantly from zero (z = -0.3988, p = 0.6901). The average effect size across the studies was slightly negative (-0.0137), but close to zero. This suggests that, on average, there is almost no relationship between Factor 5 and reading. The confidence interval includes zero, indicating that the true average effect could be slightly negative, slightly positive, or essentially zero. This suggests that the relationship is not statistically significant. The z-value is close to zero, and the p-value is 0.6901, which is higher than the conventional significance threshold (0.05), indicating that the average effect is not statistically significant and there is no strong evidence of a consistent effect across the studies. The forest plot (Chart 26) illustrates how the effect sizes from the studies always surround zero but do not depict a tendency toward one direction over the other, supporting the claims that there is no significant correlation present.

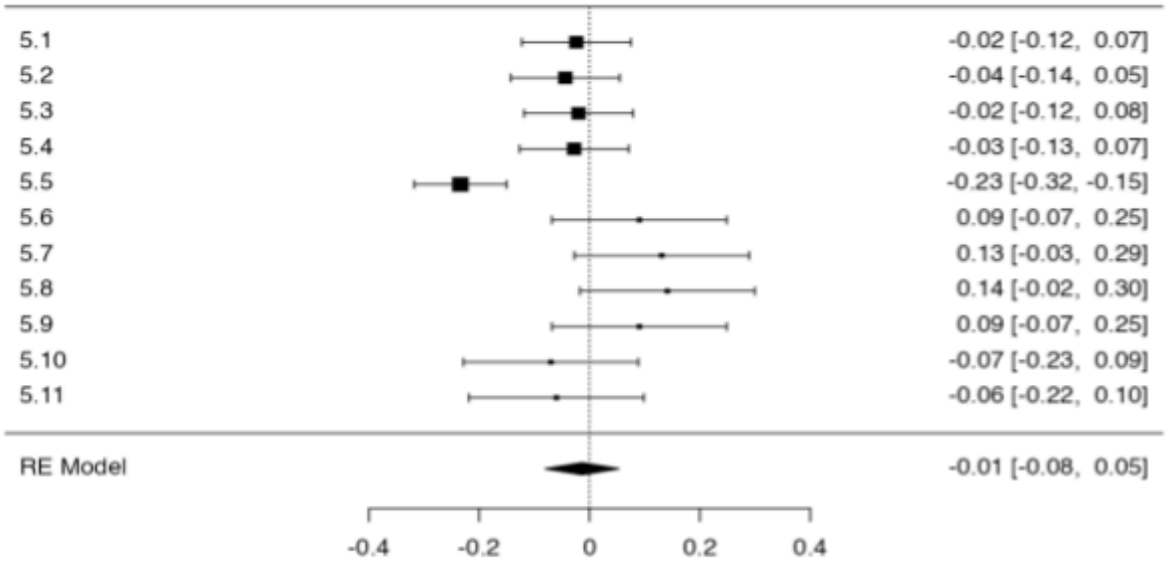

Chart 26 | A forest plot depicting the effects sizes of the meta-analysis of Behavioral Inhibition

According to the results of the heterogeneity statistics the true outcomes appear to be heterogeneous (Q(10) = 35.8429, p < 0.0001, tau2 = 0.0086, I2 = 69.7941%). The Q statistic

is significantly larger than the degrees of freedom (10), with a p-value of less than 0.0001, indicating substantial heterogeneity among the studies. This means the studies' results vary more than would be expected by chance alone. A Tau-Squared value of 0.0086 which is modest but suggests some variability in the true effect sizes across studies. An I-Squared of nearly 70% indicates that a substantial portion of the observed variability in effect sizes is due to true differences between the correlations, rather than random error. The 95% prediction interval for the true outcomes is given by -0.2072 to 0.1799 which includes both negative and positive values, suggesting that while the average effect size is slightly negative, some studies could show a positive effect, and others could show a more pronounced negative effect. The residuals revealed one study (5.5) that had a value larger than ±2.8376 and therefore suggesting the presence of an outlier in this model. According to Cook's distances, there are no studies that should be considered overly influential. The Regression test indicates some evidence of funnel plot asymmetry (p = 0.0073), suggesting potential publication bias. However, the Rank Correlation test does not indicate significant funnel plot asymmetry (p = 0.1107), so the evidence for publication bias is mixed.

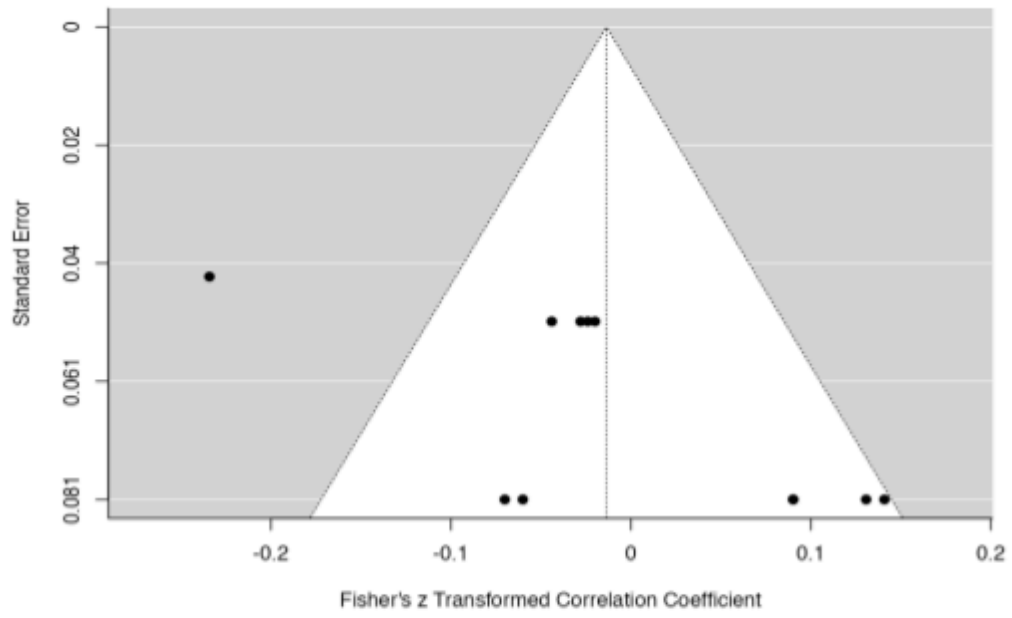

Chart 27 | Funnel plot for Behavioral Inhibition - Reading Meta-Analysis

Discussion

Peer-reviewed literature was systematically screened to create our inventory of temperamental dimensions observed to correlate with either math or reading ability. A total of 226 correlations were extrapolated, identifying 29 temperamental dimensions from which five temperamental factors were created based on the theoretical work and existing measures to reflect children's profiles. The information provided by these factors allowed for us to investigate the interaction between children's temperamental profiles and learning abilities in math and reading specifically, as well as provided an opportunity to identify some challenges and potential for future directions.

Hyperactivation & Distractibility

Our first factor, Hyperactivation & Distractibility consisted of eight temperamental dimensions; Activity, Impulsivity, Distractibility, Low Task Orientation, Emotionality, Reactivity, Surgency and Intensity. We had hypothesized that a child with high scores in these temperamental dimensions would have lower scores in both math and reading. We based this hypothesis off of previous literature that suggested that individually each of these traits was negatively correlated to either math or reading or both. A child with high Activity levels may display constant movement and restlessness which can interfere with the ability to concentrate, leading to missed instructions, incomplete tasks, and reduced practice time (Rudasill et al., 2010). Impulsivity may cause a child to rush through reading or math problems not separating the necessary steps, leading to mistakes and lower accuracy. High Distractibility means frequent shifts in attention resulting in a lack of deeper engagement with academic material. A Low Task Orientation, similarly to Distractibility, indicates a propensity for disruptive behaviors and difficulty staying focused on completing a task, often resulting in unfinished assignments, skipping challenging parts, or avoiding tasks altogether

(Bulotsky-Shearer et al., 2011). High Emotionality and Reactivity alike depict a tendency to become easily frustrated or anxious when faced with challenging math problems or difficult reading passages, potentially leading to avoidance behaviors and decreased persistence. Surgent children prefer interactive and dynamic environments that might make it harder for them to engage in solitary, focused academic tasks. Overall, children with high scores in Hyperactivation & Distractibility could face challenges interfering with their ability to focus, stay on task, manage emotions, and persist through difficulty, which can lead to lower scores in both math and reading as these subjects require sustained attention, careful processing, and the ability to manage frustration and maintain focus over time. Additionally, a study by Chauhan et al. (2019) depicted how children with Attention Deficit Hyperactivity Disorder (ADHD) were high on Activity level, Intensity of reaction, Approach, and Distractibility and low on Persistence and threshold of Responsiveness in their examination of the relationship between temperamental dimensions and ADHD. Their depiction of the temperament of a child with ADHD was similar to our definition of a Hyperactivation & Distractibility child, and according to Loe & Feldman (2007) (ADHD) is associated with poor grades, poor reading and math standardized test scores.

Our meta-analytic model depicts negative correlations between Hyperactivation & Distractibility and both math and reading skills. We obtained very similar results for both, the meta-analytic correlation coefficient for math was -0.1409 while for reading it was -0.1407. These results indicate that the temperamental dimensions defining Hyperactivation & Distractibility have a balanced and consistent effect on learning abilities. So we can speculate that these dimensions influence a specific competence underlying the learning abilities.

We can make the assumption from these results, and also based on the fact that previous literature also did not denote that any of these temperaments affect neither the cognitive abilities nor the motivations that are specific to either math or reading. The effect

that our first temperamental factor, Hyperactivation & Distractibility, has upon academic achievement may be a general one. The child's profile that is created by this factor describes an individual who has trouble sitting still, maintaining focus and controlling one's impulses, emotions and reactions, all of these characteristics combined depict a child who may struggle in a structured classroom environment, more so than one troubled in the development of math or reading skills. However, over time the hardship of the classroom may build up and result in these children actually falling behind in their development of mathematical and reading abilities if an appropriate intervention is not conducted in time. For example, a study conducted by DuPaul and colleagues (2006) observed equally beneficial effects of both Individualized Academic Intervention (data-based decision-making model that involved ongoing feedback to teachers) and Generic Academic Intervention (consultant-teacher collaboration) on children's mathematics and reading achievement. An example of an Individualized Academic Intervention that has been depicted as being an effective strategy in order to improve academic skills of pupils with ADHD is the Computerized Assisted Instruction (CAI) (Botsas & Grouios, 2017). However, even interventions as small as specific temperament-related strategies, such as exercise for children with high Activity levels can be beneficial in promoting active participation in school activities (Rothbart & Jones, 1999). Additionally, in both the cases of reading and math the majority of the correlation coefficients were negative, but both meta-analyses contained ranges that additionally depicted weak positive correlation coefficients for Hyperactivation & Distractibility. This suggested that despite a child pertaining to a high score regarding the temperamental dimensions of Hyperactivation & Distractibility not in every case will result in a decrease in math and reading ability. This may be a result of some children developing compensatory strategies or any sort of moderator error.

Goal Orientation & Regulation

Goal Orientation & Regulation is a factor created as a reversal of Hyperactivation & Distractibility, it has 3 dyads, Effortful Control and Attention, Persistence and Perseverance, and Self-Regulation and Sensory Regulation. The assumption or hypothesis for this factor anticipated a positive correlation with both math and reading. However, based on the majority of the literature supporting the importance of Self-Regulation and Persistence for math we point to a stronger correlation for mathematical ability, this is due to the fact that from the first moments in which basic math skills are taught to the final days of school math problems remain challenging as the mathematical taught become more complicated the more it is learned. This is why Self-Regulation and Persistence are beneficial for math for the entirety of schooling. However, reading is difficult whilst being learned and practiced but children reach a point in which this becomes an automated process, and from that point onward it is only the contents in the reading material that increases in difficulty. This is why we predict that performance tested for reading in the first year of primary school when learning to read is a competence that necessitates Attention and Persistence will be more strongly correlated to Factor 2 than it will be in later years. However, our meta-analysis does not divide the age group and academic levels, therefore this is something interesting to be addressed in the future. The results show a difference in meta-analytic correlation coefficients with reading results in 0.1621 and math 0.1876, despite the difference following the trend we expect there is not a significant difference between the correlation coefficients. The correlation coefficients themselves were significant, despite being on what could usually be considered the lower bound for significance, it is important to keep in mind how large our overall sample is. The meta-analytic model used to observe the relationship between Goal Orientation & Regulation and math included 52 correlations but just within these correlations 3542 participants were included. This could also be an explanation for the fact that for both math

and reading the lower bound for Fisher r-to-z transformed correlation coefficients ranged from weak negatives to the stronger positives in which the majority of the distribution lied. The results for both math and reading depicted how a child with a profile of an individual who can maintain focus, regulate their emotions and reactions and who can motivate oneself despite ambiguous feedback to continue will persevere in learning abilities.

Negative Tendency

In the overall collection of our data we found three temperamental dimensions that reflected a tendency toward negative emotions, the three dimensions Negative Affect, Negative Emotionality and Anger comprise Negative Tendency. The hypothesis for these correlations with math and reading were that there would be a negative relationship, because of the way that these traits can impact a child's ability to focus, process information, and engage with academic tasks. High levels of Negative Affect and Negative Emotionality can lead to specific types of academic anxiety, such as math anxiety or reading anxiety which can impair cognitive functioning, particularly working memory, which is essential for solving math problems and comprehending reading materials (Pelegrina et al., 2020)(Katzir et al., 2018). Children with high levels of Anger and Negative Emotionality may struggle to manage frustration when faced with challenging tasks which may lead to a lack of Persistence and avoidance behaviors. This was correct for both cases, however these analyses did have a substantial difference with math having an estimated average correlation coefficient of -0.1066 for 2594 participants while reading resulted in a correlation coefficient of -0.0812 with a sample of 2769. These results support our hypothesis that Negative Tendency would negatively correlate with academic performance, but it did depict that this relationship seemed to be stronger for math.

The implications of this could be due to the fact that math is an academic domain that relies more heavily on memory, rather than reading which later during the academic path becomes an automatic process. Previous research has depicted a negative correlation between negative emotions, specifically anger, anxiety and shame and math learning memory (Syawal et al., 2019). These results could be applied to interventions which utilize techniques such as mindfulness in school to help students decrease negative emotions and improve mathematical learning.

Positive Tendency & Social Skills

Positive Tendency & Social Skills is also, somewhat of a reversed factor of Negative Tendency consisting of temperamental dimensions that reflect positive emotions. However, this factor does include traits related to sociability as the majority of our theoretical background provided profiles of children who had a tendency towards positive emotions which improved their social relationships with peers and teachers alike. Positive Tendency & Social Skills consists of Adaptability/Agreeableness, Positive Affect, Positive Mood, Positive Reactivity, Mood and Resilience, Sociability, Affiliation and Soothability. Our hypothesis for Positive Tendency & Social Skills was supported by the results, in which the estimated average correlation coefficient for math and reading were positive and pretty similar. However, the correlation coefficient for reading was 0.1453 for 15 correlations with a sample of 2556. Meanwhile the correlation Coefficient for math was 0.1447 for 40 correlations across 2589 participants. The larger number of correlations provides a higher statistical power, increasing the likelihood of detecting true effects and provides more narrow confidence intervals for a more precise estimate of the pooled effect size. The fact that both of these meta-analyses had similar results despite the difference in sample size depicts the stability and consistency of the relationship between Positive Tendency & Social Skills and

the different domains of academic performance. This could imply that the effect is outside of a specific learning ability but rather the general influence of positivity and sociability on school performance.

Behavioral Inhibition

Behavioral Inhibition consists of three temperamental dimensions, Shyness, Social Inhibition and Inhibition. This factor is based on 23 correlations, from five studies, for both math and reading. This factor brought the most uncertainty in the process of setting our hypotheses as most of the previous literature explored factors such as Shyness and Inhibition in relationship with cognitive abilities, Theory of Mind, a child's social capacities and their motivations, but not much could be found concerning learning domains, such as math or reading abilities. Our prediction for this factor was that it would be negatively correlated to both math and reading. This was due to previous literature that illustrate that Shyness and behavioral inhibition are signs of psychosocial challenges and anxiety that has a vast impact on child's wellbeing, social networks and academic performance (Kalutskaya et al., 2015). The results of our meta-analyses were two non-significant negative correlations, indicating that despite the relationship between Positive Tendency & Social Skills and both math and reading being the expected negative, the relationship was not strong enough for us to eliminate that this was not just a matter of chance.

Limitations

The results of these meta-analyses should be interpreted with concern for certain limitations. Firstly, the large amount of initial search results gathered through the electronic databases, despite the limitations, in future research potential searches of gray literature should be included, as even we did find a dissertation addressing temperament and reading ability (Wilson, 2011). Second, the current state of non-uniformity when discussing

temperament may have caused us to miss articles that would have met our inclusion criteria. Despite the numerous and specific search terms used it is likely that due to a lack of common vocabulary in this research field, authors use different terms to refer to similar constructs somewhat interchangeably making it harder to conduct a comprehensive research . Thirdly, the present meta-analysis contains a great amount of sources of heterogeneity (which will be discussed below) whether that be concerning age groups, learning measures, definitions and measurement tools related to temperament. Finally, this meta-analysis did not analyze the effect of the covariates that could act as potential moderators in the investigated relationship, such as gender, age, SES, reporter bias, familial composition, maternal/paternal education and language/bilingualism.

Sources of Heterogeneity

Temperamental Dimensions Definitions

Despite the substantial number of articles that had been published in the last years regarding temperament and its relationship with academic achievement, math and reading abilities the lack of consistency and uniformity surrounding the temperament construct could be a potential limitation to the meta-analyses. Firstly, the amount of different temperamental dimensions (28) that we collect may not have been a true representation, as aforementioned, we did find that some dimensions with different names had overlapping definitions, or dimensions with different names used the same tool to measure what was essentially the same behaviors. Secondly, a variety of instruments were used to assess these temperamental traits, each looking for very specific tell-tale signs of behavioral correlates of temperament. This reflects the specific frameworks of reference of the questionnaires based on different theories. In order to control for this potential bias due to the heterogeneity of the reported temperamental dimensions, macro factors were created in which the dimensions with the

same basic construct but different labels were combined. This operation was done by basing the choice of dimensions to be combined on the comparison of definitions in the literature. This pre-processing was the only way to have the power and the number needed to conduct the analyses, but it was not without risk. Obviously, since the reference theories of temperament are different, even the temperamental dimensions do not seem to be perfectly comparable, this has inevitably introduced a part of error and heterogeneity.

Reporter Bias

Another potential limitation of this analysis is related to the reporters conducting the temperamental assessments. Throughout the 19 articles and 226 we have four different options for reporters used across all the data, the reports are usually either; Parents, Teachers, Self-Report or Psychologists. There was only one article which utilized trained psychologists (Blackson, 1995), which is an additional expense but may have served as a way to limit biasing results, specifically if conducted in laboratory settings (Vroman et al., 2014). The majority of the studies included parents as reporters (48.8%), but prior studies suggest that parental assessments of temperament are generally biased (Olino et al., 2020 & Kroes et al., 2006). This is due to the fact that parents' reports may be guided by anything from their own habituation to their children's behaviors, to their own psychopathologies, letting their own depressive or anxious tendencies influence how they perceive their child and their child's behavior. The second most common reporter throughout our studies was a teacher report (35.5%) of a child's temperament, however, a study by Tatalović Vorkapić (2017) depicted how teacher's assessments of a child's temperament are not independent from their own personality traits. The results of this study depicted how teachers who were highly extraverted, or teachers who had very little experience in teaching preschoolers, reported their students as being significantly more emotional. Additionally, preschool teachers who were found to be more emotionally unstable reported that their students were significantly less

active and sociable. This once more depicts how the assessment of a child's temperament may be altered by the personality or psychopathologies of the rater. Finally, seeing as we had a very wide age range of participants, the studies which we included that observed highschool students' temperaments utilized self-report methods (16.1%). The self-administered questionnaire method can only be used with older children, each specific instrument defines an age from which the questionnaire can be filled in directly by the child. For example, the EATQ questionnaire (Capaldi & Rothbart, 1992) can only be completed from 13 years onwards. Despite the fact that in this article there was a good convergence between parent and self-report of early adolescents, and the conclusion stating that the self-report provided a reliable and valid assessment, not much more advantages could be found concerning the use of self-reports of temperament.

Standardized Test Scores

The 19 studies included in this study come from six different countries and seven different states within the United States of America. This provides our meta-analysis with a level of diversity that is beneficial for supporting our claim that temperament, outside of being a stable trait, can also be a universal indicator of reading and mathematical ability. However, an issue that does arise from this cross-cultural data collection we acquired is that there is a great variability in learning assessment tests used. Despite this consideration, six out of the 19 studies included, have used subtests of the Woodcock-Johnson (WJ) Tests of Cognitive Abilities, specifically, the Psychoeducational Battery or the Woodcock Reading Mastery Tests, and the versions used depended on the time of publication of the article in question.

Moreover, another potential issue is represented by the fact that the majority of the other studies look at learning as a parallel to academic achievement through their own standardized assessments or grades. This meta-analysis includes five Finnish studies which

use their own school assessments of reading and math by simply reporting the end of year grade. Even outside of the Finnish studies, math grades were a commonly used assessment tool used to determine mathematical ability for a child, however, an end of year grade may not depict as much a child's specific math abilities but rather their performance as a student over the past year in their math class. This may be one of the reasons for the high levels of heterogeneity amongst this meta-analysis. In this study we have decided to consider learning domains separately by analyzing mathematics and reading in a disjoint way. This is because different skills underlie the acquisition of these specific abilities and we believe that they can be influenced differently by the facets of temperament. Simply using the final grade of the subject loses much of the specificity underlying the skills acquired and consequently the power to detect possible relationships with specific temperamental traits.

Age

This meta-analysis not only included a very large total of participants, 9847 to be exact, but included a very large age range. Seeing as we wanted to include school aged children, we have participants ranging from preschool (4 years old) to 12th graders (17 years old). This large variation in participant ages can explain the increased heterogeneity in effect sizes across studies. These age differences might have had an influence on the relationship between temperament and learning abilities as the results may vary significantly between preschool-aged children and adolescents, causing variability in the strength or direction of the observed effects. Despite this variability not being due to our dimensions as temperament is considered stable, all of the theories on which understanding of temperament is based acknowledge the influence of the environment. The difference between the social and educational environment between the first and last years of school could account for some of this variability, the fact that teaching becomes less individualistic and social relationships

become more complex can affect children's performance. Additionally, even though temperament remains stable, our two learning abilities do not. Mathematical ability and teaching follow somewhat of an exponential trajectory over the years, as children master more math skills the problems they face become harder and the following math skill will be more complex with new knowledge required. Additionally, as mathematical learning progresses mathematics change requiring new cognitive processes, as children move away from arithmetic, towards geometry, trigonometry and even calculus individual differences in processing may make their effect. Meanwhile reading follows a different trajectory, there are two main phases to consider, the phase of learning to read and the phase where reading is already automated. The influence of temperament may differ in a child depending on what phase of reading acquisition they are in.

Additionally, age can interact with other variables, such as gender, socioeconomic status and developmental stages, further increasing heterogeneity. This could have made it challenging to isolate the effect of our primary interest. The differences that occur at different developmental stages, whether that being cognitive, emotional, or social stage can affect how children perform in school or how certain temperamental traits manifest. For example, sociability and maintaining strong relationships with peers is more important for adolescents rather than for preschool aged children, therefore this dimension may be more pronounced and prioritized during these years. Future studies may want to perform separate analyses for different age groups to account for these developmental differences, or use age as a covariate to examine how it affects the relationship between temperament and academic achievement.

Conclusion

In conclusion, our systematic review and meta-analysis for temperamental dimensions and mathematical or reading ability identifies 19 articles, 226 correlations published over the last 29 years that depicted how different temperamental profiles of school aged children related to learning. To the best of our knowledge there is only one other meta-analysis that has been conducted which investigates the relationship between three temperamental dimensions (Effortful Control, Negative Affect and Surgency) and the general area of school achievement (Nasvytienė & Lazdauskas, 2021). In this review, Factor 1 contain Surgency as a dimension which unlike the work of Nasvytienė & Lazdauskas (2021) did depict a negative correlation with math and reading achievement, however our Factor 2 which contained Effortful Control and Factor 3 with Negative Affect supported their results. Our interest was to see if a child's temperament was related to their performance in reading and mathematical, the results depict trends of academic performance for all temperamental profiles created except for Factor 5. The main difference between Nasvytienė & Lazdauskas and our meta-analysis is the expanded observation of temperament. Their study solely focused on the aforementioned three dimensions while we broadened the allowed effect of temperament to the five factors which each included 3-9 narrower attributes of temperament. Despite the expansion of our perspective of temperament, future work should more broadly investigate temperamental dimensions as understanding the specific relationship between temperamental traits and academic learning could allow us to target the temperamental characteristics and try to adapt environmental strategies in order to help children acquire specific skills. Additionally, our focus on learning rather than academic achievement could not only allow for a better understanding of the milestones of acquired skills that lead to the development of mathematical and reasoning abilities, but could also aid in the understanding of academic achievement. Academic achievement is somewhat of an umbrella term as it involves

cognitive and non-cognitive factors such as intellectual level, personality, temperament, motivation, skills, interests, study habits, self-esteem, social and student-teacher relationships (Lamas, 2015). Finally, our meta-analysis did not perform a moderator analysis as Nasvytienė & Lazdauskas, which would be beneficial for the future. Overall, our study depicts the importance of temperament in the acquisition of academic learning as it offers stable profiles to observe the differences in acquisition of learning skills, the effects of school environment at different time points and the interaction with developmental stages.

Bibliography

Alatupa, S., Pulkki-Råback, L., Hintsanen, M., Ravaja, N., Raitakari, O. T., Telama, R., et al. (2010). School performance as a predictor of adulthood obesity: A 21-year follow- up study. European Journal of Epidemiology, 25, 267–274.

Auerbach, J. G., Zilberman-Hayun, Y., Berger, A., & Atzaba-Poria, N. (2019). Longitudinal Pathways to Reading Achievement at age 7 years: Child and environmental influences. *Reading Psychology*, *40*(3), 269–292. https://doi.org/10.1080/02702711.2019.1614128

Blackson T. C. (1995). Temperament and IQ mediate the effects of family history of substance abuse and family dysfunction on academic achievement. *Journal of clinical psychology*, *51*(1), 113–122. https://doi.org/10.1002/1097-4679(199501)51:1<113::aid-jclp2270510118>3.0.co;2-

Botsas, G., & Grouios, G. (2017). COMPUTER ASSISTED INSTRUCTION OF STUDENTS WITH ADHD AND ACADEMIC PERFORMANCE: A BRIEF REVIEW OF STUDIES CONDUCTED BETWEEN 1993 AND 2016, AND COMMENTS.

Bulotsky-Shearer, R. J., Fernandez, V., Dominguez, X., & Rouse, H. L. (2011). Behavior Problems in Learning Activities and Social Interactions in Head Start Classrooms and Early Reading, Mathematics, and Approaches to Learning. *School Psychology Review*, *40*(1), 39–56. https://doi.org/10.1080/02796015.2011.12087727

Calkins, S., & Degnen, K. (2005). Temperament in Early Development. In *Comprehensive Handbook of Personality and Psychopathology, Child Psychopathology* (p. 512). John Wiley & Sons.

Capaldi, D. M., & Rothbart, M. K., (1992) Development and Validation of an Early Adolescent Temperament Measure, The Journal of Early Adolescence, 12, 2, https://doi.org/10.1177/0272431692012002002

Caspi A, Shiner RL. (2006). Personality development. In: *Handbook of child psychology:Vol. 3. Social, emotional, and personality development.* 6th ed. Wiley; New York. pp. 300–364. Damon W, Lerner R, Eisenberg N.

Caspi, A., Wright, B. R. E., Moffitt, T. E., & Silva, P. A. (1998). Early failure in the labor market: Childhood and adolescent predictors of unemployment in the transition to adulthood. American Sociological Review, 63, 424–451.

Carey W. B. (1970). A simplified method for measuring infant temperament. *The Journal of pediatrics*, *77*(2), 188–194. https://doi.org/10.1016/s0022-3476(70)80322-5

Checa, P., Rodríguez‐Bailón, R., & Rueda, M. R. (2008). Neurocognitive and temperamental systems of self‐regulation and early adolescents' social and academic outcomes. *Mind, Brain, and Education*, *2*(4), 177–187. https://doi.org/10.1111/j.1751-228x.2008.00052.x

Chen, X., Rubin, K. H., & Li, B. -S. (1995). Depressed mood in Chinese children: Relations with school performance and family environment. Journal of Consulting and Clinical Psychology, 63, 938–947.

Chauhan, N., Shah, R., Padhy, S., & Malhotra, S. (2019). Relation between temperament dimensions and attention-deficit/hyperactivity disorder symptoms. *Industrial psychiatry journal*, *28*(1), 58–62. https://doi.org/10.4103/ipj.ipj_74_19

Coplan, R. J., Barber, A. M., & Lagacé-Séguin, D. G. (1999). The role of child temperament as a predictor of early literacy and numeracy skills in preschoolers. *Early Childhood Research Quarterly*, *14*(4), 537–553. https://doi.org/10.1016/s0885-2006(99)00025-3

Coll, C. G., Kagan, J., & Reznick, J. S. (1984). Behavioral inhibition in young children. *Child Development*, *55*(3), 1005. https://doi.org/10.2307/1130152

Cloninger C. R. (1987). A systematic method for clinical description and classification of personality variants. A proposal. *Archives of general psychiatry*, *44*(6), 573–588. https://doi.org/10.1001/archpsyc.1987.01800180093014

Cloninger, C. R. (1993). A psychobiological model of temperament and character. *Archives of General Psychiatry*, *50*(12), 975. https://doi.org/10.1001/archpsyc.1993.01820240059008

Cloninger (1994). *The temperament and character inventory (TCI): A guide to its development and use*. St. Louis, MO: Center for Psychobiology of Personality, Washington University. ISBN 978-0-9642917-1-3.

Cloninger, C. R., Cloninger, K. M., Zwir, I., & Keltikangas-Järvinen, L. (2019). The complex genetics and biology of human temperament: a review of traditional concepts in relation to new molecular findings. *Translational Psychiatry*, *9*(1). https://doi.org/10.1038/s41398-019-0621-4

Curtindale, L., Laurie-Rose, C., Bennett-Murphy, L., & Hull, S. (2007). Sensory modality, temperament, and the development of sustained attention: A vigilance study in children and

adults. *Developmental Psychology*, *43*(3), 576–589. https://doi.org/10.1037/0012-1649.43.3.576

De Fruyt, F., Bartels, M., Van Leeuwan, K. G., De Clercq, B., Mervielde, I., (2006) Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology.* 538–552.

DuPaul, G. J., Jitendra, A. K., Volpe, R. J., Tresco, K. E., Lutz, J. G., Junod, R. E., Cleary, K. S., Flammer, L. M., & Mannella, M. C. (2006). Consultation-based academic interventions for children with ADHD: Effects on reading and mathematics achievement. *Journal of Abnormal Child Psychology*, *34*(5), 633–646. https://doi.org/10.1007/s10802-006-9046-7

Dyson, M. W., Olino, T. M., Durbin, C. E., Goldsmith, H. H., Bufferd, S. J., Miller, A. R., & Klein, D. N. (2015). The structural and rank-order stability of temperament in young children based on a laboratory-observational measure. *Psychological assessment*, *27*(4), 1388–1401. https://doi.org/10.1037/pas0000104

Fauzi, A. & Widjajanti, D. B., (2018). Self-regulated learning: the effect on student's mathematics achievement. Journal of Physics: Conference Series, Volume 1097, The 5th International Conference on Research, Implementation, & Education of Mathematics and Sciences 7–8 May 2018, Yogyakarta, Indonesia.

Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the Revised Infant Behavior Questionnaire. Infant Behavior and Development, 26 (1), 64-86.

Guglielmi, R. S. (2008). Native language proficiency, English literacy, academic achievement, and occupational attainment in limited-English-proficient students: A latent growth modeling perspective. Journal of Educational Psychology, 100, 322–342.

Gutierrez-Zotes, A., Labad, J., Martorell, L., Gaviria, A., Bayón, C., Vilella, E., & Cloninger, C. R. (2015). The revised Temperament and Character Inventory: normative data by sex and age from a Spanish normal randomized sample. *PeerJ*, *3*, e1481. https://doi.org/10.7717/peerj.1481

Hipson, W. & Séguin, D. (2017). Goodness of Fit Model. 10.1007/978-3-319-28099-8_757-1.

Josefsson, K., Jokela, M., Cloninger, C. R., Hintsanen, M., Salo, J., Hintsa, T., Pulkki-Råback, L., & Keltikangas-Järvinen, L. (2013). Maturity and change in personality: developmental trends of temperament and character in adulthood. *Development and psychopathology*, *25*(3), 713–727. https://doi.org/10.1017/S0954579413000126

Kagan, J. (1969). The three faces of continuity in human development. In D. A. Goslin(Ed.), Handbook of socialization theory and research (pp. 26–74). Cambridge, MA: HarvardUniversity Press.

Kagan, J. (1997). Temperament and the Reactions to Unfamiliarity. *Child Development*, *68*(1), 139–143. https://doi.org/10.2307/1131931

Kagan, J., & Snidman, N. (1999). Early childhood predictors of adult anxiety disorders. *Biological psychiatry*, *46*(11), 1536–1541. https://doi.org/10.1016/s0006-3223(99)00137-7

Kalutskaya I, Archbell KA, Moritz Rudasill K, Coplan RJ. (2015) Shy children in the classroom: From research to educational practice. *Translational Issues in Psychological Science*.;1(2):149–57.

Katzir, T., Kim, Y.-S. G., & Dotan, S. (2018). Reading self-concept and reading anxiety in second grade children: The roles of word reading, emergent literacy skills, working memory and gender. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01180

Keogh, B. K. (2003). Temperament in the Classroom: Understanding Individual Differences. Baltimore, MD: Paul H Brookes Publishing

Kiff, C. J., Lengua, L. J., & Zalewski, M. (2011). Nature and nurturing: parenting in the context of child temperament. *Clinical child and family psychology review*, *14*(3), 251–301. https://doi.org/10.1007/s10567-011-0093-4

Lafontaine, M. P., Knoth, I. S., Lippé, S., (2020) Learning Abilities, In: Gallagher, A., Bulteau, C., Cohen, D., Michaud, J. L., Handbook of Clinical Neurology, Elsevier, Vol. 173, 241-254

Lamas, H. (2015). School Performance. Propósitos y Representaciones, 3(1), 313-386. doi: http:// dx.doi.org/10.20511/pyr2015.v3n1.74

Lavigne, J. V., Hopkins, J., Gouze, K. R., & Russo, J. (2023). Age 4 predictors of age 5 academic achievement: A multi-domain model of contextual, parent, and Child effects. *Child & Youth Care Forum*, *52*(5), 1145–1178. https://doi.org/10.1007/s10566-022-09728

Loe, I., M., & Feldman, H., M. (2007). Academic and educational outcomes of children with ADHD. *Journal of Pediatric Psychology*, *32*(6), 643–654. https://doi.org/10.1093/jpepsy/jsl054

Luby, J. L., Svrakic, D. M., McCallum, K., Przybeck, T. R., & Cloninger, C. R. (1999). The Junior Temperament and Character Inventory: Preliminary validation of a child self-report

measure. *Psychological Reports, 84*(3, Pt 2), 1127–1138. https://doi.org/10.2466/PR0.84.3.1127-1138*amp;*

Marjoribanks, K. (2005). Family background, academic achievement, and educational aspirations as predictors of Australian young adult's educational attainment. Psychological Reports, 96, 751–754.

Martin, R. P., Olejnik, S., & Gaddis, L. (1994). Is temperament an important contributor to schooling outcomes in elementary school? Modeling effects of temperament and scholastic ability on academic achievement. In W. B. Carey & S. C. McDevitt (Eds.), *Prevention and early intervention: Individual differences as risk factors for the mental health of children: A festschrift for Stella Chess and Alexander Thomas* (pp. 59–68). Brunner/Mazel.

Martin, R.P. (1994), Temperament assessment battery for children (TABC). Infant Ment. Health J., 15: 238-238. https://doi.org/10.1002/1097-0355(199422)15:2<238::AID-IMHJ2280150212>3.0.CO;2-K

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* 4:1. doi: 10.1186/2046-4053-4-1

Nasvytienė D, Lazdauskas T. (2021) Temperament and Academic Achievement in Children: A Meta-Analysis. *European Journal of Investigation in Health, Psychology and Education*. 11(3):736-757. https://doi.org/10.3390/ejihpe11030053

Newman, J., Noel, A., Chen, R., & Matsopoulos, A. S. (1998). Temperament, selected moderating variables and early reading achievement. *Journal of School Psychology*, *36*(2), 215–232. https://doi.org/10.1016/s0022-4405(98)00006-5

Olino, T. M., Guerra, K., Hayden, E. P., & Klein, D. (2020). Evaluating maternal psychopathology biases in reports of child temperament: An investigation of measurement invariance. *Psychological Assessment* , *32*(11). https://doi.org/10.31234/osf.io/tnwqs

Pelegrina, S., Justicia-Galiano, M. J., Martín-Puga, M. E., & Linares, R. (2020). Math anxiety and working memory updating: Difficulties in retrieving numerical information from working memory. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00669

Putnam, S. P., Gartstein, M.A., & Rothbart, M. K, (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. Infant Behavior and Development, 29 (3), 386-401.

Putnam, S.P., Rothbart, M.K. and Gartstein, M.A. (2008), Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. Inf. Child Develop., 17: 387-405. https://doi.org/10.1002/icd.582

Rothbart, M. K. (1981). Measurement of temperament in infancy. Child Development, 52, 569-578.

Rothbart, M. K. (2011). Becoming who we are: Temperament and personality in development. New York: The Guilford Press.

Rothbart, M. K., & Ahadi, S. A. (1994). Temperament and the Development of Personality. *Journal of Abnormal Psychology*, *103*.

Rothbart, M. K., Derryberry, D. (1981). “Development of individual differences in temperament,” in Advances in Developmental Psychology, Vol. 1, eds M. E. Lamb and A. L. Brown (Mahwah, NJ: Lawrence Erlbaum Associates), 37–86

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development*, *72*(5), 1394–1408. https://doi.org/10.1111/1467-8624.00355

Rothbart, M., K., Jones, L., B. (1999) Temperament: Developmental perspectives. In: Gallimore R, Bernheimer LP, et al., editors. *Developmental perspectives on children with high-incidence disabilities.* Mahwah, NJ: Lawrence Erlbaum Associates. pp. 33–53

Rudasill, K. M., Gallagher, K. C., & White, J. M. (2010). Temperamental attention and activity, classroom emotional support, and academic achievement in third grade. *Journal of School Psychology*, *48*(2), 113–134. https://doi.org/10.1016/j.jsp.2009.11.002

Schraeder, B. D., Heverly, M. A., & Rappaport, J. (1990). Temperament, behavior problems, and learning skills in very low birth weight preschoolers. *Research in Nursing & Health*, *13*(1), 27–34. https://doi.org/10.1002/nur.4770130106

Simonds, J. (2006). THE ROLE OF REWARD SENSITIVITY AND RESPONSE EXECUTION IN CHILDHOOD EXTRAVERSION. *Universal Dissertation, University of Oregon*

Svens-Liavåg, C., Korhonen, J., & Arnolds-Granlund, S.-B. (2023). School performance and behaviour profiles among Finnish adolescents, and their relations with temperament traits: A person-centred approach. *Scandinavian Journal of Educational Research*, *68*(4), 605–618. https://doi.org/10.1080/00313831.2023.2175244

Syawal Amran, M., Surat, S., & Rahman, S. (2019). Emotions in Learning Mathematics and Its relationship to Memory: Insight from Neuroscience to Classroom findings. *Religación*, *4*(18), 210-215. Retrieved from https://revista.religacion.com/index.php/religacion/article/view/421

Tatalović Vorkapić, S. (2017). Is the evaluation of children's temperament independent from Pre-school Teachers' personality? *The European Proceedings of Social and Behavioural Sciences*. https://doi.org/10.15405/epsbs.2017.01.02.10

Thomas, A. (1963). *Behavioral individuality in early childhood*. New York University Press.

Thomas, A., & Chess, S. (1977). *Temperament and development.* Brunner/Mazel.

Valiente, C., Doane, L. D., Clifford, S., Grimm, K. J., & Lemery-Chalfant, K. (2021). School readiness and achievement in early elementary school: Moderation by students' temperament. *Journal of Applied Developmental Psychology*, *74*, 101265. https://doi.org/10.1016/j.appdev.2021.101265

Viljaranta, J., Aunola, K., Mullola, S., Virkkala, J., Hirvonen, R., Pakarinen, E., & Nurmi, J. (2015). Children's temperament and academic skill development during first grade: Teachers' interaction styles as mediators. *Child Development*, *86*(4), 1191–1209. https://doi.org/10.1111/cdev.12379

von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental medicine and child neurology*, *49*(11), 868–873. https://doi.org/10.1111/j.1469-8749.2007.00868.x

Vroman, L. N., Lo, S. L., Durbin, C. E., (2014) Structure and Convergent Validity of Children's Temperament Traits as Assessed by Experimenter Ratings of Child Behaviour, Journal of Research in Personality, 52, 6-12

Wang, Z., Soden, B., Deater‐Deckard, K., Lukowski, S. L., Schenker, V. J., Willcutt, E. G., Thompson, L. A., & Petrill, S. A. (2015). Development in reading and math in children from

different ses backgrounds: The moderating role of child temperament. *Developmental Science*, *20*(3). https://doi.org/10.1111/desc.12380

Wilson, S. B. (2011). *Preschool Externalizing Behavior Mediates the Relation Between Temperament and Emergent Literacy* (dissertation).

Windle, M. (1992). Revised dimensions of temperament survey (DOTS--r): Simultaneous group confirmatory factor analysis for adolescent gender groups. *Psychological Assessment*, *4*(2), 228–234. https://doi.org/10.1037//1040-3590.4.2.228

Xiao, F., & Sun, L. (2021). Students' motivation and affection profiles and their relation to mathematics achievement, persistence, and behaviors. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.533593