# Self-Overestimation in Early Childhood and Beyond

Mengtian Xia

# Self-Overestimation in Early Childhood and Beyond

Mengtian Xia

ISBN: 978-94-6473-282-5

Cover: Jiaji Wang

Printing: Ipskamp Printing

DOI: https://doi.org/10.33540/2006

**Self-Overestimation in Early Childhood and Beyond**

Zelfoverschatting in de Vroege Kindertijd en Daarna

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht

op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling,

ingevolge het besluit van het college voor promoties

in het openbaar te verdedigen

op vrijdag 17 november des middags 2023 te 12.15 uur

door

Mengtian Xia

geboren op 13 april 1993

te Beijing, China

**Promotor:**

Prof. dr. S.C.E Thomaes

**Copromotor:**

Dr. A.M.G. Poorthuis

**Beoordelingscommissie:**

Prof. dr. E.H. de Bree

Dr. E. Brummelman

Prof. dr. J.J.S. Dubas

Prof. dr. T.A.J.M. van Gog

Prof. dr. C. Sedikides

## Contents

# Chapter 1

# General Introduction

Young children often exhibit striking self-confidence in the face of challenges. “I can remember all ten words here!” “I know exactly how it works!” “I can throw this ball to the very edge of the field!” Indeed, experts have argued that young children are generally overconfident in their abilities, understanding, and knowledge. For example, preschool and early school-age children have been found to perceive themselves as having greater motor skills (e.g., Plumert, 1995; Plumert & Schwebel, 1997), greater memory spans (e.g., Flavell et al., 1970; Yussen & Levy, 1975), a better understanding of how things work (Mills & Keil, 2004), greater depth of knowledge (Kominsky & Keil, 2014), and better mathematical abilities (Miller et al., 1988), than their objective abilities or performances warrant.

Thus, it seems that children are born great, in almost every sense of the word, at least in their own eyes. However, many questions about self-overestimation in childhood remain unanswered. For example, how is it associated with age, from early childhood to middle and late childhood? How universal is children’s self-overestimation? What psychological processes contribute to children’s self-overestimation? This dissertation aims to yield an in-depth understanding of children's self-overestimation in terms of its variation across tasks, age, culture, and historical time, as well as its psychological underpinnings.

## Variation in Children's Self-Overestimation Across Types of Tasks

Children's self-overestimation has been studied using a variety of tasks and activities (e.g., motor tasks, cognitive tasks); however, there is a lack of evidence as to whether or how children's self-overestimation varies across tasks.

Theoretically, the process of estimating performance on tasks that focus on motor abilities is somewhat different from the process of estimating performance on tasks that focus on cognitive abilities. Tasks that focus on motor abilities require primarily physical actions that involve the use of muscles and the coordination of bodily movements. Examples of such tasks that have been used in research into children’s self-overestimation include long jumps and ball throwing (Almeida, et al., 2017; Schneider, 1998). Compared to cognitive tasks, motor tasks are more dependent on perceptual information about the environment and the integration of sensory information from different sources, such as vision, touch, and proprioception (Gibson, 2014). Cognitive tasks primarily require mental activities such as attention, thinking, and decision-making. Examples of such tasks include memory and recall tasks and other complex tasks, such as solving math problems or listing differences in word pairs. Compared to motor tasks, completing cognitive tasks tends to be less connected to children’s external worlds, and mostly appeals to abstract processes in the mind (Wilson, 2002). Given these differences, it is possible that the extent to which children overestimate their performance will vary across types of tasks. However, this possibility has not yet been systematically evaluated.

## Age-Specificity of Self-Overestimation in Children

As children grow up, the ability to generate accurate self-estimates becomes increasingly important, as it enables them to make informed decisions and effectively accomplish their goals. Does children's self-overestimation disappear with age, over the course of childhood?

In a classic study (Yussen & Levy, 1975), participants in kindergarten, third grade, and adult groups all overestimated their performance on a memory task, but the extent of their self-overestimation suggested an age-related trend: by calculating the ratio of an individual's self-estimate to the corresponding measure of their actual performance, the self-overestimation effect was as high as 2.42 for 4-year-olds, dropped to 1.59 for 8-year-olds and further dropped to 1.06 for adults (i.e., approaching accuracy). Some later studies involving children in more than one age group found similar evidence, suggesting age-related change in self-overestimation on various types of tasks (Kominsky & Keil, 2014; Schneider, 1998; Shin et al., 2007; Was & Al-Harthy, 2018). Thus, it is possible that, as children grow up, they come to exhibit a lower degree of self-overestimation. Such an age-related trend would dovetail with improved (meta-)cognitive abilities, improved access to information to facilitate accurate self-estimates, and decreased adaptiveness of self-overestimation when children grow older.

Children's (meta-)cognitive abilities develop and improve gradually over the course of childhood (Coutinho et al., 2005; Krueger & Mueller, 2002). The older children are, the better they will be able to learn from experience, to monitor their performance, and to incorporate competency-related information into their performance estimates—acquired skills that, at least in theory, should reduce children's tendency to overestimate their task performance.

As children become older and progress in school, they will be exposed to more sources of information about their competence and achievements. For example, their performances in school are more frequently evaluated (e.g., in the form of regular normative assessment results, individualized feedback from teachers; Konold et al., 2004; Stipek & Iver, 1989). Also, in middle childhood, children learn to engage in social comparisons with peers, and to incorporate the outcomes of such comparisons into their self-views (Ruble et al., 1980; Van der Aar et al., 2018). These sources of information can impact children's self-views, and should enable and encourage children to generate more realistic self-estimates (Dweck, 2002).

It has been argued that self-overestimation may have adaptive value for young children, and optimizes children's opportunities to learn and develop (Bjorklund, 1997; Bjorklund & Green, 1992). Given that children often lack experience and expertise in most of the activities they pursue, having accurate self-estimates could potentially lead them to be reluctant to take on new challenges, or to give up easily on difficult tasks. Accordingly, it has been theorized that self-overestimation can provide children with a sense of competence and efficacy that motivates them to persist in the face of difficulty or failure, thus offering opportunities for growth (Shin et al., 2007).

Similarly, according to Bandura's view of self-efficacy (Bandura, 1982), one's self-perception of efficacy (or perceived mastery of action) drives action, regardless of its veracity. Thus, if children hold favorable views of themselves and their competencies, this may benefit their actual performance, even if these self-views are out of touch with reality.

Notably, it is possible that while self-overestimation may have some benefits for younger children, its adaptive value decreases as children grow older. For older children, accurate self-views should help children to choose tasks that are appropriate to their developmental level, to learn on what tasks and learning domains they need to invest extra effort, and more generally, to effectively cultivate their abilities (Escribano & Díaz-Morales, 2014; Gresham et al., 2000).

## Self-Overestimation of Children Across Cultures and Time

Developmental studies have traditionally focused rather narrowly on children growing up in Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies (Henrich et al., 2010). Much less is known about children growing up in other societies. Given that children from WEIRD societies constitute a relatively narrow subset of the global population, and one that possibly is rather peculiar in important respects, the breadth and diversity of human development has not yet been captured as well as it should (Henrich et al., 2010; Kline et al., 2018). This may be especially pertinent to the topic of children's self-views, which is the focus of the present dissertation. It has been observed that there are striking differences in self-construal and self-enhancement between individuals from different parts of the world, and it is possible that these differences are already manifest in the self-views of young children (Song & Wang, 2020; Wang, 2004).

**Independent Versus Interdependent Self-Views**

Studies in adults have shown that Westerners (i.e., North-Americans, Northern and Western-Europeans, Australians) tend to have more independent and less interdependent self-concepts than people living in other parts of the world (Hofstede 1980; Morling & Lamoreaux, 2008; Oyserman et al., 2002). In the West, society places more emphasis on the self, on the ways in which people differ from others, and on the importance of being assertive and self-reliant. Other societies, such as many Asian societies, place more emphasis on how people relate to others, on the ways in which people are interdependent, and on the importance of living in harmony with others (Markus & Kitayama, 1991; Triandis, 1989).

In the West, the normative imperative of culture emphasizes the discovery and expression of one's own unique attributes, independently of others (Hofstede et al., 2005; Johnson, 1985). People construct the self as an individual whose behavior is organized and motivated primarily by reference to his or her own internal repertoire of thoughts, feelings, and actions, rather than those of others. In short, Western society's view of the self derives from a belief in the wholeness and uniqueness of the individual's internal configuration of attributes, and accordingly promotes self-actualization and the development of one's distinct potential (Markus & Kitayama, 1991).

In contrast, the normative imperative in many non-Western cultures emphasizes the fundamental connectedness between individuals and the maintenance of interdependence (Gardner et al., 1999; Miller, 1988). People construct the self as part of an ensemble of social relationships—as such, the self is dependent upon the thoughts, feelings, and actions of the people they hold relationships with. This view of the self is thus relatively embedded in social context, and people are accordingly driven to fit in rather than stand out.

**Self-Enhancement Versus Modesty**

There is some evidence to suggest that self-enhancement (i.e., the tendency to view oneself more positively, or less negatively, than objective circumstances warrant; Alicke & Sedikides, 2009) is more pronounced in Western cultures than in non-Western cultures. For example, studies have found that Westerners are more likely to endorse positive self-views, while people from East Asian countries are more likely to show self-effacement (i.e., the tendency to minimize one's differences from others; Heine & Hamamura 2007; Mezulis et al., 2004).

This claim, however, is not without debate. Other scholars have argued that self-enhancement is an innate human drive that manifests in different ways, depending on cultural context. For example, according to Sedikides et al. (2003, 2007), individuals from Eastern cultures tend to engage in stronger self-enhancement for attributes and abilities that are valued in collectivist societies, whereas individuals from Western cultures tend to engage in stronger self-enhancement for attributes and abilities that are valued in individualistic societies. Thus, from this perspective, self-enhancement is strategically pursued by elevating oneself on significant attributes, but not on less significant ones (Alicke, 1985; Brown & Kobayashi, 2002).

The conceptual counterpart of self-enhancement is modesty, which refers to a tendency not to gloat about one's own performances, not to express confidence in one's own abilities, or to even engage in self-effacing or other-enhancing behaviors. Modesty can be seen as a self-presentational tactic that, around the world, manifests itself in a propensity toward non-boastful, attention-avoiding, and gracious social behavior (Cai et al., 2011; Chen et al., 2009). Although modesty is valued predominantly positively in most parts of the world, the modesty norm is especially potent and pervasive in East Asian cultures, including China—for example, it is reflected in Chinese curricula and teaching and parenting practices (Wang & Ollendick, 2001; Zhu & Chang, 2019). Chinese children are habitually socialized to avoid self-aggrandizement, not to brag about personal achievements, and to engage in self-effacement such as by minimizing their own competent or virtuous behavior and accomplishments (Lee et al., 1997; Wang & Ollendick, 2001).

**Cultural Variation in Child Socialization**

In early childhood, the parent-child relationship is a key context in which socialization occurs. Parents hold an influential role as socializers due to their primary caregiving responsibilities, control over their children's resources, and ability to shape their children's (social) environments (Grusec, 2011; Kuczynski, 2003; Stayton et al., 1971). For example, in a study involving third-grade children (mostly 8 years of age) and their mothers, it was found that children's sense of pride, a positive self-evaluative emotion, was fostered by mothers who placed a high priority on experiencing and expressing positive emotion. On the other hand, children whose mothers were more likely to emphasize negative emotions inhibited pride in their children (Hagan et al., 2021).

Socialization influences the extent to which children assimilate a culture's norms, beliefs, and values. Cultural socialization refers to the process through which parents communicate or convey cultural norms, beliefs, values (and associated customs and behaviors) to their children. From a young age, children internalize and adopt these messages, which thus influence them to conform to cultural expectations (Lee et al., 2006). Chen and French (2008) argue that cultural norms and values shape the ways in which children think, feel, and behave. This includes children’s self-representations: cultures prescribe different self-views as desirable, such cultural prescriptions transpire in the socialization messages (e.g., from parents, educators) that children are exposed to, which may then shape the ways in which children think of themselves from a young age (Luo et al., 2013; Wu et al., 2002; Xu et al., 2005).

For these reasons, it seems important to investigate whether children who grow up in cultures that stress the social norm of modesty, exhibit self-overestimation to a similar extent as their Western counterparts. We will do so in children growing up in China.

Cultural variation can be studied from a geographical perspective (i.e., by comparing groups of children growing up in different societies), which will be the focus of the research reported in this dissertation, but it can also be studied from a historical perspective (i.e., by comparing groups of children growing up in different times; e.g., Gentile et al., 2010). The past few decades have seen an increase in individualistic values and practices around the world, which translate into the ways in which adults (e.g., parents, teachers) communicate with and treat their children (Kashima & Kashima, 2003; Oishi, 2010; Santos et al., 2017). The first studies on children's self-overestimation already emerged from the late 1960s, which allows for an exploration of whether these historical time trends are reflected in the extent to which children have overestimated themselves across recent historical time.

## Psychological Underpinnings of Children's Self-Overestimation

Over the past decades, researchers have sought to identify the factors that contribute to children's self-overestimation, primarily focusing on (meta)cognitive developmental limitations. However, the evidence has been inconsistent. Considering the fundamental human drive for self-enhancement, one might ask whether children's self-overestimation is motivated, at least in part, by a *desire* to perform well and demonstrate competence.

### Immature (Meta)Cognition

One early explanation for children's self-overestimation centered on their lack of (meta)cognitive ability (reviewed in Bjorklund & Green, 1992). Children's self-estimation of task performance is based on their (meta)cognitive ability to process information from various sources, including their experience of performance on similar tasks, their general self-efficacy beliefs, and motivational factors (Harter, 2015). Metacognition refers to individuals' knowledge of their own cognition and the factors that influence it, involving awareness of their own abilities and use of strategies, as well as the monitoring and evaluation of their problem-solving (Bjorklund, 1997; Bjorklund & Green, 1992). According to this interpretation, children's limited cognitive and metacognitive abilities prevent them from effectively tracking and monitoring their own performance, and from incorporating ability- and performance-related information into their self-estimates (Flavell, 1979).

Initially, researchers tended to emphasize children's "monitoring deficiency" to account for self-overestimation, arguing that young children are not yet fully capable to monitor, realistically perceive, or retain information on their own performance and abilities (Schneider, 1998). Indeed, several empirical studies that involved multiple trials of a task partially supported this interpretation, finding that children's self-overestimation of task performance did not decline with experience across trials (Lipko et al., 2009; Schneider, 1998; Shin et al., 2007).

Other studies, however, have found that children appear to have competent monitoring abilities, yet they do not consistently incorporate relevant information (e.g., on their task performance) into their self-representations and adjust their self-overestimation. For example, when asked to recall their performance shortly after performing a task, even 4-year-old children were able to recall that information accurately, yet they remained overconfident when predicting their performance on the next trial immediately afterward (Lipko et al., 2009; Schneider, 1998). In specific situations, however, children do manage to retain relevant information and incorporate it into their self-estimates. In Lipko-Speed's (2013) experiment, when children engaged in a memory task in which they were provided with the exact same stimulus material across trials, they overestimated their performance on the second trial less than on the first trial. Furthermore, one study found that children who watched their peers fail on a set of motor tasks, subsequently made more conservative self-estimates than children who saw their peers succeed (Plumert & Schwebel, 1997). This suggests that children are able—at least in some conditions—to track and use the information conveyed by social comparison cues to adjust their own estimates of performance. Together, although the evidence is not conclusive, these findings do suggest that self-overestimation in childhood may partially stem from immature (meta)cognition, in particular incorporation inconsistency.

**Wishful Thinking**

One of the most widely endorsed assumptions about the self is that people are motivated to view themselves positively (Henrich et al., 2010; Zell et al., 2020). This is no exception among children, who exhibit multiple forms of self-enhancement (Harter, 1996, 2006; Trzesniewski et al., 2011). Self-overestimation can be seen as a form of self-enhancement by which children exaggerate the favorability of their self-views, as reflected in them estimating their abilities and task performance more positively than objective benchmarks justify. Thus, another perspective on the underpinnings of children's self-overestimation, emphasizes motivational processes and argues that children's self-overestimation comes (in part) from their desire to be and appear competent. This perspective has been labeled by some scholars as the "wishful thinking hypothesis," suggesting that children provide self-estimates based on how well they *want* to perform rather than how well they *are able to* perform (Schneider, 1998; Stipek et al., 1984).

As a self-enhancing motivation, it is logical that the desire for good performance should be reserved only for one's own performance, and should not extend to the performance of other people. A number of studies have supported this notion, finding that children tend to overestimate their own, but not their peers' performance (Lipko et al., 2009; Schneider, 1998; Stipek & Hoffman, 1980).

According to the wishful thinking hypothesis, children's self-overestimation should be malleable and dependent upon how much children desire to perform well. Unfortunately, there is a lack of direct evidence to test this assumption. However, one insightful study found that when receiving a reward was conditional on the good performance of peers (i.e., a manipulation designed to enhance children's desire for the good performance of their peers), children provided more inflated estimates of their peers' performance. In fact, they overestimated their peers' performance to the same degree as they overestimated their own performance (Stipek et al., 1984).

## Aims of This Dissertation

The purpose of this dissertation is to extend current knowledge of children's self-overestimation of performance and achievement. More specifically, the research in this dissertation is organized around the following five sets of aims.

### Aim 1. Do Children Overestimate Their Task Performance? To What Extent Do They Do So? Do They Overestimate Their Task Performance Equally Across Different Tasks?

This dissertation aims to determine the magnitude and robustness of children's self-overestimation. To do so, we (1) conduct two comparative cultural studies that investigate the robustness and magnitude of young children's self-overestimation (Chapter 2 and 3), and (2) synthesize the empirical evidence base on children's self-estimates and their corresponding actual task performance using a meta-analytic approach (Chapter 4). It is not yet known whether children self-overestimate to a similar degree on tasks requiring different domains of knowledge and ability. In the same meta-analysis, we will thus explore whether children's self-overestimation differs across motor and cognitive tasks (i.e., memory tasks, and other cognitive tasks; Chapter 4).

### Aim 2. How Does Self-Overestimation Vary by Age in Childhood?

There is suggestive evidence that children demonstrate diminished self-overestimation the older they are (Powel & Toni, 1994; Shin et al., 2007; Was & Al-Harthy, 2018). Nevertheless, the robustness and shape of this effect remain to be tested. Therefore, in the dissertation, we will (1) test if self-overestimation in childhood decreases with age, and (2) explore the possibility that children's self-overestimation decreases more sharply across certain ages (e.g., from preschool to school age) as compared to other ages. We will examine linear and nonlinear effects of age on children's self-overestimation effect using a meta-analytic approach (Chapter 4).

**Aim 3. How Does Children's Self-Overestimation Vary Across Historical Time?**

Over the past few decades, social changes such as economic wealth, the transition from agrarian to industrial and post-industrial economies, and the mobility of populations, have driven an increase in individualistic values and practices around the world (Kashima & Kashima, 2003; Oishi, 2010; Santos et al., 2017). These changes may affect children's development through changes in socialization practices, and could be reflected in how they perceive their abilities and performance. We will examine this possibility by exploring whether the extent to which children self-overestimate has changed over historical time—i.e., the past five decades (Chapter 4).

**Aim 4. Do Children in Non-Western Societies Overestimate Their Task Performance?**

Almost all research in the field so far has been conducted in children growing up in WEIRD societies. It is unknown whether children's self-overestimation generalizes to the different (i.e., non-WEIRD) cultural context of China. This dissertation explores this topic through two comparative cultural studies, which investigates cross-cultural differences in the magnitude of self-overestimation, as well as cross-cultural differences in the mechanisms underlying self-overestimation (Chapters 2 and 3) among Dutch and Chinese children.

**Aim 5. Why Do Children Overestimate Their Task Performance?**

This dissertation aims to examine the potential psychological underpinnings of self-overestimation in childhood: what psychological factors account for children's self-overestimation? One candidate factor is the immaturity of children's (meta)cognition, one other is children's motivation for being and appearing competent.

Children's estimates of their task performance depend on their cognitive abilities, especially their metacognitive abilities. Accordingly, it is possible that children's self-overestimation is (partly) due to their immature (meta)cognition. Some studies have found that, at least in some conditions, 4-year-olds can already monitor and retain information about their performance with relative accuracy (Lipko et al., 2009; Schneider, 1998). In this dissertation, we will test whether children use performance information and task experience to adjust their performance estimates toward accuracy (Chapter 2).

Another possibility, which emphasizes motivational processes, suggests that children's self-overestimation comes (in part) from their desire to perform well or, more generally, to be and appear competent. We will test this possibility in two ways. We will test whether children’s tendency to overestimate performance is limited to their own performance and does not extend to the performance of peers, which would suggest that children are actually able to make accurate performance estimates, but they just do not do so when it comes to their own performance (suggestive of a motivational drive to self-overestimate; Chapter 2). We will also test if children self-overestimate less when faced with a competing motivation to provide accurate self-estimates (i.e., when they are promised a reward for providing accurate self-estimates), which would again suggest that children's self-overestimation is not *solely* attributable to immature (meta)cognition (Chapter 3).

**Study Samples and Designs**

This dissertation reports on two primary empirical studies involving data from four samples, and a meta-analytic study involving data from 39 publications of empirical studies. For each of these studies (primary and meta-analyzed), biases in children's self-estimates were examined in a task-specific context by comparing differences between children’s subjective, prospective self-estimates and corresponding objective measures of actual performance.

**Participants**

The primary research reported on in this dissertation includes data from participants growing up in the Netherlands and their counterparts growing up in China.

The Netherlands is a country that aligns with the WEIRD profile (Henrich et al., 2010). For example, the Netherlands is a Western-European society that places a high value on individual independence (Individualism Index = 80; Hofstede et al., 2005). China is different along important cultural, political, and economic dimensions. For example, Chinese society tends to value the interdependence among individuals (Individualism Index = 20; Hofstede et al., 2005).

Chinese society is strongly influenced by Confucianism, and accordingly, emphasizes the social norm of modesty. Experts have explained the phenomenon of modesty in the Confucian context from the perspective of "saving others' face" (Bond et al., 1982; Leung, 1996). According to this perspective, individuals are expected to downplay their own accomplishments to prevent others from experiencing feelings of being threatened (i.e., saving others' face), thereby maintaining interpersonal harmony (Han, 2011). Additionally, while self-enhancement exists in China (Falbo et al., 1997; Farh et al., 1991), it is often engaged in in a strategic manner (i.e., "tactical self-enhancement"; Kanagawa et al., 2001; Kitayama et al., 1997), in light of the prevailing norm of modesty. For example, studies have found that in situations where modesty is valued or contextually encouraged, Chinese individuals tend to downplay explicit self-positivity; nevertheless, they still enhance the self, through implicit means by adopting modest attitudes or engaging in behavior that aligns with the contextual norms (Cai et al., 2011).

Children ages 4 to 5 were selected as participants for this dissertation. First, previous research has suggested that children in this age group exhibit pronounced self-overestimation. We aim to replicate this phenomenon and investigate whether it holds in non-Western populations as well (i.e., Chinese children). Second, this age group presents a unique opportunity for understanding the early manifestations of children's self-representations: young children are in the process of developing the cognitive abilities that allow them to form representations of their own attributes and skills (Harter, 2006). As such, we examine children's self-estimates at a time that they begin to form.

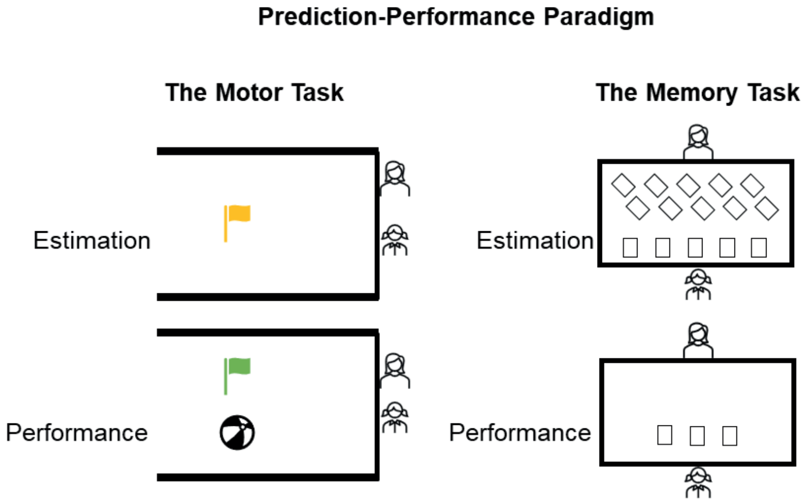
**Research Design of Primary Empirical Studies**

The primary research reported on in this dissertation will employ an adapted prediction-performance paradigm to assess children's self-estimation, applied in both a ball-throwing task and a picture-remembering task. To account for the cognitive abilities of participants at this age and the potential interference of language in cross-national studies, we assess children's self- (and other-) estimates behaviorally (Figure 1). Specifically, for the ball throwing task, children mark their estimates of performance by placing a flag where they think their (or another child's) ball will land; and for the memory task, children leave blank cards to mark their estimates of how many pictures they (or another child) will recall correctly.

We will use instructions and choose tasks within the motor and cognitive domains to ensure that the tasks are developmentally appropriate for preschool-aged children in both countries. For example, for the ball-throwing task, the experimenter models a normative posture for throwing the ball to prevent children from using ball throwing strategies that would limit their performance. For the memory task, we use pictures of objects (e.g., cat, pencil, bike) that are familiar to both Dutch and Chinese young children.

**Figure 1**

*Visualization of Task Procedures Used in Chapters 2 and 3*




The empirical studies will use a multiple-trial design to obtain multiple measures of both children's estimated and actual performance, which will allow us to explore how children's self-overestimation varies across trials (and thus, to what extent children learn from task experience and feedback). In one of the empirical studies (reported in Chapter 3), we will employ experimental research methods to obtain causal evidence on the wishful thinking hypothesis for the psychological underpinnings of children's self-overestimation. Specifically, we will experimentally manipulate a competing motivation (i.e., obtaining a reward) for children to provide accurate performance estimates. If this incentive leads children to provide more accurate self-estimates, this would suggest that their self-overestimation is at least partially motivated.

## Outline of This Dissertation

This dissertation reports on two cross-national empirical studies in Chapters 2 and 3 and a meta-analysis in Chapter 4, designed to contribute to the dissertation's aims (Table 1).

The chapters address questions about the robustness and magnitude (Chapters 2, 3, and 4), variation across age and historical time (Chapters 4), and cross-cultural generality and psychological underpinnings (Chapters 2 and 3) of self-overestimation of task performance in childhood. **Chapter 2** examines (1) whether and to what extent children overestimate their task performance, and (2) whether self-overestimation differs between Dutch and Chinese children. In addition, it scrutinizes the immature (meta)cognition hypothesis by tracking changes in children's self-estimates over multiple trials, i.e., by examining whether children are able to effectively incorporate task experience and feedback into their self-estimates. It also explores the wishful thinking hypothesis, by examining differences in children's estimates of their own performance and that of others. **Chapter 3** is also conducted in samples of Dutch and Chinese children. It builds on Chapter 2 and tests the wishful thinking hypothesis of children's self-overestimation by setting up experimental conditions to examine whether children lower their self-overestimation when a competing motivation for estimation accuracy is activated. Again, in the same chapter, self-overestimation is tracked across trials to examine if children incorporate task experience and feedback into their self-estimates. **Chapter 4** builds on Chapters 2 and 3 by synthesizing the available evidence on children's self-overestimation in task performance settings. It aims to systematically determine the magnitude and robustness of self-overestimation in childhood across various types of tasks, age, and historical time.

For the primary studies reported in this dissertation, we preregistered the research design, hypotheses, and analytical approach (Chapter 2: https://aspredicted.org/tu8ib.pdf; Chapter 3: https://aspredicted.org/j9hr2.pdf). The research materials, syntax, and data can be found at https://osf.io/kmjgz/ (Chapter 2), https://osf.io/j9kdx/ (Chapter 3), and https://osf.io/78dyw/ (Chapter 4).

**Table 1**

*Outline of Empirical Chapters and the Dissertation Aims*

| Chapter | Sample | Age | Design | Type of Tasks | Measures | Aim 1. Magnitude and Robustness | Aim 2. Age-Trends | Aim 3. Time-Trends | Aim 4. Cross-Cultural Generality | Aim 5. Psychological Underpinnings |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | Samples 1 and 2 N = 194 | 4 and 5 years | Prediction-performance paradigm; Multiple-trial | Motor task | 4 Self-estimates, 4 other-estimates, and 3 actual performances | ✓ | | | ✓ | ✓ |
| | Samples 1 and 2 N = 191 | | | Memory task | 4 Self-estimates, 4 other-estimates and 3 actual performances | ✓ | | | ✓ | ✓ |
| 3 | Samples 3 and 4 N = 187 | 4 and 5 years | Prediction-performance paradigm; Multiple-trial | Motor task | 4 self-estimates and 3 actual performances | ✓ | | | ✓ | ✓ |
| | Samples 3 and 4 N = 189 | | | Memory task | 4 self-estimates and 3 actual performances | ✓ | | | ✓ | ✓ |
| 4 | Sample 5 K = 217 | 4 to 12 years | Meta-Analysis | Motor, memory and other cognitive tasks | Overall effect size (self-overestimation effect) | ✓ | | | | |
| | | | | | Moderation by mean age of the samples | | ✓ | | | |
| | | | | | Moderation by task type | ✓ | | | | |
| | | | | | Moderation by year of study | | | ✓ | | |

# Chapter 2

## Young Children's Overestimation of Performance:

## A Cross-Cultural Comparison

Published as: Xia, M., Poorthuis, A. M. G., Zhou, Q., & Thomaes, S. (2022). Young children's overestimation of performance: A cross-cultural comparison. *Child Development, 93(2)*, e207–e221. https://doi.org/10.1111/cdev.13709

Author contributions

MX, AP, and ST conceptualized the study. MX, AP, and QZ were responsible for data collection. MX analyzed the data and wrote the first draft of the manuscript. AP and ST provided feedback on the manuscript.

**Abstract**

Western literature suggests that young children overestimate their performance across a range of tasks. Research in non-Western cultures, however, is lacking. In 2019, 101 Chinese (52% girls) and 98 Dutch (49% girls) children, ages 4 and 5, were asked to estimate how well they would perform on both a motor and a memory task. Children from both countries overestimated their performance to the same extent ($\eta_p^2$ = .077 and .027 for the motor and memory tasks, respectively). They generally persevered in doing so despite receiving realistic performance feedback. Yet, children overestimated their peers' performance about as much as their own performance, in some cases even more. This is the first demonstration of performance overestimation in children growing up in a non-Western culture.

*Keywords:* overestimation, self-perception, cognitive bias, cultural comparison

# Young Children's Overestimation of Performance: A Cross-Cultural Comparison

A Chinese proverb says that newborn calves are not afraid of tigers. Similarly, young children often seem undeterred by unfamiliar tasks and challenges. Research has shown that, in early childhood, children often feel overconfident about managing new tasks and challenges, and overestimate their competencies and performance (Lipko et al., 2009; Plumert, 1995; Shin et al., 2007; Yussen & Levy, 1975). However, this research has been conducted nearly exclusively in samples of children growing up in Western cultures, a limitation that applies to much of developmental science (Nielsen et al., 2017). Thus, the cultural generalizability of these results is yet unknown. This is important, especially in light of differences in the cultural values of self-enhancement and modesty in Western and East Asian cultures. Here, we ask to what extent young children's self-overestimation and its underlying psychological mechanisms generalize to children growing up in China. We investigate this question using a series of structured observations in young children growing up in mainland China and, as a comparison, their counterparts growing up in the Netherlands.

## Self-Enhancement and Modesty Across Cultures

In general, culture is an important source of psychological and behavioral variation, and in particular in terms of self-development (Henrich et al., 2010; Kline et al., 2018; Markus & Kitayama, 1991; Triandis, 1989; Wang, 2006). Culture prescribes what is a "good person," and cultural members, including children, try to live up to that ideal (Bornstein & Cheah, 2006; Gaertner et al., 2008; Triandis, 1989). In Western cultures (e.g., the United States, Northern Europe), social norms emphasize the importance of positive distinctiveness and personal success (e.g., Sedikides et al., 2015). Children are exposed, from a young age, to messages that convey it is ideal for them to be unique and stand out from others (Gürel & Brummelman, 2020; Thomaes et al., 2017; Young-Eisendrath, 2008). For example, such messages are communicated through mass media emphasizing the importance of being "special," adults encouraging social comparison and competition (e.g., in sports), and educational practices in schools such as singling out good performance (Gürel et al., 2020). Practices such as these both reflect and feed culturally shared ideals of attaining independence and agency in Western cultures.

In East Asian cultures (e.g., China, Japan), social norms more often emphasize the importance of interpersonal cohesion and harmony, of "fitting in" rather than "standing out" (Markus & Kitayama, 1991). Indeed, reflecting the Confucian proverb "haughtiness invites loss while modesty brings benefits," modesty is a prevailing social norm in these cultures. As a disposition, modesty reflects a tendency for individuals to downplay their abilities or achievements, or at least refrain from self-aggrandizement in order to maintain or promote social bonds (Kim et al., 2010; O'Mara et al., 2012). East Asian children are often familiarized with modesty as a social norm from an early age. For example, from early childhood, they learn not to present themselves to others in overly flattering ways, and to exercise restraint in communicating their accomplishments or good performance to others (Luo et al., 2013; Wang & Ollendick, 2001; Wu et al., 2002; Xu et al., 2005). These socialization practices reinforce culturally shared ideals of personal integration and social connection.

From middle childhood, children learn to reason about modesty as a self-presentational tactic that can benefit others' evaluations of the self (Watling & Banerjee, 2007; Yoshida et al., 1982). Cultural differences in such reasoning emerge from this age. For example, East Asian children ages 7-11 rate the modest self-presentations of their peers (portrayed in hypothetical scenarios) more favorably than their Western counterparts do (Heyman et al., 2011; Lee et al., 1997). Cultural differences also manifest in terms of actual modest behaviors. For example, in a modesty dilemma paradigm that provided children an opportunity to talk about a good deed they had done, East Asian children ages 7-11 were more likely to show modest behavior (i.e., falsely denying that they had done a good deed) as compared to their Canadian counterparts (Fu et al., 2016).

In rudimentary form, modest behavior may first appear at an even younger age. Already during the preschool years, children anticipate that they are being evaluated by others and they engage in various behavioral strategies to promote their reputational interests (Botto & Rochat, 2019; Heyman et al., 2021; Tomasello & Vaish, 2013). It is possible that the self-presentations of East Asian preschoolers are shaped by prevailing social norms and socialization practices that emphasize modesty (Luo et al., 2013; Wu et al., 2002). Indeed, one study found that from around age 4, Chinese children already describe themselves in a more neutral or modest way than Western children do, who provide more favorable self-descriptions (Wang, 2004). To be sure, this finding does not mean that Chinese children necessarily hold less positive self-concepts than Western children—rather, modest self-presentations can be tactical. Although not yet tested in children, social psychological research has identified the tendency for Chinese adults to deemphasize the positivity of the self in their self-presentations, even if they do positively evaluate themselves in response to indirect or allegedly private measures of self-evaluation (Cai et al., 2011; Kim et al., 2010).

**Self-Estimation of Competence and Task Performance in Early Childhood**

Early work on young children's self-estimation of competence and task performance focused on cognitive tasks. For example, Flavell et al. (1970) investigated how children in preschool and kindergarten (as compared to older children) estimate their performance on a memory task. They found that children in both of the youngest age groups overestimated their memory span (as compared to their actual memory span) prior to the task, more so than older children did. In fact, the overestimation effect in preschoolers and kindergarteners was 2.06 and 2.21, respectively—referring to the ratio between children's self-estimated and actual memory span. Thus, children thought they would do more than twice as well as they actually did. Other work similarly found that preschoolers and kindergarteners overestimate their cognitive competencies, such as in terms of their performance on school tasks, and their understanding of mechanical devices and procedures (Mills & Keil, 2004; Stipek, 1981).

Similar research has examined young children's self-estimation of competence and performance on motor tasks. Such self-estimation is often dependent on children's perception of affordances (Gibson, 2014)—i.e., they determine which actions are possible given their physical capabilities (e.g., body size or strength) and situational demands. The research found that young children (i.e., at least up until age 5 or 6) routinely overestimate what they are physically capable of. For example, they misjudge whether they are able to stand on steep slopes (Klevberg & Anderson, 2002), whether their hands fit through small openings (Ishak et al., 2014), whether their bodies fit through small doorways (Franchak, 2019), and whether they are capable of challenging motor tasks (e.g., removing a toy from a shelf standing on tiptoes; Plumert, 1995). Similarly, they predict they will achieve better on various motor tasks (i.e., jumping as far as possible, throwing a ball with accuracy) than they actually do (Schneider, 1998). In this latter study, the overestimation effect in preschoolers and kindergarteners was 1.4 and 1.18, respectively, for the jumping task; and 1.61 and 1.62, respectively, for the ball throwing task.

**What May Account for Self-Overestimation in Early Childhood?**

Two main, though not mutually exclusive explanations have been offered to account for the apparent pervasiveness of young children's self-overestimation. One is the "monitoring deficiency" account. According to this explanation, young children are not yet capable of reliably monitoring and retaining information on their abilities and their past performances, which means they do not have the cognitive means to accurately estimate their future performance (reviewed in Bjorklund & Green, 1992; Schneider, 1985). Another is the "wishful thinking" account. This explanation proposes that young children often fail to reliably distinguish between their wishes and expectations (Stipek et al., 1984). This would lead them to make performance predictions based on how well they would *want* to perform, rather than on how well they are actually able to perform, resulting in self-overestimation (Lipko-Speed, 2013; Schneider, 1998; Stipek et al., 1984).

Research in samples of Western children has challenged the monitoring deficiency account. For example, studies that assessed children's performance postdiction (i.e., performance recollection shortly after completing a task) found that even 4-year-olds are usually able to remember their performance on a previous task, but still, they remain overly confident when predicting their performance on a future task (Lipko et al., 2009; Schneider, 1998). This work thus suggests, different from what the monitoring deficiency account posits, that even young children typically can accurately monitor their task performance. However, they often fail to integrate this information into their estimates of their future performance. As some preschoolers stated when researchers showed them their past failures on a memory task: "If you give me a different list [of items to recall] like that, I could do it." (Yussen & Levy, 1975, p. 507).

Other research in Western children did provide partial support for the wishful thinking account. For example, a number of studies found that preschoolers' estimates of the performance of their peers are sometimes more accurate (i.e., less inflated) than their estimates of their own performance (Lipko et al., 2009; Schneider, 1998; Stipek et al., 1984). Moreover, when promised a reward for the good performance of their peers (i.e., so that good peer performance becomes desirable), preschoolers raise their estimates of their peers' performance accordingly (Stipek et al., 1984). Thus, young children often make overly optimistic performance estimates when good performance is desirable. That said, research has also found that wishful thinking is context-dependent and can account for overconfidence on some tasks, but not on others (Lipko et al., 2009; Schneider, 1998).

**The Present Study**

Research shows that young children often overestimate their competence and task performance, but evidence has been obtained virtually exclusively in Western samples. We ask to what extent self-overestimation can also be found in children growing up in an East Asian cultural context that highly values modesty. In a first cross-cultural study of its kind, we examine young children's performance estimates in samples of Chinese and Dutch children, using both a cognitive and a motor task. We do so by tracking participants' estimated and actual performance across task trials.

We also explore the psychological underpinnings of children's performance estimates, informed by the monitoring deficiency and wishful thinking accounts. We use multiple-trial tasks and make salient how children perform, to be able to test the possibility that children's performance estimates gradually become more realistic as they gain experience and receive performance feedback. We ask children to estimate both their own and an unknown peer's performance, to test whether their judgments are more realistic when they have no investment in good performance.

We test children ages 4 and 5—an age at which (Western) children typically overestimate their competence and performance. We use behavioral assessments, rather than questionnaires or interviews, to assess children's performance estimation. This allows for direct cultural comparison and minimizes potential language confounds. We calculate self-overestimation as the discrepancy between children's estimates of their performance just prior to the task, and their actual performance on the task.

We test the hypotheses that children (1) overestimate their performance on both tasks; (2) persist in overestimating their task performance across trials; and (3) overestimate their own performance more than they overestimate the performance of their peers. For each of the hypotheses, we explore potential differences between Chinese and Dutch children—our primary interest was in the overestimation of Chinese children, and we included a sample of Western children to allow direct cultural comparison.

We preregistered our hypotheses, design, targeted sample size, and analysis plan at aspredicted.org (https://aspredicted.org/tu8ib.pdf), ["A Study on the Phenomenon of Children's Overestimation" (#29787)]. In Supporting Information 1 (Appendix), we specify where and why we deviated from the preregistered analysis plan. We deviated from the preregistered analysis plan to reduce the risk of Type 1 error due to multiple testing. We conducted additional analyses to provide further evidence relevant to the hypotheses and we omitted one analysis that turned out to be superfluous in light of the research findings.

## Method

### Participants

We tested 101 children from China (52% girls) and 98 children from the Netherlands (49% girls). Participants were ages 4 and 5. We recruited participants for both samples, in the same way, using convenience sampling. We contacted (pre)schools to ask if they were interested in taking part in the study. If they were, we shared informed consent forms among parents of all students ages 4 or 5. We tested all children for whom we received consent. We conducted the study in the fall of 2019 (in both countries). The study was approved by the ethics board of the Faculty of Social Sciences, Utrecht University.

Chinese children's mean age was 4 years and 9 months (*SD* = 5.0 months, range = 50-71 months). They were recruited from a preschool in Wenzhou City, Zhejiang Province. The informed parental consent rate was 72%. Participants lived in an urban area. The school serves ethnically homogeneous, predominantly middle to upper class communities (in terms of family income and education level). Preschool education in mainland China aims to help children adapt to the school system and is mainly organized around structured and collaborative play.

Dutch children's mean age was 5 years and 0 months (*SD* = 6.1 months, range = 50-71 months). They were recruited from six primary schools across the Netherlands (in the Netherlands, most children start primary school at the age of 4). The informed parental consent rate was 63%. Participants predominantly lived in urban or suburban areas. The schools mainly serve ethnically homogeneous, middle class communities. Similar to preschool education in China, education in the first two grade years aims to help children adapt to school and mainly involves structured and collaborative play.

***Data Exclusion***

We excluded the data of five participants (*n* = 1 and *n* = 4 Chinese and Dutch children, respectively) on the motor task, and the data of eight participants (*n* = 1 and *n* = 7 Chinese and Dutch children, respectively) on the memory task, from the pertaining analyses. Following our preregistered protocol, we excluded data either because they were incomplete (*n* = 1 and *n* = 7 for the motor and memory task, respectively), or because they deviated more than 3 *SD*s from the mean (*n* = 4 and *n* = 1 for the motor and memory task, respectively). Thus, we analyzed motor task data of *n* = 100 Chinese children and *n* = 94 Dutch children; and memory task data of *n* = 100 Chinese children and *n* = 91 Dutch children.

**Procedure**

All participants performed the motor task first. To retain statistical power, we did not counterbalance the order of tasks. To limit possible carryover effects or fatigue, participants performed the memory task on another day (2–14 days later). The experimenters spoke participants' native language (i.e., Mandarin or Dutch). All task instructions and responses to potential questions were standardized, translated, and back-translated from English by bilingual speakers.

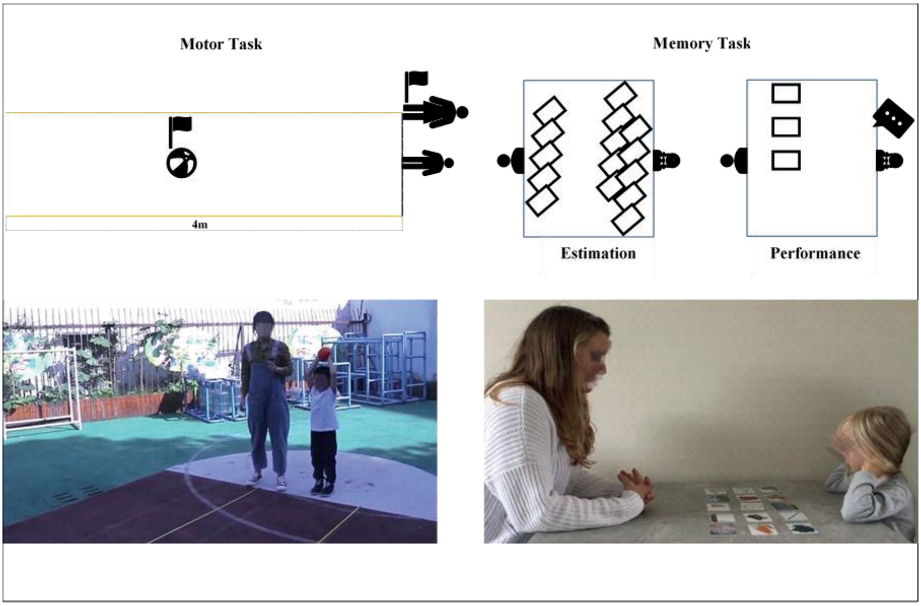
***Motor Task***

We designed the motor task for the present study purposes. We aimed to design a task that was easy to understand for children this age. We kept task difficulty constant across trials (so that performance feedback on one trial can potentially inform children's subsequent performance estimate). We tested children individually in a spacious place on school grounds (Figure 1). We instructed them to stand in front of the starting line of a throwing field, which consisted of two parallel extended 4 m rulers, placed one meter apart, to mark the boundaries of the field. The experimenter handed the ball (i.e., 11 cm in diameter and 1 kg in weight) to the child and said, "Here you go. You can briefly hold the ball so you know a bit how it feels." Then the experimenter took the ball back and asked: "How far do you think you can throw the ball? Could you put the flag somewhere on the throwing field to tell me?" After the child placed the green flag, we registered the distance from the starting line to the flag (i.e., Motor Self Estimate 1) and immediately removed the flag. We then asked the children to return to the starting line and lift the ball over their heads. The experimenter instructed the children: "When I count to three, you will throw the ball as far as possible, okay? Now, one, two, three." The second experimenter observed where the ball first landed and recorded its distance from the starting line (i.e., Motor Self Performance 1), and placed a blue flag on the spot to provide children with feedback on their performance.

Next, with the blue flag still present, we asked children to place the green flag in the throwing field again, to indicate how far they thought they would throw the ball (i.e., Motor Self Estimate 2). The experimenters then removed both flags, asked children to throw the ball as far as they could, and placed the blue flag where the ball landed (i.e., Motor Self Performance 2). We repeated this procedure until participants had made four estimates, and we had recorded three ball-throwing distances. Note that we included an estimate after the last ball throw to be able to test, for each ball throw, whether children learn from their previous performance and adjust their estimates.

**Figure 1**

*Test Setting*




Immediately after participants completed the above task, we showed them a video in which a child of about the same age, sex, and nationality performed the same task. We showed the video on a tablet computer, at the testing site. We introduced the peer with a common Chinese or Dutch name (i.e., Xiaoming or Xiaohong; Daan or Lisa) to be easily referred to. The experimenter paused the video just before the child in the video was about to throw the ball, and asked participants: "How far do you think […] can throw the ball? As before, can you put a green flag on the throwing field to tell me what you think?" After participants placed the flag, the second experimenter recorded its distance from the starting line (i.e., Motor Peer Estimate 1) and immediately removed it. Experimenters assisted participants to watch the video of the peer's performance and placed a blue flag to provide participants with feedback on the peer's performance (i.e., Motor Peer Performance 1).

We placed the blue flag where the ball landed in the corresponding trial when participants themselves took part in the ball throwing task (i.e., Motor Peer Performance 1 = Motor Self Performance 1; Motor Peer Performance *n* = Motor Self Performance *n*). This allowed us to directly compare children's performance estimates for themselves with those for their peers, unconfounded by any differences in actual performance. With the blue flag still present, we asked participants to place the green flag again to indicate how far they thought the child in the video would throw the ball the second time (i.e., Motor Peer Estimate 2). Then, the experimenters removed the two flags, assisted participants to watch the video of the peer's performance, and placed the blue flag (again, matching participants' own previous performance on the corresponding trial). We repeated this procedure until participants had made four estimates and had watched the peer in the video throw the ball three times.

***Memory Task***

We modeled the memory task after similar methodologies used in previous studies (Lipko et al., 2009; Lipko-Speed, 2013; Shin et al., 2007). Again, we made sure that the task was fairly easy to understand for children this age, and we kept task difficulty constant across trials. We tested participants individually in a quiet and private room (Figure 1). We laid out a set of 15 blank cards on the table (previous work used 10 to 15 cards; we used 15 cards to ensure ample scope for children to overestimate their performance). The experimenter sat face to face with the child and said: "Next you will try to remember the same number of cards. But those cards will have pictures on them. How many cards do you think you can remember? Just leave the number of cards that you think you can remember on the table. You can give the rest of the cards back to me." The experimenter recorded how many cards children left on the table (i.e., Memory Self Estimate 1) and then removed all cards.

Next, the experimenter showed the first of three sets of 15 picture cards. Each set contained 15 picture cards, and each picture corresponded to one of 15 themes (e.g., fruits, animals, musical instruments, toys). The experimenter laid out the picture cards on the table, one by one, and asked: "Can you tell me what it is when I show you the card?" Children were almost always able to name the pictures. If not, the experimenter informed them how to name the picture. We always followed children's own use of words—thus, if they used an incorrect word to name a picture, the experimenter did not correct them. Next, participants studied the picture cards, until the experimenter removed them after 15 s, and said: "Now you can tell me the name of each picture that you remember." Each time the child recalled a picture correctly, the experimenter placed a picture card face down on the table. The experimenter encouraged children by saying "try again" or "think about it" when participants remained silent or seemed distracted for more than 5 s. When children said that they could not recall any more pictures or remained silent or distracted for more than 20 s, the experimenter ended the trial and said: "Okay! These are the card(s) that you remembered correctly." The experimenter recorded the number of correctly recalled picture cards (i.e., Memory Self Performance 1).

Next, with the correctly recalled face-down picture card(s) still on the table, we laid out another row that consisted of 15 blank cards, to allow children to estimate their performance on the next trial. Note that each time we laid out cards on the table, we created a row with approximately equal distance between the cards to give children an intuitive understanding of how their estimate for the next trial related to their performance on the previous trial (in this way, we did not need to rely on their number sense). The experimenter told children "Now let's try again. We will use cards with different pictures on them this time." The procedure was identical. Thus, children first indicated how many picture cards they thought they could remember this time, after which they studied the new set of picture cards for 15 s, and recalled as many pictures as possible. Again, the experimenter recorded the number of correctly recalled picture cards (i.e., Memory Self Performance 2). This procedure was repeated until participants had made four estimates and we had recorded three memory performances. Again, we needed an extra estimate after the last memory performance to be able to test whether children learn from their previous performance.

Immediately after participants completed the task, we showed them a video in which a peer of about the same age, sex, and nationality performed the same task. The experimenter paused the video just before the child in the video was about to recall the picture cards, and asked: "How many picture cards do you think […] can remember? As before, just leave the same number of blank cards to indicate how many pictures you think (s)he can remember." The second experimenter recorded the number of blank cards left on the table (i.e., Memory Peer Estimate 1) and immediately removed all cards. Next, the experimenter assisted participants to watch the video of the peer's performance. They placed picture card(s) face down on the table to provide the participant with feedback on the peer's performance (i.e., Memory Peer Performance 1).

As in the motor task, the number of cards we laid out on the table to indicate the peer's performance matched the number of cards that participants themselves had correctly recalled in the corresponding trial (i.e., Memory Peer Performance 1 = Memory Self Performance 1; Memory Peer Performance *n* = Memory Self Performance *n*). With the face-down picture card(s) still on the table, we asked participants again to leave blank cards on the table to indicate how many pictures they thought the child in the video would remember (i.e., Memory Peer Estimate 2). Then, the experimenter removed all the cards, assisted participants to watch the video of the peer's performance, and again placed a number of random picture card(s) face down on the table, matching participants' own performance (i.e., Memory Peer Performance 2). We repeated this procedure until participants had made four estimates and had watched the peer in the video perform three trials.

## Results

### Analytic Strategy

We first conducted a series of descriptive analyses to test the equivalence of our samples and to explore potential sex and age effects for our main variables.

Next, to address our first hypothesis, we determined if children overestimated themselves on both tasks. We also explored potential cultural differences. We conducted a 2 (Performance Index: self-estimated or actual) × 3 (Trial: 1, 2, or 3) × 2 (Nationality: Chinese or Dutch) repeated measures ANCOVA.

To address our second hypothesis, we determined if children would update their estimates of their own performance based on how they performed on prior trials, for both tasks. We conducted a 4 (Trial: 1, 2, 3, or 4) × 2 (Nationality: Chinese or Dutch) repeated measures ANCOVA, which allowed us to examine if children's performance estimates remained the same across trials (i.e., after receiving performance feedback). We also explored potential cultural differences. Furthermore, we explored correlations between children's actual performance on task trials and their subsequent performance estimates, as an additional test of whether they used performance feedback to inform their performance estimates.

To address our third hypothesis, we tested if children overestimated their own performance more than their peer's performance, on both tasks. For Trial 1, children's estimates of their own and their peer's performance cannot be meaningfully compared: Whereas children had no reference point to estimate their own Trial 1 performance, they did have such a reference point (i.e., their own performance) to estimate their peer's Trial 1 performance. Accordingly, to address this hypothesis, we compared performance estimates for Trials 2, 3, and 4. We first conducted a 2 (Performance Index: peer-estimation or actual) × 3 (Trial: 1, 2, or 3) × 2 (Nationality: Chinese or Dutch) repeated measures ANCOVA to determine if children overestimated their peer's performance to begin with. Next, as a direct test of the third hypothesis, we conducted a 2 (Estimation Target: self or peer) × 3 (Trial: 2, 3, or 4) × 2 (Nationality: Chinese or Dutch) repeated measures ANCOVA to determine if children more strongly overestimated their own performance than their peer's performance.

Again, we explored potential cultural differences.

The tests of the three hypotheses are confirmatory; they are based on previous empirical findings and have been pre-registered. The tests of cultural differences (and the descriptive analyses) are exploratory; it is the first time that cultural differences in children's overestimation are examined.

**Descriptive Analyses**

Tables 1 and 2 present the descriptive statistics for children's performance estimates and actual performance on the motor and memory tasks.

**Table 1**

*Children's Self-Estimates, Task Performance, and Peer Estimates on the Motor Task*

| | | All Children | | Chinese Children | | Dutch Children | |
|---|---|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Trial 1 | Self-Estimate | 297.7 | 105.6 | 315.4 | 98.4 | 278.8 | 110.2 |
| | Task Performance | 116.4 | 45.7 | 106.9 | 38.1 | 126.5 | 51.0 |
| | Peer-Estimate | 235.9 | 99.2 | 232.2 | 101.3 | 239.9 | 97.4 |
| Trial 2 | Self-Estimate | 218.8 | 103.4 | 210.0 | 102.4 | 228.2 | 104.3 |
| | Task Performance | 119.0 | 45.7 | 107.2 | 42.2 | 131.5 | 46.1 |
| | Peer-Estimate | 224.1 | 102.4 | 220.0 | 102.5 | 228.4 | 102.5 |
| Trial 3 | Self-Estimate | 226.6 | 105.2 | 207.6 | 104.8 | 246.9 | 102.4 |
| | Task Performance | 129.3 | 50.8 | 117.7 | 44.9 | 141.6 | 53.9 |
| | Peer-Estimate | 226.9 | 104.2 | 226.6 | 109.3 | 227.1 | 99.2 |
| Trial 4 | Self-Estimate | 233.1 | 108.3 | 218.3 | 106.4 | 248.7 | 108.6 |
| | Peer-Estimate | 224.3 | 104.1 | 228.7 | 110.4 | 219.6 | 97.4 |

*Note:* Scores reflect distance in centimeters.

**Table 2**

*Children's Self-Estimates, Task Performance, and Peer-Estimates on the Memory Task*

| | | All Children | | Chinese Children | | Dutch Children | |
|---|---|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Trial 1 | Self-Estimate | 5.55 | 4.17 | 6.17 | 4.72 | 4.87 | 3.36 |
| | Task Performance | 4.78 | 2.15 | 4.80 | 2.25 | 4.76 | 2.04 |
| | Peer-Estimate | 7.66 | 4.36 | 8.40 | 4.63 | 6.85 | 3.91 |
| Trial 2 | Self-Estimate | 6.51 | 4.13 | 6.13 | 3.92 | 6.92 | 4.33 |
| | Task Performance | 3.67 | 1.80 | 3.91 | 1.79 | 3.41 | 1.78 |
| | Peer-Estimate | 7.96 | 4.28 | 8.90 | 4.44 | 6.93 | 3.86 |
| Trial 3 | Self-Estimate | 7.09 | 4.20 | 6.80 | 4.31 | 7.42 | 4.07 |
| | Task Performance | 3.51 | 1.94 | 3.97 | 2.01 | 3.01 | 1.74 |
| | Peer-Estimate | 8.12 | 4.26 | 9.17 | 4.22 | 6.96 | 4.02 |
| Trial 4 | Self-Estimate | 7.57 | 4.59 | 7.61 | 4.82 | 7.53 | 4.36 |
| | Peer-Estimate | 8.06 | 4.51 | 9.16 | 4.39 | 6.85 | 4.35 |

*Note:* Scores reflect the number of cards remembered (possible range 0-15).

On average, Dutch children ($M$ = 133.2) performed better than Chinese children ($M$ = 110.6) on the motor task, $F(1, 192) = 15.22$, $p < .001$, $\eta_p^2 = .073$. Conversely, Chinese children ($M$ = 4.23) performed better than Dutch children ($M$ = 3.73) on the memory task, $F(1, 189) = 4.92$, $p = .028$, $\eta_p^2 = .025$.

Older children performed better on both the motor task (*r*s = .37-.43, *p*s < .001) and the memory task (*r*s = .24-.27, *p*s ≤ .001). Children's estimations of their own performance were mostly unrelated to age. As for children's estimations of their peer's performance, however, older children made more cautious estimates for most trials on the memory task (Trial 2-4: *r*s = -.15 to -.31, *p*s ≤ .033; Trial 1: $r$ = -.06, $p$ = .404), but not the motor task (*r*s = -.01 to .05, *p*s > .517). Because Chinese participants were slightly younger (i.e., 3 months) than Dutch participants, we included age as a covariate in all subsequent analyses.

On average, boys ($M$ = 133.0) performed better than girls ($M$ = 110.8) on the motor task, $F(1, 192) = 14.62$, $p < .001$, $\eta_p^2 = .071$. We found no sex difference for children's estimations of their own performance ($F(1, 192) = 0.95$, $p = .330$, $\eta_p^2 = .005$) or that of their peer's performance ($F(1, 192) = 2.62$, $p = .107$, $\eta_p^2 = .013$) on the motor task. As for the memory task, we found no sex difference in children's performance, $F(1, 189) = 2.70$, $p = .102$, $\eta_p^2 = .014$. However, boys did make more favorable estimates of their own performance ($M = 6.99$) than girls did ($M = 5.79$), $F(1, 189) = 6.14$, $p = .014$, $\eta_p^2 = .031$. Boys also made more favorable estimations of their peers' performance ($M = 8.56$) than girls did ($M = 7.27$), $F(1, 189) = 6.29$, $p = .013$, $\eta_p^2 = .032$. Because of the sex differences we found, we also included sex as a covariate in all subsequent analyses.

**Do Children Overestimate Their Performance?**

***Confirmatory Analysis***

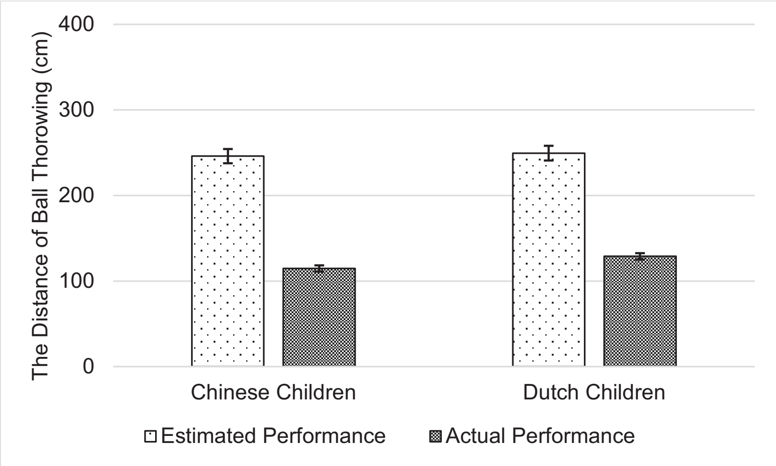
As hypothesized, children overestimated their performance on both tasks. For the motor task, there was a significant main effect of Performance Index, with children's estimates of their performance ($M = 247.8$) being more than twice as high as their actual performance ($M = 121.8$), $F(1, 190) = 15.84$, $p < .001$, $\eta_p^2 = .077$. This equals a self-overestimation effect of 2.03. For the memory task, there was a significant main effect of Performance Index as well, with children's estimates of their performance ($M = 6.38$) again being substantially higher than their actual performance ($M = 3.97$), $F(1, 187) = 5.28$, $p = .023$, $\eta_p^2 = .027$. This equals a self-overestimation effect of 1.61.

***Exploratory Analysis***

We found no evidence for a cultural difference in the extent to which children overestimated their performance (Figure 2 and 3). The Performance Index × Nationality interaction was non-significant, both on the motor task ($F(1, 190) = 1.06$, $p = .305$, $\eta_p^2 = .006$) and the memory task ($F(1, 187) = 1.58$, $p = .211$, $\eta_p^2 = .008$).

**Figure 2**

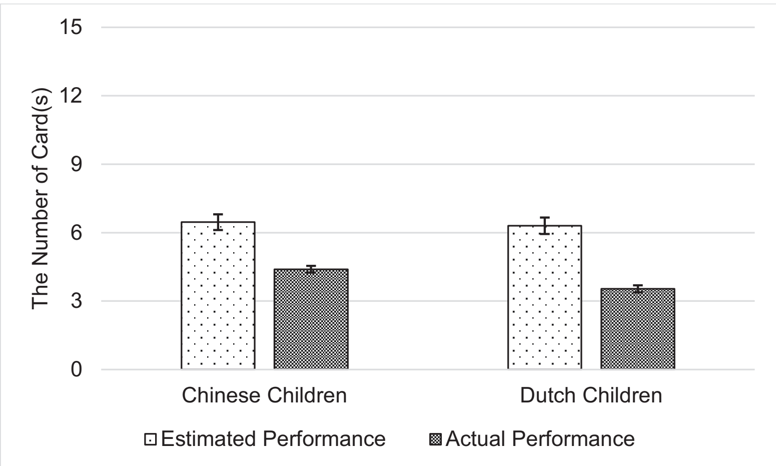
*Chinese and Dutch Children's Estimated and Actual Performance on the Motor Task*




*Note:* Error bars represent standard errors.

**Figure 3**

*Chinese and Dutch Children's Estimated and Actual Performance on the Memory Task*




*Note:* Error bars represent standard errors.

## Do Children Persist in Overestimating Their Performance Across Trials?

### *Confirmatory Analysis*

As hypothesized, we found that children's estimates of their own performance, on both tasks, were relatively stable across trials. There were no significant main effects of Trial on the motor task ($F(2.33, 442.35) = 2.81$, $p = .053$, $\eta_p^2 = .015$), nor on the memory task ($F(2.35, 439.20) = 2.62$, $p = .065$, $\eta_p^2 = .014$). This finding suggests that, overall, children made little use of performance feedback to inform their subsequent performance estimates.

### *Exploratory Analysis*

We did find cultural differences. The Trial × Nationality interactions were significant for both the motor task ($F(2.33, 442.35) = 7.85$, $p < .001$, $\eta_p^2 = .040$), and the memory task ($F(2.35, 439.20) = 4.32$, $p = .010$, $\eta_p^2 = .023$). As Figures 4 and 5 show, cultural differences pertained mainly to the change that occurred from Trial 1 to 2. For the motor task, separate analyses for each nationality showed that Chinese children's performance estimates significantly decreased from Trial 1 to 2 ($F(1, 97) = 4.89$, $p = .029$, $\eta_p^2 = .048$). This decrease was smaller and not significant for Dutch children ($F(1, 91) = 2.10$, $p = .150$, $\eta_p^2 = .023$). For the memory task, Chinese children's performance estimates did not change from Trial 1 to 2 ($F(1, 97) = 0.03$, $p = .856$, $\eta_p^2 = .000$), whereas Dutch children's performance estimates even showed an increasing (rather than decreasing) trend, although this effect was not significant ($F(1, 88) = 0.83$, $p = .365$, $\eta_p^2 = .009$). After the second trial, Chinese and Dutch children's performance estimates remained largely stable, for both tasks.

**Figure 4**

*Chinese and Dutch Children's Performance Estimates Across Trials on the Motor Task*

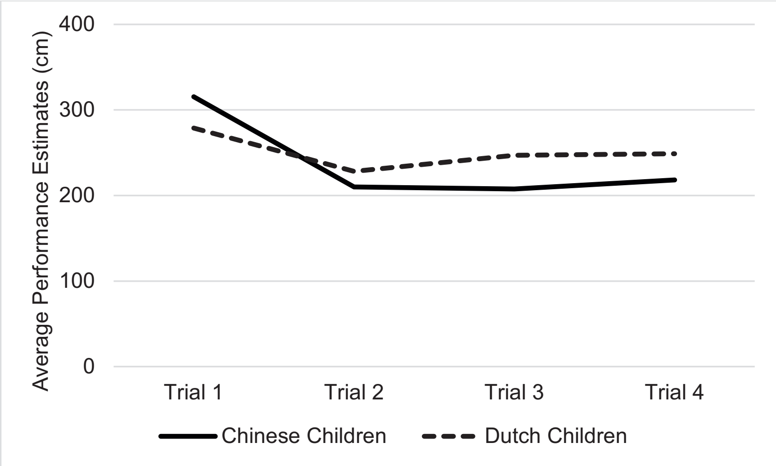
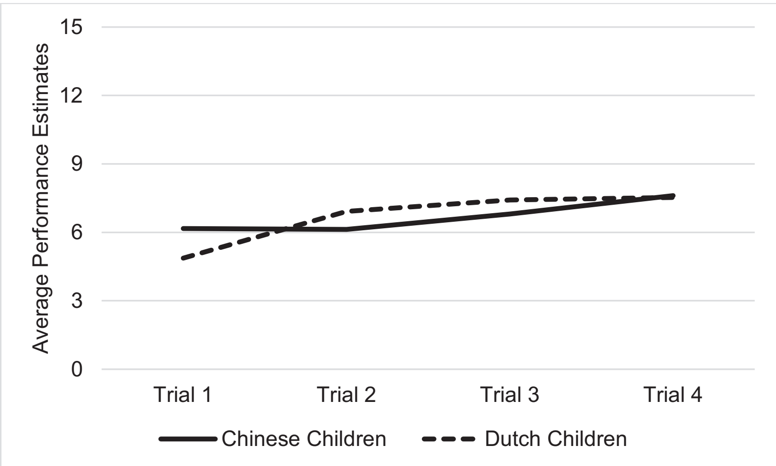



**Figure 5**

*Chinese and Dutch Children's Performance Estimates Across Trials on the Memory Task*




To further explore the extent to which participants incorporated performance feedback into the estimates of their future performance, we inspected the pattern of correlations between children's task performance and their performance estimates on later trials (with age and sex partialled out), for both tasks.

As for the motor task (Table 3), the correlations between children's actual performance on a trial and their performance estimates for the subsequent trial were moderately positive and significant. This pattern of association is consistent with the possibility that children did, at least to some extent, make use of performance feedback to update their performance estimates on this task. Here we found one difference between the Chinese and Dutch samples: The correlation between children's actual performance on the first trial and their performance estimate for the second trial was less strong in Chinese children as compared to Dutch children (Fischer's $Z = -2.97$, $p < .01$).

**Table 3**

*Correlations Between Estimates and Performance on the Motor Task*

| | Estimate2 | Estimate3 | Estimate4 | Performance 1 | Performance 2 | Performance 3 |
|---|---|---|---|---|---|---|
| Estimate1 | $.41^{***}$ $(.34^{**}/.52^{***})$ | $.25^{**}$ $(.24^{*}/.33^{**})$ | $.21^{**}$ $(.19/.28^{**})$ | $.22^{**}$ $(.23^{*}/.26^{*})$ | $.17^{*}$ $(.12/.28^{**})$ | $.16^{*}$ $(.11/.26^{*})$ |
| Estimate2 | | $.61^{***}$ $(.59^{***}/.63^{***})$ | $.63^{***}$ $(.67^{***}/.59^{***})$ | $\mathbf{.46}^{***}$ $(\mathbf{.26}^{**}/\mathbf{.60}^{***})$ | $.49^{***}$ $(.46^{***}/.52^{***})$ | $.35^{***}$ $(.28^{**}/.40^{***})$ |
| Estimate3 | | | $.73^{***}$ $(.66^{***}/.79^{***})$ | $.38^{***}$ $(.24^{*}/.46^{***})$ | $\mathbf{.46}^{***}$ $(\mathbf{.47}^{***}/\mathbf{.43}^{***})$ | $.46^{***}$ $(.43^{***}/.46^{***})$ |
| Estimate4 | | | | $.30^{***}$ $(.25^{*}/.33^{**})$ | $.43^{***}$ $(.41^{***}/.44^{***})$ | $\mathbf{.54}^{***}$ $(\mathbf{.53}^{***}/\mathbf{.53}^{***})$ |
| Performance 1 | | | | | $.59^{***}$ $(.41^{***}/.72^{***})$ | $.52^{***}$ $(.42^{***}/.57^{***})$ |
| Performance 2 | | | | | | $.66^{***}$ $(.55^{***}/.75^{***})$ |

*Note:* Bold values are the correlations between performance on Trial $N$ and estimate on Trial $N + 1$.

Correlations for Chinese and Dutch children, respectively, are reported in brackets.

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

As for the memory task (Table 4), we found no such pattern of association. Here, children's actual performance on a trial and their performance estimations for the subsequent trial were not significantly correlated. In addition, for this task, we found no differences between the Chinese and Dutch samples.

**Table 4**

*Correlations Between Estimates and Performance on the Memory Task*

| | Estimate2 | Estimate3 | Estimate4 | Performance 1 | Performance 2 | Performance 3 |
|---|---|---|---|---|---|---|
| Estimate1 | $.59^{***}$ (.71$^{***}$/.52$^{***}$) | $.35^{***}$ (.49$^{***}$/.14) | $.32^{***}$ (.37$^{***}$/.23$^{*}$) | .02 (-.01/.05) | .09 (.02/.09) | .12 (.18/-.12) |
| Estimate2 | | $.55^{***}$ (.77$^{***}$/.32$^{**}$) | $.43^{***}$ (.63$^{***}$/.22$^{*}$) | **.02** **(-.07/.13)** | .01 (-.09/.12) | -.02 (.06/-.06) |
| Estimate3 | | | $.70^{***}$ (.79$^{***}$/.59$^{***}$) | $-.15^{*}$ (-.11/-.18) | **-.09** **(-.11/-.03)** | -.08 (-.11/.01) |
| Estimate4 | | | | $-.15^{*}$ (-.17/-.12) | -.12 (-.17/-.07) | **-.03** **(-.09/.06)** |
| Performance 1 | | | | | $.41^{***}$ (.44$^{***}$/.36$^{**}$) | $.37^{***}$ (.35$^{***}$/.37$^{***}$) |
| Performance 2 | | | | | | $.51^{***}$ (.42$^{***}$/.54$^{***}$) |

*Note:* Bold values are the correlations between performance on Trial *N* and estimate on Trial *N* + 1.

Correlations for Chinese and Dutch children, respectively, are reported in brackets.

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

## Do Children Overestimate Their Own Performance More Than Their Peer's Performance?

### *Confirmatory Analysis*

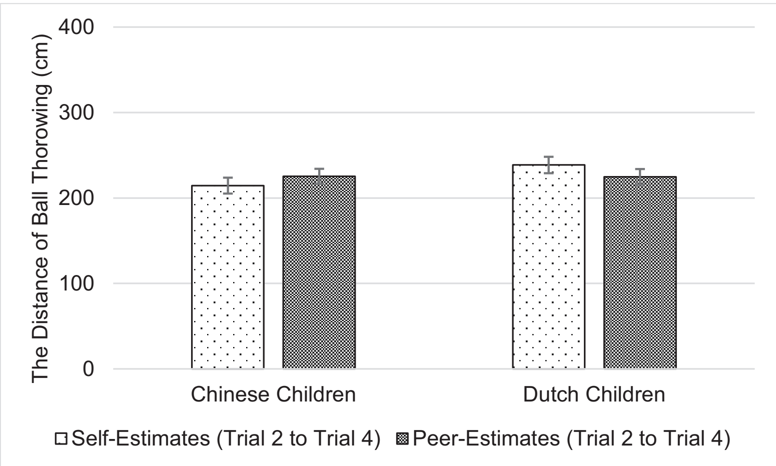
Children overestimated their peer's performance on both tasks. On the motor task, a significant main effect of Performance Index indicated that children's estimates of their peer's performance (*M* = 229.0) were higher than their peer's actual performance (*M* =

121.8), $F(1, 190) = 18.65$, $p < .001$, $\eta_p^2 = .089$. This equals a peer-overestimation effect of 1.88. Similarly, on the memory task, a significant main effect of Performance Index indicated that children's estimates of their peer's performance ($M = 7.87$) were higher than their peer's actual performance ($M = 3.97$), $F(1, 187) = 35.61$, $p < .001$, $\eta_p^2 = .160$. This equals a peer-overestimation effect of 1.98.

Importantly, we found no support for the hypothesis that children would overestimate their own performance more than their peer's performance. On the motor task, there was no significant main effect of Estimation Target, $F(1, 190) = 2.46$, $p = .119$, $\eta_p^2 = .013$. Children's estimates of their own performance were about the same ($M = 226.5$) as their estimates of their peer's performance ($M = 225.0$). On the memory task, we did find a significant main effect of estimation target, but it was in the opposite direction of the hypothesis, $F(1, 187) = 12.66$, $p < .001$, $\eta_p^2 = .063$. Children estimated their own performance ($M = 7.07$) less (not more) favorably than they estimated their peer's performance ($M = 8.01$) (Figures 6 and 7).

***Figure 6***

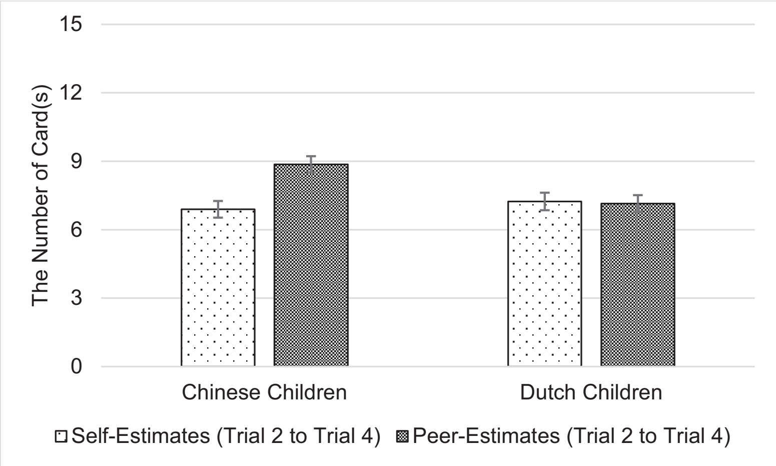
*Chinese and Dutch Children's Self- and Peer-Estimates on the Motor Task*




*Note:* Error bars represent standard errors.

**Figure 7**

*Chinese and Dutch Children's Self- and Peer-Estimates on the Memory Task*




*Note:* Error bars represent standard errors.

### *Exploratory Analysis*

We found cultural differences. On the motor task, Chinese and Dutch children differed in how they estimated their own performance relative to their peer's performance, $F(1, 190) = 4.26$, $p = .04$, $\eta_p^2 = .022$. Specifically, separate analyses for each nationality showed that whereas Chinese children estimated their own motor performance to be descriptively worse ($M = 212.0$) than the performance of their peer ($M = 225.1$), Dutch children estimated their own motor performance to be descriptively better ($M = 241.3$) than the performance of their peer ($M = 225.0$), although these differences in own versus peer performance estimates were not significant, $p$s > .08, $\eta_p^2$s < .033. We found a similar pattern for the memory task, $F(1, 187) = 13.52$, $p < .001$, $\eta_p^2 = .067$. Whereas Chinese children estimated their own memory performance to be significantly worse ($M = 6.85$) than that of their peer ($M = 9.08$), $F(1, 97) = 10.32$, $p = .002$, $\eta_p^2 = .096$, Dutch children estimated their own memory performance to be descriptively better ($M = 7.29$) than the performance of their peer ($M$ = 6.91), $F(1, 88) = 3.60$, $p = .061$, $\eta_p^2 = .039$.

## Discussion

We obtained evidence that Chinese 4- and 5-year-olds overestimate their performance on both a motor and a memory task about as much as their Dutch counterparts do. This finding suggests that young children's self-overestimation is not a uniquely Western phenomenon. It can even be observed in a culture where children are socialized, from a young age, to refrain from self-aggrandizement and show modesty (Wang, 2004; Wu et al., 2002; Xu et al., 2005). Thus, our research suggests that the factors that push young Chinese children to hold inflated expectations of their performance are more powerful than those that pull them to adhere to the modesty imperative.

We also examined psychological processes that may account for young children's self-overestimation. We found that, by and large, both Chinese and Dutch children persisted in overestimating their performance across trials, even if salient performance feedback indicated that they did not perform as well as they anticipated. Prior work has shown that children in the preschool and early school years are able to make quite accurate postdictions: They generally remember their performance on a task when asked directly afterward (Lipko et al., 2009; Schneider, 1998). Moreover, they realize that their past performance can predict their future performance on the same task (Lipko-Speed, 2013). Our findings are thus consistent with a view that despite these abilities, children do not fully incorporate performance feedback into their performance predictions (Lipko et al., 2009; Lipko-Speed, 2013; Schneider, 1998).

And yet, we found two important qualifiers to this general pattern. First, on the motor task, Chinese (but not Dutch) children did lower their performance estimates after the first trial, on which they typically performed worse than they had predicted. Second, also on the motor task, both Chinese and Dutch children's performance estimates were associated with their performance in previous trials. This latter finding suggests that even if children generally persisted in self-overestimation, they did make use of their experience to make somewhat more informed estimates of their performance on subsequent trials —at least on the motor task. We conclude that it is not monitoring deficiency, but rather, incorporation inconsistency that contributes to young children's self-overestimation: Children ages 4 and 5 fail to consistently or fully incorporate performance feedback into their performance estimates. This pattern was generally true for children from both nationalities, although Chinese children gave somewhat more evidence of realistically updating their performance estimates on the motor task than Dutch children did.

Our results are inconsistent with the wishful thinking account for young children's self-overestimation. According to this account, preschoolers and kindergarteners often fail to distinguish performance wishes and expectations. One would expect, then, that their desire to be competent should positively bias their estimates of their own performance, but not those of an unknown peer's performance (because they have little investment in the peer's success; Stipek & Hoffman, 1980; Stipek et al., 1984). In the research design that we used, it was possible for such an effect to occur. Much like young children can use their peers' performance as a reference point to make informed estimates of their own performance (Plumert & Schwebel, 1997), in this study, children could use their own performance as a reference point to realistically estimate their peer's performance. And yet, this is not what they did—instead, both Chinese and Dutch children generally overestimated their peer's performance about as much as their own. In fact, on the memory task, Chinese children overestimated their peer's performance even more (not less) than they overestimated their own.

Prior work did find that under some conditions, (Western) children make more accurate performance estimates when judging a peer than when judging themselves

(Schneider, 1998; Stipek & Hoffman, 1980; Stipek et al., 1984). And yet, some of the same work showed that such discrepancies between self-and peer-estimates are conditional upon task characteristics, and such factors as the salience of past performance (Lipko et al., 2009; Schneider, 1998; Stipek et al., 1984). Together, this evidence corroborates the view that children exhibit a general positivity bias in their judgment of attributes and abilities—at least from the preschool age, they attend to, process, and interpret information selectively to maintain positive views of both themselves and others (Boseovski, 2010). We even found that Chinese (but not Dutch) children sometimes overestimate their peer's performance more than their own performance. This finding may be another manifestation of Chinese children's tactical self-presentation—they possibly anticipated that making positive predictions about a peer would reflect well on them. Even then, our overall pattern of findings—including those in Dutch children—suggests that wishful thinking did not contribute to children's self-overestimation.

An overarching question that emerges from these findings pertains to the consequences of young children's self-overestimation. Research has demonstrated some potential negative consequences: To the extent that children more strongly overestimate, specifically, their physical ability, they may be at increased risk of accidental injury (Plumert, 1995; Plumert & Schwebel, 1997). Nevertheless, there may also be important benefits to children's self-overestimation that transcend cultural boundaries. Indeed, it has been argued that some aspects of cognitive immaturity, including self-overestimation, have adaptive value (Bjorklund, 1997; Bjorklund & Green, 1992; Schwebel & Plumert, 1999). Given that young children have little experience with most activities they engage in, they could easily become discouraged or shy away from novel challenges if they accurately perceived the limits to their ability. Self-overestimation may allow young children to feel efficacious despite their inexperience, to persist in the face of difficulty or failure, and to take on new challenges—

thereby gaining important opportunities to develop abilities and improve performance (Shin et al., 2007).

**Strengths, Limitations and Future Research**

Our research is the first to compare self-estimates of performance in children growing up in a Western (i.e., the Netherlands) and non-Western (i.e., mainland China) cultural context. The current literature on children's emerging self-evaluation is heavily skewed towards samples of Western children, which raises questions about generalizability (Nielsen et al., 2017). This research provides a first step toward building a more culturally diverse understanding of children's self-overestimation. We did so by building upon well-established performance prediction methodological paradigms. To allow direct cross-cultural comparison and avoid potential language confounds, we obtained non-verbal performance estimates (i.e., placing flags, retaining cards), and also provided performance feedback using similar non-verbal cues. We did so for both tasks, to maximize task comparability. Another methodological strength is that, in assessing peer performance estimates, we kept the alleged performance of the peer the same as the performance of the participant, to allow direct comparison unconfounded by differences in actual performance.

We also acknowledge limitations. We asked children to make self-and peer-estimates of performance in a fixed order (i.e., self-estimates always preceded peer-estimates). Indeed, our pilot study showed that it is difficult for children this age to estimate the performance of their peers with limited task experience, which is why we decided not to counterbalance. The implication, however, is that children's estimates of their peers' performance may have been somewhat colored by their own experiences with the task.

Our findings suggest that young children do not consistently incorporate performance feedback into the estimates of their future performance. A valuable step for future research would be to provide an experimental test of this mechanism by comparing the performance estimates of children who do and do not receive feedback. Moreover, future research may test the developmental specificity of self-overestimation by including older age groups in cross-cultural comparisons. Research in Western samples suggests that self-overestimation is pervasive in young children, but can sometimes be observed in older age groups as well. Future research will need to address the cultural generalizability of such observations.

In addition, research is needed to better understand both the malleability and adaptiveness of young children's self-overestimation. For example, are there situational boundary conditions to self-overestimation? To what extent is self-overestimation rooted in socialization practices by parents? How do learning environments, and the extent to which they make salient individual achievement or normative evaluation (Dweck et al., 2014; Pang & Richey, 2007; Stipek & Daniels, 1988), influence children's self-overestimation? And when or why is it adaptive for children to overestimate themselves? Insight in questions as these will be key to informing parenting experts and educators on how to help young children develop healthy views of themselves.

Finally, our findings should be interpreted in light of China's sociocultural change during the past few decades. Self-enhancement is on the rise in China, a development that has been tied to socioeconomic transformation and changing cultural values (Cai, et al., 2012; Zhang et al., 2017). Sociocultural change has been most pronounced in urban areas, where traditional cultural heritage now coexists with contemporary, individualistic values—a development which is echoed in evolving parenting practices, and has consequences for child adjustment (Chen & Li, 2012; Chen et al., 2009). We conducted our study in such an urban area—the city of Wenzhou. Thus, while the self-overestimation of the Chinese children we studied was robust and substantial, future work will need to verify to what extent it can also be observed in children growing up in rural areas.

**Coda**

Young children's self-overestimation is not a uniquely Western phenomenon. Our research finds that non-Western (i.e., Chinese) young children overestimate their task performance as much as their Western (i.e., Dutch) counterparts do. Moreover, children from both cultures persevere in overestimating themselves, despite receiving accurate performance feedback. Their rosy outlook on their own performance generalizes, though, to how they estimate the performance of their peers. In fact, Chinese children sometimes overestimate the performance of their peers even more than their own. Newborn calves are not afraid of tigers—indeed, they have high aspirations, both for themselves and for their peers.

# Chapter 3

## Why Do Young Children Overestimate Their Task Performance? A Cross-Cultural Experiment

Published as: Xia, M., Poorthuis, A. M. G., & Thomaes, S. (2023). Why do young children overestimate their task performance? A cross-cultural experiment. *Journal of Experimental Child Psychology, 226,* Article 105551. https://doi.org/10.1016/j.jecp.2022.105551

Author contributions

MX, AP, and ST conceptualized the study. MX and ST were responsible for data collection. MX analyzed the data and wrote the first draft of the manuscript. AT and ST provided feedback on the manuscript.

**Abstract**

Young children are generally overconfident in their abilities and performances, but the reasons that underlie such self-overestimation are unclear. The present cross-cultural experiment aimed to address this issue, testing the possibility that young children's overconfidence in task performance is, at least in part, motivated. We tested 89 Chinese (49% girls) and 104 Dutch (50% girls) children aged 4 and 5 years and asked them to estimate how well they would perform on both a motor task and a memory task. They were randomly assigned to either an experimental condition (in which they were promised a reward for providing accurate performance estimates) or a no-incentive control condition, and then they performed the task. The incentive lowered Chinese (but not Dutch) children's performance overestimation on the motor task. Unexpectedly, children did not overestimate their performance on the memory task. Thus, this study supports the view that young children's self-overestimation can be motivated (rather than due to cognitive immaturity alone), but also reveals task contingencies and cultural differences.

*Keywords:* Overestimation, Motivation, Cognitive immaturity, Cross-cultural comparison

# Why Do Young Children Overestimate Their Task Performance?

# A Cross-Cultural Experiment

Compared with older children and adults, young children are generally overconfident in their abilities, understanding, and knowledge. At least in early childhood, they tend to overestimate their performance on various tasks and activities. For example, they believe that they are more capable of completing motor tasks than they actually are, they overestimate their abilities when performing cognitive tasks, they are overconfident of the decisions they make, and they believe to have higher peer status than reality warrants (Boulton & Smith, 1990; Lipko et al., 2009; Mills & Keil, 2004; Piehlmaier, 2020; Plumert, 1995). Although this phenomenon has been well established, the reasons why young children tend to overestimate themselves are not well understood. The current cross-cultural experiment addressed this issue—specifically, it tested the possibility that children's self-overestimation is motivated. We considered self-overestimation to be motivated if children are able to estimate their performances more accurately (e.g., when accuracy is incentivized) than they typically do. We tested this possibility by conducting a between-participants experiment using a motor and a memory task in samples of children growing up in a Western country (the Netherlands) and a non-Western culture (China).

## Metacognitive Immaturity and Incorporation Inconsistency

One explanation that has been offered for young children's self-overestimation emphasizes that young children still lack the (meta)cognitive ability that would allow them to develop more accurate self-perceptions. Thus, according to this account, self-overestimation is a manifestation of "(meta)cognitive immaturity" (reviewed in Bjorklund & Green, 1992; Schneider, 1985). Metacognition refers to individuals' knowledge of their own cognition and the factors that affect it (Bjorklund & Green, 1992). It involves awareness of one's abilities, the use of learning or performance strategies, and the evaluation and monitoring of problem solving (Bjorklund, 1997). Children's metacognitive abilities develop and improve gradually over time (Coutinho et al., 2005; Krueger & Mueller, 2002). Accordingly, the cognitive immaturity explanation holds that young children tend to self-overestimate because they are not yet fully capable of monitoring, realistically perceiving, or retaining information on their performances and abilities (Schneider, 2008)—a form of cognitive immaturity that has been labeled "monitoring deficiency" (Schneider, 1998).

However, some studies have found that young children persist in self-overestimation even if they have the cognitive abilities that should enable them to hold more accurate self-views. For example, even 4-year-old children are able to accurately remember their performance on a previous trial when asked to make a postdiction (i.e., to recollect their performance shortly after engaging on a task), and yet they remain overconfident when predicting their performance on the next trial (Lipko et al., 2009; Schneider, 1998). Other studies have shown that even when young children persist in overestimating their task performance across trials, they sometimes do show signs of using their previous experiences with the task. For example, Lipko-Speed (2013) found that when 4- and 5-year-old children were provided with exactly the same stimulus material on a memory task, they overestimated their performance on a second trial significantly less than they did on the first trial. Adding to this evidence, research has shown that slightly older children, from 6 years of age, are able to use social information to provide more accurate estimates of their ability. For example, children who watched a same-aged peer fail on a set of physical tasks subsequently made more conservative estimates of their own abilities as compared with their counterparts who saw the peer succeed (Plumert & Schwebel, 1997).

Together, these findings suggest that cognitive immaturity (i.e., lack of ability) is not the sole explanation for children's self-overestimation: Even if young children are able to monitor and accurately process their own (as well as their peers') performance, they somehow fail to consistently or fully incorporate this information into their performance estimates. Indeed, we have previously suggested that not "monitoring deficiency" but rather "incorporation inconsistency" characterizes young children's overestimation (Xia et al., 2022).

**Motivation and Wishful Thinking**

A second and potentially complementary explanation for young children's self-overestimation emphasizes motivational processes, especially those related to wishful thinking. According to this account, young children tend to make performance estimates based on how well they would want to perform rather than on how well they are actually able to perform (Lipko-Speed, 2013; Schneider, 1998; Stipek et al., 1984). Young children's self-overestimation, then, is assumed to be due to their desire for performing well or, more generally, to their desire to be competent individuals.

There is some evidence consistent with the wishful thinking account. For example, studies have found that children's overestimation of performance is limited to their own performance, and when they estimate the performance of a peer (for whom they are less likely to desire good performance), their estimates can be more accurate (Lipko et al., 2009; Schneider, 1998; Stipek et al., 1984; Stipek & Hoffman, 1980). Moreover, when children are promised a reward for the good performance of a peer (so that good performance of the peer becomes more desirable), they tend to provide more inflated estimates of their peer's performance (Stipek et al., 1984). Importantly, some inconsistent findings have been obtained as well, suggesting that wishful thinking may account for self-overestimation on some tasks more than on others and in some cultures more than in others (Boseovski, 2010; Xia et al., 2022). Still, the body of literature is consistent with the view that young children make overly optimistic performance estimates to the extent that good performance is desirable—and, thus, that their self-overestimation may be motivated, rather than due to cognitive immaturity per se. A direct test of the possibility that young children can estimate their own performance more accurately if they want to, however, is still lacking.

**The Development of Self-Representations Across Cultures**

Children's self-perceptions are socially constructed and vary across cultures (Brummelman & Thomaes, 2017; Wang, 2006). The content and valence of children's self-representations are shaped by the sociocultural context in which they grow up (Harter, 1998; Wang, 2004). Research has shown that cultural socialization is embedded in the daily interactions between children and their social environments. For young children, parents play a central role as socializing agents such as in the context of family conversations, parent-directed meaning-making, and disciplinary practices. Such exchanges convey culturally specific norms and beliefs to children, which shape their self-representations over time (Brummelman & Thomaes, 2017; Fung & Chen, 2001; Ng et al., 2019; Wang, 2001).

In Western cultures, social norms emphasize the importance of positive distinctiveness and personal success (Markus & Kitayama, 1991; Sedikides et al., 2015). Children are exposed, from a young age, to messages that suggest it is desirable to be unique, to be self-reliant, or to stand out (Gürel & Brummelman, 2020; Thomaes et al., 2017; Young-Eisendrath, 2008). Conversely, in Eastern Asian cultures, social norms more often emphasize the importance of interpersonal cohesion and harmony-of "fitting in" rather than "standing out" (Markus & Kitayama, 1991). For example, notwithstanding within-cultural differences (Cai et al., 2012; Zhang et al., 2017), Chinese children are more often socialized to exercise restraint in how favorably they present themselves to others or how they communicate about their good performances (Luo et al., 2013; Wang & Ollendick, 2001; Wu et al., 2002; Xu et al., 2005). Indeed, the traditional cultural norm of modesty, which is socialized from a young age, shapes the way in which children present themselves to others in Eastern Asian cultures (Kim et al., 2010; Wang, 2004; Yamagishi et al., 2012).

Despite these cultural differences, our previous work in fact found important similarities in how Western (i.e., Dutch) and Eastern (i.e., Chinese) preschoolers and kindergarteners estimate their task performance; children from both countries overestimated their task performance to a similar extent and largely continued to do so even after obtaining performance feedback (Xia et al., 2022). Yet, whether the reasons why children self-overestimate are also similar across cultures is still unknown.

**The Current Experiment**

In the current cross-cultural, between-participants experiment we addressed the overarching question of whether young children's self-overestimation is motivated. We tested whether young children estimate their task performance more accurately when they are promised a reward for accuracy. Such a finding would suggest that their self-overestimation is at least partly motivated and not due to cognitive inability alone. This would help to explain previous evidence that even when children are able to accurately monitor their performance, they do not reliably incorporate this information into the estimates of their future performance; it may often be more appealing or rewarding for them to be overly optimistic about their performance.

As in our earlier work, we invited Chinese and Dutch children to work on both a cognitive task and a motor task, and we tracked their estimated and actual performances across trials. We hypothesized that (1) children would overestimate their performance on both tasks, (2) children in the reward condition would overestimate themselves less than children in the control condition (an effect that may become progressively stronger across trials), and (3) the association between children's performance on one trial and their performance estimate for the next trial would be stronger for children in the reward condition than for those in the control condition. For each of the hypotheses, we explored potential differences between Chinese and Dutch children.

We preregistered our hypotheses, design, targeted sample size, and analysis plan with aspredicted.org (https://aspredicted.org/j9hr2.pdf), ["Can children make accurate performance estimates?" (#59771)].

## Method

### Participants

We tested 104 children from the Netherlands (50% girls) and 89 children from Mainland China (49% girls). Participants were 4 and 5 years of age. We recruited participants for both samples in the same way. We contacted (pre)schools to ask whether they were interested in taking part in the study. If they were, we shared informed consent forms among parents of children aged 4 or 5 years. We conducted the study in the spring of 2021 in both countries. The study was approved by the ethics board of the Faculty of Social and Behavioural Sciences at Utrecht University.

Dutch children's mean age was 5 years and 1 month (SD = 5.8 months, range = 48-71 months). They lived in central regions in the Netherlands, predominantly in suburban areas. Participants were recruited from the first two grade years of five primary schools. These schools serve communities that are relatively homogeneous in ethnicity (White) and middle-class. The informed consent rate was 57.9%. Education in the first two grade years in the Netherlands aims to help children adapt to school and mainly involves structured and collaborative play.

Chinese children's mean age also was 5 years and 1 month (SD = 3.9 months, range = 52-71 months). They lived in the urban area of Suzhou City, Jiangsu Province. Participants were recruited from a preschool that serves communities that are homogeneous in ethnicity (Han Chinese) and middle to upper class. The informed consent rate was 51.0%. Similar to the early grade years in Dutch primary schools, preschool education in mainland China is mainly organized around structured and collaborative play as the main activity and aimed at helping children adapt to the school system.

We excluded the data of seven participants on the motor task and the data of four participants on the memory task (all from China) from the pertaining analyses. Following our preregistered protocol, we excluded data either because participants in the experimental group failed to pass the control question after two attempts (*n* = 1 for both the motor and the memory tasks), because they provided incomplete data (*n* = 2 for the memory task), or because one or more of their estimated or actual performance scores deviated more than three standard deviations from the mean (*n* = 5 and *n* = 1 for the motor and memory task, respectively). Thus, we analyzed motor task data of *n* = 103 Dutch children and *n* = 84 Chinese children[1] and memory task data of *n* = 104 Dutch children and *n* = 85 Chinese children[2].

**Procedure**

All participants performed the motor task first and then performed the memory task later the same day, with at least one hour between the two tasks. We administered the motor and memory tasks in a fixed order because our pilot study showed that it is easier for children of this age to meaningfully estimate their memory performance if they have experience with similar performance estimation tasks. Thus, we decided to start with the motor task, which was perceived as easier and more familiar. Participants were randomly assigned to either the experimental or control groups. The experimenters spoke participants' native language (i.e., Dutch or Mandarin). All task instructions and  responses to potential questions were

[1] The composition of the final sample for the motor task was similar to the original sample (Dutch children: 51% girls, $M_{age}$ = 5 years and 1 month, *SD* = 5.8 months, range = 48-71 months; Chinese children: 52% girls, $M_{age}$ = 5 years and 1 month, *SD* = 3.9 months, range = 52-71 months).

[2] The composition of the final sample for the memory task was similar to the original sample (Dutch children: 50% girls, $M_{age}$ = 5 years and 1 month, *SD* = 5.8 months, range = 48-71 months; Chinese children: 51% girls, $M_{age}$ = 5 years and 1 month, *SD* = 4.0 months, range = 52-71 months).

standardized, translated, and back-translated from English by bilingual speakers.

***Motor Task***

We modeled the procedures and instructions for the motor task after a task that we developed previously (Xia et al., 2022). In the current study, we implemented an experimental manipulation in the context of the task.

We tested children individually in a spacious place on their school grounds. We told them that the task would involve throwing a ball as far as possible. We invited children to stand in front of the starting line of a 4-m long and 1-m wide throwing field. The experimenter handed the ball (11 cm in diameter and 1 kg in weight) to the children and said, "Here you go. You can briefly hold the ball, so you know a bit how it feels.” The experimenter then took the ball back and asked children to place a green flag on the throwing field to estimate how far they thought they could throw the ball. At this point, the experimenter told children in the experimental condition: "If you put the flag in the right place, so if you tell me precisely where the ball will land, you will get a surprise gift!" Children in the control condition were not given this instruction.

To ensure that the children in the experimental condition understood the instruction, we asked them a control question. The experimenter introduced a hypothetical child of about the same age, sex, and nationality who had performed the same task. The experimenter placed a green flag on the ground to index the peer’s estimated performance and a blue flag on the same spot to index the peer’s actual performance. The flags were placed away from the actual throwing field. Then the experimenter said: “[name of hypothetical child] put the flag where he/she thought he/she would throw the ball. Then he/she threw the ball as far as possible, and the ball landed here. Do you think [name of hypothetical child] got a gift?” Children who answered incorrectly were given an explanation, and they then answered the question again with the flags placed at a different spot. Only the data of children who passed the question in two attempts (99.5%) were included in the analyses.

Next, we registered the distance from the starting line to the green flag that children had placed to estimate their performance (i.e., Motor Estimate 1) and removed the flag immediately. We then asked children to return to the starting line and lift the ball over their heads. The experimenter instructed children: “When I count to three, you will throw the ball as far as possible, okay? Now, one, two, three.” The experimenter observed where the ball first landed, placed a blue flag on the spot to provide children with performance feedback, and recorded the distance from the starting line (i.e., Motor Performance 1).

Next, with the blue flag still present, we asked children to place the green flag again to estimate their performance on the next trial (i.e., Motor Estimate 2). The experimenter then removed both flags, asked children to throw the ball as far as they could, and placed the blue flag where the ball landed (i.e., Motor Performance 2). We repeated this procedure until participants had made four estimates and we had recorded three ball-throwing distances. Note that we included an estimate after the last ball throw to be able to test whether children learn from their previous performance and adjust their estimates. The experimenter reminded participants in the experimental condition of the reward for accuracy each time they made a performance estimate.

On completion of the motor task, the experimenter told each participant in the experimental condition, “Because you could tell me so well how far you would throw the ball, you will receive a small gift!” and gave them a sticker. The experimenter told each participant in the control condition, “Because you worked so hard, you will receive a small gift!” and gave them a sticker as well. Next, we brought children back to their regular classrooms.

***Memory Task***

We also modeled the procedures and instructions for the memory task also after a task that we developed previously (Xia et al., 2022). However, in the current study, we simplified the task materials (i.e., the picture cards to be memorized) to shorten the task duration and ensure its equivalence to the motor task, and we implemented the experimental manipulation.

We tested participants individually in a quiet and private room at their school. We told them that the task would involve remembering as many picture cards as possible. We first laid out a set of 15 blank cards on the table and told children that in the actual task there would be pictures on the cards. The experimenter sat face to face with the children and said, “How many cards do you think you can remember? Just leave the number of cards that you think you can remember on the table. You can give the rest of the cards back to me.” Here, the experimenter told participants in the experimental condition, "If the number of cards that you leave on the table is correct, so if you tell me precisely how many cards you will be able to remember, you will get a surprise gift!" Children in the control condition were not given this instruction.

The experimenter recorded how many cards children left on the table (i.e., Memory Estimate 1) and then removed all cards. Next, the experimenter showed a set of 15 picture cards. These cards had pictures of common objects that children this age are familiar with (e.g., fruits, animals, musical instruments, toys). The experimenter laid out the picture cards on the table one by one, read their names aloud (e.g., “cat,” “pencil”), and asked children to repeat them. Only when children repeated the name of the picture correctly could they continue to the next card. Next, participants studied the picture cards, until the experimenter removed them after 15 s and said, "Now you can tell me the name of each picture that you remember.” Each time a child recalled a picture correctly, the experimenter placed the pertaining picture card face down on the table. When children said that they could not recall any more pictures, or remained silent or distracted for more than 20 s, the experimenter ended the trial and recorded the number of correctly recalled picture cards (i.e., Memory

Performance 1).

Next, with the correctly recalled face-down picture cards still on the table, children were asked to estimate their performance on the next trial (i.e., Memory Estimate 2). Again, the experimenter told children in the experimental condition (only), “Remember you will get a surprise gift later when you tell me precisely how many cards you will be able to remember.” Note that each time we laid out cards on the table, we created a row with approximately equal distance between the cards to give children an intuitive understanding of how their estimate for the next trial related to their performance on the previous trial (thus, we did not need to rely on children’s number sense). We used the same stimulus materials (i.e., the same set of 15 picture cards) and naming procedure throughout the experiment. Children were then asked again to study the cards and recall as many pictures as possible (i.e., Memory Performance 2). This procedure was repeated until participants had made four memory estimates and we had recorded three memory performances. Each time children in the experimental condition made a performance estimate, they were reminded of the reward.

On completion of the memory task, the experimenter prepared an array of small gifts (e.g., erasers, glitter pens) for participants to choose from. The experimenter told children in the experimental condition, “Because you could tell me so well how far you would throw the ball and how many pictures you would remember, you can pick a small gift!”. The experimenter told children in the control group, “Because you worked hard at both tasks, you can pick a small gift!”. After children chose a gift, we thanked them for their participation and brought them back to their regular classrooms.

## Results

### Analytic Strategy

We first conducted a series of descriptive analyses, tested the equivalence of our samples and conditions, and explored potential sex and age effects for the main study variables.

To address Hypotheses 1 and 2 and the exploratory research question, for both tasks we conducted a 2 (Performance Index: estimated or actual) × 2 (Condition: experimental or control) × 3 (Trial: 1, 2, or 3) × 2 (Nationality: Dutch or Chinese) repeated-measures analysis of covariance (ANCOVA). We conducted follow-up analyses to interpret significant three-way interactions.

To address Hypothesis 3 and the exploratory research question, for both tasks we computed correlations between children's actual task performance and their performance estimates on subsequent trials. We compared the strength of the pertaining correlations between children in the accuracy reward and control conditions using the *cocor* program (Diedenhofen & Musch, 2015). To explore cultural specificity, we inspected the same pattern of correlations for children from the Netherlands and China separately.

The tests of the three hypotheses are confirmatory and were preregistered. The tests of cultural differences (and the descriptive analyses) are exploratory; this study provides the first cross-cultural test of the motivated nature of young children's self-overestimation.

**Descriptive Analyses**

Tables 1 and 2 present the descriptive statistics for children's estimated and actual performances on the motor and memory tasks.

**Table 1**

*Children's Estimates and Performances on the Motor Task*

| | | All Children | | Experimental Condition | | Control Condition | | Dutch Children | | Chinese Children | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Trial 1 | Estimate | 265.8 | 107.0 | 241.6 | 103.9 | 288.7 | 105.4 | 297.3 | 98.7 | 227.1 | 104.6 |
| | Performance | 151.2 | 62.0 | 144.6 | 59.5 | 157.4 | 64.0 | 145.7 | 63.7 | 157.8 | 59.6 |
| Trial 2 | Estimate | 221.6 | 90.3 | 208.4 | 90.2 | 234.2 | 89.0 | 241.7 | 96.5 | 197.0 | 75.5 |
| | Performance | 158.6 | 63.6 | 146.2 | 61.6 | 170.3 | 63.7 | 148.6 | 60.8 | 170.8 | 65.3 |
| Trial 3 | Estimate | 226.4 | 90.7 | 204.0 | 91.6 | 247.7 | 85.0 | 238.5 | 97.2 | 211.5 | 80.1 |
| | Performance | 160.7 | 69.8 | 143.5 | 69.1 | 177.0 | 66.8 | 149.4 | 70.9 | 174.6 | 66.3 |
| Trial 4 | Estimate | 235.1 | 102.5 | 203.4 | 102.6 | 265.2 | 93.4 | 250.4 | 112. | 216.3 | 86.5 |

**Table 2**

*Children's Estimates and Performances on the Memory Task*

| | | All Children | | Experimental Condition | | Control Condition | | Dutch Children | | Chinese Children | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Trial 1 | Estimate | 7.22 | 3.92 | 7.34 | 4.24 | 7.09 | 3.61 | 7.96 | 4.11 | 6.31 | 3.50 |
| | Performance | 6.01 | 2.06 | 6.02 | 2.12 | 6.00 | 2.02 | 5.47 | 1.67 | 6.67 | 2.31 |
| Trial 2 | Estimate | 7.53 | 3.46 | 7.51 | 3.63 | 7.56 | 3.31 | 7.78 | 3.58 | 7.24 | 3.30 |
| | Performance | 6.84 | 2.06 | 6.73 | 2.00 | 6.94 | 2.13 | 6.38 | 1.87 | 7.39 | 2.16 |
| Trial 3 | Estimate | 7.67 | 3.43 | 7.09 | 3.34 | 8.23 | 3.44 | 7.93 | 3.43 | 7.34 | 3.41 |
| | Performance | 7.07 | 2.48 | 6.92 | 2.46 | 7.21 | 2.51 | 6.60 | 2.20 | 7.65 | 2.69 |
| Trial 4 | Estimate | 8.00 | 3.51 | 7.67 | 3.51 | 8.32 | 3.50 | 8.12 | 3.39 | 7.86 | 3.66 |

Children in both conditions did not differ in age or sex distribution (*p*s ≥ .130), suggesting that random assignment to conditions was effective.

Children's estimates of performance were mostly unrelated to age, although on the motor task older children made more favorable estimates on Trial 3 ($r = .17$, $p = .017$, all other *p*s ≥ .104). Older children performed better on both the motor task (*r*s ≥ .20, *p*s ≤ .007) and the memory task (*r*s ≥ .15, *p*s ≤ .045). We included age as a covariate in all subsequent analyses.

On the motor task, boys made more favorable estimates ($M = 257.8$) than girls ($M = 217.8$), $F(1, 185) = 12.71$, $p < .001$, $\eta_p^2 = .064$, and boys also performed better ($M = 169.1$)

than girls ($M$ = 145.2), $F$(1, 185) = 8.10, $p$ = .005, $\eta_p^2$ = .042. On the memory task, boys made more favorable estimates ($M$ = 8.17) than girls ($M$ = 7.05), $F$(1, 187) = 8.17, $p$ = .005, $\eta_p^2$ = .042, but there was no sex difference in performance, $F$(1, 187) = 0.04, $p$ = .837, $\eta_p^2$ < .001. We also included sex as a covariate in all subsequent analyses.

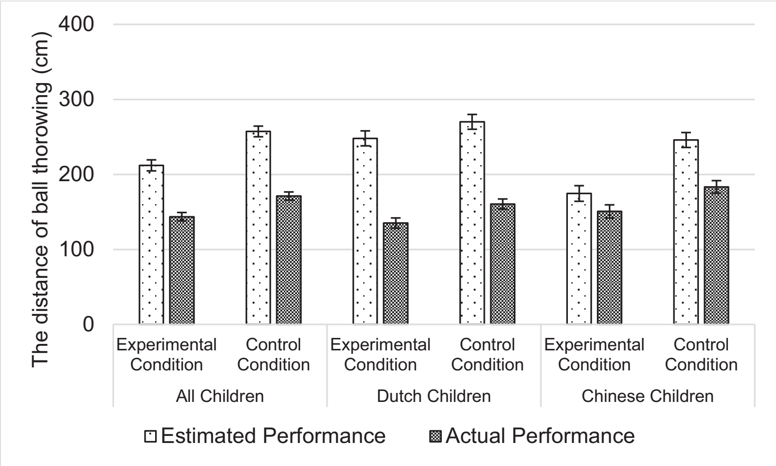
**Does the Accuracy Reward Reduce Children's Overestimation?**

For the motor task, there was a significant main effect of Performance Index; children estimated that they would perform better ($M$ = 234.7) than they actually did ($M$ = 157.4), $F$(1, 181) = 9.64, $p$ = .002, $\eta_p^2$ = .051. The hypothesized Performance Index × Condition interaction was significant, $F$(1, 181) = 4.47, $p$ = .036, $\eta_p^2$ = .024, indicating that the accuracy incentive caused children to overestimate their performance less.

We also found cultural differences. The Performance Index × Nationality interaction was significant, $F$(1, 181) = 66.16, $p$ < .001, $\eta_p^2$ = .268, indicating that Dutch children overestimated their performance more than Chinese children. Furthermore, and importantly, the effectiveness of the accuracy reward in reducing performance overestimation differed among children from the Netherlands and China (Figure 1). The Performance Index × Condition × Nationality interaction was significant, $F$(1, 181) = 6.31, $p$ = .013, $\eta_p^2$ = .034. We conducted follow-up analyses for children from both countries separately to interpret the interaction. For Chinese children, the accuracy reward led to reduced self-overestimation (i.e., $M$ = 23.9 vs $M$ = 62.6 in the experimental vs control conditions, respectively), $F$(1, 80) = 13.04, $p$ = .001, $\eta_p^2$ = .140. For Dutch children, however, the accuracy reward did not lead to reduced self-overestimation (i.e., $M$ = 113.0 vs $M$ = 109.7 in the experimental vs control conditions, respectively), $F$(1, 99) = 0.07, $p$ = .790, $\eta_p^2$ = .001 (Figure 1).

**Figure 1**

*Children's Estimated and Actual Performance on the Motor Task in Both Conditions*

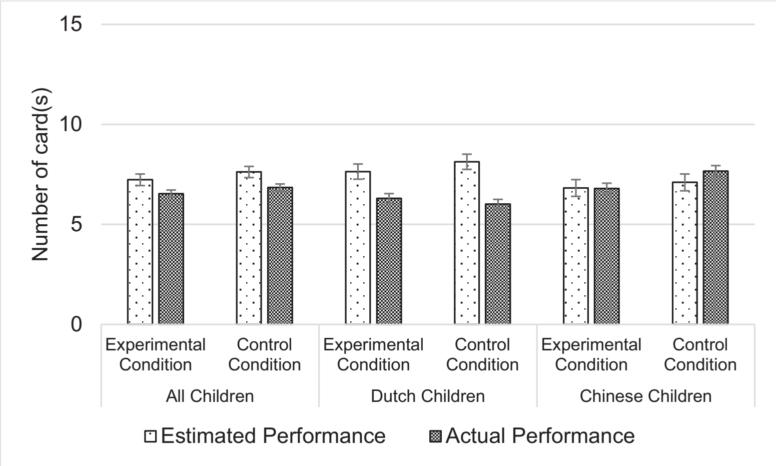



*Note.* Error bars represent standard errors.

For the memory task, surprisingly, we found no significant main effect of Performance Index, $F(1, 183) = 0.82$, $p = .366$, $\eta_p^2 = .004$. Thus, across countries and conditions, children did not estimate that they would perform better than they actually did. There was a significant interaction of Performance Index × Nationality, $F(1, 183) = 26.99$, $p < .001$, $\eta_p^2 = .129$, indicating that Dutch children overestimated more ($M = 1.73$; $M_{Estimate} = 7.89$, $M_{Performance} = 6.16$) than Chinese children ($M = -0.27$; $M_{Estimate} = 6.96$, $M_{Performance} = 7.23$) (Figure 2). However, follow-up analysis, which repeated the analysis for the Dutch sample only, found that even Dutch children did not significantly overestimate their performance, as indicated by a non-significant main effect of performance index, $F(1, 100) = 0.84$, $p = .363$, $\eta_p^2 = .008$.

**Figure 2**

*Children's Estimated and Actual Performance on the Memory Task in Both Conditions*




*Note.* Error bars represent standard errors.

**Do the Effects of the Accuracy Reward Become Progressively Stronger Across Trials?**

Because children did not overestimate their performance on the memory task, we addressed this question for the motor task only. Here, we found a significant interaction of Performance Index × Trial, $F(1.74, 314.11) = 5.58$, $p = .006$, $\eta_p^2 = .030$. Specifically, children's self-overestimation decreased significantly from Trial 1 to Trial 2, $F(1, 181) = 4.68$, $p = .032$, $\eta_p^2 = .025$, but not from Trial 2 to Trial 3, $F(1, 181) = 1.29$, $p = .258$, $\eta_p^2 = .007$. We found no evidence that the self-overestimation of children in the experimental condition decreased more steeply across trials compared with children in the control condition (which would have indicated a learning effect, such that it would have taken some time for the effect of the reward to manifest). Neither the Performance Index × Trial × Condition interaction nor the Performance Index × Trial × Condition × Nationality interaction was significant ($ps \geq .079$).

**Does the Accuracy Reward Increase the Strength of Association Between Children's Performance on a Trial and Their Subsequent Performance Estimate?**

To examine the extent to which participants incorporated performance feedback into the estimates of their future performance, we inspected the pattern of correlations between children's task performance and their performance estimates on later trials (with age and sex partialled out) for both tasks (Tables 3 and 4).

As for the motor task (Table 3), the correlations between children's actual performance on a trial and their performance estimate for the subsequent trial were strongly positive and significant. We found no differences in the strength of the pertaining correlations between children in the accuracy reward and control conditions, $p$s ≥ .114.

**Table 3**

*Correlations Between Estimates and Performances on the Motor Task*

| | Estimate2 | Estimate3 | Estimate4 | Performance1 | Performance2 | Performance3 |
|---|---|---|---|---|---|---|
| Estimate1 | .49*** | .37*** | .36*** | .30*** | .21** | .15* |
| | (.54***/.40***) | (.34**/.32**) | (.30**/.32**) | (.26*/.28**) | (.07/.23*) | (.07/.11) |
| Estimate2 | | .68*** | .52*** | **.53*** ** | .48*** | .38*** |
| | | (.71***/.62***) | (.61***/.40***) | **(.49***/.54***)** | (.46***/.45***) | (.35**/.36***) |
| Estimate3 | | | .76*** | .42*** | **.53*** ** | .56*** |
| | | | (.84***/.64***) | (.35**/.45***) | **(.48***/.52***)** | (.59***/.45***) |
| Estimate4 | | | | .34*** | .45*** | **.60*** ** |
| | | | | (.41***/.24*) | (.43***/.40***) | **(.64***/.48***)** |
| Performance1 | | | | | .70*** | .57*** |
| | | | | | (.69***/.68***) | (.60***/.53***) |
| Performance2 | | | | | | .75*** |
| | | | | | | (.65***/.82***) |

*Note.* Bold values are the correlations between performance on Trial n and estimate on Trial n + 1. Correlations under the experimental and control conditions, respectively, are reported in parentheses.

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

As for the memory task (Table 4), the correlations between children's actual performance on a trial and their performance estimates for the subsequent trial were moderately positive and significant. Again, the strength of the correlations did not differ for children in the accuracy reward and the control conditions, *p*s ≥ .096.

**Table 4**

*Correlations Between Estimates and Performances on the Memory Task*

| | Estimate2 | Estimate3 | Estimate4 | Performance1 | Performance2 | Performance3 |
|---|---|---|---|---|---|---|
| Estimate1 | .41*** | .36*** | .29*** | .12 | .18* | .11 |
| | (.37***/.49***) | (.36***/.39***) | (.28**/.31**) | (.14/.10) | (.18/.19) | (.16/.06) |
| Estimate2 | | .45*** | .48*** | **.25**** | .30*** | .04 |
| | | (.47***/.47***) | (.54***/.42***) | **(.36**/.13)** | (.41***/.20) | (.10/-.01) |
| Estimate3 | | | .63*** | .19* | **.40***** | .25** |
| | | | (.61***/.64***) | (.29**/.10) | **(.50***/.32**)** | (.39***/.11) |
| Estimate4 | | | | .19** | .33*** | **.27***** |
| | | | | (.24*/.16) | (.28**/.39***) | **(.27*/.28**)** |
| Performance1 | | | | | .59*** | .47*** |
| | | | | | (.53***/.65***) | (.34**/.58***) |
| Performance2 | | | | | | .63*** |
| | | | | | | (.59***/.66***) |

*Note.* Bold values are the correlations between performance on Trial n and estimate on Trial n + 1. Correlations under the experimental and control conditions, respectively, are reported in parentheses.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Thus, across cultures, we found no indications that children were more likely to incorporate performance feedback into their performance estimates when they were rewarded for accuracy.

Next, we inspected the same pattern of correlations for children from the Netherlands and China separately. For Dutch children, we found no significant differences between the pertaining correlations in the accuracy reward and control conditions for both tasks. For

Chinese children, we found two significant differences. On the motor task, the correlation between children's actual performance on Trial 3 and their estimated performance on Trial 4 was less strong in the control condition as compared with the accuracy reward condition (Fischer's $z = 2.58$, $p = .001$). Similarly, on the memory task, the correlation between children's actual performance on Trial 1 and their estimated performance on Trial 2 was less strong in the control condition as compared with the accuracy reward condition (Fischer's $z = 2.10$, $p = .035$). Thus, for Chinese children, we found some indications that children were more likely to incorporate performance feedback into their performance estimates when they were rewarded for accuracy.

**Robustness Analyses**

Our main research questions can also be addressed by using an alternative analytical approach that relies on a single index of children's overestimation. Although we did not preregister this approach, we conducted the analyses nonetheless. The results are reported in the Supporting Information 2 (Appendix).

Specifically, we computed an overestimation index by dividing children's estimated performance by their corresponding actual performance. For the motor task, the results replicate the finding that the accuracy incentive causes reduced overestimation in Chinese children but not in Dutch children. For the memory task, the alternative approach does allow for evaluating effects of the accuracy incentive (because it does not hinge on children's actual and estimated performance being significantly different). Here, we found little evidence that the accuracy incentive influenced the degree to which Chinese and Dutch children overestimated themselves.

## Discussion

This preregistered cross-cultural experiment examined to what extent young children's frequently observed overestimation of performance is motivated rather than due to cognitive inability alone. We did so by testing whether the promise of a reward for providing accurate performance estimates would reduce preschoolers' and kindergarteners' performance overestimation. On the motor task that we conducted, we found that it did, but only for Chinese children (not for Dutch children). Chinese children in the accuracy reward condition provided more accurate motor performance estimates than those in the control condition. Unexpectedly, on the memory task, children did not overestimate their performance to begin with, and so we were unable to examine the putative effects of the accuracy reward for this task using our preregistered approach.

Even if our experimental findings pertain to the motor task only and were not consistent across cultures, they indicate that young children's self-overestimation *can* be motivated. That is, Chinese children proved to be able to estimate their performance more accurately when it was desirable for them to do so. This is not what we would have found if they were unable to monitor their performances or calibrate their performance estimates altogether. Experts have argued that young children often engage in wishful thinking; their desires color their perceptions and beliefs about reality, which may lead them to overestimate their abilities and performances (Bernard et al., 2016; Stipek et al., 1984). They *want* to be able to perform well, and so they *think* they will be able to perform well. The current study suggests that this process is relatively malleable and shows that subsets of young children can estimate their performance more accurately when doing so is incentivized.

Our findings can be understood in light of the motive for self-enhancement—the universal tendency for individuals to view themselves favorably. The self-enhancement motive explains why individuals are often concerned with achieving well and earning social approval or acclaim; these experiences allow for experiencing oneself as a competent person worthy of approval (Alicke & Sedikides, 2009; James, 1950; Sedikides & Gregg, 2008). The motive for self-enhancement can be observed from a young age, such as in children's drive for reputation management, tendency to make self-serving attributions, or sensitivity to experiencing failure (Kelsey et al., 2018; Thomaes et al., 2017; Trzesniewski et al., 2011). As such, we propose that the motivated nature of self-overestimation that we observed in Chinese children can be seen as a manifestation of self-enhancement; even if they are able to view themselves relatively accurately, the desire to view themselves favorably positively biases their performance estimation.

What explains why we observed motivated self-overestimation in Chinese children but not in Dutch children? We speculate that providing accurate (rather than excessively flattering) self-estimates of performance is compatible with the prevalent social norm of modesty in China. From a young age, Chinese children are aware of the norm for modest self-presentation, more so than children growing up in Western societies (Luo et al., 2013; Wang & Ollendick, 2001; Wu et al., 2002; Xu et al., 2005). As such, it might not be surprising that they were more responsive than Dutch children to the incentive to estimate their performance accurately. Dutch children may have been less responsive to that incentive because of the social norm for positive distinctiveness that they are more familiar with (Gürel & Brummelman, 2020; Thomaes et al., 2017; Young-Eisendrath, 2008).

Importantly, although the accuracy reward reduced Chinese children's self-overestimation, it did not eliminate it entirely. One possible explanation is that the reward that we used to incentivize accuracy was not powerful enough to fully override the motivational appeal of providing favorable performance estimates. Another explanation, however, is that children's self-overestimation is only partly motivated. Of course, cognitive immaturity may still contribute to children's self-overestimation even if it does not fully account for it. Such cognitive immaturity effects may be especially pronounced for tasks that children are unfamiliar with and, thus, that require the allocation of limited mental resources. For example, engaging in a motor task and estimating how well one will perform is an unfamiliar and relatively effortful activity that might make the metacognitive information processing that is needed to form accurate performance predictions more challenging (Bjorklund & Green, 1992).

This may also account for why we found no overestimation on the memory task, which was relatively easy. Whereas in our previous work we asked participants to memorize new sets of pictures for each trial (Xia et al., 2022), in the current study we asked participants to memorize the same set of pictures for each trial. Not only did participants recall more pictures in the current study (i.e., an average difference of 2.6 cards), they also estimated their memory performance more accurately. This dovetails with previous evidence that children's performance overestimation is higher when they need to memorize new sets of pictures as compared with familiar ones (Lipko-Speed, 2013). Thus, the relative ease of the memory task that we used may explain why children did not self-overestimate on this task; it may have allowed children to more effectively engage in the metacognitive processing that is needed to form accurate performance estimates.

**Strengths, Limitations, and Future Research**

Our research provides the first causal test of the psychological underpinnings of young children's self-overestimation. It did so by comparing samples of children who grow up in Western (i.e., the Netherlands) and non-Western (i.e., Mainland China) countries, which allowed us to establish cultural differences. We used well-established performance prediction methodological paradigms and adopted behavioral assessments of estimated and actual performances to avoid language confounds and allow direct cross-cultural comparisons.

We also acknowledge limitations. We slightly adjusted our memory task as compared with the one we used in our previous work (Xia et al., 2022)—we used identical rather than different sets of picture cards for each trial—to make the memory and motor task procedures identical. This adjustment made the memory task easier, which may have been the reason why children did not overestimate their memory performance. This finding illustrates that although children often overestimate themselves, such overestimation is not absolute or unavoidable; in fact, task characteristics influence whether and to what extent children overestimate themselves.

We examined young children's overestimation of performance, assuming that they would not strategically adjust their actual performance to match their estimated performance. In other words, we assumed that children would always perform as well as they could. However, we found that, for the motor task, participants in the experimental condition did perform slightly worse than those in the control condition. Although we did not observe children in the experimental condition to deliberately perform worse than they could, we cannot rule out the possibility that some of them did. Thus, the results on the task should be interpreted with this caveat in mind. For future work, one approach to identify children's putative use of "opportunistic underperformance strategies" would be to measure their pre-manipulation task performance (i.e., their task performance before receiving experimental instructions that incentivize accurate performance estimates). This would allow researchers to detect improbable discrepancies between children's baseline and post-manipulation performance.

Finally, we administered the motor and memory tasks in a fixed order. Although the fixed order was necessary for our experiment, we do acknowledge the inherent downsides to such an approach. For example, we cannot rule out the possibility that, following the motor task, some children in the experimental condition realized that they were rewarded regardless of how well they did (i.e., some participants received a sticker even if their performance estimates had not been very accurate). This might have influenced how some participants approached the memory task. Future work could replicate the present study findings using a counterbalanced design.

## Conclusions

Our research provides the first causal and cross-cultural evidence that young children's self-overestimation can be motivated. Chinese children (but not Dutch children) overestimated their motor performance less when they were incentivized to do so, which suggests that cognitive immaturity (i.e., lack of ability) is not the sole explanation for the self-overestimation in which young children often engage. Moreover, our findings demonstrate that young children's self-overestimation and its psychological underpinnings are not set in stone but rather malleable and dependent on task characteristics and cultural differences.

# Chapter 4

## Children's Overestimation of Performance Across Age, Task, and Historical Time: A Meta-Analysis

Submitted as: Xia, M., Poorthuis, A. M. G., & Thomaes, S. (2023). Children's overestimation of performance across age, task, and historical time: A meta-analysis. *Manuscript submitted for publication.*

Author contributions

MX, AP, and ST conceptualized the study. MX conducted the literature search. MX and AP were responsible for the coding of the data. MX analyzed the data and wrote the first draft of the manuscript. AP and ST provided feedback on the manuscript.

**Abstract**

Children tend to overestimate their performance on a variety of tasks and activities. The present meta-analysis examines the specificity of this phenomenon across age, tasks, and five decades of historical time. We established the ratio between children's self-estimates of task performance and their actual (i.e., objectively measured) task performance, and examined its moderating factors (i.e., sample age, type of task, year of data collection). We included 217 effect sizes from 39 published articles (3,879 participants recruited from pre- and primary schools; 89.7% North American/European descent; 49.5% females). Children’s self-overestimation was robust across tasks, with their estimates of performance being 1.3 times their actual performance. In addition, children's self-overestimation decreased with sample age and increased with the year of data collection.

*Keywords:* overestimation, task performance, childhood, overconfidence, metacognition, self-enhancement

# Children's Overestimation of Performance Across Age, Task, and Historical Time: A Meta-Analysis

Children can be strikingly confident, if not *over*confident, of their abilities, personal attributes, and knowledge. This is particularly striking when it comes to how they estimate their performance on various tasks and activities they engage in. Research has repeatedly shown that young children, especially, tend to overestimate how well they perform on cognitive, motor, or other types of tasks. Initial research evidence suggests this is true for children from both Western (i.e., North American, Western European) and Eastern (i.e., Chinese, Omani) cultures, suggesting that self-overestimation is not just a cultural phenomenon (Was & Al-Harthy, 2018; Xia et al., 2022; Yussen & Levy Jr, 1975).

At the same time, several important questions are still unanswered. In particular, does children's self-overestimation decrease gradually over the course of childhood? What is the relative magnitude of children's self-overestimation across different types of tasks? And might children's self-overestimation be, in part, a sign of the times, perhaps reflecting an increased cultural emphasis on individual competence or "standing out"? The present meta-analysis synthesizes the literature to examine systematic variation in children's self-overestimation and addresses each of these central questions.

## Self-Enhancement and Self-Overestimation

Self-enhancement refers to various psychological phenomena that involve taking a tendentiously positive view of oneself (Sedikides & Gregg, 2008). For example, adults tend to think of themselves as above average in desirable attributes (i.e., *better than average*, Zell et al., 2020). They also overestimate their ability to have control over events (Thompson, 1999), have overly charitable views of their own kindness and selflessness (*holier than thou*, Epley & Dunning, 2000), and believe they can complete tasks with more ease, and faster, than they actually do (*planning fallacy*, Buehler et al., 1994).

Similar self-enhancement tendencies can be found in children (Harter, 1996, 2006; Trzesniewski et al., 2011). For example, in the early primary school years, children expect their own desirable attributes (e.g., being generous) to stay the same in the future, while they expect their less desirable attributes (e.g., being selfish) to change for the better (Diesendruck & Lindenbaum, 2009; Lockhart et al., 2002). Primary school-aged children also tend to make self-serving attributions (i.e., they attribute positive events to themselves and negative events to external factors), even more so than adolescents and adults (Marsh, 1986; Mezulis et al., 2004). Furthermore, they mostly attribute desirable traits to themselves, about 4 times more than neutral attributes, and about 11 times more than non-desirable attributes (Thomaes et al., 2017).

Children’s self-overestimation is another form of self-enhancement, reflecting a tendency to estimate one’s own task performance or ability more positively than reality (e.g., their objective task performance) warrants. Children’s self-estimates are based on information from various sources, including their performance experience on similar tasks (e.g., social feedback, test results), their general self-efficacy beliefs (e.g., attributions of personal competence and effectiveness), and motivational factors (e.g., effort attributions, wishful thinking; Harter, 2015; Schneider, 1998). Even if some of these sources may provide realistic competence-relevant information, which should enable children to make relatively accurate self-estimates, children’s self-estimations tend to be positively biased (Trzesniewski et al., 2011; Xia et al., 2022).

Studies typically establish such bias in children’s self-estimates by comparing two units of information: children’s (subjective and prospective) self-estimates of performance on some task or activity, and an (objective) measure of their actual performance. Accordingly, children's self-overestimation is typically studied in performance contexts that involve a specific task or activity for which children’s performance can be measured. For example, on recall tasks, children are asked to estimate how many items (e.g., of a set of pictures) they will be able to recall, which is then compared to how many items they actually recall.

**The Development of Self-Overestimation**

Children overestimate themselves from a young age, and across various performances and abilities. For example, in one classic study (Yussen & Levy Jr, 1975), participants in various age groups overestimated how well they would perform on a memory task, but children (i.e., both kindergarteners and third graders) overestimated their performance more than adults did. Later research found that kindergarteners and primary school aged children overestimate themselves on a wide range of performances and abilities. For example, as compared to objective performance criteria, they overestimate their physical abilities (e.g., walking backward on a balance beam; Almeida et al., 2017; Schneider, 1998), their problem-solving abilities (e.g., performance on mathematical or spelling tasks; Heath & Glen, 2005; Stipek & Tannatt, 1984), and their knowledge and understanding (e.g., knowledge about the meaning of difficult words; Bain & McCallum, 1986; Kominsky & Keil, 2014). Children's self-overestimation is not just limited to tasks or performance contexts they are unfamiliar with. In fact, they generally are persistent in overestimating themselves across multiple trials of the same task, even when receiving accurate performance feedback with each trial (Lipko et al., 2009; Xia et al., 2022; Yussen & Levy Jr, 1975). This is different for adults, who typically become more realistic, or even underestimate themselves, after gaining task-related experience (Finn & Metcalfe, 2007).

Cognitive developmental theory suggests that children's self-overestimation should decrease over the course of childhood. In particular, children's metacognitive ability (i.e., the ability to reflect upon their own cognitive processes and behavior) develops from early to late childhood (Kuhn, 2000; Lyons & Ghetti, 2010). Older children are better able to monitor, realistically perceive, and retain information about their abilities (Schneider, 1998).

Accordingly, they become more effective at learning from performance experience and feedback (Franchak, 2019; Plumert, 1995), allowing them to incorporate relevant information into their performance estimates. Notably, theory also suggests that it may be beneficial for young children to overestimate themselves: like other forms of cognitive immaturity (e.g., immature information-processing abilities that set the stage for young children's rapid language acquisition; Bjorklund, 1997), overestimation may have adaptive value. In particular, it has been argued that overestimation enables young children to take on new and challenging tasks, and to persevere in the face of setbacks, providing them with critical opportunities for learning and growth (Bjorklund, 1997; Bjorklund & Green, 1992).

What does empirical evidence suggest about how children's self-overestimation changes with age? A number of studies have explored if children from different age groups differ in the extent to which they overestimate themselves. This work found that preschoolers and early primary school aged children generally overestimate themselves more than older children; on memory tasks (Bertrand et al., 2017; Cottini et al., 2021; Shin et al., 2007; Was & Al-Harthy, 2018; Yussen & Levy Jr, 1975), but also on other types of cognitive tasks (Kominsky & Keil, 2014), and on motor tasks (Powel & Toni, 1994). Thus, evidence suggests that, over the course of childhood, children's self-estimation generally become less positively biased. In fact, some studies found that older children do not overestimate themselves at all, or even underestimate themselves (Lee et al., 1988; Schneider et al., 2000; Worden & Sladewski-Awig, 1982).

However, inconsistent findings have been obtained as well. For example, some studies have found that older primary school aged children do not provide more realistic self-estimates than kindergarteners, although this differed across types of tasks (Yussen & Berman, 1981). Importantly, most relevant research has used single age-group designs, and those studies that did use multiple age-group designs covered relatively narrow age ranges.

Accordingly, systematic evidence on age trends in self-overestimation over the course of childhood is still lacking.

**Children's Overestimation Across Types of Tasks**

Children's overestimation has been studied using various tasks and activities, including tasks requiring motor skills and cognitive skills. We define motor tasks as tasks that primarily require movement of the muscles of the body to be fulfilled. Motor tasks that have been used in research on children's overestimation include, for example, jumping (e.g., jumping as far as possible), throwing (e.g., throwing as accurately as possible), and cycling (e.g., trying bike challenges, such as riding in a circle). We define cognitive tasks as tasks that primarily require some cognitive activity, such as decision-making, problem-solving, memorizing, paying attention, or making judgments, to be fulfilled. Cognitive tasks that have been used in overestimation research typically include memory tasks (e.g., memorizing and then recalling stimuli such as pictures, words, etc.). However, other cognitive tasks have been used as well. We define them as tasks that primarily require the retrieval and application of information from existing knowledge to solve problems, such as listing differences between words or solving math problems.

The processes needed for children to estimate their performance differ, at least in part, for motor and cognitive tasks. For example, in estimating their motor performance, children will often rely on perceptual information about the environment (e.g., *how big is the opening of the bucket I need to throw the balls in?*). In estimating their cognitive performance, children will primarily rely on what they know about their own mental functioning (e.g., *how good am I with words?*). Given these differences, it is possible that the extent to which children overestimate themselves will vary across types of tasks. For example, in one study that used identical procedures to assess children's estimates of their memory and motor performance, children overestimated themselves on the motor task, but not on the memory task (Xia et al., 2023). We will synthesize the literature to explore possible differences in the extent to which children overestimate themselves on different types of tasks.

**Historical Trends**

Over the past several decades, individualistic values and practices (e.g., that emphasize independence, self-directedness, and uniqueness) have increased—not just in Western societies, but in most societies around the world (Santos et al., 2017). Historical trends in individualism have complex and societally diverse causes but are often driven by socioeconomic factors such as economic wealth, transitions from agricultural to industrial and post-industrial economies, and residential mobility (Kashima & Kashima, 2003; Oishi, 2010; Santos et al., 2017). Individualistic trends are reflected in practices, including socialization practices, with downstream consequences for how children develop (Greenfield, 2009; Kagitcibasi, 2017; Keller, 2022). In increasingly individualistic societies, children are more often exposed to messages (e.g., from parents, at school, or in the media) that emphasize the importance and normativity of being unique, personally competent or successful, or to "stand out from the crowd" (Gürel & Brummelman, 2020; Thomaes et al., 2017; Twenge & Campbell, 2009). It is possible that such messages are reflected in how children, from young age, view themselves and their abilities and performances. Therefore, another aim of the meta-analysis was to synthesize evidence to test whether children overestimate themselves more in more recently conducted studies.

**The Current Meta-Analysis**

The current meta-analysis synthesizes the empirical evidence base on children's self-estimated and actual task performance to determine the magnitude and robustness of children's self-overestimation across sample age, type of task, and five decades of historical time. We included studies that sampled children from preschools or primary schools. We calculated the accuracy of children's self-estimated task performance (which can theoretically range from self-underestimation to self-overestimation) as the ratio of children’s self-estimated and actual (i.e., objectively measured) performance on the same task.

We hypothesized, first, that children's self-estimated task performance would be more favorable than their actual task performance, reflected in an overestimation effect (i.e., ratio) greater than one. Second, we hypothesized that young children would overestimate themselves more than older children. Third, while we did not specify a hypothesis, we explored whether there are systematic differences in the extent to which children overestimate themselves across tasks (i.e., motor tasks, memory tasks, and other cognitive tasks). Fourth, we hypothesized that children would overestimate themselves more in more recently conducted studies.

## Method

Research materials (i.e., data, search strategy) are available on the Open Science Framework at https://osf.io/78dyw/.

### Literature Search

We searched PsycINFO, Eric, and Web of Science to identify articles on children's self-estimated and actual task performance. We used the following Boolean string: ((perform* adj4 estimat*) OR (perform* adj4 predict*) OR (perform* adj4 judge*) OR (abilit* adj4 estimat*) OR (abilit* adj4 predict*) OR (abilit* adj4 judge*) OR (competence adj4 estimat*) OR (competence adj4 predict*) OR (competence adj4 judge*) OR overestimat* OR overconfiden* OR underestimat* OR underconfiden* OR (estimat* adj2 accura*) OR (predict* adj2 accura*) OR (judge* adj2 accura*)) AND (activit* OR experiment* OR task OR test) AND (toddler* OR child* OR pupil* OR kindergarten* OR nurser* OR preschool* OR (primary school) OR (elementary school). We considered articles published until October 2022.

**Inclusion and Exclusion Criteria**

To be eligible for inclusion, articles had to meet the following criteria:

**Accessibility.** Articles had to (a) be accessible online or through interlibrary loan, or by the author(s) upon request, (b) be written in English, and (c) be an empirical article published in a peer-reviewed journal.

**Participants.** Articles had to (d) report on research that sampled children from preschools or primary schools. If the information on school type was lacking, participants' mean age could not exceed 13.0 years, and the sample could not include any individual aged 15.0 years or older. Articles had to (e) report on research in which participants were not recruited based on a diagnosis of a psychological disorder (e.g., ADHD, ASD, or learning disorder).

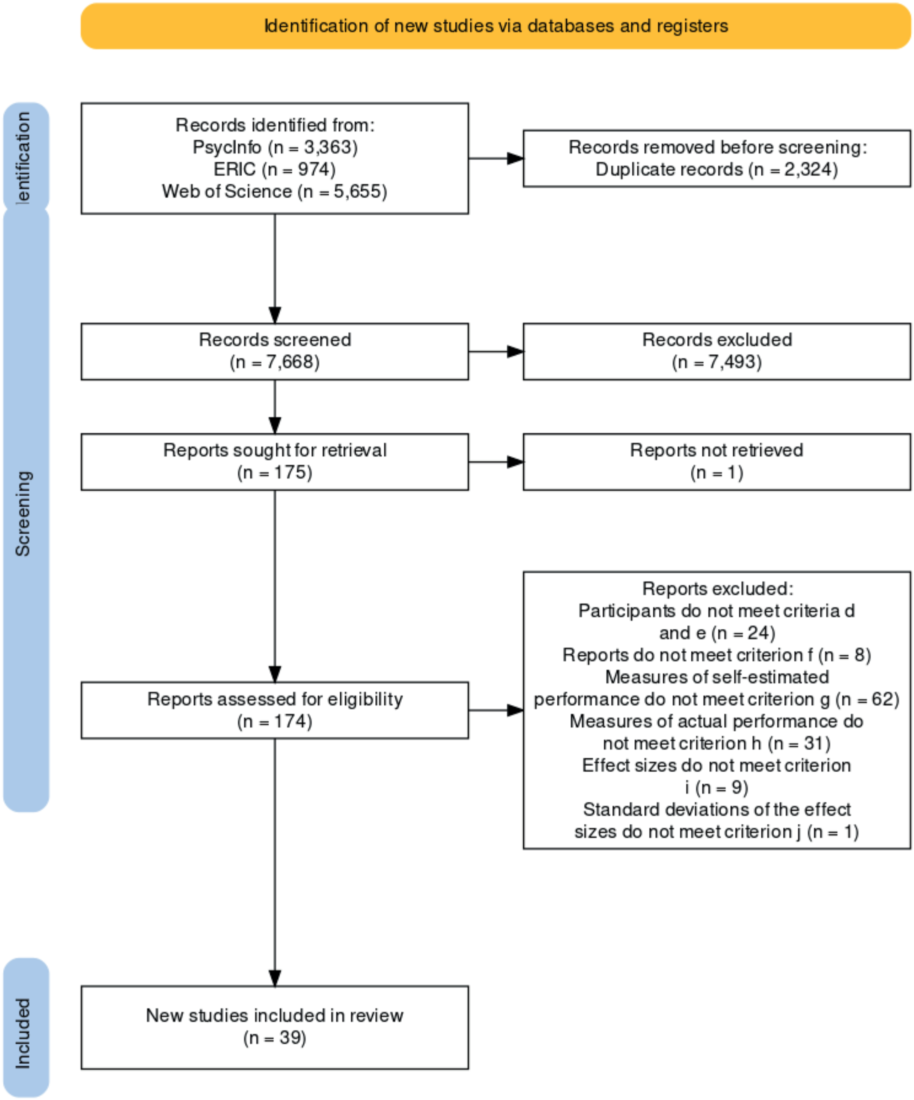
**Measurements.** Articles had to (f) report on research in which measures of both self-estimated and actual performance were obtained, where (g) the self-estimated performance measure refers to participants' estimation (prediction or judgment) of their upcoming performance on a specific task (activity, experiment, or test), and (h) the actual performance measure refers to participants' objective performance on the same task (e.g., standardized test performance, number of memorized items, distance of ball throw).

**Effect sizes.** Articles had to (i) report effect size(s) (or other information from which effect size(s) could be computed, or effect sizes were provided by the author(s) upon request), and (j) report the standard deviation(s) of the effect size(s) (or these standard deviations were provided by the author(s) upon request or could be imputed).

Our systematic literature search identified 7,668 articles published in peer-reviewed journals. Based on a screening of titles and abstracts, we examined the full texts of 174 articles to assess eligibility. A total of 39 articles (217 effect sizes, 3,879 participants) met the inclusion criteria (see Figure 1, Haddaway et al., 2022).

**Figure 1**

*Article Selection Flowchart*




More than half of the articles reported on studies conducted in North America (*N* = 22, 56.4%), one-third reported on studies conducted in Europe (*N* = 13, 33.3%), and the remaining articles reported on studies conducted in New Zealand, Turkey, Oman, and China

($N = 4$, 10.3%). For the articles that reported the sex composition of their samples (i.e., 87.2% of articles), the mean proportion of female participants was 49.5%.

**Obtaining Effect Sizes**

We extracted participants' estimates of their performance on the upcoming task as "self-estimated performance," and the corresponding, objectively measured task performance as "actual performance." Using Ratio of Means (RoM; Friedrich et al., 2008), we operationalized the accuracy of children's self-estimation by calculating the ratio between the children's self-estimated and actual performance (i.e., self-estimated performance / actual performance). Values greater than one indicate overestimation and values smaller than one indicate underestimation (a value of one indicates perfectly accurate self-estimation). We used RoM because it allows for comparing overestimation indices across tasks, regardless of the specific measurement units they are based on (i.e., regardless of the measurement units that indexed self-estimated and actual performance).

We included effect sizes (i.e., RoM indices) as follows: (1) We included all relevant effect sizes from each article. This includes effect sizes for subsamples of different ages (e.g., four-year-olds and six-year-olds; Schneider, 1998), different grades (e.g., second graders and fourth graders; Miller et al., 1988), different nationalities (e.g., Chinese children and Dutch children; Xia et al., 2022), and different sexes (e.g., boys and girls; Kolovelonis & Goudas, 2018). For studies that used within-subjects designs, we included all relevant effect sizes for different tasks (e.g., *Face* and *Building* picture tasks; Elmose & Happé, 2014), different task difficulties (e.g., tossing beanbags into a basket from a four, six, and eight feet distance; Powel & Toni, 1994), different trials using the same materials (e.g., first attempt and second attempt; Lipko-Speed, 2013), and different trials using different materials (e.g., different word lists for each trial; Bird, 1984). (2) We excluded effect sizes from (sub)samples of children with psychological diagnoses (e.g., autism; Elmose & Happé, 2014), or other special conditions (e.g., experienced gymnasts; Peker et al., 2021). (3) For studies with experimental designs with a control group, we included effect sizes from the control group only (e.g., *no-norm* control group; Yussen & Levy Jr, 1975). (4) For studies in which children were assigned to groups based on their task performance, we included the overall effect size only (and ignored effect sizes for different groups; Almeida et al., 2017). (5) For studies in which children provided multiple self-estimates for a single performance, we averaged self-estimates (Wojcik et al., 2022).

**Extraction and Coding of Moderators**

We coded samples' mean age. If this was not reported, we coded the grade levels that participants were sampled from and imputed the average ages of children in those grade levels in the pertaining country as the sample's mean age.

We coded task type as motor task (e.g., throwing, jumping, kicking, walking backward, cycling), memory task (e.g., memorizing and recalling words, pictures, or sentences), or other cognitive task (e.g., solving math problems, mastering basic concepts, listing differences of word pairs, spelling).

We extracted the year of data collection for each included study. If this information was missing from the article (which it mostly was), we requested the information from the authors. If we did not receive the information ($N = 12$ studies, 30.8%), we calculated year of data collection as two years before the year of publication (Gentile et al., 2009; Oliver & Hyde, 1993).

A random selection of roughly two-thirds of the selected studies (i.e., $N = 27$ studies, 69.2%) was coded by two independent coders. Cohen's κ was satisfactory (all κs > 0.97). Discrepancies were discussed and resolved between coders.

An overview of study characteristics and coding for each included report is provided in the Supporting Information 3 (Appendix).

**Imputation for Missing Standard Deviations**

Standard deviations of self-estimated and/or actual performance were missing for 76 effect sizes (35%). We imputed these with the pooled standard deviations of other included studies that used the same type of task (i.e., motor, memory, other cognitive) and the same measurement scale.

**Analytic Strategy**

We conducted meta-analytic computations in R using the metafor package (Schwarzer et al., 2015; Viechtbauer, 2010). We applied a three-level meta-analytic model. This model does not assume homogeneity of effect sizes within studies and allows us to use all relevant effect sizes to achieve maximum statistical power (Assink & Wibbelink, 2016). We adopted a random-effects model, as the number of participants in the included studies was unbalanced (i.e., ranging from $N = 10$ to $N = 303$; Borenstein et al., 2021). Our three-level random-effects model thus took into account the heterogeneity of samples (sampling variance, level 1), the variance between effect sizes within studies (within-study variance, level 2), and the variance between effect sizes across studies (between-study variance, level 3).

Ratio of Mean (RoM) scores come from a skewed distribution. A log transformation was therefore used to transform the RoM data to approximately conform to normality. We computed the natural logarithm of each RoM score and its standard error (SE) using the following equations (Friedrich et al., 2011).

$$\ln(\mathrm{RoM}) = \ln\left(\frac{\mathrm{mean_{est}}}{\mathrm{mean_{per}}}\right)$$

$$\mathrm{SE}[\ln(\mathrm{RoM})] = \sqrt{\frac{1}{\mathrm{n}}\left(\frac{\mathrm{SD_{est}}}{\mathrm{mean_{est}}}\right)^2 + \frac{1}{\mathrm{n}}\left(\frac{\mathrm{SD_{per}}}{\mathrm{mean_{per}}}\right)^2}$$

We obtained the sampling variance by squaring the standard error (SE).

Prior to the analyses, we checked for outlying effect sizes—i.e., standardized *z*-values (of the natural logarithm of RoM) larger than 3.29 or smaller than -3.29 (Tabachnick &

Fidell, 2013). We identified four outliers in one study (Kominsky & Keil, 2014), and replaced them with the nearest observation that was not an outlier (i.e., winsorizing; Blaine, 2018).

To answer our research questions, we, first, estimated the overall accuracy of children's self-estimation (i.e., the overall overestimation effect) by fitting a three-level random-effects model to the data. We used the REstricted Maximum Likelihood estimation method (REML) to estimate parameters (Hox et al., 2017; Viechtbauer, 2005). We back-transformed the results to obtain the overall effect and 95% confidence interval (CI), as follows (Friedrich et al., 2011):

95% CI = $\exp\{[\ln(\mathrm{RoM})] \pm 1.96 \times \mathrm{SE}[\ln(\mathrm{RoM})]\}$

Second, we conducted two separate log-likelihood-ratio tests to determine whether the within-study variance (level 2) and between-study variance (level 3) were significant. We examined how the total variance was distributed over the three levels of the meta-analytic model (Cheung, 2014).

Third, we examined potential moderator effects by the sample mean age, task type, and year of data collection. Following Hox et. al. (2017), we first tested the three potential moderators in separate models, and then tested all significant moderators in a single model. We first fitted linear meta-regression models. For continuous moderators (sample mean age and year of data collection), we further fitted nonlinear models (Restricted Cubic Spline model; Stone & Koo, 1985), which allowed us more flexibility in examining associations between observed effects (lnRoM) and moderators, and allowed us to identify floor or ceiling effects (by choosing the number and position of "knots" where the piecewise cubic polynomials are connected). Finally, we compared models based on their log-likelihood and information criteria.

We performed sensitivity analyses to determine if our findings were affected by decisions regarding effect size extraction, standard deviation imputation, and treatment of outliers. Regarding effect size extraction, we tested if the findings were different if we averaged the effect sizes across trials (resulting in 161 effect sizes), or across trials and within-subjects conditions (unless tasks were completely different, e.g., when children worked on *Math, Triangles, Similarities*, and *Beanbag* tasks; Miller et al., 1988) (resulting in 134 effect sizes). We also tested if findings were different if we excluded (rather than imputed) effect sizes for which standard deviations were unavailable, and if we used a different imputation method (i.e., if we imputed the largest or smallest, rather than pooled, standard deviation of studies using the same type of task and measurement scale). Finally, we tested if findings were different if we kept all outliers or removed outliers from the analyses (rather than winsorized outliers).

We inspected funnel plots to identify potential publication bias and quantitatively assessed asymmetry with a modified Egger's regression test (Nakagawa & Santos, 2012). Specifically, we included a measure of precision (i.e., standard error) as a predictor in our three-level model. We determined the "small-study" effect sizes that contributed to an asymmetrical funnel plot and removed them. We then computed the overall effect size and tested moderation effects again (Griffin et al., 2021).

## Results

### Overall Effect

We found strong support for the hypothesis that children overestimate themselves. The overall RoM was significantly larger than one (RoM = 1.332, 95% CI = [1.199, 1.481]), $t(216) = 5.332$, $p < .001$. Overall, children estimated that their task performance would be one-third better than it turned out to be.

### Heterogeneity in Effect Sizes

We found significant variability in effect sizes within studies ($p < .001$) and between studies ($p < .001$). We examined how the total variance was distributed over the three levels of our meta-analytic model. We found that 3.01% of the total variance could be attributed to sampling variance (i.e., level 1), 34.11% could be attributed to differences between effect sizes within studies (i.e., level 2), and 62.88% could be attributed to differences between effect sizes across studies (i.e., level 3). We concluded that there was substantial variation in effect sizes within and between studies, making the following moderator analyses meaningful.

## Moderator Analyses

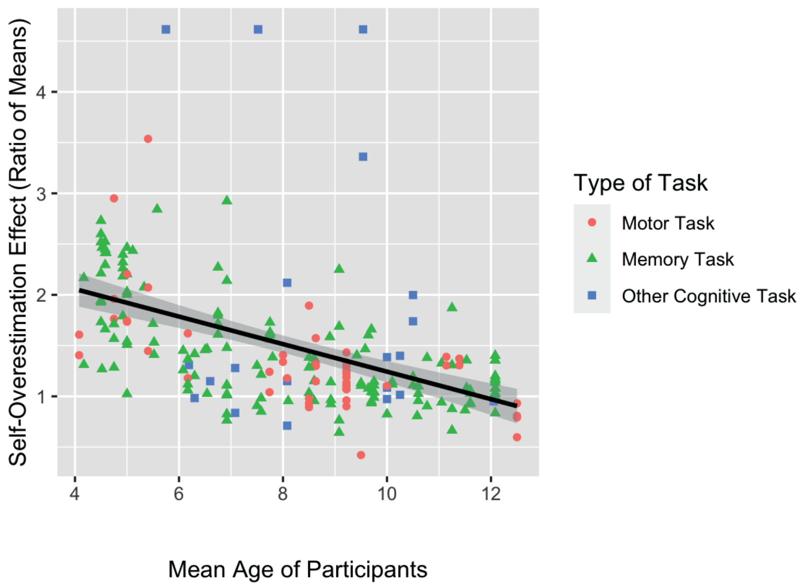
### *Age*

The mean ages of the included samples ranged from 4.08 to 12.50 (pooled sample mean age = 8.19). As hypothesized, sample mean age was a significant moderator of the overall effect, $F(1, 215) = 43.721$, $p < .001$. We found less overestimation in older samples, $B = -0.079$, $t(215) = -6.612$, $p < .001$ (Figure 2). In fact, in the oldest samples, with a sample mean age of around 11 to 12, children's self-estimates approached accuracy (RoMs ≈ 1). By contrast, in the youngest samples, with a sample mean age of around four to five, children's self-estimates were most strikingly inaccurate (i.e., positively biased), with RoMs greater than two (i.e., indicating that participants thought they would perform twice as well as they actually did).

We additionally conducted nonlinear analyses by fitting three- and four-knots Restricted Cubic Spline models. We found that the linear model ($AIC = 55.001$) provided better fit for the data than the non-linear models ($AIC$s > 81.267; See Supporting Information 4 in the Appendix). Thus, our findings suggest that children's self-overestimation gradually decreases with age (from 4 to 12); we found no evidence that children's self-overestimation decreases more pronouncedly at certain ages than at others.

**Figure 2**

*Self-Overestimation (Ratio of Means) by Sample Mean Age*

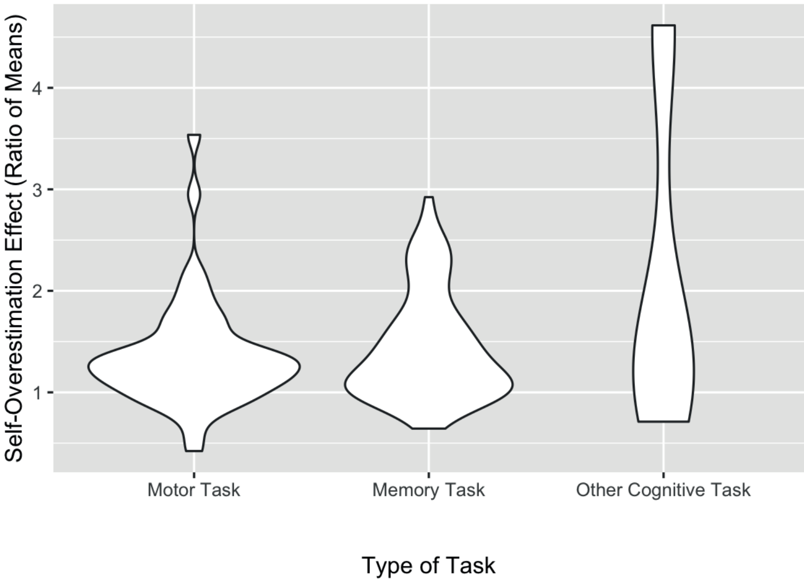



***Task Type***

Of the 39 included studies, 8 studies (20.5%) involved motor tasks, 21 studies (53.8%) involved memory tasks, 7 studies (17.9%) involved other cognitive tasks, and 3 studies (7.7%) involved more than one type of task. Of the 217 effect sizes, 49 (22.6%) pertained to motor tasks, 146 (67.3%) pertained to memory tasks, and 22 (10.1%) pertained to other cognitive tasks. In our analyses, we used the motor task as the reference category to compare the other categories (i.e., memory and other cognitive tasks) with. We found no significant moderating effect of task type on the overall effect, $F(2, 214) = 0.204$, $p = .816$ (Figure 3). Thus, there was no evidence that children's self-overestimation depends on task type.

**Figure 3**

*Children's Self-Overestimation (Ratio of Means) and Its Distribution Across Task Types*




### *Data Collection Year*

The years in which studies were conducted ranged from 1968 to 2019. We found a significant moderating effect for year of data collection on the overall effect, $F(1, 215) = 5.365$, $p = .021$ (Figure 4). As hypothesized, we found higher levels of overestimation in more recently conducted studies, $B = 0.008$, $t(215) = 2.316$, $p = .021$. We again conducted nonlinear analyses by fitting three- and four-knots Restricted Cubic Spline models. The linear model ($AIC = 88.135$) fitted the data better than the non-linear models ($AIC$s > 159.905), suggesting that children's self-overestimation has gradually increased over the past several decades (See Supporting Information 4 in the Appendix).

**Figure 4**

*Changes in Children's Self-Overestimation (Ratio of Means) by Year of Data Collection*


***Multiple Moderator Model***

Because we found that participants' mean age and year of data collection were significant moderators, we simultaneously added these variables to our model. We first confirmed that at least one of the regression coefficients of the moderators significantly deviated from zero, as indexed by a significant omnibus test, $F(2, 214) = 24.377$, $p < .001$. The regression coefficients for both participants' mean age ($B = -0.077$, $t(214) = -6.442$, $p < .001$) and year of data collection ($B = 0.006$, $t(214) = 2.011$, $p = .046$) significantly deviated from zero. We thus concluded that both child age and year of data collection have robust and unique moderating effects.

**Sensitivity Analyses**

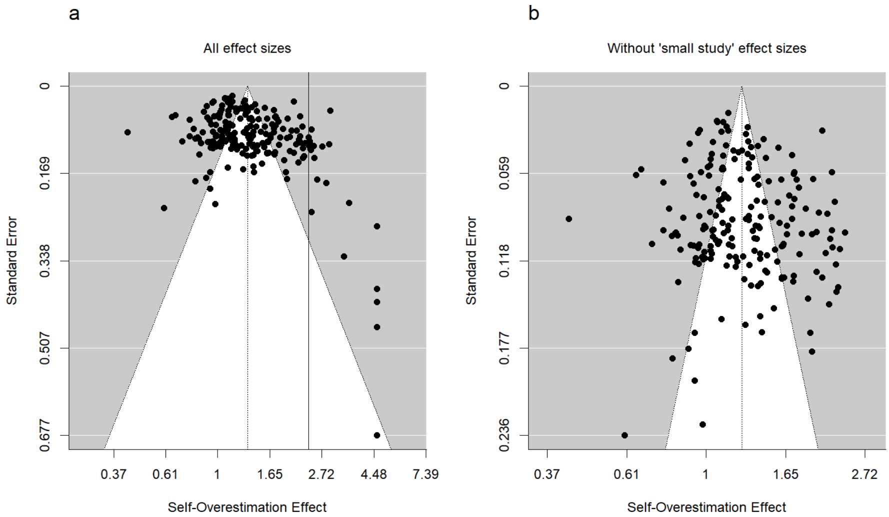
Our main findings were not affected by decisions regarding effect size extraction, standard deviation imputation, or treatment of outliers. With regard to the overall overestimation effect, children's estimates of their task performance consistently were at least 1.3 times their actual performance (RoMs ranged from 1.303 to 1.402, $t$s ≥ 4.667, $p$s < .001). Moreover, the moderating effects for the sample mean age (-0.081 ≤ $B$s ≤ -0.073, $t$s ≤ -4.989, $p$s < .001) and year of data collection (0.009 ≤ $B$s ≤ 0.007, $p$s ≤ .038) were consistently significant. The moderating effect for task type was consistently non-significant, $p$s ≥ .588 (see Supporting Information 5 in the Appendix).

**Publication Bias**

Visual inspection of the distribution of obtained effect sizes in a funnel plot suggested asymmetry (Figure 5a). There were a few small studies (i.e., as signified by the large standard error) with relatively large effect sizes (i.e., which can be seen in the middle- and bottom-right of the plot), and no equivalent small studies with relatively small effect sizes (i.e., which would have otherwise been seen in the bottom-left of the plot). We used a modified Egger's Regression test which confirmed that there was significant funnel plot asymmetry ($B$ = 2.82, $SE$ = 0.56, 95% CI = [1.72, 3.93], $p$ < .001). After removing extreme effect sizes (i.e., the effect sizes in the middle- and bottom-right of the plot, $n$ = 20), we no longer observed funnel plot asymmetry ($B$ = 1.45, $SE$ = 0.75, 95% CI = [-0.02, 2.92], $p$ = .053; see Figure 5b). The overall effect size was only slightly lower (RoM = 1.250, 95% CI = [1.155, 1.354], $p$ < .001), and still indicated that children overestimated their task performance. Thus, our overall finding that children overestimate themselves is not just due to publication bias. Moreover, the moderating effects for sample mean age ($B$ = -0.064, $t$(195) = -5.992, $p$ < .001) and year of data collection ($B$ = 0.006, $t$(195) = 2.369, $p$ = .019) remained significant, while the moderating effect for task type continued to be non-significant, $F$(2, 194) = 0.045, $p$ = .956.

**Figure 5**

*Funnel Plot Displaying Self-Overestimation Effect Sizes as a Function of Precision (i.e., Standard Error)*




*Note*. Funnel plots display standardized effect size estimates (circles) as a function of precision (i.e., standard error).

(a) All effect sizes ($k = 217$) were plotted. The funnel plot was centered on the overall effect size for children's self-overestimation effect (RoM = 1.332) indicated by a vertical dashed line. As shown, there was significant asymmetry ($p < .001$), with effect sizes in the middle- and bottom-right of the plot contributing to small-study effects.

(b) Upon removal of these effects, the funnel plot showed symmetry (which indicates no presence of small-study effects, $p = .053$), and the overall self-overestimation effect size was only slightly smaller (RoM = 1.250).

## Discussion

The present meta-analysis provided the first quantitative review of children's (over)estimation of their own task performance. We synthesized data from 39 published articles (i.e., 217 effect sizes, 3,879 participants), to examine the magnitude and robustness of children's overestimation across age, types of task, and historical time.

First, our findings confirmed that children overestimate their task performance. Specifically, aggregated across samples of preschool and primary school children, diverse types of tasks, and recent as well as older studies, children estimated their task performance to be one-third better than they actually performed. Second, our results suggest that children's self-overestimation gradually decreases with age, from early to late childhood. We found no evidence for accelerated development towards more accurate self-estimation at certain ages. Third, we found no evidence that children's self-overestimation systematically varies across different types of tasks. Children showed similar levels of self-overestimation on motor tasks, memory tasks, or other types of cognitive tasks. Fourth, we found that children overestimated themselves more strongly in studies that were more recently conducted.

**Theoretical and Applied Considerations**

What may be the consequences of children's self-overestimation? It seems possible that their self-overestimation predisposes children to make mistakes or experience setbacks on tasks that are harder than they anticipate. Indeed, they may insufficiently acknowledge what it takes to successfully complete or perform well on various tasks. Relatedly, children's self-overestimation may be associated with risk taking, which may be especially consequential in the physical domain. For example, research has found that young children who overestimate their motor abilities more than other children are more likely to experience accidental injuries (Plumert, 1995; Plumert & Schwebel, 1997).

At the same time, children's self-overestimation may have adaptive value as well. Young children are universal novices (Shin et al., 2007), who struggle on the many unfamiliar tasks and challenges that they face in their everyday lives, from tying their shoelaces and learning to swim, to paying attention in class and learning to count (Bjorklund, 1997; Bjorklund & Green, 1992). To the extent that they think they will be better able to complete these tasks than they are, they may more actively seek out, engage with, and persist on these tasks, even if they prove to be challenging. As such, self-overestimation may ultimately benefit children's learning and growth. This is consistent with Bandura's (1982) theorizing on self-efficacy which, he argued, encourages individuals to invest effort needed for optimal performance and learning. In early and middle childhood, it may thus be beneficial for children to hold positive illusions of their ability and task performance. Indeed, there is some evidence to suggest that, over time, children who more strongly overestimate themselves learn more and perform better as compared to counterparts with more realistic self-perceptions (Shin et al., 2007).

What accounts for the finding that children's self-overestimation decreases with age? Self-overestimation is rooted, in part, in young children's immature metacognitive abilities, which develop over time (Coutinho et al., 2005; Krueger & Mueller, 2002). In particular, older children are better able to monitor and incorporate relevant information into their performance estimates. Changes in educational context may matter as well. As children progress in school, they are more frequently evaluated (e.g., in the form of teacher feedback, grades), they have learned to compare their own competencies and performances with those of their peers, and they are better able to incorporate such evaluations and social comparisons into their self-perceptions—these developments may jointly help them to develop more realistic self-estimates (Dweck, 2002; Harter, 2015).

What does it mean that children's self-overestimation generalizes across different types of tasks? We propose that it might indicate that children's overestimation is partially driven by motivational factors (e.g., driven by the pursuit of good performance), rather than being exclusively the result of some (meta)cognitive deficit—indeed, children overestimated themselves on motor tasks as much as on cognitive tasks, even if motor tasks require less cognitive processing. Of course, it is possible that children's self-overestimation does vary across task dimensions that we did not assess. Task difficulty could be one such dimension: to the extent that it is hard for (young) children to estimate the difficulty of the tasks they engage with (Gweon et al., 2017), it is possible that they are more likely to overestimate their performance on challenging tasks that tax their abilities (rather than easy tasks that match their abilities). Indeed, in one study, 4- and 6-year-olds made more accurate performance estimates on the tasks they considered to be the least difficult (Schneider, 1998). Although we considered coding task difficulty in the present research, we thought it would not be feasible to do so (i.e., most method sections of the articles we included lack the level of detail needed for coding of task difficulty by non-involved researchers). Future work could experimentally manipulate task difficulty to test whether it influences the accuracy of children's performance estimation.

Our finding that children overestimated themselves more in more recently conducted studies is reminiscent of claims that young people, at least in Western cultures, may have developed more favorable self-views over the past few decades (Gentile et al., 2010; Twenge et al., 2017; but see Trzesniewski et al., 2008). This trend might reflect global increases in individualism, manifest in social norms emphasizing the importance of personal competence, self-prominence, and favorable self-views (Hamamura, 2012; Santos et al., 2017; Thomaes et al., 2017; Zhang et al., 2017). From young age, children are exposed to messages that convey it is important or desirable to think well of themselves and their abilities. For example, through mass media, children are often exposed to role models that radiate self-assurance, and many parents and teachers seek to help children to think well of themselves (Brummelman et al., 2015; Hewitt, 1998; Thomaes et al., 2017). To the extent that young children internalize such messages, or seek to live up to social norms, they will be tempted to perceive themselves as positively as they can, or to express confidence, even overconfidence, in their own abilities.

Our research may allow parents, educators, and other professionals working with children to better understand children's self-perceptions of performance and ability. In young children (in the preschool and early primary school years), self-overestimation is normative, and there is little reason to assume that its manifestations, such as in classroom settings, should pose a reason for concern or remedial action (even though high levels of overestimation of motor abilities are associated with risk of accidental injury). For older children (in the late grade school years), it becomes increasingly normative to estimate task performance relatively accurately. It is unknown whether or how children who continue to hold inflated self-perceptions well into late childhood would benefit from intervention. We speculate that non-normative levels of overestimation could, in some cases, be early signs of deviant self-development trajectories—for example, narcissistic self-development trajectories (Thomaes & Brummelman, 2016). Future research will need to explore the potentially maladaptive nature of self-perceptions that continue to be inflated in late childhood (or beyond), and what can be done to help children develop more accurate self-perceptions.

**Strengths, Limitations, and Future Research**

Our meta-analysis examined (positive) bias in children's self-perceived task performance and ability. We included studies that allowed for a direct comparison between children's performance estimates on specific tasks and their actual, objectively measured performance on the same tasks. We used the ratio of means (*RoM*) method to operationalize bias in children's self-estimation, which allowed us to consider studies from various research fields and using various types of tasks. As such, our study is the first to systematically examine not only the magnitude of children's self-overestimation, but also the extent to which it varies across children's age, types of tasks, and (historical) time.

We also acknowledge limitations. Meta-analyses do not allow for precision tests of age differences. We indexed age based on sample mean ages. While informative, sample mean ages do not reflect underlying sample age ranges, which renders tests of age differences less rigorous. Future studies could use alternative meta-analytic methods that include individual participant data (IPDMA) to allow for more precise estimates of age effects (Smith et al., 2016). Furthermore, in the absence of relevant longitudinal research, we relied on cross-sectional studies to examine how children's self-overestimation varies with age. A priority for future research is to conduct longitudinal research to examine within-person changes in self-overestimation over time.

The bulk of research included in our meta-analysis was conducted with children growing up in Western (i.e., North American, Western European) countries. Previous work has found evidence to suggest that children's self-overestimation is not a uniquely Western phenomenon (Was & Al-Harthy, 2018; Xia et al., 2022, 2023). Still, the relative lack of studies conducted in non-Western countries means we were unable to systematically examine potential cultural differences in children's self-overestimation (and its age, task, and historical time contingencies). Consistent with recent calls to diversify and globalize developmental science, we encourage increased sampling of children growing up in non-Western countries to improve cultural generalizability (Nielsen et al., 2017).

**Coda**

"I am sure I will recall all 10 pictures you just showed me!" The present meta-analysis showed that children, and especially young children (i.e., preschoolers, kindergarteners, and early primary school aged children), routinely overestimate their performance on a wide range of tasks. These illusory self-perceptions are even more pronounced in recent cohorts of children than they were in children who grew up several decades ago. We hope that our study will spur further research into children's inflated self-perceptions of performance and ability, their causes and consequences, and their cross-cultural variation.

# Chapter 5

# General Discussion

Children often show a striking confidence when faced with challenges, and they tend to be overconfident in their abilities, understanding, and knowledge. This is reflected, for example, in how they estimate their task performance. The purpose of this dissertation was to expand our knowledge of self-overestimation of task performance in childhood, in terms of its variation across tasks, age, and historical time, its cross-cultural generality, and its psychological underpinnings.

## Summary of Findings and Implications

The findings and implications of this dissertation are drawn from two cross-national empirical studies and a meta-analysis, organized in three chapters, and centered around five overarching aims. Table 1 summarizes the main findings of the chapters in this dissertation.

### Aim 1. Do Children Overestimate Their Task Performance? To What Extent Do They Do So? Do They Overestimate Their Task Performance Equally Across Different Tasks?

Children's overestimation of task performance can be operationalized by comparing their self-estimates of task performance with corresponding measures of their actual task performance. In Chapter 2, young children's self-estimates of their task performance were significantly higher than their actual performance on both a motor and a memory task. In Chapter 3, young children's self-estimates were again significantly higher than their actual performance on the motor task, but their estimated and actual performance did not differ on the memory task (which was modified to be easier compared to the memory task used in Chapter 2). In Chapter 4, a meta-analysis revealed that children's (ages 4 to 12) self-estimates were on average 1.3 times higher than their actual performance, as established across various types of tasks. Thus, overall, this dissertation supports the view that children tend to overestimate their task performance.

**Table 1**

*Findings on Self-Overestimation in Childhood, Organized by Chapter and Dissertation Aim*

| Chapter | Aim | Findings |
|---|---|---|
| 2 | 1 | Children overestimated their performance on both motor and memory tasks. |
| | 4 | Both Dutch and Chinese children overestimated their performance on motor and memory tasks, and they did so to a similar extent. |
| | 5 | Children persisted in self-overestimation across trials, on both motor and memory tasks. |
| | 5 | On the motor task, children's actual performance on one trial and their self-estimate on the next trial were positively associated. No such positive association was found on the memory task. |
| | 5 | On the motor task, Chinese (but not Dutch) children' self-overestimation decreased significantly from Trial 1 to 2. On the memory task, neither Chinese nor Dutch children's self-overestimation changed across trials. |
| | 5 | Children overestimated their peers' performance on the motor task about as much as they overestimated their own. |
| | 5 | Children overestimated their peers' performance on the memory task even more than they overestimated their own. Tests of cross-national differences found that this was true for Chinese children, but not for Dutch children. |
| 3 | 1 | Children overestimated their performance on the motor task, but not on the memory task. |
| | 5 | On the motor task, children's self-overestimation decreased from Trials 1 to 2, but not from Trials 2 to 3. |
| | 5 | Children's actual performance on one trial and their self-estimates on the next trial were positively associated, on both motor and memory tasks. |
| | 5 | On the motor task, children's self-overestimation decreased when they were promised a reward for providing accurate self-estimates. |
| | 5 | Tests of cross-national differences found that this effect was true for Chinese children, but not for Dutch children. |
| 4 | 1 | Children overestimated their task performance. Their self-estimates were 1.3 times higher than their actual performance. |
| | 1 | Children's self-overestimation did not differ between memory, motor, and other cognitive tasks. |
| | 2 | The magnitude of children's self-overestimation decreased with age. |
| | 3 | The magnitude of children's self-overestimation increased with year of data collection (ranging from 1968 to 2019). |

***Implications***

Acknowledging self-overestimation as a common bias in children's self-representations offers valuable insight for parents, educators, and other professionals working with children. The finding that young children's self-overestimation is pervasive, and generalizes across tasks and cultures, may suggest there is no reason for undue concern when self-overestimation is observed in an individual child. Although the consequences of children's self-overestimation were not examined in the present dissertation, other work does speak to such consequences. Self-overestimation is considered to have adaptive value for young children (Bjorklund & Beer, 2016; Shin et al., 2007). Specifically, self-overestimation is assumed to encourage young children to seek out and persist on new challenges, maximizing opportunities for learning and growth. Bandura's self-efficacy theory suggests that self-overestimation allows children to feel efficacious even if they lack experience on the tasks and challenges they face, and motivates them to exert the necessary effort to achieve optimal performance (Bandura, 1982). Indeed, research by Shin et al. (2007) found that children with higher levels of self-overestimation showed better recall performance on subsequent trials (i.e., they showed greater gains or fewer losses), compared to children with lower levels of self-overestimation. That said, self-overestimation, in particular of motor skills, may also come with certain risks. For example, there is some evidence that children who overestimate their abilities are more likely to attempt activities beyond their skill level and to engage in risky behaviors, which can lead to serious injury (Plumert, 1995; Plumert & Schwebel, 1997).

We found little evidence that children's self-overestimation varied across types of tasks. This might suggest that children's estimates of task performance are based on similar cognitive processes, regardless of the type of mental or physical activity these tasks require. In addition, suggestive clues to a possible effect of task difficulty on children's self-overestimation were unexpectedly obtained in Chapters 2 and 3. Some previous work already suggested that children self-overestimate less when they perceive tasks to be easier or more familiar (Lipko-Speed, 2013; Schneider, 1998). Our finding that children overestimated their performance on the memory task in Chapter 2, but not in Chapter 3, is consistent with this possibility. The memory task used in Chapter 3 is a simplified version of the memory task used in Chapter 2: It used the exact same pictures to be memorized in all three trials (rather than different pictures for each trial, as in Chapter 2). Thus, our pattern of findings on both versions of the memory task suggests that children are more likely to make accurate self-estimates when the tasks they work on are relatively easy.

**Aim 2. How Does Self-Overestimation Vary by Age in Childhood?**

In Chapter 4, we identified an age trend in self-overestimation of task performance across childhood: Children’s level of self-overestimation decreased with age, ranging from around 2.0 at age 4 (i.e., indicating that children this age expect to perform twice as well as they actually do), to about 1.0 at age 12 (i.e., indicating accuracy). This decline was gradual and did not show more pronounced changes at specific ages.

***Implications***

The finding that younger children overestimate their task performance more than older children, mirrors developmental trends in children's (meta)cognition. Children's (meta)cognitive abilities (i.e., planning, monitoring, and evaluating; Escolano-Pérez et al., 2019; Veenman et al., 2006) play a key role in children's self-estimation, and these abilities develop over time (Coutinho et al., 2005; Krueger & Mueller, 2002). As a result, older children are better able to use relevant information to form a realistic sense of their abilities and performances, which is reflected in the accuracy of their self-estimates. In addition, as children get older, more sources of information about their abilities and performance are available. For example, they receive more frequent normative assessment results, performance feedback from teachers, and social comparison information, and this information contributes to the accuracy of their self-estimates (Butler, 1989; Stipek & Daniels, 1988). Notably, this age trend is consistent with the assumed decreasing adaptiveness of self-overestimation across childhood, which posits that optimal development of older children requires the formation of more accurate self-views (Escribano & Díaz-Morales, 2014; Gresham et al., 2000). In addition, recent research has observed that age-related changes in children's self-overestimation also coincide with similar changes in children's wishful thinking bias: older children are less likely to exhibit wishful thinking than younger children (Wente et al., 2020).

**Aim 3. How Does Children's Self-Overestimation Vary Across Historical Time?**

In Chapter 4, we found evidence that children overestimated their task performance more in more recently conducted studies. This historical increase in self-overestimation was gradual, rather than more pronounced during certain time periods. As times have progressed and societies have changed, children growing up in contemporary society show more self-overestimation compared to previous generations of children.

***Implications***

The finding that children overestimate their task performance more in recently conducted studies is novel. At the same time, it is reminiscent of some evidence suggesting that individuals of other ages (i.e., older children, adolescents, and young adults) have also developed more favorable views of themselves (e.g., higher levels of self-esteem and, possibly, narcissism) over the past decades (Gentile et al., 2010; Twenge et al., 2017). During this time, many societies worldwide have increasingly adopted a more urban, individualistic orientation (Greenfield, 2009), which is reflected in social norms that increasingly emphasize the importance of individual competence, self-prominence, and favorable self-views (Hamamura, 2012; Santos et al., 2017; Thomaes et al., 2017; Zhang et al., 2017). This historical trend has been observed for Western societies, but there is evidence that it has generalized to other parts of the world as well (Santos et al., 2017). As a result, children growing up in contemporary societies around the world may be increasingly exposed to messages that convey it is important or desirable for them to think well of oneself. We propose that our finding that young children increasingly overestimate their task performance may reflect this historical trend.

**Aim 4. Do Children in Non-Western Societies Overestimate Their Task Performance?**

The dissertation explored differences in children's self-overestimation in Western and non-Western societies. Specifically, Dutch society emphasizes individualistic values and tends to promote self-prominence, self-reliance, and uniqueness, whereas Chinese society traditionally places more emphasis on collectivistic values and tends to promote modesty (Kim et al., 2010; O'Mara et al., 2012; Santos et al., 2017; Sedikides et al., 2015). These cultural differences are reflected in the socialization process, especially in the goals and values that parents (and other socializing agents) convey to children (Tamis-LeMonda et al., 2007). And yet, we found only few differences between Dutch and Chinese children's self-estimation of task performance. In Chapter 2, both Dutch and Chinese young children overestimated their task performance, they did so to a similar extent, and they both continued to do so even after receiving accurate performance feedback. In Chapter 3, on the motor task, both Dutch and Chinese children overestimated their task performance—the only difference being that Chinese (but not Dutch) children provided more accurate self-estimates when rewarded for doing so. In Chapter 4, we were unable to address cultural differences due to a lack of cultural diversity in the studies we meta-analyzed, i.e., only 10.3% of articles were not from North America or Europe. The evidence from this dissertation is by no means definitive, but it does suggest that self-overestimation is not a uniquely Western phenomenon.

### *Implications*

What does it mean that young children growing up in non-Western societies (or, at least, in China) also overestimate their task performance, even when they come from a culture where children are traditionally socialized from an early age to avoid self-aggrandizement and show modesty (Wang, 2004; Wu et al., 2002; Xu et al., 2005)? One possibility is that the tendency for children to overestimate themselves is widespread and relatively independent of cultural variation, which would be consistent with the view of self-overestimation as a functional adaptation that benefits young children's learning and growth (Bjorklund & Beer, 2016; Shin et al., 2007). At the same time, it is also possible that individualistic trends have permeated modern-day China (and perhaps especially the urban areas where we conducted our research) to such an extent, that the Chinese children we tested have been exposed to quite similar norms and socialization messages as their Dutch counterparts—more than we anticipated at the outset of the research (Bian et al., 2022; Chen et al., 2005; Santos et al., 2017). Thus, while we consider our cross-cultural findings to be valuable, caution is needed in interpreting these findings.

## Aim 5. Why Do Children Overestimate Their Task Performance?

### *Immature (Meta)Cognition*

One explanation for children's self-overestimation refers to their immature (meta)cognitive abilities, which may prevent them from effectively monitoring their own performance and incorporating performance-relevant information into their self-views (i.e., the immature (meta)cognition hypothesis; Flavell, 1979). This dissertation addressed this hypothesis by (1) examining potential changes in children's self-estimates across task trials (i.e., after receiving performance feedback), and (2) testing associations between children's actual task performance and their estimated task performance on the next trial. We found several indications that 4- and 5-year-old children, in fact, are able to monitor their performance and incorporate performance-relevant information into their performance estimates—they just do not consistently or fully do so. For example, on the motor task in Chapter 3, while children continued to overestimate their task performance across trials, we did find a significant decrease in self-overestimation from Trial 1 to 2 (but not from Trial 2 to 3). Moreover, on the motor task in Chapters 2 and 3, and the memory task in Chapter 3, children's actual performance on one trial and their self-estimate on the next trial were significantly positively associated. Thus, our findings suggest that it is not metacognitive deficiency, but rather, metacognitive inconsistency that influences young children's self-overestimation: They fail to consistently or fully incorporate performance-relevant information into their performance estimates.

***Wishful Thinking***

A complementary explanation for children's self-overestimation emphasizes motivational processes, suggesting that children's self-overestimation comes (in part) from their desire to being and appearing competent, and from not distinguishing between their desires and their expectations (i.e., the wishful thinking hypothesis; Schneider, 1998; Stipek et al., 1984). This dissertation addressed this hypothesis by (1) examining the effect of accuracy rewards on children's self-overestimation, and (2) testing potential differences between children's estimates of their own and their peers' performance. The results provide inconsistent support for the wishful thinking hypothesis. In Chapter 3, the promise of receiving a reward for providing accurate self-estimates reduced the extent to which Chinese children, but not Dutch children overestimated their performance on the motor task (for the memory task, children did not overestimate their performance in the first place). This finding lends some support for the view that children's self-overestimation can be at least partially motivated, consistent with the wishful thinking hypothesis. However, in Chapter 2, we did not find that children's self-estimates were more favorable than their other-estimates, which is inconsistent with this hypothesis. Future work will need to employ alternative research methods to scrutinize if, when, and how wishful thinking may drive young children's self-overestimation.

***Implications***

The findings reported in this dissertation suggest that while both (meta)cognitive and motivational factors may contribute to children's self-overestimation, neither of these factors alone can sufficiently account for children's self-overestimation. Consistent with previous work (Lipko et al., 2009; Schneider, 1998), the findings show that while children are able to monitor and incorporate their actual task performance into their performance estimates, they do so inconsistently or only partly, exhibiting some degree of cognitive immaturity. Furthermore, the findings show that while children's desire to perform well *can* color their perceptions of their task performance (cf., Bernard et al., 2016; Stipek et al., 1984), wishful thinking does not always or fully account for their self-overestimation. Speculatively, (meta)cognitive and motivational factors may jointly account for children's self-overestimation. For example, it is possible that children's strong desire to perform well drives their failure to consistently or fully make use of metacognitive abilities that would otherwise lead them to form realistic performance estimates. Future research is needed to further scrutinize the joint cognitive and motivational processes that explain why young children overestimate their performance.

## Strengths, Limitations, and Future Research

The dissertation contributes to our understanding of self-overestimation in childhood through cross-cultural and experimental research as well as meta-analysis. It reports the first empirical synthesis of research on children's self-overestimation, allowing for an accurate estimation of the degree to which children overestimate their task performance, and its variation across factors such as types of task and children's age, among others. The studies reported in Chapters 2 and 3 are the first to directly compare the self-estimates of children growing up in Western and non-Western societies. This is important, given that the existing literature is heavily biased toward children growing up in Western contexts. It provides a first step in building a more culturally diverse and representative understanding of children's self-estimation of their task performance.

This dissertation also contributes insight into the psychological underpinnings of children's self-overestimation by examining both (meta)cognitive and motivational factors. The study reported on in Chapter 3 provides the first causal test of the motivational nature of children's self-overestimation. Finally, this dissertation revealed, using rigorous meta-analytic methods, that the extent to which children overestimate their task performance has increased over the past decades—a finding that contributes to our knowledgebase of how people's self-views have changed in recent history, possibly as a corollary of widespread trends in individualism (Greenfield, 2016; Inglehart & Oyserman, 2004; Santos et al., 2017).

The dissertation has limitations as well. Our samples of Chinese children were drawn from urban areas in China, which have seen impactful socio-cultural changes in recent years. As such, our findings from the Chinese samples cannot be assumed to be representative for Chinese children in general. Self-enhancement is on the rise, perhaps especially in urban areas of China, with consequences for parenting styles and children's adjustment (Chen et al., 2009; Chen & Li, 2012). Future research in other, more rural parts of China is needed to provide a more comprehensive understanding of cross-cultural and within-cultural differences. Moreover, research in other non-WEIRD societies (i.e., non-East Asian societies) is needed to build a global evidence-base, and provide a more comprehensive test of the universality of children's self-overestimation.

The meta-analysis in Chapter 4 examined how children's self-overestimation co-varies with age during childhood. While valuable, meta-analytic tests of age effects are relatively imprecise. That is, these tests rely on sample mean ages, rather than individual participant ages. Moreover, we meta-analyzed cross-sectional studies, but not longitudinal studies (which were not available). Thus, our meta-analytic findings do not speak to change in self-overestimation over time. Future research could use longitudinal methods to examine the development of self-(over)estimation in children over time, and possibly across a wider age range. In particular, we found for the oldest samples in our meta-analysis (with a mean age of approximately 12) that self-overestimation effects hardly existed anymore or, in fact, shifted to self-*under*estimation. Future research could explore developmental trends in individual self-estimation, not just in children but across the lifespan, to provide a comprehensive understanding of developmental change in individuals' self-estimation.

Theory suggests that self-overestimation in early childhood may be adaptive, but empirical evidence is virtually non-existent. This dissertation does not speak to the adaptiveness of self-overestimation. Future research could examine the consequences of children's self-overestimation across various ages, to test the possibility that it is specifically young children who benefit from overestimating themselves. Such research could examine short term consequences of self-overestimation, such as in terms of children's task motivation, persistence, and performance. It could also examine longer term consequences, such as the development of competence and its downstream consequences for social and emotional development (e.g., self-esteem, well-being).

## Conclusion

Children overestimate their task performance, and they do so to a similar degree for different types of tasks. Such self-overestimation is not a uniquely Western phenomenon: Dutch and Chinese children tend to show similar levels of self-overestimation. That said, we also found some cultural differences, in that Dutch children tend to be tenacious in their self-overestimation, while Chinese children are more inclined to realistically adjust their self-overestimation if the situation calls for it. Younger children overestimate their task performance more than older children, and recent generations of children have shown greater self-overestimation than previous generations. Why do children overestimate their task performance? According to the findings of this dissertation, children's self-overestimation is rooted in both (meta)cognitive and motivational factors. Cognitively, children are not always effective in incorporating performance-related information in their performance estimates; and motivationally, children's performance estimates are sometimes colored by a desire for good performance. This dissertation thus contributes deeper understanding of whether, when, where, and why children overestimate their task performance.

# Appendix

# Supporting Information

**Supporting Information 1 – Deviations From Preregistered Analysis Plan (Chapter 2)**

The analyses reported in Chapter 2 ('Young Children's Overestimation of Performance: A Cross-Cultural Comparison') deviate from the analysis plan that we preregistered at aspredicted.org (#29787) in four ways (https://aspredicted.org/tu8ib.pdf). We list and explain the changes that we made here. We made change #1 before data analysis had begun. We made changes #2-4 after data analysis had begun.

Please note that what we labeled 'hypothesis 2' in the preregistration, is labeled 'hypothesis 3' in the article (and vice versa).

| Change | Where? | What? | Why? |
|---|---|---|---|
| 1 | Results | To test our preregistered hypotheses 1-3, we used repeated measures ANOVA analyses, rather than a series of T-Tests (preregistered hypotheses 1 and 2) or one-way ANOVAs (preregistered hypothesis 3). | We decided to use repeated measures ANOVA analyses to address our preregistered hypotheses 1-3 for two reasons. First, these analyses allow us to test the hypotheses both across samples (i.e., for Chinese and Dutch children together) and between samples (i.e., comparing Chinese and Dutch children). Second, these analyses provide a more conservative test of our hypotheses. We would have needed to test our hypotheses using separate T-Tests or one-way ANOVAs for each trial, thus increasing the risk of type 1 error. By using repeated measures ANOVAs, we kept the risk of Type 1 error low. |
| 2 | Results | We added a 2 (Performance Index: peer-estimation or actual) × 3 (Trial: 1, 2, or 3) × 2 (Nationality: Chinese or Dutch) repeated measures ANOVA before we tested our preregistered hypothesis 2. | In the preregistration, we assumed that children would overestimate the performance of their peers. We later decided that we would need to conduct an independent analysis to confirm this assumption. |
| 3 | Results | For our preregistered hypothesis 3, we did not analyze if children's overestimation of their peers' | In the preregistration, we assumed that children would make more accurate estimations of their |

| | | | |
|---|---|---|---|
| | | performance would gradually decrease across trials. | peers' performance as compared to their own performance. We anticipated doing follow-up analyses to learn more about this discrepancy and test the possibility that it would be due to children gradually making more accurate estimations of their peers' performance (but not their own performance) across trials. However, we found no support for the hypothesis that children would make more accurate estimations of their peers' performance in the first place. This finding made further analysis superfluous. |
| 4 | Results | We added correlational analyses as an additional test of our preregistered hypothesis 3. Specifically, we inspected correlations between children's actual performance on task trials and their performance estimations on subsequent trials. | We conducted these correlational analyses to provide further insight into whether children use performance feedback to inform their performance estimations. In doing so, we conform to an analytic approach that has been used before to address this question (Lipko et al., 2009). |

**Supporting Information 2 – Robustness Analyses (Chapter 3)**

Children's overestimation can be computed as a single index. We wanted to establish to what extent our main findings generalize if we rely on such an overestimation index.

**Analytic Strategy**

We computed an overestimation index by dividing children's estimated performance by their corresponding actual performance. We then log-transformed that outcome. For both tasks, we conducted a 2 (Condition: Experimental or Control) × 3 (Trial: 1, 2, or 3) × 2 (Nationality: Dutch or Chinese) repeated measures ANCOVA. As in the analyses reported in the manuscript, we used age and sex as covariates in the analyses. To explore cultural specificity, we conducted follow-up 2 (Condition: Experimental or Control) × 3 (Trial: 1, 2, or 3) repeated measures ANCOVA for children from both countries separately.

**Descriptive Analyses**

Tables 1 and 2 present the descriptive statistics for children's overestimation on the motor and memory tasks.

**Table 1**

*Children's Overestimation on the Motor Task*

| | All Children | | Experimental Group | | Control Group | | Dutch Children | | Chinese Children | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Overestimation | 2.03 | 1.3 | 1.91 | 1.2 | 2.15 | 1.3 | 2.43 | 1.5 | 1.55 | 0.7 |
| Overestimation | 1.51 | 0.7 | 1.55 | 0.7 | 1.46 | 0.6 | 1.75 | 0.7 | 1.21 | 0.4 |
| Overestimation | 1.54 | 0.7 | 1.56 | 0.8 | 1.52 | 0.6 | 1.76 | 0.8 | 1.27 | 0.4 |

*Note.* Scores reflect children's overestimation, indexed by their estimated performance divided by their actual performance.

**Table 2**

*Children's Overestimation on the Memory Task*

| | All Children | | Experimental Group | | Control Group | | Dutch Children | | Chinese Children | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Overestimation | 1.34 | 0.9 | 1.38 | 1.0 | 1.29 | 0.7 | 1.55 | 0.9 | 1.08 | 0.8 |
| Overestimation | 1.17 | 0.6 | 1.16 | 0.6 | 1.17 | 0.6 | 1.27 | 0.6 | 1.04 | 0.6 |
| Overestimation | 1.18 | 0.7 | 1.05 | 0.5 | 1.31 | 0.9 | 1.28 | 0.7 | 1.06 | 0.7 |

*Note.* Scores reflect children's overestimation, indexed by their estimated performance divided by their actual performance.

**Children's Overestimation on the Motor Task**

For the motor task, we found no difference between conditions in overestimation, $F(1, 181) = 0.16$, $p = .690$, $\eta_p^2 = .001$. Thus, across countries, the accuracy reward did not cause children to overestimate their performance less. This was independent of trial, as evident from a non-significant Trial × Condition interaction, $F(1.70, 308.21) = 1.76$, $p = .178$, $\eta_p^2 = .010$.

We did find cultural differences, however. Dutch children showed more pronounced overestimation ($M = 1.75$) than Chinese children ($M = 1.24$), $F(1, 181) = 61.36$, $p < .001$, $\eta_p^2 = .253$. We conducted follow-up analyses for children from both countries separately. For Chinese children, the accuracy reward led to reduced self-overestimation, $F(1, 80) = 5.52$, $p = .021$, $\eta_p^2 = .065$. For Dutch children, however, the accuracy reward did not lead to reduced self-overestimation, $F(1, 99) = 2.44$, $p = .122$, $\eta_p^2 = .024$. There was no significant interaction of Trial × Condition in the Chinese sample ($F(1.67, 133.54) = 1.25$, $p = .286$, $\eta_p^2 = .015$), nor in the Dutch sample ($F(1.71, 169.37) = 0.60$, $p = .524$, $\eta_p^2 = .006$).

Thus, notwithstanding some differences pertaining to the aggregate sample, these results are largely consistent with the main results for the motor task reported in the manuscript. On the motor task, the accuracy incentive causes reduced overestimation in Chinese children, but not in Dutch children.

**Children's Overestimation on the Memory Task**

For the memory task, we also found no difference between conditions in overestimation, $F(1, 182) = 0.60$, $p = .440$, $\eta_p^2 = .003$. Thus, across countries, the accuracy reward did not influence the degree to which children overestimated their performance. Again, this was independent of trial, as evident from a non-significant Trial × Condition interaction, $F(1.87, 341.13) = 1.24$, $p = .291$, $\eta_p^2 = .007$.

Again, we did find cultural differences. Dutch children showed more pronounced overestimation ($M = 1.20$) than Chinese children ($M = 0.89$), F(1, 182) = 26.92, $p < .001$, $\eta_p^2 = .129$. We conducted follow-up analyses for children from both countries separately. There was no difference between conditions in the Dutch sample ($F(1, 100) = 3.31$, $p = .072$, $\eta_p^2 = .032$), nor in the Chinese sample ($F(1, 80) = 0.32$, $p = .576$, $\eta_p^2 = .004$). In the Dutch (but not the Chinese) sample, however, the Trial × Condition interaction was significant, $F(1.96, 195.85) = 4.35$, $p = .015$, $\eta_p^2 = .042$. Indeed, from Trial 2 to Trial 3, the overestimation of Dutch children decreased somewhat in the experimental group (from 1.09 to 0.97), but rebounded in the control group (from 1.19 to 1.35), $F(1, 100) = 4.92$, $p = .029$, $\eta_p^2 = .047$.

Together, these results for the memory task add to those reported in Chapter 3. Recall that the analyses reported in that chapter did not provide evidence that children's estimated performance significantly differed from their actual performance, and so we were not able to test the effect of the accuracy incentive. Here, using an alternative analytical approach, we found little evidence that the accuracy incentive influenced children's overestimation, and this was mostly independent of whether children were Chinese or Dutch (with the exception of change in overestimation from Trial 2 to 3 in the latter group).

## Supporting Information 3 – Articles Included in the Meta-Analysis (Chapter 4)

**Table 1**

*Characteristics of 39 Articles Included in the Meta-Analysis*

| Article | *N* | Sex | Country | Age | Task Type | Data Collection | RoM |
|---|---|---|---|---|---|---|---|
| Almeida et al., 2017 | 303 | unknown | Portugal | 8.63 | motor | 2014 | 1.148 |
| Almeida et al., 2017 | 303 | unknown | Portugal | 8.63 | motor | 2014 | 1.574 |
| Almeida et al., 2017 | 303 | unknown | Portugal | 8.63 | motor | 2014 | 1.328 |
| Almeida et al., 2017 | 303 | unknown | Portugal | 8.63 | motor | 2014 | 1.299 |
| Bain & McCallum, 1986 | 72 | 51.4 | USA | 6.19 | other cognitive | 1985 | 1.308 |
| Bain et al., 1989 | 63 | 52.4 | USA | 6.60 | other cognitive | 1988 | 1.149 |
| Bertrand et al., 2017 | 19 | 57.9 | France | 5.11 | memory | 2015** | 2.436 |
| Bertrand et al., 2017 | 20 | 60.0 | France | 7.80 | memory | 2015** | 1.386 |
| Bird, 1984 | 18 | 100 | New Zealand | 9.70 | memory | 1982** | 1.043 |
| Bird, 1984 | 18 | 100 | New Zealand | 9.70 | memory | 1982** | 0.950 |
| Bird, 1984 | 18 | 0 | New Zealand | 9.70 | memory | 1982** | 1.661 |
| Bird, 1984 | 18 | 0 | New Zealand | 9.70 | memory | 1982** | 0.933 |
| Bird, 1984 | 16 | 100 | New Zealand | 11.60 | memory | 1982** | 0.938 |
| Bird, 1984 | 16 | 100 | New Zealand | 11.60 | memory | 1982** | 1.028 |
| Bird, 1984 | 16 | 0 | New Zealand | 11.60 | memory | 1982** | 1.062 |
| Bird, 1984 | 16 | 0 | New Zealand | 11.60 | memory | 1982** | 0.924 |
| Bird, 1984 | 10 | 100 | New Zealand | 12.08 | memory | 1982** | 1.158 |
| Bird, 1984 | 10 | 100 | New Zealand | 12.08 | memory | 1982** | 1.026 |
| Bird, 1984 | 10 | 0 | New Zealand | 12.08 | memory | 1982** | 1.202 |
| Bird, 1984 | 10 | 0 | New Zealand | 12.08 | memory | 1982** | 1.135 |
| Bird, 1984 | 10 | 100 | New Zealand | 12.08 | memory | 1982** | 1.174 |
| Bird, 1984 | 10 | 100 | New Zealand | 12.08 | memory | 1982** | 1.083 |
| Bird, 1984 | 10 | 0 | New Zealand | 12.08 | memory | 1982** | 1.034 |
| Bird, 1984 | 10 | 0 | New Zealand | 12.08 | memory | 1982** | 0.993 |
| Bird, 1984 | 10 | 100 | New Zealand | 12.08 | memory | 1982** | 0.835 |

| Article | *N* | Sex | Country | Age | Task Type | Data Collection | RoM |
|---|---|---|---|---|---|---|---|
| Bird, 1984 | 10 | 100 | New Zealand | 12.08 | memory | 1982** | 0.994 |
| Bird, 1984 | 10 | 0 | New Zealand | 12.08 | memory | 1982** | 1.077 |
| Bird, 1984 | 10 | 0 | New Zealand | 12.08 | memory | 1982** | 1.402 |
| Bird, 1984 | 10 | 100 | New Zealand | 12.08 | memory | 1982** | 0.98 |
| Bird, 1984 | 10 | 100 | New Zealand | 12.08 | memory | 1982** | 0.987 |
| Bird, 1984 | 10 | 0 | New Zealand | 12.08 | memory | 1982** | 1.356 |
| Bird, 1984 | 10 | 0 | New Zealand | 12.08 | memory | 1982** | 1.356 |
| Cottini et al., 2021a | 37 | 51.4 | Italy | 9.58 | memory | 2018 | 1.466 |
| Cottini et al., 2021b | 48 | 51.0 | Italy | 5.58 | memory | 2019 | 2.842 |
| Cottini et al., 2021b | 34 | 54.3 | Italy | 9.42 | memory | 2019 | 1.405 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.136 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.072 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.101 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.225 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.333 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.110 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 0.961 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.188 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 0.902 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.252 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.432 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.297 |
| Ducheyne et al., 2012 | 115 | 47.8 | Belgium | 9.22 | motor | 2010 | 1.070 |
| Elmose & Happé, 2014 | 21 | 0 | Denmark | 10.10 | memory | 2012** | 1.143 |
| Elmose & Happé, 2014 | 21 | 0 | Denmark | 10.10 | memory | 2012** | 1.118 |
| Ewers & Wood, 1993 | 19 | 100 | USA | 10.50* | other cognitive | 1991** | 1.740 |
| Ewers & Wood, 1993 | 19 | 0 | USA | 10.50* | other cognitive | 1991** | 1.999 |
| Gaskill & Murphy, 2004 | 19 | 63.2 | USA | 7.58 | memory | 2000 | 0.851 |
| Gaskill & Murphy, 2004 | 19 | 63.2 | USA | 7.58 | memory | 2000 | 0.982 |
| Gaskill & Murphy, 2004 | 19 | 63.2 | USA | 7.58 | memory | 2000 | 1.215 |

| Article | *N* | Sex | Country | Age | Task Type | Data Collection | RoM |
|---|---|---|---|---|---|---|---|
| Gilpin & Boyden, 1978 | 28 | 0 | USA | 8.00 | motor | 1975 | 1.407 |
| Gilpin & Boyden, 1978 | 28 | 100 | USA | 8.00 | motor | 1975 | 1.339 |
| Heath & Glen, 2005 | 39 | 43.6 | Canada | 12.05 | other cognitive | 2000 | 0.950 |
| Kolovelonis & Goudas, 2018 | 62 | 0 | Greece | 11.39 | motor | 2015 | 1.372 |
| Kolovelonis & Goudas, 2018 | 76 | 100 | Greece | 11.39 | motor | 2015 | 1.305 |
| Kolovelonis & Goudas, 2018 | 116 | 0 | Greece | 11.14 | motor | 2015 | 1.390 |
| Kolovelonis & Goudas, 2018 | 120 | 100 | Greece | 11.14 | motor | 2015 | 1.304 |
| Kominsky & Keil, 2014 | 25 | 52.0 | USA | 5.75 | other cognitive | 2010 | 13.875 |
| Kominsky & Keil, 2014 | 25 | 52.0 | USA | 5.75 | other cognitive | 2010 | 56.774 |
| Kominsky & Keil, 2014 | 30 | 56.7 | USA | 7.52 | other cognitive | 2010 | 4.615 |
| Kominsky & Keil, 2014 | 30 | 56.7 | USA | 7.52 | other cognitive | 2010 | 11.215 |
| Kominsky & Keil, 2014 | 28 | 42.9 | USA | 9.54 | other cognitive | 2010 | 3.361 |
| Kominsky & Keil, 2014 | 28 | 42.9 | USA | 9.54 | other cognitive | 2010 | 18.071 |
| Krebs & Roebers, 2012 | 161 | unknown | Switzerland | 9.64 | memory | 2010 | 1.603 |
| Krebs & Roebers, 2012 | 161 | unknown | Switzerland | 9.64 | memory | 2010 | 1.079 |
| Krebs & Roebers, 2012 | 121 | unknown | Switzerland | 11.53 | memory | 2010 | 1.356 |
| Krebs & Roebers, 2012 | 121 | unknown | Switzerland | 11.53 | memory | 2010 | 1.085 |
| Kvavilashvili & Ford, 2014 | 80 | 50.0 | UK | 5.33 | memory | 2004 | 2.075 |
| Lavis & Mahy, 2021 | 24 | 53.2 | Canada | 4.52 | memory | 2019 | 1.269 |
| Lavis & Mahy, 2021 | 23 | 53.2 | Canada | 4.52 | memory | 2019 | 2.466 |
| Lavis & Mahy, 2021 | 21 | 48.8 | Canada | 5.52 | memory | 2019 | 1.411 |
| Lavis & Mahy, 2021 | 20 | 48.8 | Canada | 5.52 | memory | 2019 | 1.532 |
| Lavis & Mahy, 2021 | 22 | 58.1 | Canada | 6.54 | memory | 2019 | 1.027 |
| Lavis & Mahy, 2021 | 21 | 58.1 | Canada | 6.54 | memory | 2019 | 1.467 |
| Lee & Austin, 1986 | 40 | 50.0 | USA | 9.50* | motor | 1984** | 0.422 |
| Lee et al., 1988 | 10 | 0 | USA | 8.50* | motor | 1986 | 1.894 |
| Lee et al., 1988 | 10 | 0 | USA | 8.50* | motor | 1986 | 0.930 |
| Lee et al., 1988 | 10 | 100 | USA | 8.50* | motor | 1986 | 0.978 |
| Lee et al., 1988 | 10 | 100 | USA | 8.50* | motor | 1986 | 0.893 |
| Lee et al., 1988 | 10 | 0 | USA | 12.50* | motor | 1986 | 0.792 |

| Article | *N* | Sex | Country | Age | Task Type | Data Collection | RoM |
|---|---|---|---|---|---|---|---|
| Lee et al., 1988 | 10 | 0 | USA | 12.50* | motor | 1986 | 0.808 |
| Lee et al., 1988 | 10 | 100 | USA | 12.50* | motor | 1986 | 0.597 |
| Lee et al., 1988 | 10 | 100 | USA | 12.50* | motor | 1986 | 0.931 |
| Lipko et al., 2009 | 21 | 57.1 | USA | 5.00 | memory | 2005 | 1.509 |
| Lipko et al., 2009 | 21 | 57.1 | USA | 5.00 | memory | 2005 | 2.205 |
| Lipko et al., 2009 | 21 | 57.1 | USA | 5.00 | memory | 2005 | 2.003 |
| Lipko et al., 2009 | 32 | 62.5 | USA | 4.58 | memory | 2005 | 1.663 |
| Lipko et al., 2009 | 32 | 62.5 | USA | 4.58 | memory | 2005 | 2.413 |
| Lipko et al., 2009 | 32 | 62.5 | USA | 4.58 | memory | 2005 | 2.448 |
| Lipko et al., 2009 | 32 | 62.5 | USA | 4.58 | memory | 2005 | 2.525 |
| Lipko et al., 2009 | 32 | 62.5 | USA | 4.58 | memory | 2005 | 2.295 |
| Lipko et al., 2009 | 32 | 45.3 | USA | 4.92 | memory | 2005 | 1.791 |
| Lipko et al., 2009 | 32 | 45.3 | USA | 4.92 | memory | 2005 | 2.322 |
| Lipko et al., 2009 | 32 | 45.3 | USA | 4.92 | memory | 2005 | 2.183 |
| Lipko et al., 2009 | 32 | 45.3 | USA | 4.92 | memory | 2005 | 2.397 |
| Lipko et al., 2009 | 32 | 45.3 | USA | 4.92 | memory | 2005 | 2.263 |
| Lipko et al., 2012 | 22 | 45.5 | USA | 6.17 | memory | 2007 | 1.365 |
| Lipko et al., 2012 | 22 | 45.5 | USA | 6.17 | memory | 2007 | 1.061 |
| Lipko et al., 2012 | 22 | 45.5 | USA | 6.17 | memory | 2007 | 1.118 |
| Lipko et al., 2012 | 47 | 44.7 | USA | 8.92 | memory | 2007 | 1.140 |
| Lipko et al., 2012 | 47 | 44.7 | USA | 8.92 | memory | 2007 | 0.924 |
| Lipko et al., 2012 | 47 | 44.7 | USA | 8.92 | memory | 2007 | 0.966 |
| Lipko et al., 2012 | 33 | 48.5 | USA | 6.92 | memory | 2007 | 1.479 |
| Lipko et al., 2012 | 33 | 48.5 | USA | 6.92 | memory | 2007 | 1.111 |
| Lipko et al., 2012 | 33 | 48.5 | USA | 6.92 | memory | 2007 | 1.108 |
| Lipko et al., 2012 | 33 | 48.5 | USA | 6.92 | memory | 2007 | 1.011 |
| Lipko et al., 2012 | 33 | 48.5 | USA | 6.92 | memory | 2007 | 1.022 |
| Lipko-Speed, 2013 | 27 | unknown | USA | 4.50 | memory | 2008 | 1.942 |
| Lipko-Speed, 2013 | 27 | unknown | USA | 4.50 | memory | 2008 | 1.733 |
| Lipko-Speed, 2013 | 27 | unknown | USA | 4.50 | memory | 2008 | 2.213 |

| Article | *N* | Sex | Country | Age | Task Type | Data Collection | RoM |
|---|---|---|---|---|---|---|---|
| Lipko-Speed, 2013 | 27 | unknown | USA | 4.50 | memory | 2008 | 1.926 |
| Lipko-Speed, 2013 | 27 | unknown | USA | 4.50 | memory | 2008 | 2.730 |
| Lipko-Speed, 2013 | 27 | unknown | USA | 4.50 | memory | 2008 | 2.600 |
| Lipko-Speed, 2013 | 27 | unknown | USA | 4.50 | memory | 2008 | 2.520 |
| Lipko-Speed, 2013 | 27 | unknown | USA | 4.50 | memory | 2008 | 1.945 |
| McGee, II et al., 1970 | 20 | 50.0 | USA | 6.30 | other cognitive | 1968** | 0.983 |
| Miller et al., 1988 | 46 | 45.7 | USA | 8.08 | other cognitive | 1987 | 2.119 |
| Miller et al., 1988 | 47 | 51.1 | USA | 10.00 | other cognitive | 1987 | 1.386 |
| Miller et al., 1988 | 46 | 45.7 | USA | 8.08 | other cognitive | 1987 | 1.150 |
| Miller et al., 1988 | 47 | 51.1 | USA | 10.00 | other cognitive | 1987 | 1.086 |
| Miller et al., 1988 | 46 | 45.7 | USA | 8.08 | other cognitive | 1987 | 0.711 |
| Miller et al., 1988 | 47 | 51.1 | USA | 10.00 | other cognitive | 1987 | 0.973 |
| Miller et al., 1988 | 46 | 45.7 | USA | 8.08 | motor | 1987 | 1.179 |
| Miller et al., 1988 | 47 | 51.1 | USA | 10.00 | motor | 1987 | 1.103 |
| Peker et al., 2021 | 27 | unknown | Turkey | 7.74 | motor | 2019 | 1.041 |
| Peker et al., 2021 | 27 | unknown | Turkey | 7.74 | motor | 2019 | 1.241 |
| Powel & Toni, 1994 | 22 | 45.5 | USA | 5.41 | motor | 1992** | 1.447 |
| Powel & Toni, 1994 | 22 | 45.5 | USA | 5.41 | motor | 1992** | 2.072 |
| Powel & Toni, 1994 | 22 | 45.5 | USA | 5.41 | motor | 1992** | 3.537 |
| Pressley & Ghatala, 1989 | 18 | 50.0 | USA | 7.08 | other cognitive | 1987** | 0.837 |
| Pressley & Ghatala, 1989 | 18 | 50.0 | USA | 7.08 | other cognitive | 1987** | 1.279 |
| Pressley & Ghatala, 1989 | 18 | 49.1 | USA | 10.25 | other cognitive | 1987** | 1.015 |
| Pressley & Ghatala, 1989 | 18 | 49.1 | USA | 10.25 | other cognitive | 1987** | 1.400 |
| Schneider et al., 2000 | 17 | 47.1 | Germany | 8.60 | memory | 1998 | 1.374 |
| Schneider et al., 2000 | 17 | 47.1 | Germany | 8.60 | memory | 1998 | 1.152 |
| Schneider et al., 2000 | 17 | 60.0 | Germany | 10.60 | memory | 1998 | 1.104 |
| Schneider et al., 2000 | 18 | 60.0 | Germany | 10.60 | memory | 1998 | 0.956 |
| Schneider et al., 2000 | 16 | 71.9 | Germany | 6.30 | memory | 1998 | 1.201 |
| Schneider et al., 2000 | 16 | 71.9 | Germany | 6.30 | memory | 1998 | 1.418 |
| Schneider et al., 2000 | 16 | 46.9 | Germany | 8.60 | memory | 1998 | 1.325 |

| Article | *N* | Sex | Country | Age | Task Type | Data Collection | RoM |
|---|---|---|---|---|---|---|---|
| Schneider et al., 2000 | 16 | 46.9 | Germany | 8.60 | memory | 1998 | 0.955 |
| Schneider et al., 2000 | 16 | 40.6 | Germany | 10.60 | memory | 1998 | 1.098 |
| Schneider et al., 2000 | 16 | 40.6 | Germany | 10.60 | memory | 1998 | 1.026 |
| Schneider, 1998 | 24 | unknown | Germany | 4.08 | motor | 1996 | 1.405 |
| Schneider, 1998 | 24 | unknown | Germany | 6.17 | motor | 1996 | 1.179 |
| Schneider, 1998 | 24 | unknown | Germany | 4.08 | motor | 1996 | 1.608 |
| Schneider, 1998 | 24 | unknown | Germany | 6.17 | motor | 1996 | 1.620 |
| Schneider, 1998 | 24 | unknown | Germany | 4.17 | memory | 1996 | 1.312 |
| Schneider, 1998 | 24 | unknown | Germany | 6.08 | memory | 1996 | 1.259 |
| Schneider, 1998 | 24 | unknown | Germany | 4.17 | memory | 1996 | 2.167 |
| Schneider, 1998 | 24 | unknown | Germany | 6.08 | memory | 1996 | 1.451 |
| Shin et al., 2007 | 28 | 46.9 | USA | 6.75 | memory | 2005 | 1.716 |
| Shin et al., 2007 | 28 | 46.9 | USA | 6.75 | memory | 2005 | 1.822 |
| Shin et al., 2007 | 28 | 46.9 | USA | 6.75 | memory | 2005 | 1.798 |
| Shin et al., 2007 | 28 | 46.9 | USA | 6.75 | memory | 2005 | 1.610 |
| Shin et al., 2007 | 28 | 46.9 | USA | 6.75 | memory | 2005 | 2.270 |
| Shin et al., 2007 | 35 | 52.8 | USA | 7.75 | memory | 2005 | 1.626 |
| Shin et al., 2007 | 35 | 52.8 | USA | 7.75 | memory | 2005 | 1.563 |
| Shin et al., 2007 | 35 | 52.8 | USA | 7.75 | memory | 2005 | 1.629 |
| Shin et al., 2007 | 35 | 52.8 | USA | 7.75 | memory | 2005 | 1.728 |
| Shin et al., 2007 | 35 | 52.8 | USA | 7.75 | memory | 2005 | 1.607 |
| Shin et al., 2007 | 26 | 53.8 | USA | 9.75 | memory | 2005 | 1.109 |
| Shin et al., 2007 | 26 | 53.8 | USA | 9.75 | memory | 2005 | 1.147 |
| Shin et al., 2007 | 26 | 53.8 | USA | 9.75 | memory | 2005 | 1.037 |
| Shin et al., 2007 | 26 | 53.8 | USA | 9.75 | memory | 2005 | 1.045 |
| Shin et al., 2007 | 26 | 53.8 | USA | 9.75 | memory | 2005 | 0.988 |
| Swanson, 1983 | 12 | 33.3 | USA | 8.10 | memory | 1981** | 0.955 |
| Swanson, 1983 | 12 | 33.3 | USA | 10.58 | memory | 1981** | 0.807 |
| Was & Al-Harthy, 2018 | 62 | 45.3 | Oman | 5.50 | memory | 2015 | 1.716 |
| Was & Al-Harthy, 2018 | 60 | 50.0 | Oman | 6.50 | memory | 2015 | 1.458 |

| Article | *N* | Sex | Country | Age | Task Type | Data Collection | RoM |
|---|---|---|---|---|---|---|---|
| Was & Al-Harthy, 2018 | 61 | 42.6 | Oman | 7.50 | memory | 2015 | 1.303 |
| Was & Al-Harthy, 2018 | 55 | 50.9 | Oman | 8.50 | memory | 2015 | 1.034 |
| Was & Al-Harthy, 2018 | 60 | unknown | Oman | 8.50 | memory | 2015 | 1.286 |
| Was & Al-Harthy, 2018 | 60 | unknown | Oman | 9.50 | memory | 2015 | 1.131 |
| Was & Al-Harthy, 2018 | 61 | unknown | Oman | 10.50 | memory | 2015 | 1.110 |
| Was & Al-Harthy, 2018 | 60 | unknown | Oman | 8.50 | memory | 2015 | 1.383 |
| Was & Al-Harthy, 2018 | 60 | unknown | Oman | 9.50 | memory | 2015 | 1.105 |
| Was & Al-Harthy, 2018 | 61 | unknown | Oman | 10.50 | memory | 2015 | 1.194 |
| Wojcik et al., 2022 | 20 | 20.0 | UK | 11.05 | memory | 2008 | 0.941 |
| Wojcik et al., 2022 | 20 | 20.0 | UK | 11.05 | memory | 2008 | 1.326 |
| Worden & Sladewski-Awig, 1982 | 24 | 50.0 | USA | 5.00* | memory | 1980** | 1.541 |
| Worden & Sladewski-Awig, 1982 | 24 | 50.0 | USA | 7.50* | memory | 1980** | 0.907 |
| Worden & Sladewski-Awig, 1982 | 24 | 50.0 | USA | 10.00* | memory | 1980** | 0.824 |
| Worden & Sladewski-Awig, 1982 | 24 | 50.0 | USA | 11.50* | memory | 1980** | 0.864 |
| Xia et al., 2022 | 100 | 52.0 | China | 4.75 | motor | 2019 | 2.950 |
| Xia et al., 2022 | 100 | 52.0 | China | 4.75 | motor | 2019 | 1.959 |
| Xia et al., 2022 | 100 | 52.0 | China | 4.75 | motor | 2019 | 1.764 |
| Xia et al., 2022 | 94 | 49.0 | Netherlands | 5.00 | motor | 2019 | 2.204 |
| Xia et al., 2022 | 94 | 49.0 | Netherlands | 5.00 | motor | 2019 | 1.735 |
| Xia et al., 2022 | 94 | 49.0 | Netherlands | 5.00 | motor | 2019 | 1.744 |
| Xia et al., 2022 | 100 | 52.0 | China | 4.75 | memory | 2019 | 1.285 |
| Xia et al., 2022 | 100 | 52.0 | China | 4.75 | memory | 2019 | 1.568 |
| Xia et al., 2022 | 100 | 52.0 | China | 4.75 | memory | 2019 | 1.713 |
| Xia et al., 2022 | 91 | 49.0 | Netherlands | 5.00 | memory | 2019 | 1.023 |
| Xia et al., 2022 | 91 | 49.0 | Netherlands | 5.00 | memory | 2019 | 2.029 |
| Xia et al., 2022 | 91 | 49.0 | Netherlands | 5.00 | memory | 2019 | 2.465 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 6.92 | memory | 1978 | 2.140 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 9.08 | memory | 1978 | 1.688 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 11.25 | memory | 1978 | 1.308 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 6.92 | memory | 1978 | 2.923 |

| Article | *N* | Sex | Country | Age | Task Type | Data Collection | RoM |
|---|---|---|---|---|---|---|---|
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 9.08 | memory | 1978 | 2.249 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 11.25 | memory | 1978 | 1.870 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 6.92 | memory | 1978 | 0.825 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 9.08 | memory | 1978 | 0.764 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 11.25 | memory | 1978 | 0.876 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 6.92 | memory | 1978 | 0.763 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 9.08 | memory | 1978 | 0.643 |
| Yussen & Berman, 1981 | 38 | 50.0 | USA | 11.25 | memory | 1978 | 0.665 |
| Yussen & Levy Jr, 1975 | 24 | 50.0 | USA | 4.60 | memory | 1973** | 2.416 |
| Yussen & Levy Jr, 1975 | 24 | 50.0 | USA | 8.90 | memory | 1973** | 1.585 |
| Yussen & Paquette, 1978 | 32 | 50.0 | USA | 10.77 | memory | 1976** | 1.380 |
| Yussen & Paquette, 1978 | 32 | 50.0 | USA | 10.77 | memory | 1976** | 0.904 |

*Note*. * We imputed the average ages of children in those grade levels in the pertaining country as the sample's mean age.

** We calculated year of data collection as two years before the year of publication (Gentile et al., 2009; Oliver & Hyde, 1993).

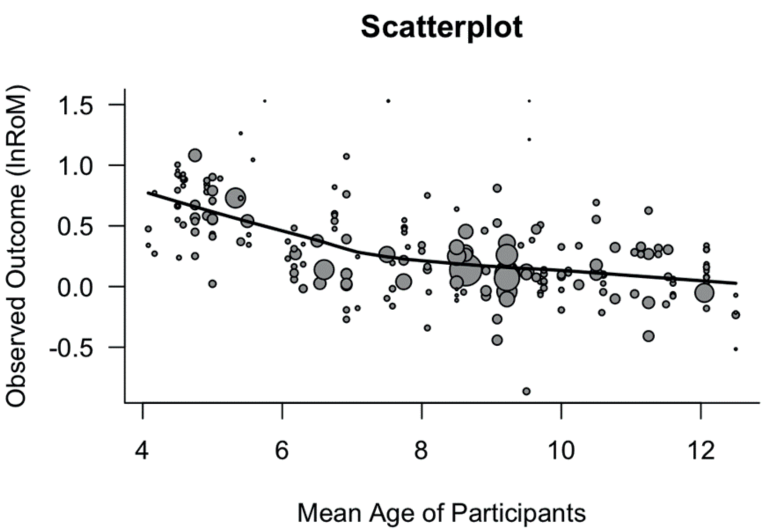
**Supporting Information 4 – Nonlinear Analyses (Chapter 4)**

**Age**

We analyzed the association between the observed effect sizes and sample mean ages in a way that does not assume a linear relation. We first examined a scatterplot of the data, with effect sizes along the y-axis and sample mean ages along the x-axis (Figure 1).

**Figure 1**

*Scatterplot of the Observed Outcome (lnRoM) and Sample Mean Age*




*Note*. The datapoint size is proportional to the inverse of the standard error (i.e., more precise estimates are reflected by larger datapoints).

We also plotted the locally weighted scatterplot smoothing (LOWESS) curve in the scatterplot. We then fitted a spline model consisting of a series of piecewise cubic polynomials (Restricted Cubic Spline model, Stone & Koo, 1985). We chose the number of knots (i.e., where the piecewise cubic polynomial connects) to fit the data. We fitted 3-knot and 4-knot models to test for potential increases or decreases that are more pronounced at some ages than others. The exact positions of the knots were identified by the model. We also directly selected the knot positions ourselves based on the curve shown in Fig. 1, as well as on research evidence suggesting differences in self-overestimation between preschoolers and primary school-aged children (i.e., Shin et al., 2007; Was & Al-Harthy, 2018). We chose (age *4, 7, 13*) as a 3-knots model and (age *4, 6, 8, 13*) as a 4-knots model to fit the data. We compared the models in terms of their log likelihoods and information criteria (Deviance, AIC and BIC, Table 1). We found that the linear model provided a better fit (i.e., a higher log likelihood value) and a better balance between the increased fit and the increased model complexity (as indicated by the lower values of the information criteria Deviance, AIC and BIC) than the non-linear models. We conclude that children's self-overestimation gradually decreases with age (from age 4 to 12).

**Table 1**

*Log Likelihoods and Information Criteria for Different Models of the Association Between Children's Self-Overestimation Effect and Sample Age*

| | Linear model | 3-knots model | 4-knots model | 3 knots model (age 4, 7, 13) | 4 knots model (age 4, 6, 8, 13) |
|---|---|---|---|---|---|
| LogLik | -23.501 | -36.881 | -36.209 | -36.633 | -36.587 |
| Deviance | 47.001 | 73.763 | 72.419 | 73.267 | 73.173 |
| AIC | 55.001 | 81.763 | 82.419 | 81.267 | 83.173 |
| BIC | 68.484 | 95.227 | 99.225 | 94.731 | 99.980 |

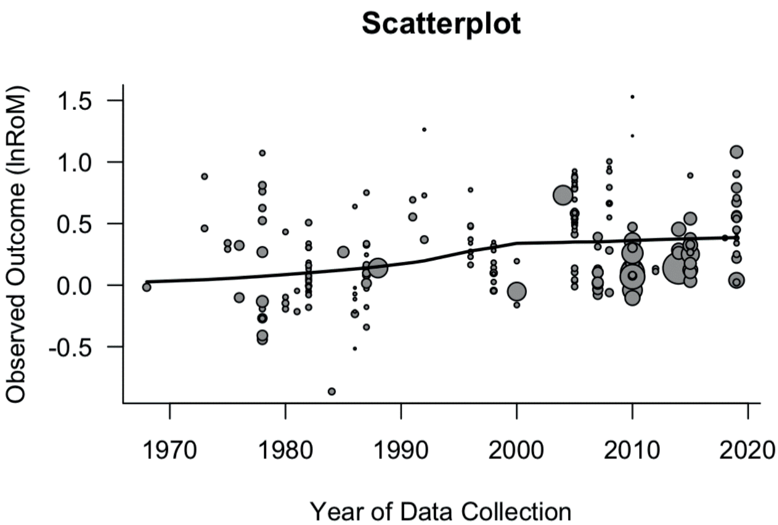
*Note.* Loglik = Log likelihood; AIC = Akaike information criterion; BIC = the Bayesian information criterion

**Data Collection Year**

Similarly, we modeled the association between the observed effect sizes and the year of data collection by fitting the Restricted Cubic Spline models. We first examined a scatterplot of the data, with effect sizes along the y-axis and year of publication along the x-axis (Figure 2).

**Figure 2**

*Scatterplot of the Observed Outcome (lnRoM) and Year of Data Collection*




*Note*. The datapoint size is proportional to the inverse of the standard error (i.e., more precise estimates are reflected by larger datapoints).

We also plotted the locally weighted scatterplot smoothing (LOWESS) curve in the scatterplot. We fitted 3-knot and 4-knot Restricted Cubic Spline models. The positions of the knots were identified by the model. Based on the curve shown in Fig. 2, and in the absence of relevant theory or empirical evidence, we chose not to identify the knot positions ourselves. We compared the models in terms of their log likelihoods and information criteria (Table 2).

The linear model provided a better fit (i.e., a higher log likelihood value) and a better balance between the increased fit and the increased model complexity (as indicated by the lower values of the information criteria Deviance, AIC and BIC). We conclude that children's self-overestimation has gradually increased over the past several decades.

**Table 2**

*Log Likelihoods and Information Criteria for Different Models of the Association Between Children's Self-Overestimation Effect and Data Collection Year*

| | Linear | 3 knots | 4 knots |
|---|---|---|---|
| LogLik | -40.068 | -75.952 | -75.164 |
| Deviance | 80.135 | 151.905 | 150.329 |
| AIC | 88.135 | 159.905 | 160.329 |
| BIC | 101.618 | 173.369 | 177.135 |

*Note.* Loglik = Log likelihood; AIC = Akaike information criterion; BIC = the Bayesian information criterion

## Supporting Information 5 – Sensitivity Analyses (Chapter 4)

### Overall Effect of Self-Overestimation

Table 1 shows the outcomes of sensitivity analyses for the overall effect of children's self-overestimation. The Ratio of Means (RoM) and 95% Confidence Intervals were only slightly affected by decisions regarding effect size extraction, standard deviation imputation, and treatment of outliers. Sensitivity analyses confirmed that overall children overestimated their performance 1.3 to 1.4 times.

**Table 1**

*Sensitivity Analyses for the Overall Effect of Self-Overestimation*

| Decisions | *RoM* | *t* | 95% confidence interval |
|---|---|---|---|
| Averaged the effect sizes across trials (k = 161) | 1.330*** | 5.338 | 1.199-1.475 |
| Averaged the effect sizes across trials and within-subjects conditions (k = 134) | 1.303*** | 5.935 | 1.193-1.424 |
| Excluded effect sizes for which standard deviations were unavailable (k = 141) | 1.402*** | 5.533 | 1.244-1.580 |
| Imputed the largest standard deviation of studies using the same type of task and same measurement scale (k = 217) | 1.331*** | 5.342 | 1.197-1.480 |
| Imputed the smallest standard deviation of studies using the same type of task and same measurement scale (k = 217) | 1.334*** | 5.344 | 1.200-1.483 |
| Retained all effect sizes (k = 217) | 1.353*** | 4.667 | 1.191-1.536 |
| Removed outliers (k = 213) | 1.323*** | 5.605 | 1.200-1.459 |

*Note.* *** $p < .001$

### Moderator Analyses

Here we report the outcomes of sensitivity analyses for the moderating effects of age, type of task, and year of data collection on overestimation.

#### *Averaged the Effect Sizes Across Trials*

**Heterogeneity in Effect Sizes.** We found significant variability in effect sizes within studies ($p < .001$) and between studies ($p < .001$). We concluded that the following moderator analyses are meaningful.

**Age.** Sample mean age was a significant moderator of the overall effect, $F(1, 159) = 33.572$, $p < .001$. We found less overestimation in older samples, $B = -0.076$, $t(159) = -5.794$, $p < .001$.

**Task Type.** We found no significant moderating effect of task type on the overall effect, $F(2, 158) = 0.210$, $p = .811$.

**Data Collection Year.** We found a significant moderating effect for year of data collection on the overall effect, $F(1, 159) = 5.287$, $p = .023$, such that more recently conducted studies reported higher levels of overestimation, $B = 0.008$, $t(159) = 2.299$, $p = .023$.

***Averaged the Effect Sizes Across Trials and Within-Subjects Conditions***

**Heterogeneity in Effect Sizes.** We found significant variability in effect sizes within studies ($p < .001$) and between studies ($p = .001$). We concluded that the following moderator analyses are meaningful.

**Age.** Sample mean age was a significant moderator of the overall effect, $F(1, 132) = 34.056$, $p < .001$. We found less overestimation in older samples, $B = -0.073$, $t(132) = -5.836$, $p < .001$.

**Task Type.** We found no significant moderating effect of task type on the overall effect, $F(2, 131) = 0.247$, $p = .782$.

**Data Collection Year.** We found a significant moderating effect for year of data collection on the overall effect, $F(1, 132) = 7.781$, $p = .006$, such that more recently conducted studies reported higher levels of overestimation, $B = 0.007$, $t(132) = 2.789$, $p = .006$.

***Excluded Effect Sizes for Which Standard Deviations were Unavailable***

**Heterogeneity in Effect Sizes.** We found significant variability in effect sizes within studies ($p < .001$) and between studies ($p < .001$). We concluded that the following moderator analyses are meaningful.

**Age.** Sample mean age was a significant moderator of the overall effect, $F(1, 139) = 24.887$, $p < .001$. We found less overestimation in older samples, $B = -0.077$, $t(139) = -4.989$, $p < .001$.

**Task Type.** We found no significant moderating effect of task type on the overall effect, $F(2, 138) = 0.275$, $p = .760$.

**Data Collection Year.** We found a significant moderating effect for year of data collection on the overall effect, $F(1, 139) = 4.403$, $p = .038$, such that more recently conducted studies reported higher levels of overestimation, $B = 0.008$, $t(139) = 2.098$, $p = .038$.

***Imputed the Largest Standard Deviation***

**Heterogeneity in Effect Sizes.** We found significant variability in effect sizes within studies ($p < .001$) and between studies ($p < .001$). We concluded that the following moderator analyses are meaningful.

**Age.** Sample mean age was a significant moderator of the overall effect, $F(1, 215) = 41.687$, $p < .001$. We found less overestimation in older samples, $B = -0.079$, $t(215) = -6.457$, $p < .001$.

**Task Type.** We found no significant moderating effect of task type on the overall effect, $F(2, 214) = 0.211$, $p = .810$.

**Data Collection Year.** We found a significant moderating effect for year of data collection on the overall effect, $F(1, 215) = 5.404$, $p = .021$, such that more recently conducted studies reported higher levels of overestimation, $B = 0.008$, $t(215) = 2.325$, $p =$

.021.

***Imputed the Smallest Standard Deviation***

**Heterogeneity in Effect Sizes.** We found significant variability in effect sizes within studies ($p < .001$) and between studies ($p < .001$). We concluded that the following moderator analyses are meaningful.

**Age.** Sample mean age was a significant moderator of the overall effect, $F(1, 215) = 45.216$, $p < .001$. We found less overestimation in older samples, $B = -0.080$, $t(215) = -6.724$, $p < .001$.

**Task Type.** We found no significant moderating effect of task type on the overall effect, $F(2, 214) = 0.188$, $p = .828$.

**Data Collection Year.** We found a significant moderating effect for year of data collection on the overall effect, $F(1, 215) = 5.238$, $p = .023$, such that more recently conducted studies reported higher levels of overestimation, $B = 0.008$, $t(215) = 2.289$, $p = .023$.

***Retained All Effect Sizes***

**Heterogeneity in Effect Sizes.** We found significant variability in effect sizes within studies ($p < .001$) and between studies ($p < .001$). We concluded the following moderator analyses are meaningful.

**Age.** Sample mean age was a significant moderator of the overall effect, $F(1, 215) = 40.045$, $p < .001$. We found less overestimation in older samples, $B = -0.081$, $t(215) = -6.328$, $p < .001$.

**Task Type.** We found no significant moderating effect of task type on the overall effect, $F(2, 214) = 0.532$, $p = .588$.

**Data Collection Year.** We found a significant moderating effect for year of data collection on the overall effect, $F(1, 215) = 4.541$, $p = .034$, such that more recently conducted studies reported higher levels of overestimation, $B = 0.009$, $t(215) = 2.131$, $p = .034$.

***Removed Outliers***

**Heterogeneity in Effect Sizes.** We found significant variability in effect sizes within studies ($p < .001$) and between studies ($p < .001$). We concluded that the following moderator analyses are meaningful.

**Age.** Sample mean age was a significant moderator of the overall effect, $F(1, 211) = 46.781$, $p < .001$. We found less overestimation in older samples, $B = -0.079$, $t(211) = -6.840$, $p < .001$.

**Task Type.** We found no significant moderating effect of task type on the overall effect, $F(2, 210) = 0.071$, $p = .932$.

**Data Collection Year.** We found a significant moderating effect for year of data collection on the overall effect, $F(1, 211) = 5.591$, $p = .019$, such that more recently conducted studies reported higher levels of overestimation, $B = 0.007$, $t(211) = 2.365$, $p = .019$.

In sum, the substantive conclusions that we draw regarding the moderating effects of age, type of task, and year of data collection on overestimation are not affected by decisions regarding effect size extraction, standard deviation imputation, and treatment of outliers.

# References

References marked with an asterisk indicate studies included in the meta-analysis (Chapter 4).

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of personality and social psychology, 49*(6), 1621.

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology, 20*(1), 1–48.

*Almeida, G., Luz, C., Martins, R., & Cordovil, R. (2017). Do children accurately estimate their performance of fundamental movement skills? *Journal of Motor Learning and Development, 5*(2), 193–206.

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

*Bain, S. K., Holliman, B., & McCallum, R. S. (1989). Children's self-predictions and teachers' predictions of basic concept mastery: Effects of socioeconomic status, locus of control, and achievement. *Journal of Psychoeducational Assessment, 7*(3), 235–245.

*Bain, S. K., & McCallum, R. S. (1986). Student and teacher predictions of kindergartner's basic concept mastery: How valid are they? *Journal of Psychoeducational Assessment, 4*(2), 163–168. https://doi.org/10.1177/073428298600400208

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147. https://doi.org/10.1037/0003-066X.37.2.122

Bernard, S., Clément, F., & Mercier, H. (2016). Wishful thinking in preschoolers. *Journal of Experimental Child Psychology*, *141*, 267–274. https://doi.org/10.1016/j.jecp.2015.07.018

*Bertrand, J. M., Moulin, C. J. A., & Souchay, C. (2017). Short-term memory predictions across the lifespan: monitoring span before and after conducting a task. *Memory, 25*(5), 607–618. https://doi.org/10.1080/09658211.2016.1200625

Bian, Q., Chen, Y., Greenfield, P. M., & Yuan, Q. (2022). Mothers' experience of social change and individualistic parenting goals over two generations in urban China. *Frontiers in Psychology, 12*, 487039.

*Bird, J. E. (1984). Gender differences in expectations of children when anticipated feedback and anonymity are varied. *Genetic Psychology Monographs, 110*(2), 307–325.

Bjorklund, D. F. (1997). The role of immaturity in human development. *Psychological Bulletin*, *122*(2), 153–169. https://doi.org/10.1037/0033-2909.122.2.153

Bjorklund, D. F., & Beer, C. (2016). The adaptive value of cognitive immaturity: Applications of evolutionary developmental psychology to early education. *Evolutionary perspectives on chld development and education*, 3-32.

Bjorklund, D. F., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist*, *47*(1), 46–54. https://doi.org/10.1037/0003-066X.47.1.46

Blaine, B. (2018). Winsorizing. In B. E. Blaine, *The SAGE encyclopedia of educational research, measurement, and evaluation* (p. 1817). SAGE Publications.

Bond, M. H., Leung, K. & Wan, K. C. (1982). The social impact of self-effacing attributions: The Chinese case. *The Journal of Social Psychology, 118*, 157–166.

Bornstein, M. H., & Cheah, C. S. L. (2006). The place of "culture and parenting" in the ecological contextual perspective on developmental science. In *Parenting beliefs, behaviors, and parent–child relations* (pp. 3–33).

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. (2021). *Introduction to meta-analysis*. John Wiley & Sons.

Boseovski, J. J. (2010). Evidence for "Rose-colored glasses": An examination of the positivity bias in young children's personality judgments. *Child Development*

*Perspectives*, *4*(3), 212–218. https://doi.org/10.1111/j.1750-8606.2010.00149.x

Botto, S. V., & Rochat, P. (2019). Evaluative audience perception (EAP): How children come to care about reputation. *Child Development Perspectives, 13*(3), 180–185. https://doi.org/10.1111/cdep.12335

Boulton, M. J., & Smith, P. K. (1990). Affective bias in children's perceptions of dominance relationships. *Child Development*, *61*(1), 221–229.

Brown, J. D., & Kobayashi, C. (2002). Self‐enhancement in Japan and America. *Asian Journal of Social Psychology, 5*(3), 145-168.

Brummelman, E., & Thomaes, S. (2017). How children construct views of themselves: a social‐developmental perspective. *Child Development, 88*(6), 1763–1773. https://doi.org/10.1111/cdev.12961

Brummelman, E., Thomaes, S., Nelemans, S. A., De Castro, B. O., Overbeek, G., & Bushman, B. J. (2015). Origins of narcissism in children. *Proceedings of the National Academy of Sciences, 112*(12), 3659–3662. https://doi.org/10.1073/pnas.1420870112

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67*(3), 366–381.

Butler, R. (1989). Mastery versus ability appraisal: A developmental study of children's observations of peers' work. *Child Development*, 1350-1361.

Cai, H., Kwan, V. S., & Sedikides, C. (2012). A sociocultural approach to narcissism: The case of modern China. *European Journal of Personality, 26*(5), 529-535.

Cai, H., Sedikides, C., Gaertner, L., Wang, C., Carvallo, M., Xu, Y., O'Mara, E. M., & Jackson, L. E. (2011). Tactical self-enhancement in China: is modesty at the service of self-enhancement in East Asian culture? *Social Psychological and Personality Science*, *2*(1), 59–64. https://doi.org/10.1177/1948550610376599

Chen, S., Bond, M. H., Chan, B., Tang, D., & Buchtel, E. E. (2009). Behavioral manifestations of modesty. *Journal of Cross-Cultural Psychology, 40*(4), 603-626.

Chen, X., Cen, G., Li, D., & He, Y. (2005). Social functioning and adjustment in Chinese children: The imprint of historical time. *Child development, 76*(1), 182-195.

Chen, X., & French, D. C. (2008). Children's social competence in cultural context. *Annu. Rev. Psychol., 59*, 591-616.

Chen, X., & Li, D. (2012). Parental encouragement of initiative-taking and adjustment in Chinese children from rural, urban, and urbanized families. *Journal of Family Psychology*, *26*(6), 927–936. https://doi.org/10.1037/a0030708

Chen, X., Wang, L., & Wang, Z. (2009). Shyness-sensitivity and social, school, and psychological adjustment in rural migrant and urban children in China. *Child Development*, *80*(5), 1499–1513. https://doi.org/10.1111/j.1467-8624.2009.01347.x

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structual equation modeling approach. *Psychological Methods, 19,* 211-229.

*Cottini, M., Basso, D., & Palladino, P. (2021). Improving prospective memory in school-aged children: Effects of future thinking and performance predictions. *Journal of Experimental Child Psychology, 204*, 105065. https://doi.org/10.1016/j.jecp.2020.105065

*Cottini, M., Basso, D., Pieri, A., & Palladino, P. (2021). Metacognitive monitoring and control in children’s prospective memory. *Journal of Cognition and Development, 22*(4), 619–639. https://doi.org/10.1080/15248372.2021.1916500

Coutinho, S., Wiemer-Hastings, K., Skowronski, J. J., & Britt, M. A. (2005). Metacognition, need for cognition and use of explanations during ongoing learning and problem solving. *Learning and Individual Differences*, *15*(4), 321–337. https://doi.org/10.1016/j.lindif.2005.06.001

Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE, 10*(4), 1–12. https://doi.org/10.1371/journal.pone.0121945

Diesendruck, G., & Lindenbaum, T. (2009). Self-protective optimism: Children's biased beliefs about the stability of traits. *Social Development, 18*(4), 946–961. https://doi.org/10.1111/j.1467-9507.2008.00494.x

*Ducheyne, F., De Bourdeaudhuij, I., Lenoir, M., & Cardon, G. (2012). Test-retest reliability and validity of a child and parental questionnaire on specific determinants of cycling to school. *Pediatric Exercise Science, 24*(2), 289–311. https://doi.org/10.1123/pes.24.2.289

Dweck, C. S. (2002). The development of ability conceptions. In A. Wigfield, J. Eccles, U. Schiefele, R. Roeser, & P. Davis-Kean, *Development of achievement motivation* (pp. 57-88). John Wiley & Sons.

Dweck, C., Walton, G. M., & Cohen, G. L. (2014). Academic tenacity: Mindset and skills that promote long-term learning. In *Bill & Melinda Gates Foundation*.

*Elmose, M., & Happé, F. (2014). Being aware of own performance: How accurately do children with autism spectrum disorder judge own memory performance? *Autism Research, 7*(6), 712–719. https://doi.org/10.1002/aur.1421

Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self-or social prediction? *Journal of Personality and Social Psychology, 79*(6), 861–875. https://doi.org/10.1037/0022-3514.79.6.861

Escolano-Pérez, E., Herrero-Nivela, M. L., & Anguera, M. T. (2019). Preschool metacognitive skill assessment in order to promote educational sensitive response from mixed-methods approach: complementarity of data analysis. *Frontiers in psychology, 10*, 1298.

Escribano, C., & Díaz-Morales, J. F. (2014). Are self-reported grades a good estimate of academic achievement?/¿ Son las notas auto-informadas una buena estimación del rendimiento académico?. *Estudios de Psicología, 35*(1), 168-182.

*Ewers, C. A., & Wood, N. L. (1993). Sex and ability differences in children's math self-efficacy and prediction accuracy. *Learning and Individual Differences, 5*(3), 259–267. https://doi.org/10.1016/1041-6080(93)90006-E

Falbo, T., Poston Jr, D. L., Triscari, R. S., & Zhang, X. (1997). Self-enhancing illusions among Chinese schoolchildren. *Journal of Cross-Cultural Psychology, 28*(2), 172-191.

Farh, J. L., Dobbins, G. H., & Cheng, B. S. (1991). Cultural relativity in action: A comparison of self‐ratings made by Chinese and US workers. *Personnel psychology, 44*(1), 129-147.

Finn, B., & Metcalfe, J. (2007). The role of memory for past test in the underconfidence with practice effect. *Journal of Experimental Psychology: Learning Memory and Cognition, 33*(1), 238–244. https://doi.org/10.1037/0278-7393.33.1.238

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist, 34*(10), 906.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology*, *1*(4), 324–340. https://doi.org/10.1016/0010-0285(70)90019-8

Franchak, J. M. (2019). Development of affordance perception and recalibration in children and adults. *Journal of Experimental Child Psychology*, *183*, 100–114. https://doi.org/10.1016/j.jecp.2019.01.016

Friedrich, J. O., Adhikari, N. K. J., & Beyene, J. (2008). The ratio of means method as an alternative to mean differences for analyzing continuous outcome variables in meta-analysis: A simulation study. *BMC Medical Research Methodology, 8*(32). https://doi.org/10.1186/1471-2288-8-32

Friedrich, J. O., Adhikari, N. K. J., & Beyene, J. (2011). Ratio of means for analyzing continuous outcomes in meta-analysis performed as well as mean difference methods. *Journal of Clinical Epidemiology, 64*(5), 556–564. https://doi.org/10.1016/j.jclinepi.2010.09.016

Fu, G., Heyman, G. D., & Lee, K. (2016). Learning to be unsung heroes: Development of reputation management in two cultures. *Child Development*, *87*(3), 689–699. https://doi.org/10.1111/cdev.12494

Fung, H., & Chen, E. C. H. (2001). Across time and beyond skin: Self and transgression in the everyday socialization of shame among Taiwanese preschool children. *Social Development*, *10*(3), 419–437. https://doi.org/10.1111/1467-9507.00173

Gardner, W. L., Gabriel, S., & Lee, A. Y. (1999). "I" value freedom, but "we" value relationships: Self-construal priming mirrors cultural differences in judgment. *Psychological science, 10*(4), 321-326.

Gaertner, L., Sedikides, C., & Chang, K. (2008). On pancultural self-enhancement: Well-adjusted Taiwanese self-enhance on personally valued traits. *Journal of Cross-Cultural Psychology*, *39*(4), 463–477.

*Gaskill, P. J., & Murphy, P. K. (2004). Effects of a memory strategy on second-graders' performance and self-efficacy. *Contemporary Educational Psychology, 29*(1), 27–49. https://doi.org/10.1016/S0361-476X(03)00008-0

Gentile, B., Grabe, S., Dolan-Pascoe, B., Twenge, J. M., Wells, B. E., & Maitino, A. (2009). Gender differences in domain-specific self-esteem: A meta-analysis. *Review of General Psychology, 13*(1), 34–45. https://doi.org/10.1037/a0013689

Gentile, B., Twenge, J. M., & Campbell, W. K. (2010). Birth cohort differences in self-esteem, 1988-2008: A cross-temporal meta-analysis. *Review of General Psychology, 14*(3), 261–268. https://doi.org/10.1037/a0019919

Gibson, J. J. (2014). *The ecological approach to visual perception: classic edition.* Psychology Press.

*Gilpin, A. R., & Boyden, J. G. (1978). Children's conceptual tempo and level of aspiration. *Perceptual and Motor Skills, 46*(2), 587–593. https://doi.org/10.2466/pms.1978.46.2.587

Greenfield, P. M. (2009). Linking social change and developmental change: Shifting pathways of human development. *Developmental Psychology, 45*(2), 401–418. https://doi.org/10.1037/a0014726

Greenfield, P. M. (2016). Social change, cultural evolution, and human development. *Current Opinion in Psychology, 8*, 84-92.

Gresham, F. M., Lane, K. L., MacMillan, D. L., Bocian, K. M., & Ward, S. L. (2000). Effects of positive and negative illusory biases: Comparisons across social and academic self-concept domains. *Journal of school psychology, 38*(2), 151-175.

Griffin, J. W., Bauer, R., & Scherf, K. S. (2021). A quantitative meta-analysis of face recognition deficits in autism: 40 years of research. *Psychological Bulletin, 147*(3), 268–292. https://doi.org/10.1037/bul0000310

Grusec, J. E. (2011). Socialization processes in the family: Social and emotional development. *Annual review of psychology, 62*, 243-269.

Gürel, Ç., & Brummelman, E. (2020, April 7). *The Problem with Telling Children They're Better Than Others*. Retrieved from Scientific American: https://www.scientificamerican.com/article/the-problem-with-telling-children-theyre-better-than-others/

Gürel, Ç., Brummelman, E., Sedikides, C., & Overbeek, G. (2020). Better than my past self: Temporal comparison raises children's pride without triggering superiority goals. *Journal of Experimental Psychology: General*, *149*(8), 1554–1566. https://doi.org/10.1037/xge0000733

Gweon, H., Asaba, M., & Bennett-Pierre, G. (2017). Reverse-engineering the process: Adults' and preschoolers' ability to infer the difficulty of novel tasks. *Cognitive Science*.

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews, 18*(2), 1–12. https://doi.org/10.1002/cl2.1230

Hagan, C. A., Halberstadt, A. G., & Leary, K. A. (2021). Socialization of children's experience and expression of pride. *Infant and Child Development, 30*(4), e2230.

Hamamura, T. (2012). Are cultures becoming individualistic? A cross-temporal comparison of individualism-collectivism in the United States and Japan. *Personality and Social Psychology Review, 16*(1), 3–24. https://doi.org/10.1177/1088868311411587

Han, K. H. (2011). The self-enhancing function of Chinese modesty: From a perspective of social script. *Asian Journal of Social Psychology, 14*(4), 258-268.

Harter, S. (1996). Historical roots of contemporary issues involving self-concept. In B. A. Bracken, *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 1–37). John Wiley & Sons.

Harter, S. (1998). The development of self-representations. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 553–617). John Wiley & Sons, Inc.

Harter, S. (2006). The self. In W. Damon, & R. Lerner, *Handbook of child psychology* (pp. 505-570). John Wiley & Sons.

Harter, S. (2015). *The construction of the self: Developmental and sociocultural foundations*. Guilford Publications.

*Heath, N. L., & Glen, T. (2005). Positive illusory bias and the self-protective hypothesis in children with learning disabilities. *Journal of Clinical Child and Adolescent Psychology,* *34*(2), 272–281. https://doi.org/10.1207/s15374424jccp3402

Heine, S. J., & Hamamura, T. (2007). In search of East Asian self-enhancement. *Personality and Social Psychology Review, 11*(1), 4-27.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–135. https://doi.org/10.1017/S0140525X0999152X

Hewitt, J. (1998). *The myth of self-esteem: Finding happiness and solving problems in America*. St Martin's Press.

Heyman, G. D., Compton, A. M., Amemiya, J., Ahn, S., & Shao, S. (2021). Children's reputation management: Learning to identify what is socially valued and acting upon it. *Current Directions in Psychological Science*, *30*(4), 315–320. https://doi.org/10.1177/09637214211009516

Heyman, G. D., Itakura, S., & Lee, K. (2011). Japanese and American children's reasoning about accepting credit for prosocial behavior. *Social Development*, *20*(1), 171–184. https://doi.org/10.1111/j.1467-9507.2010.00578.x

Hofstede, G. (1980). Culture and organizations. *International studies of management & organization, 10*(4), 15-41.

Hofstede, G., Hofstede, G. J., & Minkov, M. (2005). *Cultures and organizations: Software of the mind* (Vol. 2). New York: Mcgraw-hill.

Hox, J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.

Inglehart, R., & Oyserman, D. (2004). Individualism, autonomy, self-expression: The human development syndrome. In *Comparing cultures* (pp. 73-96). Brill.

Ishak, S., Franchak, J. M., & Adolph, K. E. (2014). Perception-action development from infants to adults: Perceiving affordances for reaching through openings. *Journal of*

*Experimental Child Psychology*, *117*(1), 92–105. https://doi.org/10.1016/j.jecp.2013.09.003

James, W. (1950). *Social Identity. The Principles of Psychology, Vol. 1 and 2*.

Johnson, F. (1985). The Western concept of self. *Culture and self: Asian and Western perspectives*, 91-138.

Kagitcibasi, C. (2017). *Family, Self, and Human Development Across Cultures*. Routledge.

Kanagawa, C., Cross, S. E., & Markus, H. R. (2001). "Who am I?" The cultural psychology of the conceptual self. *Personality and Social Psychology Bulletin, 27*(1), 90-103.

Kashima, Y., & Kashima, E. S. (2003). Individualism, GNP, climate, and pronoun drop: Is individualism determined by affluence and climate, or does language use play a role? *Journal of Cross-Cultural Psychology, 34*(1), 125–134. https://doi.org/10.1177/0022022102239159

Keller, H. (2022). *Cultures of infancy*. Routledge.

Kelsey, C., Grossmann, T., & Vaish, A. (2018). Early reputation management: Three-year-old children are more generous following exposure to eyes. *Frontiers in Psychology*, *9*, 698. https://doi.org/10.3389/fpsyg.2018.00698

Kim, Y. H., Chiu, C. Y., Peng, S., Cai, H., & Tov, W. (2010). Explaining East-West differences in the likelihood of making favorable self-evaluations: The role of evaluation apprehension and directness of expression. *Journal of Cross-Cultural Psychology*, *41*(1), 62–75. https://doi.org/10.1177/0022022109348921

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: self-enhancement in the United States and self-criticism in Japan. *Journal of personality and social psychology, 72*(6), 1245.

Klevberg, G. L., & Anderson, D. I. (2002). Visual and haptic perception of postural affordances in children and adults. *Human Movement Science*, *21*(2), 169–186.

https://doi.org/10.1016/S0167-9457(02)00100-8

Kline, M. A., Shamsudheen, R., & Broesch, T. (2018). Variation is the universal: Making cultural evolution work in developmental psychology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1743). https://doi.org/10.1098/rstb.2017.0059

*Kolovelonis, A., & Goudas, M. (2018). The relation of physical self-perceptions of competence, goal orientation, and optimism with students' performance calibration in physical education. *Learning and Individual Differences, 61*, 77–86. https://doi.org/10.1016/j.lindif.2017.11.013

*Kominsky, J. F., & Keil, F. C. (2014). Overestimation of knowledge about word meanings: The "misplaced meaning" effect. *Cognitive Science, 38*(8), 1604–1633. https://doi.org/10.1111/cogs.12122

Konold, K. E., Miller, S. P., & Konold, K. B. (2004). Using teacher feedback to enhance student learning. *Teaching Exceptional Children, 36*(6), 64-69.

*Krebs, S. S., & Roebers, C. M. (2012). The impact of retrieval processes, age, general achievement level, and test scoring scheme for children's metacognitive monitoring and controlling. *Metacognition and Learning, 7*(2), 75–90. https://doi.org/10.1007/s11409-011-9079-3

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology*, *82*(2), 180–188. https://doi.org/10.1037/0022-3514.82.2.180

Kuczynski, L. (2003). Beyond bidirectionality. *Handbook of dynamics in parent-child relations*, 3-24.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9*(5), 178–181.

*Kvavilashvili, L., & Ford, R. M. (2014). Metamemory prediction accuracy for simple prospective and retrospective memory tasks in 5-year-old children. *Journal of Experimental Child Psychology, 127*, 65–81. https://doi.org/10.1016/j.jecp.2014.01.014

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

*Lee, A. M., & Austin, M. (1986). Sex referent cues, success estimations, and performance of children on a lateral jumping task. *Journal of Genetic Psychology, 147*(3), 371–378. https://doi.org/10.1080/00221325.1986.9914511

Lee, K., Cameron, C. A., Xu, F., Fu, G., & Board, J. (1997). Chinese and Canadian children's evaluations of lying and truth telling: Similarities and differences in the context of pro- and antisocial behaviors. *Child Development*, *68*(5), 924–934.

Lee, R. M., Grotevant, H. D., Hellerstedt, W. L., & Gunnar, M. R. (2006). Cultural socialization in families with internationally adopted children. *Journal of family psychology, 20*(4), 571.

*Lee, A. M., Nelson, K., & Nelson, J. K. (1988). Success estimations and performance in children as influenced by age, gender, and task. *Sex Roles, 18*(11–12), 719–726. https://doi.org/10.1007/BF00288056

Leung, M. (1996). Social network and self-enhancement in Chinese children: A comparison of self-reports and peer reports of group membership. *Social Development, 5*, 146–157.

*Lipko, A. R., Dunlosky, J., Lipowski, S. L., & Merriman, W. E. (2012). Young children are not underconfident with practice: The benefit of ignoring a fallible memory heuristic. *Journal of Cognition and Development, 13*(2), 174–188. https://doi.org/10.1080/15248372.2011.577760

*Lipko, A. R., Dunlosky, J., & Merriman, W. E. (2009). Persistent overconfidence despite practice: The role of task experience in preschoolers' recall predictions. *Journal of Experimental Child Psychology*, *103*(2), 152–166. https://doi.org/10.1016/j.jecp.2008.10.002

*Lipko-Speed, A. R. (2013). Can young children be more accurate predictors of their recall performance? *Journal of Experimental Child Psychology*, *114*(2), 357–363. https://doi.org/10.1016/j.jecp.2012.09.012

Lockhart, K. L., Chang, B., & Story, T. (2002). Young children's beliefs about the stability of traits: Protective optimism? *Child Development, 73*(5), 1408–1430. https://doi.org/10.1111/1467-8624.00480

Luo, R., Tamis-LeMonda, C. S., & Song, L. (2013). Chinese parents' goals and practices in early childhood. *Early Childhood Research Quarterly*, *28*(4), 843–857. https://doi.org/10.1016/j.ecresq.2013.08.001

Lyons, K., & Ghetti, S. (2010). Metacognitive development in early childhood: New questions about old assumptions. In A. Efklides, & P. Misailidi, *Trends and prospects in metacognition research* (pp. 259-278). Springer Science & Business Media.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253. https://doi.org/10.1037//0033-295x.98.2.224

Marsh, H. W. (1986). Self-serving effect (bias?) in academic attributions. Its relation to academic achievement and self-concept. *Journal of Educational Psychology, 78*(3), 190–200. https://doi.org/10.1037/0022-0663.78.3.190

*McGee, II, J. P., Milgram, N. A., & Riedel, W. W. (1970). Prediction of performance by retardates and nonretarded children. *American Journal of Mental Deficiency, 74*(5), 637–644.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin, 130*(5), 711–747.

Miller, J. G. (1988). Bridging the content-structure dichotomy. *The cross-cultural challenge to social psychology.* Newbury Park, CA: Sage Publications.

*Miller, S. A., Harris, Y., & Blumberg, R. (1988). Children's judgments of peers' abilities. *Merrill-Palmer Quarterly, 34*(4), 421–436.

Mills, C. M., & Keil, F. C. (2004). Knowing the limits of one's understanding: The development of an awareness of an illusion of explanatory depth. *Journal of Experimental Child Psychology*, *87*(1), 1–32. https://doi.org/10.1016/j.jecp.2003.09.003

Morling, B., & Lamoreaux, M. (2008). Measuring culture outside the head: A meta-analysis of individualism—collectivism in cultural products. *Personality and social psychology review, 12*(3), 199-221.

Nakagawa, S., & Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology, 26*(5), 1253–1274. https://doi.org/10.1007/s10682-012-9555-5

Ng, F. F. Y., Pomerantz, E. M., Lam, S. fong, & Deng, C. (2019). The role of mothers' child-based worth in their affective responses to children's performance. *Child Development*, *90*(1), e165–e181. https://doi.org/10.1111/cdev.12881

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Oishi, S. (2010). The psychology of residential mobility: Implications for the self, social relationships, and well-being. *Perspectives on Psychological Science, 5*(1), 5–21.

https://doi.org/10.1177/1745691609356781

Oliver, M. B., & Hyde, J. S. (1993). Gender differences in sexuality: A meta-analysis. *Psychological Bulletin, 114*(1), 29–51. https://doi.org/10.1037/0033-2909.114.1.29

O'Mara, E. M., Gaertner, L., Sedikides, C., Zhou, X., & Liu, Y. (2012). A longitudinal-experimental test of the panculturality of self-enhancement: Self-enhancement promotes psychological well-being both in the west and the east. *Journal of Research in Personality*, *46*(2), 157–163. https://doi.org/10.1016/j.jrp.2012.01.001

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological bulletin, 128*(1), 3.

Pang, Y., & Richey, D. (2007). Preschool education in China and the United States: A personal perspective. *Early Child Development and Care*, *177*(1), 1–13. https://doi.org/10.1080/14797580500252712

*Peker, A. T., Erkmen, N., Kocaoglu, Y., Bayraktar, Y., Arguz, A., Wagman, J. B., & Stoffregen, T. A. (2021). Perception of affordances for vertical and horizontal jumping in children: Gymnasts versus non-athletes. *Research Quarterly for Exercise and Sport, 92*(4), 770–778. https://doi.org/10.1080/02701367.2020.1775768

Piehlmaier, D. M. (2020). Overconfidence among young decision-makers: Assessing the effectiveness of a video intervention and the role of gender, age, feedback, and repetition. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-61078-z

Plumert, J. M. (1995). Relations between children's overestimation of their physical abilities and accident proneness. *Developmental Psychology*, *31*(5), 866–876. https://doi.org/10.1037/0012-1649.31.5.866

Plumert, J. M., & Schwebel, D. C. (1997). Social and temperamental influences on children's overestimation of their physical abilities: Links to accidental injuries. *Journal of*

*Experimental Child Psychology*, *67*(3), 317–337. https://doi.org/10.1006/jecp.1997.2411

*Powel, W. D., & Toni, M. (1994). Performance and confidence estimates in preschool and young grade-school children. *Child Study Journal, 24*(1), 23–35.

*Pressley, M., & Ghatala, E. S. (1989). Metacognitive benefits of taking a test for children and young adolescents. *Journal of Experimental Child Psychology, 47*(3), 430–450. https://doi.org/10.1016/0022-0965(89)90023-4

Ruble, D. N., Boggiano, A. K., Feldman, N. S., & Loebl, J. H. (1980). Developmental analysis of the role of social comparison in self-evaluation. *Developmental Psychology, 16*(2), 105.

Santos, H. C., Varnum, M. E. W., & Grossmann, I. (2017). Global increases in individualism. *Psychological Science, 28*(9), 1228–1239. https://doi.org/10.1177/0956797617700622

Schneider, W. (1985). Developmental trends in the metamemory- memory behavior relationship: An integrative review. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance. Volume 1: Theoretical perspectives* (pp. 57–109). Orlando: Academic Press.

*Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science*, *1*(2), 291–297. https://doi.org/10.1111/1467-7687.00044

Schneider, W. (2008). The development of metacognitive knowledge in children and adolescents: Major trends and implications for education. *Mind, Brain, and Education*, *2*(3), 114–121. https://doi.org/10.1111/j.1751-228X.2008.00041.x

*Schneider, W., Visé, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children’s memory monitoring: Evidence from a judgment-of-learning task. *Cognitive Development, 15*(2), 115–134. https://doi.org/10.1016/s0885-2014(00)00024-1

Schwarzer, G., Carpenter, J., & Rücker, G. (2015). *Meta-analysis with R*. Cham: springer.

Schwebel, D. C., & Plumert, J. M. (1999). Longitudinal and concurrent relations among temperament, ability estimation, and injury proneness. *Child Development*, *70*(3), 700–712. https://doi.org/10.1111/1467-8624.00050

Sedikides, C., Gaertner, L., & Cai, H. (2015). On the panculturality of self-enhancement and self-protection motivation: The case for the universality of self-esteem. In *Advances in motivation science* (*Vol. 2*, pp. 185–241). https://doi.org/10.1016/bs.adms.2015.04.002

Sedikides, C., Gaertner, L., & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of personality and social psychology, 84*(1), 60.

Sedikides, C., Gaertner, L., & Vevea, J. L. (2007). Evaluating the evidence for pancultural self‐enhancement. *Asian Journal of Social Psychology, 10*(3), 201-203.

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science*, *3*(2), 102–116. https://doi.org/10.1111/j.1745-6916.2008.00068.x

*Shin, H. E., Bjorklund, D. F., & Beck, E. F. (2007). The adaptive nature of children's overestimation in a strategic memory task. *Cognitive Development*, *22*(2), 197–212. https://doi.org/10.1016/j.cogdev.2006.10.001

Smith, K. A., Cipriani, A., & Geddes, J. R. (2016). The usefulness and interpretation of systematic reviews. *BJPsych Advances, 22*(2), 132–141. https://doi.org/10.1192/apt.bp.114.013128

Song, Q., & Wang, Q. (2020). Mother–child reminiscing about peer experiences in European American and Chinese immigrant families: The impact on children's relational self-concepts. *Asian American Journal of Psychology, 11*(1), 40.

Stayton, D. J., Hogan, R., & Ainsworth, M. D. S. (1971). Infant obedience and maternal behavior: The origins of socialization reconsidered. *Child Development*, 1057-1069.

Stipek, D., & Iver, D. M. (1989). Developmental change in children's assessment of intellectual competence. *Child development*, 521-538.

Stipek, D. J. (1981). Children's perceptions of their own and their classmates' ability. *Journal of Educational Psychology*, *73*(3), 404–410. https://doi.org/10.1037/0022-0663.73.3.404

Stipek, D. J., & Daniels, D. H. (1988). Declining perceptions of competence: A consequence of changes in the child or in the educational environment? *Journal of Educational Psychology*, *80*(3), 352–356. https://doi.org/10.1037/0022-0663.80.3.352

Stipek, D. J., & Hoffman, J. M. (1980). Development of children's performance-related judgments. *Child Development*, *51*(3), 912–914.

Stipek, D. J., Roberts, T. A., & Sanborn, M. E. (1984). Preschool-age children's performance expectations for themselves and another child as a function of the incentive value of success and the salience of past performance. *Child Development*, *55*(6), 1983–1989.

Stipek, D. J., & Tannatt, L. M. (1984). Children's judgments of their own and their peers' academic competence. *Journal of Educational Psychology, 76*(1), 75–84. https://doi.org/10.1037//0022-0663.76.1.75

Stone, C. J., & Koo, C.-Y. (1985). Additive splines in statistics. *Proceedings of the American Statistical Association*, 45–48.

*Swanson, H. L. (1983). Relations among metamemory, rehearsal activity and word recall of learning disabled and non-disabled readers. *The British Journal of Educational Psychology, 53*, 186–194. https://doi.org/10.1111/j.2044-8279.1983.tb02549.x

Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics: Pearson new international edition*. Pearson Higher Ed.

Tamis-LeMonda, C. S., Way, N., Hughes, D., Yoshikawa, H., Kalman, R. K., & Niwa, E. Y. (2007). Parents' goals for children: The dynamic coexistence of individualism and collectivism in cultures and individuals. *Social Development, 17,* 183–209.

Thomaes, S., & Brummelman, E. (2016). Narcissism. In D. Cicchetti, *Developmental* *psychopathology: Maladaptation and psychopathology* (pp. 679–725). John Wiley & Sons.

Thomaes, S., Brummelman, E., & Sedikides, C. (2017). Why most children think highly of themselves. *Child Development*, *88*(6), 1873–1884. https://doi.org/10.1111/cdev.12937

Thompson, S. C. (1999). Illusions of control: How we overestimate our personal influence. *Current Directions in Psychological Science, 8*(6), 187–190. https://doi.org/10.1111/1467-8721.00044

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology, 64*, 231–255. https://doi.org/10.1146/annurev-psych-113011-143812

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *96*(3), 506–520. https://doi.org/10.1037/0033-295X.96.3.506

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2008). Do today's young people really think they are so extraordinary? An examination of secular trends in narcissism and self-enhancement. *Psychological Science, 19*(2), 181-188.

Trzesniewski, K. H., Kinal, M. P.-A., & Donnellan, M. B. (2011). Self-enhancement and self protection in a developmental context. In M. D. Alicke & C. Sedikides (Eds.), *Handbook of self-enhancement and self-protection* (pp. 341–357). Guilford Press.

Twenge, J., & Campbell, W. (2009). *The narcissism epidemic: Living in the age of entitlement*. Simon and Schuster.

Twenge, J. M., Carter, N. T., & Campbell, W. K. (2017). Age, time period, and birth cohort differences in self-esteem: Reexamining a cohort-sequential longitudinal study. *Journal of Personality and Social Psychology, 112*(5), e9–e17. https://doi.org/10.1037/pspp0000122.supp

Van der Aar, L. P. E., Peters, S., & Crone, E. A. (2018). The development of self-views across adolescence: Investigating self-descriptions with and without social comparison using a novel experimental paradigm. *Cognitive Development, 48*, 256-270.

Veenman, M. V., Van Hout-Wolters, B. H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and learning, 1*, 3-14.

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics, 30*(3), 261–293. https://doi.org/10.3102/10769986030003261

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

Wang, Q. (2001). “Did you have fun?”: American and Chinese mother-child conversations about shared emotional experiences. *Cognitive Development*, *16*(2), 693–715. https://doi.org/10.1016/S0885-2014(01)00055-7

Wang, Q. (2004). The emergence of cultural self-constructs: Autobiographical memory and self-description in European American and Chinese children. *Developmental Psychology*, *40*(1), 3–15. https://doi.org/10.1037/0012-1649.40.1.3

Wang, Q. (2006). Culture and the development of self-knowledge. *Current Directions in Psychological Science*, *15*(4), 182–187. https://doi.org/10.1111/j.1467-8721.2006.00432.x

Wang, Y., & Ollendick, T. H. (2001). A cross-cultural and developmental analysis of self-esteem in Chinese and Western children. *Clinical Child and Family Psychology Review*, *4*(3), 253–271. https://doi.org/10.1023/A:1017551215413

*Was, C. A., & Al-Harthy, I. S. (2018). Persistence of overconfidence in young children: Factors that lead to more accurate predictions of memory performance. *European Journal of Developmental Psychology, 15*(2), 156–171.

https://doi.org/10.1080/17405629.2016.1264936

Watling, D., & Banerjee, R. (2007). Children's understanding of modesty in front of peer and adult audiences. *Infant and Child Development*, *16*(3), 227–236. https://doi.org/10.1002/icd

Wente, A. O., Goddu, M. K., Garcia, T., Posner, E., Fernández Flecha, M., & Gopnik, A. (2020). Young children are wishful thinkers: The development of wishful thinking in 3- to 10-year-old children. *Child development, 91*(4), 1166-1182.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review, 9*, 625-636.

*Wojcik, D. Z., Moulin, C. J. A., & Souchay, C. (2022). Memory and metamemory for actions in children with autism: Exploring global metacognitive judgements. *Research in Developmental Disabilities, 124*, 104195. https://doi.org/10.1016/j.ridd.2022.104195

*Worden, P. E., & Sladewski-Awig, L. J. (1982). Children's awareness of memorability. *Journal of Educational Psychology, 74*(3), 341–350. https://doi.org/10.1037/0022-0663.74.3.341

Wu, P., Robinson, C. C., Yang, C., Hart, C. H., Olsen, S. F., Porter, C. L., Jin, S., Wo, J., & Wu, X. (2002). Similarities and differences in mothers' parenting of preschoolers in China and the United States. *International Journal of Behavioral Development*, *26*(6), 481–491. https://doi.org/10.1080/01650250143000436

Xia, M., Poorthuis, A. M. G., & Thomaes, S. (2023). Why do young children overestimate their task performance? A cross-cultural experiment. *Journal of Experimental Child Psychology, 226*, 105551. https://doi.org/10.1016/j.jecp.2022.105551

*Xia, M., Poorthuis, A. M. G., Zhou, Q., & Thomaes, S. (2022). Young children's overestimation of performance: A cross-cultural comparison. *Child Development*, *93*(2), e207–e221. https://doi.org/10.1111/cdev.13709

Xu, Y., Farver, J. A. M., Zhang, Z., Zeng, Q., Yu, L., & Cai, B. (2005). Mainland Chinese parenting styles and parent-child interaction. *International Journal of Behavioral Development*, *29*(6), 524–531. https://doi.org/10.1080/01650250500147121

Yamagishi, T., Hashimoto, H., Cook, K. S., Kiyonari, T., Shinada, M., Mifune, N., Inukai, K., Takagishi, H., Horita, Y., & Li, Y. (2012). Modesty in self-presentation: A comparison between the USA and Japan. *Asian Journal of Social Psychology*, *15*(1), 60–68.

Yoshida, T., Kojo, K., & Kaku, H. (1982). A study on the development of self-presentation in children. *Japanese Journal of Educational Psychology*.

Young-Eisendrath, P. (2008). *The self-esteem trap: Raising confident and compassionate kids in an age of self-importance.* Hachette+ ORM.

*Yussen, S. R., & Berman, L. (1981). Memory predictions for recall and recognition in first-, third-, and fifth-grade children. *Developmental Psychology, 17*(2), 224–229. https://doi.org/10.1037/0012-1649.17.2.224

*Yussen, S. R., & Levy Jr, V. M. (1975). Developmental changes in predicting one's own span of short-term memory. *Journal of Experimental Child Psychology, 19*(3), 502–508. https://doi.org/10.1016/0022-0965(75)90079-X

*Yussen, S. R., & Paquette, N. S. (1978). Developmental changes in predicting recognition memory for semantically related and unrelated sentences. *Developmental Psychology, 14*(1), 107–113. https://doi.org/10.1037/0012-1649.14.1.107

Zell, E., Strickhouser, J. E., Sedikides, C., & Alicke, M. D. (2020). The Better-Than-Average Effect in Comparative Self- Evaluation: A Comprehensive Review and Meta-Analysis. *Psychological Bulletin, 146*(2), 118–149. https://doi.org/10.1037/bul0000218

Zhang, R., Noels, K. A., Guan, Y., & Weng, L. (2017). Making sense of positive self-evaluations in China: The role of sociocultural change. *European Journal of Social*

*Psychology*, *47*(1), 36–52.

Zhu, N., & Chang, L. (2019). Education and parenting in China. *School systems, parent behavior, and academic achievement: An international perspective*, 15-28.

# Samenvatting / Summary in Dutch

# Zelfoverschatting in de Vroege Kindertijd en Daarna

Jonge kinderen geven vaak blijk van een opvallend zelfvertrouwen, zelfs als ze geconfronteerd worden met moeilijke taken. “Ik kan alle woorden onthouden”, “Ik weet precies hoe het werkt”, “Ik kan deze bal helemaal tot de rand van het veld gooien”. Volgens experts zijn jonge kinderen vaak overmoedig als het aankomt op hun competenties, begrip en kennis. Uit onderzoek is bijvoorbeeld gebleken dat kleuters en kinderen in de vroege basisschoolleeftijd van zichzelf denken dat ze een betere motoriek (Plumert, 1995; Plumert & Schwebel, 1997), een beter geheugen (Flavell et al., 1970; Yussen & Levy, 1975), een beter begrip van hoe dingen werken (Mills & Keil, 2004), meer kennis (Kominsky & Keil, 2014) en betere wiskundige vaardigheden (Miller et al., 1988) hebben dan hun objectieve competenties of prestaties rechtvaardigen.

Er zijn dus aanwijzingen dat jonge kinderen hun prestaties op diverse taken overschatten. Toch zijn nog veel vragen over zelfoverschatting in de kindertijd onbeantwoord: Hoe hangt zelfoverschatting samen met leeftijd, van de vroege kindertijd tot en met de midden- en late kindertijd? Hoe universeel is kinderlijke zelfoverschatting? Welke psychologische processen dragen bij aan kinderlijke zelfoverschatting? In dit proefschrift beogen we kennis op te doen over de zelfoverschatting van kinderen op verschillende taken, op verschillende leeftijden, in verschillende culturen en in verschillende tijdsperioden. Ook onderzoeken we de psychologische mechanismen die bijdragen aan zelfoverschatting.

We beschrijven in dit proefschrift zowel primair onderzoek (twee studies met data op basis van vier steekproeven) als secundair onderzoek (een meta-analyse met data uit 39 publicaties). Voor elk van deze studies werden vertekeningen onderzocht in de zelfinschatting van kinderen in een taak-specifieke context. Die vertekeningen werden geoperationaliseerd als het verschil tussen de subjectieve, prospectieve zelfinschattingen van kinderen op een taak en objectieve metingen van hun werkelijke prestaties op die taak.

In de twee empirische studies gebruikten we een aangepaste versie van het *prediction-*

*performance* paradigma om de zelfinschatting van vier- en vijfjarige kinderen vast te stellen. We ontwikkelden een werptaak en een geheugentaak. Voor de werptaak vroegen we kinderen om een bal zo ver mogelijk te gooien. Kinderen voorspelden hun prestatie door een vlag te plaatsen op de plek waar zij dachten dat hun bal zou landen. Voor de geheugentaak vroegen we kinderen om zoveel mogelijk afbeeldingen (op een set kaarten) te onthouden. Kinderen voorspelden hun prestatie door het aantal (blanco) kaarten neer te leggen dat ze dachten zich te kunnen herinneren. Deze taken werden meerdere keren uitgevoerd om mogelijke fluctuaties in zelfoverschatting vast te stellen (bijvoorbeeld in reactie op taakervaring of feedback). In een van de empirische studies (Hoofdstuk 3) gebruikten we experimentele onderzoeksmethoden om inzicht te krijgen in het oorzakelijke psychologische proces dat verklaart waarom kinderen zichzelf overschatten.

## Doelstellingen en Belangrijkste Bevindingen

Met het onderzoek in dit proefschrift beoogden we om de huidige kennis over zelfoverschatting van kinderen te vergroten. We formuleerden vijf (gecombineerde) doelen om nog onbeantwoorde vragen te adresseren.

### Doel 1. Overschatten kinderen hun taakprestaties? In welke mate doen ze dat? Overschatten ze hun taakprestaties in dezelfde mate voor verschillende taken?

Om de mate van de zelfoverschatting van kinderen te bepalen hebben we twee cultureel vergelijkende studies (Hoofdstuk 2 en 3) en een meta-analyse (Hoofdstuk 4) uitgevoerd. Het is nog niet bekend of kinderen zichzelf in vergelijkbare mate inschatten op taken die een beroep doen op verschillende kennis of vaardigheden. In de meta-analyse hebben we daarom onderzocht of de zelfoverschatting van kinderen verschilt tussen motorische taken en cognitieve taken (geheugentaken en andere cognitieve taken; Hoofdstuk 4).

Uit ons onderzoek blijkt dat jonge kinderen hun taakprestaties overschatten. In

Hoofdstuk 2 schatten jonge kinderen hun taakprestaties significant hoger in dan hun daadwerkelijke prestaties op zowel een motorische als een geheugentaak. In Hoofdstuk 3 vonden we ditzelfde op de motorische taak, maar niet op de geheugentaak (een vereenvoudigde versie van de geheugentaak die in Hoofdstuk 2 werd gebruikt). In Hoofdstuk 4 bleek uit de meta-analyse dat kinderen (leeftijd 4 tot 12 jaar) hun taakprestaties gemiddeld 1.3 keer zo hoog inschatten als hun werkelijke prestaties, zoals vastgesteld met verschillende soorten taken.

**Doel 2. Hoe hangt kinderlijke zelfoverschatting samen met leeftijd?**

Er zijn aanwijzingen dat kinderen zichzelf minder overschatten naarmate ze ouder worden (Powel & Toni, 1994; Shin et al., 2007; Was & Al-Harthy, 2018). Toch is nog onvoldoende bekend over de exacte aard van dit leeftijdseffect. Daarom hebben we onderzocht of zelfoverschatting (1) afneemt gedurende de kindertijd, en (2) op bepaalde leeftijden (bijvoorbeeld van kleuter- tot schoolleeftijd) sterker afneemt dan op andere leeftijden. We onderzochten zowel lineaire als niet-lineaire leeftijdseffecten (Hoofdstuk 4).

We vonden ondersteuning voor het verwachte leeftijdseffect. De mate waarin kinderen zichzelf overschatten nam af met leeftijd, variërend van een ratio van ongeveer 2.0 op een leeftijd van 4 jaar (kinderen van deze leeftijd verwachten twee keer zo goed te presteren als ze daadwerkelijk doen) tot ongeveer 1.0 op een leeftijd van 12 jaar (kinderen van deze leeftijd overschatten zichzelf niet of nauwelijks meer). Deze daling was geleidelijk en was niet sterker, of minder sterk, op bepaalde leeftijden.

**Doel 3. Is de mate waarin kinderen zichzelf overschatten veranderd in de loop van de recente geschiedenis?**

In de afgelopen decennia hebben wijdverbreide maatschappelijke veranderingen, zoals toenemende economische rijkdom, de overgang van agrarische naar (post)industriële economieën, en migratiestromen, geleid tot een wereldwijde toename van individualistische waarden (Kashima & Kashima, 2003; Oishi, 2010; Santos et al., 2017). Deze veranderingen hebben hun weerslag op de ontwikkeling van kinderen. Zo is het mogelijk dat veranderende socialisatiepraktijken die in toenemende mate worden gekenmerkt door individualisering, van invloed zijn op hoe kinderen hun competenties en prestaties waarnemen. We onderzochten deze mogelijkheid door te toetsen of de mate waarin kinderen zichzelf overschatten is veranderd in de loop van de recente geschiedenis (namelijk de afgelopen vijf decennia; Hoofdstuk 4).

We vonden inderdaad empirisch bewijs dat recente generaties kinderen zichzelf meer overschatten dan generaties kinderen van enkele decennia geleden—we vonden meer zelfoverschatting in recentere studies. Deze toename in zelfoverschatting heeft geleidelijk plaatsgevonden en was niet sterker in bepaalde tijdperioden.

**Doel 4. Overschatten kinderen in niet-westerse samenlevingen hun taakprestaties?**

Vrijwel al het onderzoek tot nu toe werd uitgevoerd bij kinderen die opgroeien in samenlevingen die kunnen worden aangeduid als "WEIRD", dat wil zeggen *Western* (westers), *Educated* (hoogopgeleid), *Industrialized* (geïndustrialiseerd), *Rich* (rijk), en *Democratic* (democratisch). Het is onbekend in hoeverre kinderlijke zelfoverschatting ook in andere culturele contexten voorkomt, zoals in China. In dit proefschrift bestudeerden we cross-culturele verschillen in de mate van zelfoverschatting en de mechanismen die ten grondslag liggen aan zelfoverschatting (Hoofdstuk 2 en 3) bij Nederlandse en Chinese kinderen in de leeftijd van 4 en 5 jaar.

Hoewel we maar twee culturen vergeleken, toont ons onderzoek aan dat zelfoverschatting in ieder geval niet een uniek westers fenomeen is. We vonden veel overeenkomsten in de zelfinschatting van taakprestaties van Nederlandse en Chinese kinderen. In Hoofdstuk 2 vonden we dat zowel Nederlandse als Chinese kinderen hun taakprestaties overschatten; ze deden dit in dezelfde mate, en ze bleven dit allebei doen, zelfs na het ontvangen van feedback op hun presteren. In Hoofdstuk 3 overschatten zowel Nederlandse als Chinese kinderen hun taakprestaties op de motorische taak, maar niet op de geheugentaak. Een cultureel verschil dat we vonden, was dat alleen Chinese kinderen zichzelf accurater inschatten wanneer ze daarvoor werden beloond. In Hoofdstuk 4 konden we geen culturele verschillen onderzoeken vanwege een gebrek aan culturele diversiteit in de geïncludeerde studies (slechts een klein deel van de studies was niet uitgevoerd in Noord-Amerika of Europa).

**Doel 5. Waarom overschatten kinderen hun taakprestaties?**

We onderzochten ook de psychologische mechanismen die bijdragen aan kinderlijke zelfoverschatting. Een mogelijk mechanisme is dat de (meta)cognitie van kinderen nog onvoldoende is ontwikkeld. Een ander mogelijk mechanisme is dat kinderen een sterke behoefte hebben, of gemotiveerd zijn, om competent te zijn of te lijken.

Zelfinschattingen van taakprestaties zijn afhankelijk van cognitieve vaardigheden, vooral metacognitieve vaardigheden (het vermogen om na te denken over de eigen cognitieve processen en bijhorend gedrag). Mogelijk is de zelfoverschatting van kinderen (deels) te verklaren vanuit hun nog beperkt ontwikkelde (meta)cognitie. Eerdere studies hebben aangetoond dat, in sommige omstandigheden, vierjarigen al informatie over hun prestaties kunnen monitoren en onthouden (Lipko et al., 2009; Schneider, 1998). In dit proefschrift hebben we onderzocht of kinderen gebruik maken van informatie over hun eerdere presteren om de inschatting van hun presteren op een toekomstige taak (of taak *trial*) te verbeteren (Hoofdstuk 2).

Een andere mogelijkheid is dat de zelfoverschatting van kinderen (deels) voortkomt uit hun verlangen om competent te zijn of te lijken. Dit wordt de *Wishful Thinking* hypothese genoemd. We hebben deze mogelijkheid op twee manieren onderzocht. We hebben onderzocht of de neiging van kinderen om prestaties te overschatten beperkt is tot hun eigen prestaties en niet generaliseert naar de prestaties van leeftijdsgenoten. Dit zou betekenen dat kinderen wel in staat zijn om correcte prestatieschattingen te maken, maar dat ze dit niet doen als het hun eigen prestaties betreft (wat zou wijzen op een motivationele reden voor zelfoverschatting; Hoofdstuk 2). We hebben ook onderzocht of kinderen zichzelf minder overschatten wanneer ze worden worden gemotiveerd om accurate inschattingen te maken (wanneer hen een beloning wordt beloofd als ze hun prestaties goed inschatten). Dit zou betekenen dat de zelfoverschatting van kinderen niet *alleen* te verklaren is vanuit beperkt ontwikkelde (meta)cognitieve vaardigheden (Hoofdstuk 3).

We vonden meerdere aanwijzingen dat 4- en 5-jarige kinderen in staat zijn om hun prestaties te monitoren en om hun prestatieschattingen te baseren op informatie over hun eerdere presteren—ze doen dit alleen niet op een consistente manier. In Hoofdstuk 3 vonden we bijvoorbeeld een significante afname van zelfoverschatting op de motorische taak tussen *trial* 1 en 2 (maar niet tussen *trial* 2 en 3). Desondanks bleven kinderen hun taakprestaties overigens wel overschatten. Verder vonden we zowel in Hoofdstuk 2 (voor beide taken) als in Hoofdstuk 3 (voor de geheugentaak) dat de prestaties van kinderen op de ene *trial* en hun zelfschatting op de volgende *trial* significant positief geassocieerd waren. Deze bevindingen suggereren dat de zelfoverschatting van jonge kinderen niet voortkomt uit metacognitieve deficiëntie, maar eerder uit metacognitieve *inconsistentie*: kinderen slagen er niet in om prestatie-relevante informatie consistent of volledig op te nemen in de inschatting van hun presteren.

We vonden beperkte ondersteuning voor de *Wishful Thinking* hypothese. In Hoofdstuk 3 verminderde de in het vooruitzicht gestelde beloning voor accurate zelfinschatting de mate waarin Chinese kinderen—maar niet Nederlandse kinderen—hun prestaties op de motorische taak overschatten. Deze bevinding suggereert dat de zelfoverschatting van kinderen deels gemotiveerd kan zijn, in overeenstemming met de

*Wishful Thinking* hypothese. In Hoofdstuk 2 vonden we echter niet dat kinderen hun eigen presteren meer overschatten dan dat van leeftijdgenoten, wat niet in overeenstemming is met deze hypothese. Verder onderzoek is nodig om vast te stellen of, wanneer en hoe de zelfoverschatting van kinderen is geworteld in *wishful thinking*.

## Conclusie

Kinderen overschatten hun taakprestaties en ze doen dit in vergelijkbare mate voor verschillende soorten taken. Deze zelfoverschatting is geen uniek westers fenomeen: Nederlandse en Chinese kinderen vertonen een grotendeels vergelijkbare mate van zelfoverschatting. Desondanks vonden we enkele culturele verschillen; Chinese kinderen zijn meer dan Nederlandse kinderen geneigd om hun zelfoverschatting aan te passen als de situatie daarom vraagt. Jongere kinderen overschatten hun taakprestaties meer dan oudere kinderen, en recente generaties kinderen overschatten hun taakprestaties meer dan eerdere generaties.

Waarom overschatten kinderen hun taakprestaties? Onze bevindingen suggereren dat de zelfoverschatting van kinderen is geworteld in zowel (meta)cognitieve als motivationele factoren. Kinderen zijn niet altijd effectief in het opnemen van prestatie-relevante informatie in de inschattingen van hun presteren. Ook worden de prestatie inschattingen van kinderen soms gekleurd door een verlangen competent te zijn.

Het onderzoek in dit proefschrift biedt inzichten voor ouders, opvoeders en andere professionals die met kinderen werken. De belangrijkste is dat zelfoverschatting een veel voorkomende vertekening is in het zelfbeeld van kinderen, die generaliseert over taken en culturen. Dit suggereert dat er geen reden is voor bezorgdheid wanneer zelfoverschatting wordt waargenomen bij een individueel kind.

Experts vermoeden dat zelfoverschatting in de vroege kindertijd adaptief kan zijn en het leerpotentieel van kinderen ten goede kan komen, maar er is nog vrijwel geen empirisch bewijs voor deze aanname. In dit proefschrift is de adaptiviteit van kinderlijke zelfoverschatting niet onderzocht. Toekomstig onderzoek zou de gevolgen van de zelfoverschatting van kinderen op verschillende leeftijden in kaart kunnen brengen, en de mogelijkheid toetsen dat juist jonge kinderen baat hebben bij zelfoverschatting.

# Acknowledgements

**"It's not always necessary to be strong, but to feel strong." —** Jon Krakauer, 1996. *Into the wild.*

This dissertation would not have been possible without the help of these amazing people in my life. Despite all the inadequacies, frustrations and confusion along the way, I feel empowered thanks to your companionship.

First, I would like to thank those who gave me the courage and hope to consider and actually engage in research before I pursued my PhD. Without you, while academia may not have lost an important person, I certainly would have lost the most important interest and passion of my life. I would like to thank Dr. Lim Kam Ming for guiding me to open the magical door of research; Dr. Lynn Turner for introducing me to the beauty of research in clinical supervision; and Prof. Dr. Marcel van Aken for making it a possibility for me to actually embark on my research journey. I would also like to thank all of my supportive peers, supervisors, and clients who have given me invaluable experience in clinical work and a strong sense of curiosity.

Secondly, I would like to express my heartfelt gratitude to my most faithful companions and most reliable dependents in my research adventures, as well as to what I consider to be the most outstanding researchers in the field - Dr. Astrid Poorthuis and Prof. Dr. Sander Thomaes. I have always claimed that I had the most perfect team during my PhD studies, and indeed, you have inspired and motivated me so much that I am now able to appreciate with immense pride this dissertation that we have accomplished together.

Sander, thank you for being my role model and for always keeping that light on to enlighten me along the way. I admire your diligence, intelligence, and top-notch critical thinking and will always follow your example. I am fortunate to have you as my supervisor; you are open to our differences, flexible to our collaborations, and you give me effective feedback that allows me to be aware of my flaws while also making me feel good about myself. Over the years, with your constant encouragement, I have become almost a competent researcher.

Astrid, thank you for being my favorite teacher, both academically and as a person, you have energized my life in so many ways. I appreciate your conscientiousness, curiosity, and enthusiasm for research, even if it sometimes leads us to arguments. You have been the most creative researcher, the most patient tutor and the most considerate friend I have met during my PhD. I heartily enjoyed every conversation we had; sparks and joy were everywhere, driving me to be a better researcher and a better person.

I would also like to thank my colleagues at OWP and CAS. I am proud to be working in such a wonderful department and have gained a sense of belonging. I am grateful for every exchange of ideas, whether in a classroom, at a meeting or by the coffee machine; I am grateful for the support I have received, whether in the form of work advice or practical tips; and I am grateful for the opportunity to share in the joy of every departmental outing as well as Sinterklaas and Christmas celebrations.

**“Happiness is only real when shared.” —** Jon Krakauer, 1996. *Into the wild.*

I would like to thank my peers, family, and friends, as well as all those who trusted me (and science) and contributed their time and energy to the research for this dissertation. Even greater than making this dissertation possible, you have made it an unsurpassed happiness in my life.

I would like to thank my office roommates, Yue, Eva, Danni and Xiaohuan, who listened to my worries and complaints and made me less bitter when I was having difficulties at work. I would like to thank the Jonkies (Yue, Lysanne, Rogier, Sophie, Nagila, Eva, Judith, Anouk, Tessa, Danni, Yixin, Stefan, Jenna, Saira, Xiaohuan, and Tengwei), who have been the most invaluable resource during my PhD studies and have given me immeasurable help and support. I would also like to thank the colleagues from other departments and universities (Jan de Boer, Dr. Caspar van Lissa, Dr. Rebecca Kuiper and many others) as well as all the support staff for helping me with various difficulties, and the MSc and BSc students from China and the Netherlands for their contribution to the data collection.

I want to thank my family. Mom, thank you for always supporting me and believing in me no matter what I was trying to do. Grandma, although you can't read this message, I want you to know that with your love and influence, I am slowly realizing my dreams. Thank you to my other family members and cousins who wholeheartedly agree with my choices and support me in different ways.

I would like to thank my friends. Thanks to Ziyi, who helped me at the most critical moment of data collection; thanks to Jiaji, who designed the cover of this dissertation; and thanks to Sheryl, who was my 24/7 bad mood trash can and positive energy pump. I would like to thank my friends in China, Jiakun, Chen, Yu and all the others, who always made me feel cared for despite the long distance and the annoying time difference; and Grace, Luel, Julia and all my friends in the Netherlands, I truly appreciate your company and the joy you brought me.

Lastly, I would like to thank Lex for his unlimited supply of love and patience. Thank you for always giving me confidence and cheering me up, and thank you for making me feel at home here.

Thank you to everyone in my life. For better or for worse, thank you for allowing me to feel true, to gain courage, wisdom and love in my never-ending exploration of the world and of myself.

**“I now walk into the wild.” —** Jon Krakauer, 1996. *Into the wild.*

# Curriculum Vitae

**About the Author**

Mengtian Xia (1993) obtained her Bachelor's degree in Applied Psychology from Wenzhou Medical University (2014). During her studies, she accomplished clinical psychology internships at the China Rehabilitation Research Center and Beijing Anding Hospital, and obtained the license to practice as an intermediate counselor in mainland China. She completed her undergraduate thesis on the topic of college students' self-identity and published her first academic paper.

Thereafter, she obtained her Master's Degree in Counseling and Guidance from Nanyang Technological University (2016). During this time, she completed counseling internships in preschools, schools, universities and community centers for young children, children, young adults and seniors. Under the supervision of Dr. Kam Ming Lim, she published her master's thesis on the topic of self-esteem and pro-social behavior in elementary school students.

During 2016-2018, she worked as a child and adolescent psychologist, providing psychological assessment and psychological counseling and therapy to school-aged children and their families. She is licensed to practice as a senior counselor.

With the blessings and queries that her clinical work has brought her, she started her PhD at Utrecht University (2019). She started her research work on children's self-estimation in the Department of Developmental Psychology under the supervision of dr. Astrid Poorthuis and prof. dr. Sander Thomaes. She managed to collect data for two behavioral observational and experimental studies both in mainland China and the Netherlands. She presented her work at national and international conferences. In addition, she was involved in supervising several undergraduate and master students.

Mengtian Xia is currently a postdoctoral researcher in the Department of Developmental Psychology at Radboud University (2023).

## Publications


Xia, M., Poorthuis, A. M., Zhou, Q., & Thomaes, S. (2022). Young children's overestimation of performance: A cross-cultural comparison. *Child Development*, *93*(2), e207-e221.

Xia, M., Poorthuis, A. M., & Thomaes, S. (2023). Why do young children overestimate their task performance? A cross-cultural experiment. *Journal of Experimental Child Psychology, 226*, 105551.

Xia, M., Poorthuis, A. M. G., & Thomaes, S. (2023). Children's overestimation of performance across age, task, and historical time: A meta-analysis. *Manuscript submitted for publication.*