Journal of Cognition and Development

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hjcd20

An Investigation of the Associations Between Counterfactual Reasoning and Executive Functions in Chinese Preschool Children

Yanwen Wu

To cite this article: Yanwen Wu (27 Nov 2024): An Investigation of the Associations Between Counterfactual Reasoning and Executive Functions in Chinese Preschool Children, Journal of Cognition and Development, DOI: 10.1080/15248372.2024.2430194

To link to this article: https://doi.org/10.1080/15248372.2024.2430194

Published online: 27 Nov 2024.

JOURNAL OF COGNITION AND DEVELOPMENT
https://doi.org/10.1080/15248372.2024.2430194

An Investigation of the Associations Between Counterfactual Reasoning and Executive Functions in Chinese Preschool Children

Yanwen Wu

Faculty of Education, University of Cambridge, Cambridge, UK

ABSTRACT

Counterfactual reasoning involves reasoning about what could have happened but did not happen in the past. It assists people to learn from experience to improve future performance. Different from English with subjunctive forms (e.g. if someone *had done* something), Chinese does not apply subjunctive forms. In Chinese, people often decide that a premise is counterfactual based on their knowledge of the past, which involves working memory. This study investigated counterfactual reasoning performance and the associations between counterfactual reasoning and executive functions in 257 Chinese 4- to 6-year-olds. Children completed counterfactual reasoning, working memory, inhibitory control, and receptive language tasks. Results showed that Chinese 4-year-olds achieved above-chance counterfactual reasoning performance. Also, working memory was significantly associated with counterfactual reasoning. Hence, Chinese children might solve counterfactual reasoning tasks without complex subjunctive forms relatively early, and working memory might assist counterfactual reasoning in Chinese.

Introduction

Counterfactual reasoning involves "mentally representing how the world would be now if things had been different in the past" (Rafetseder & Perner, 2014, p. 54). The present study focuses on counterfactual reasoning about social actions, which helps people learn from the past to prepare for the future. For example, a person spilled water when pouring water carelessly. This person might think "If I had poured water more carefully, I would not have spilled water." Counterfactual thoughts activate intentions of taking corrective behaviors (e.g., I want to pour water more carefully next time), fueling behavioral regulation (e.g., pouring water more carefully next time) (Epstude & Roese, 2008). To understand the cognitive underpinnings of counterfactual reasoning and ways in which counterfactual reasoning supports learning and decision-making, it is helpful to study when and how children acquire counterfactual reasoning (Beck & Riggs, 2014). Current research in English- and German-speaking children has shown that counterfactual reasoning emerges around 6 years old (Gautam et al., 2019) and is associated

CONTACT Yanwen Wu yw422@cam.ac.uk Faculty of Education, Selwyn College, University of Cambridge, Grange Road, Cambridge, CB3 9DQ UK

with executive functions (Beck & Riggs, 2014). To complement studies in English- and German-speaking children, this study investigated counterfactual reasoning performance and the associations between counterfactual reasoning and executive functions in children speaking Chinese, which does not use subjunctive forms to directly signal counterfactuality.

Models for counterfactual reasoning

Amsel and Smalley (2000) proposed a simple model for counterfactual reasoning. In this model, a counterfactual outcome is deduced by copying a past event sequence, editing it according to a counterfactual premise, and imagining the consequences of the edited change. For example, a child played outside (event 1), came into the room with dirty shoes (event 2), and the floor got dirty (event 3). To reason from the counterfactual premise "if the child had taken off the dirty shoes," reasoners copy event 1 and change event 2 to derive that the floor would have been clean.

When editing a past event sequence, reasoners follow the nearest possible world constraint. This constraint stipulates that reasoners only change the past event causally dependent on the counterfactual premise while preserving all other past events causally independent of the counterfactual premise (Rafetseder & Perner, 2014). For example, when child A and child B played outside (event 1), child A came into the room with dirty shoes (event 2), child B came into the room with dirty shoes (event 3), and the floor got dirty (event 4). To reason about whether the floor would have been dirty if child B had taken off the dirty shoes, reasoners change event 3 while retaining event 2 to deduce that the floor would still have been dirty. As shown by this example, counterfactual reasoners consider the actual past event that two children came into the room to infer that if one child had taken off the shoes, the floor would still have been dirty. In contrast, basic conditional reasoners ignore reality and reason solely by general patterns in the world (Perner & Rafetseder, 2011). In the above example, as taking off the shoes tends to keep the floor clean, basic conditional reasoners infer that if one child had taken off the shoes, the floor would have been clean. In short, while basic conditional reasoners reason by general regularities, counterfactual reasoners follow the nearest possible world constraint in constructing alternative past scenarios.

Development of counterfactual reasoning

In both English- and Chinese-speaking children, most 3-year-olds passed counterfactual reasoning tasks solvable by general knowledge, such as a child dirtying the floor by shoes. If the child had taken off the shoes, would the floor be dirty? (correct answer: No) (Chen et al., 2007; Harris et al., 1996). Also, 4-year-olds tackled tasks such as a person at home going out to put out a fire. Where would the person be if there had been no fire? (correct answer: Home) (Hu, 2011; Riggs et al., 1998). Tasks like this are solvable by general expectations, such as without the accident, one would not be at the accident site. Children pass counterfactual reasoning tasks insolvable by general knowledge or expectations at a later age. For example, two children wore dirty shoes. If one child had taken the shoes off, would the floor be dirty? (correct answer: Yes) (Rafetseder et al., 2013). For such tasks insolvable by basic conditional reasoning, some studies suggested that children did not pass them until adolescence (Rafetseder et al. 2010; 2013). However, when reducing information complexity, English- and German-speaking children mostly gave correct answers by 5–6 years old (Rafetseder & Perner, 2010; 2018). Thus, the growth from basic conditional reasoning to counterfactual reasoning is likely to occur during the preschool years (3–6 years old).

As reviewed above, studies examining children's performance on counterfactual reasoning tasks insolvable by basic conditional reasoning have focused on children speaking English and German. These languages apply subjunctive forms (e.g., if someone *had done* something) as counterfactual markers or linguistic structures that mark the entry into the counterfactual mode of thought (Bloom, 1981). Subjunctive forms are poorly understood by preschool children (Badger et al., 2020) and are introduced into the school curriculum well beyond preschool years (e.g., Department for Education, 2014). Thus, in preschool children, errors on counterfactual reasoning tasks might reflect difficulties with comprehending subjunctive forms in counterfactual questions rather than counterfactual reasoning per se. While studies in English and German speakers may underestimate children's counterfactual reasoning abilities, children may show counterfactual reasoning abilities earlier when counterfactual conditionals are grammatically less complex (Nyhout & Ganea, 2019).

To examine this possibility, it is especially helpful to study Chinese children. In Chinese, which does not use subjunctive forms, counterfactuality is often inferred from pragmatic contexts (Y. Wang, 2012). For example, "ruguo meiyou xiayu – if not-have rain – if it had not rained" is read as counterfactual from the discourse context that people are talking about yesterday's rain. When pragmatic contexts clearly show what happened in the past, contextual inference of counterfactuality is likely to be efficient. For example, in the location change task (Francis, 2019), children watch the researcher enact a story, where a boy called Jacob hurt himself at the swimming pool. They are then asked what would happen if Jacob had not got hurt. In this task, pragmatic contexts, such as story narratives and the visual scene of the hurt boy at the swimming pool, clearly suggest that Jacob got hurt. Hence, based on contextual information, Chinese children can efficiently judge that the premise "if Jacob had not got hurt" is counterfactual. In contrast, in English and German, subjunctive forms as a syntactic structure are a key component of counterfactual sentences, which directly convey counterfactuality. When hearing a counterfactual sentence, children may devote attention to subjunctive forms as an important sentence component while attending less to more pheripheral contextual information. Thus, when children do not understand subjunctive forms, even when there is contextual information related to the past story, they may still get confused about what counterfactual sentences are about. In this case, they may resort to general knowledge with wide applicability, leading to errors caused by basic conditional reasoning. Thus, when pragmatic contexts clearly show what happened in the past, Chinese-speaking preschool children may readily infer counterfactuality based on contextual information, while English- and German-speaking preschool children may struggle with understanding complex subjunctive forms. In this condition, it is possible that Chinese-speaking children might interpret and reason from counterfactual premises at an earlier age than English- and German-speaking children.

Hence, to study when children show counterfactual reasoning abilities, it is helpful to study counterfactual reasoning performance in not only children speaking languages with subjunctive forms such as English and German but also children speaking languages without subjunctive forms such as Chinese. Existing studies that focused on Chinese preschool children tended to apply tasks solvable by basic conditional reasoning (e.g., Chen et al., 2007; Hu, 2011). To address this limitation, the first aim of the present study was to examine the performance of Chinese preschool children on a counterfactual reasoning task insolvable by basic conditional reasoning. It may inform theories on counterfactual reasoning development that have been based largely on studies in English- and German-speaking children and inspire hypotheses on whether and how this development differs across languages.

Counterfactual reasoning and executive functions

Executive functions are likely to underpin counterfactual reasoning. Executive functions are higher-order cognitive processes that support goal-directed behaviors (Baggetta & Alexander, 2016). Although it is not clear whether executive function is unidimensional or composed of two factors in the preschool years (working memory and inhibitory control; Karr et al., 2018), many studies have supported the two-factor model (e.g., Lerner & Lonigan, 2014; Miller et al., 2012). Thus, in the present study, the researcher applied independent working memory and inhibitory control tasks. As noted below, working memory and inhibitory control might differentially contribute to counterfactual reasoning.

Counterfactual reasoning and inhibitory control

Inhibitory control refers to the ability "to control one's attention, behavior, thoughts, and/or emotions to override a strong internal predisposition or external lure" (Diamond, 2013, p. 137). For example, to touch a cat picture when hearing a dog bark, children inhibit the automatic response of touching the dog picture (Willoughby et al., 2010). In counterfactual reasoning, there are two steps of inhibitory control. In the first step, reasoners overcome the tendency to report the real state of the world. For example, a doctor read in the park, received an emergency call, ran to the hospital to fetch the first-aid bag, and hurried to the swimming pool to attend to a hurt child (Rafetseder & Perner, 2010). To imagine where the doctor would be if the child had not got hurt, reasoners disengage from the doctor's current location at the swimming pool. After inhibiting reality, reasoners are prone to responding according to salient general knowledge, which is inapplicable in counterfactual reasoning tasks insolvable by basic conditional reasoning. To pass these tasks, the second step of inhibitory control is to resist the tendency to rely on well-practiced general knowledge (Francis, 2019). For example, to infer that if the child had not got hurt, the doctor would be in the park, reasoners suppress the assumption that doctors typically work in hospitals if there are not any emergent situations. As children can disengage from reality from as young as 2 years old (Harris et al., 1993), the second step of inhibiting general knowledge as a prepotent response may be particularly important for counterfactual reasoning growth in preschool years.

Current studies have found that inhibitory control was positively correlated with counterfactual reasoning tasks solvable by basic conditional reasoning (Beck et al., 2009; Drayton et al., 2011; Hu, 2011). Also, children's accuracy on such tasks was higher with than without a delay before question-answering to prevent one from responding quickly and impulsively (Beck et al., 2011; Carroll et al., 2021). The present study intended to extend current research by examining the association between inhibitory control and a counterfactual reasoning task insolvable by basic conditional reasoning.

Counterfactual reasoning and working memory

Working memory refers to the ability to mentally store and process information to support ongoing cognitive tasks such as problem-solving (Diamond, 2013). It consists of short-term memory, which temporarily stores speechlike and visuospatial information, and the central executive, which coordinates and attributes attention for information encoding, retrieval, and processing (Baddeley & Hitch, 1974). Both components are involved for maintaining and processing past and counterfactual information in counterfactual reasoning. For example, in the above doctor story, to deduce where the doctor would be if the child had not got hurt, working memory retains the past event that the child got hurt and contrasts it with the premise "if the child had not got hurt" to suggest that this premise is counterfactual. Also, working memory mentally holds both past events (e.g., the doctor read in the park before receiving the emergency call) and the counterfactual premise (e.g., "if the child had not got hurt") and integrates them for creating the counterfactual scenario (e.g., the doctor would keep reading in the park). Hence, to solve counterfactual reasoning tasks in Chinese, working memory facilitates both the derivation of premise counterfactuality and the deduction of counterfactual outcomes. Limited working memory capacity can lead children to omit past contexts to resort to less effortful basic conditional reasoning (Rafetseder et al., 2013).

Echoing the importance of working memory for counterfactual reasoning in Chinese, Chinese 3- to 5-year-olds with higher working memory capacity achieved higher accuracy on counterfactual reasoning tasks solvable by basic conditional reasoning (Zhu, 2014). Also, children's accuracy on such tasks was higher when causal chains were shorter with fewer events to recall and manipulate in working memory (Chen et al., 2007; Zhu, 2014). In contrast, studies in Canadian, Greek, UK, and US 3- to 5-year-olds reported inconsistent findings regarding whether working memory was associated with counterfactual reasoning (Beck et al., 2009, 2010; Drayton et al., 2011; German & Nichols, 2003; Müller et al., 2007). The present study intended to extend the research in Chinese children by examining the association between working memory and a counterfactual reasoning task insolvable by basic conditional reasoning in Chinese preschool children.

Counterfactual reasoning, inhibitory control, and working memory

Inhibitory control and working memory support each other and often co-occur (Diamond, 2013). On the one hand, active goal maintenance in working memory prevents one from emitting automatic responses (Tiego et al., 2018). Studies have shown that people with lower working memory spans made more inhibitory errors (Kane & Engle, 2003), and inhibitory control was partially dependent on working memory (Tiego et al., 2018). On the other hand, inhibiting irrelevant information protects the working memory workspace (Diamond, 2013). As each executive function might explain part of the association between counterfactual reasoning and the other executive function, including both as predictors is crucial for assessing the unique association between inhibitory control/working memory and counterfactual reasoning.

Embracing both executive functions is also helpful for comparing the relative contributions of inhibitory control and working memory to counterfactual reasoning. As Y. Wang (2012) noticed, two key processes in solving counterfactual reasoning tasks in Chinese are the inference of premise counterfactuality and the deduction of counterfactual outcomes. Counterfactual premise interpretation requires reasoners to recollect what really happened and detect the inconsistency between real and counterfactual states of affairs. Counterfactual outcome deduction requires reasoners to retain and integrate both the real event sequence and the counterfactual antecedent, as well as to inhibit reality-inconsistent general knowledge. Hence, inhibitory control mainly aids counterfactual outcome deduction, while working memory facilitates both counterfactual premise interpretation and counterfactual outcome deduction. Therefore, relative to inhibitory control, working memory might be more important for solving counterfactual reasoning tasks in Chinese. Also, the activation of working memory in counterfactual premise interpretation can decrease the need for inhibitory control in counterfactual outcome deduction. To interpret a counterfactual premise (e.g., "if the child had not got hurt"), children may activate past events in working memory (e.g., the child got hurt). This past recall can spread to related events (e.g., in the park, the doctor received the emergency call about the child getting hurt), reducing the prepotent tendency to rely on general knowledge. Thus, in Chinese, counterfactual reasoning might be more strongly associated with working memory than inhibitory control.

Currently, there have been two studies that applied independent inhibitory control and working memory tasks to examine the associations between counterfactual reasoning and each of the two executive functions (Beck et al., 2009; Drayton et al., 2011). Drayton et al. (2011) examined the relation between counterfactual reasoning and inhibitory control and the relation between counterfactual reasoning and working memory separately. Beck et al. (2009) was the only study that assessed whether each of the two executive functions was uniquely associated with counterfactual reasoning. They showed that in UK 3- to 4-year-olds, inhibitory control but not working memory was uniquely associated with counterfactual reasoning tasks solvable by basic conditional reasoning. The lack of an association between working memory and counterfactual reasoning might be because the counterfactual reasoning tasks could be solved by general knowledge or assumptions, which did not require children to consider both the real and counterfactual worlds, reducing the working memory demands. To address this limitation, this study intended to examine the unique association between a counterfactual reasoning task insolvable by basic conditional reasoning and each of inhibitory control and working memory in Chinese preschool children.

Rationale of this study

This study investigated the performance of Chinese preschool children on a counterfactual reasoning task insolvable by basic conditional reasoning and the associations between counterfactual reasoning, inhibitory control, and working memory. It aimed to 1) examine how counterfactual reasoning accuracy changed with age in Chinese preschool children and 2) investigate the unique association between counterfactual reasoning and each of the two executive functions (inhibitory control and working memory). With a lack of studies on the associations between counterfactual reasoning tasks insolvable by basic conditional reasoning and executive functions in Chinese children, this study was largely exploratory. Nevertheless, considering the importance of working memory for counterfactual reasoning in Chinese, one hypothesis was made that counterfactual reasoning was positively associated with working memory.

Apart from executive functions, there are other abilities that might support counterfactual reasoning, such as temporal reasoning, or the ability to represent and relate events in time (Hoerl & McCormack, 2019), and the understanding of the nearest possible world constraint (Rafetseder et al., 2013). These abilities were not considered because first, this study did not aim to investigate all possible factors that might be associated with counterfactual reasoning. Rather, it focused on executive functions. Second, there is a lack of suitable tasks for testing these abilities in Chinese preschool children.

Methods

Participants

Mandarin-speaking children were recruited from two public kindergartens and one private kindergarten in an eastern Chinese city. The sample consisted of 257 50–78-month-olds ($M = 65.56$ months, $SD = 7.65$ months; 124 females). The sample size was larger than what was required according to power analysis. In G* Power (Faul et al., 2007), linear multiple regression with R^2 increases was selected, as this study tested whether each of the two executive functions was uniquely associated with counterfactual reasoning. The effect size was set to $f^2 = 0.085$, as existing studies revealing associations between counterfactual reasoning and executive functions reported at least small-to-medium correlations. In addition, $p = 0.05$, and power = 0.8 were selected. Furthermore, before data collection, it was expected that data on the following variables would be obtained: age, gender, receptive language (included as a covariate as explained later), inhibitory control, and working memory. Hence, the total number of predictors was five. The number of tested predictors was one (working memory/inhibitory control). The power analysis showed that the required sample size was $n = 95$.

There were 67 4-year-olds ($M = 55.70$ months, $SD = 2.64$ months; 30 females), 116 5-year-olds ($M = 65.41$ months, $SD = 3.58$ months; 57 females), and 74 6-year-olds ($M = 74.73$ months, $SD = 2.14$ months; 37 females). Parent-reported mother education levels (data missing for two children) were : 84 (32.68%) high school or below, 72 (28.02%) junior college, and 99 (38.52%) university and beyond. An additional 27 children were excluded due to teacher-reported developmental disorders ($n = 6$), pilot study participation ($n = 2$), kindergarten absence ($n = 3$), inattention characterized by non-responding or leaving the seat to play during any cognitive task ($n = 5$), and control question incorrectness ($n = 11$). Informed parental consent was obtained for all children.

Procedures

This study was approved by the author's university. Data collection took place in two phases (102 and 155 children respectively), between which kindergartens were closed due to Covid situations and the winter holiday. In both phases, each child received a battery of cognitive tasks, which were administered in two 15–35-min sections in a quiet classroom in the child's kindergarten. This study adopted a fixed task order to enhance the consistency of task delivery. In this task order, no two consecutive tasks assessed the same ability, and more engaging tasks were put between less engaging tasks to avoid cognitive fatigue. Table 1 shows the tasks that were administered in Phase 1 and 2, respectively.

It can be seen from Table 1 that between Phase 1 and 2, the major difference in study procedures was that one working memory task (Working Memory Span), one inhibitory control task (Animal Go/No-Go), and two stories in the counterfactual reasoning task were

Table 1. Cognitive tasks administered in phases 1 and 2.

Section	Task 1	Task 2	Task 3	Task 4
Phase 1				
1	Spatial Conflict Arrows (IC)	Working Memory Span (WM)	Location Change (CR, with four stories)	Pick the Picture (WM)
2	Silly Sounds Stroop (IC)	Picture Vocabulary (RL)	a task irrelevant to this study	Animal Go/No-Go (IC)
Phase 2				
1	Spatial Conflict Arrows (IC)	a task irrelevant to this study	Location Change (CR, with two stories)	Pick the Picture (WM)
2	Silly Sounds Stroop (IC)	Picture Vocabulary (RL)	a task irrelevant to this study	

Note. IC = Inhibitory Control. WM = Working Memory. CR = Counterfactual Reasoning. RL = Receptive Language.

dropped in Phase 2. These tasks were dropped because they had low difficulty (>90%) or validity issues. Validity issues mean that a task was not correlated with any other tasks assessing the same ability or loaded poorly on the factor that it was devised to measure. As Working Memory Span and Animal Go/No-Go were excluded from data analysis, due to space constraints, they will not be described in full in this article. Readers may refer to Willoughby et al. (2012) for further details of these tasks.

After dropping the tasks with difficulty or validity issues, the remaining counterfactual reasoning, executive function, and receptive language tasks were administered in both phases. For all these tasks, scores had similar ranges in both phases. Also, mean scores did not significantly differ between phases ($p > 0.05$ on the Mann-Whitney U-test for non-normal data). In addition, the pattern of associations between counterfactual reasoning and executive functions did not differ between phases. Hence, data from both phases were combined to maximize the sample size. In data analysis, phase was added as a predictor to control for the effect of inter-phase differences in study procedures.

Measures

Counterfactual reasoning

This study applied the location change task suitable for 4- to 6-year-olds (Francis & Gibson, 2023; Rafetseder & Perner, 2010). The original location change task included four stories. For each story, there were four conditions. Two conditions only included two locations. In these conditions, counterfactual questions could be answered by choosing the only different-from-reality option based on the general assumption that changing a past event is likely to change the event outcome including the character's final location. The other two conditions included three locations: a character's typical workplace, a location other than the workplace (i.e., an atypical location), and a location where an emergency occurred. For example, in the story where a doctor was the main character, there was a hospital as the doctor's typical workplace, a park as an atypical location, and a swimming pool where a child got hurt. In the 2-Typical condition, before work, the doctor went to the park to read a book. Then he worked in the hospital. In the hospital, he received an emergency call. He got his first-aid bag and ran to the swimming pool to attend to the hurt child. In the 2-Atypical condition, after work, the doctor went to the park to read a book. In the park, he received an emergency call. He ran to the hospital to get his first-aid bag and ran to the swimming pool to attend to the hurt child. In both conditions, the counterfactual question asked where the doctor would be if there had been no emergency. In the 2-Typical condition, the correct answer was the hospital as the doctor's typical workplace. In the 2-Atypical condition, the correct answer was the park as the atypical location. Hence, counterfactual questions in the 2-Typical condition but not the 2-Atypical condition could be solved by basic conditional reasoning. As the focus of the present study was counterfactual reasoning not basic conditional reasoning, it adopted the 2-Atypical condition of the location change task.

Independent professional translators translated English scripts in Francis (2019) into Chinese and back-translated Chinese scripts into English. Names, onomatopoetic words, and locations familiar to Chinese children were adopted, and complex English clauses were decomposed into simple Chinese sentences. Scripts were reviewed by Chinese kindergarten teachers and Chinese-English bilingual developmental psychologists to ensure cultural appropriateness. Table 2 and Figure 1 show the stories in the location change task and an example of task props. All four stories were administered in Phase 1. After Phase 1 data collection, it was found that the firefighter story was too easy for Chinese preschool children (accuracy >90%) and the teacher story had a low factor loading on the counterfactual

Table 2. Stories in the location change task.

Story	Event 1	Event 2	Event 3	Event 4
Doctor	Reading in the park after work	Receiving an emergency call	Running to the hospital to get the first-aid bag	Running to the swimming pool to attend to the hurt child
Policeman	Buying food in the supermarket after work		Running to the police station to get the bike	Driving to the car park to handle the accident
Firefighter	Watching TV at home after work		Running to the fire station to get the fire extinguisher	Running to the forest to put out the fire
Teacher	Going to the park with children	A child getting hurt	Taking the child back to the kindergarten for emergent treatment	Taking the child to the hospital

Figure 1. Props in the doctor Story.

reasoning factor. Hence, these two stories were excluded from data analysis and were not administered in Phase 2. The counterfactual reasoning score was based on the doctor story and the policeman story as described later.

At the task start (the doctor story as an example), the researcher introduced the three places and asked the child to point to each of them. If the child could not identify all places correctly, the place introduction was repeated. Next, the researcher enacted the story using props. After the story enactment, five questions were asked. Firstly, the *now* control question was asked: "Where is the doctor now?" (correct answer: Swimming Pool). Secondly, the counterfactual question was asked: "If the child had not got hurt, where would the doctor be now?" (correct answer: Park). If the child did not respond, choices (park and hospital) were provided in a counterbalanced way. Thirdly, a "Why?" question was newly added to the Chinese version to understand why children made reasoning errors. Phase 1 data showed that this question was not efficient for elucidating reasoning processes (results presented later). Hence, it was dropped in Phase 2 to reduce task length. Fourthly, the *before* control question was asked: "Where was the doctor at the very beginning?" (correct answer: Park). As this question could be interpreted as where the doctor was before going off work, if a child answered "hospital," another question followed: "Where did he go afterwards?" (correct answer: Park). Finally, a *motivation* control question was asked: "Why did the doctor leave the park?" (correct answer: Any response related to the accident). This question was newly added to the Chinese version to ensure that children understood the reason for location changes. Following Francis (2019), children were given two opportunities for all control questions.

The scoring of the location change task was based on the doctor story and the policeman story. In each story, there was a counterfactual question and three control questions. Children were included in the data analysis when they passed the *now* and *before* control questions. They were not excluded when they answered the *motivation* control question incorrectly, as errors on this explanation question were mostly due to children's reluctance to talk to the researcher. Control questions did not contribute to the task score. The task score was based on counterfactual questions. There was one counterfactual question in each of the two (doctor and policeman) stories. Children got one mark for answering one counterfactual question correctly. A score of 0 was for children who answered both counterfactual questions incorrectly. A score of 1 was for children who answered the counterfactual question in either the doctor or policeman story correctly. A score of 2 was for children who answered both counterfactual questions correctly.

Inhibitory control

On an iPad, children completed two inhibitory control tasks from the EF Touch battery (Willoughby et al., 2010, 2012), which has been applied to assess Chinese 3- to 6-year-olds' executive functions (J. Wang et al., 2019). In *Spatial Conflict Arrows*, there were two buttons on the left and right sides of the screen respectively, and an arrow appeared above one of the two buttons. Children were asked to touch the button to which the arrow pointed. After six training items, children received 12 congruent items where the arrow pointed to the same direction as its location on the screen left or right, 12 incongruent items where the arrow pointed to the direction opposite to its location on the screen left or right, and 12 mixed items (7 congruent and 5 incongruent items). The task score was the mean accuracy across the 17 incongruent items. In *Silly Sounds Stroop*, children saw a dog and a cat on the left and right sides of the screen, respectively. The locations of the two pictures changed randomly across items. Children heard an animal's utterance and were asked to touch the dog when hearing a cat meow and touch the cat when hearing a dog bark. There were two training items and 17 test items. The task score was the mean accuracy across all test items. Scores for *Spatial Conflict Arrows* and *Silly Sounds Stroop* were significantly correlated with each other ($rho = 0.26$, $p < 0.001$) and were averaged into an inhibitory control score.

Working memory

On an iPad, children completed the working memory task, *Pick the Picture*, from the EF Touch battery. Children saw an array of two to six pictures on the screen. They were initially asked to touch any picture. After this touch, the same pictures appeared in different locations. Children were asked to touch a new picture, after which picture locations changed randomly. Children were asked to touch a new picture, after which picture locations changed randomly. This process repeated until the number of touches reached the total number of pictures in the array. There were two training items and 32 test items. The task score was the mean accuracy across all test items. Being self-ordered, this task required children to monitor past choices and organize subsequent responses (Cragg & Nation, 2007).

Receptive language

"Receptive language is the ability to accurately comprehend what is said, written, or signed by others" (Frazier, 2011, p. 1228). In this study, receptive language was included as a covariate. As explained earlier, the location change task was narrative-based. Hence, children's ability to comprehend linguistic input might affect their performance on the location change task. In investigating the associations between counterfactual reasoning and executive functions, to control for receptive language as a potential covariate, children completed a receptive language task. In the *Picture Vocabulary Test*, children saw four pictures arranged in a grid and pointed to the picture that showed the meaning of a word said twice by the researcher. The test started from the item from which children passed eight consecutive items and ended when children made six errors in eight consecutive items. The task score (0-120) was the number of correctly answered items plus the number of items before the starting point. In Chinese 3- to 8-year-olds, Gong and Guo (1984) showed that this task had excellent reliability and was associated with language achievements and other verbal intelligence tests.

Statistical analysis

Firstly, answers to the "Why" question in Phase 1 were analyzed. Next, descriptive statistics for cognitive tasks and Kendall's tau-b rank correlation between study variables were estimated. Afterwards, for achieving the first research aim, counterfactual reasoning accuracy and the proportion of children getting each score were plotted for 4-, 5-, and 6-year-olds respectively. The Kruskal–Wallis test assessed whether counterfactual reasoning scores differed between age groups. Also, the one-sample Wilcoxon signed-rank test examined whether counterfactual reasoning scores in each age group differed by chance. For achieving the second research aim, regression analysis was conducted for predicting counterfactual reasoning scores. Before data collection, it was expected that linear regression would be used. This expectation was based on a finding by Francis (2019) who, using the English version of the location change task, showed that scores for the 2-Atypical condition did not significantly deviate from normality. However, in the current sample, the data violated the assumptions of linear regression (curvature in the normal P-P plot). Instead of linear regression, for counterfactual reasoning scores with only three discrete values (0, 1, and 2), ordinal regression and multinomial regression were potentially applicable. Multinomial regression was adopted because data did not fulfil the proportional odds assumption ($p < 0.05$ on the test of parallel lines). As slope coefficients were not the same across different response categories, multinomial regression was preferred over ordinal regression.

Multinomial regression "breaks the outcome variable down into a series of comparisons between two categories" (Field et al., 2012, p. 346). There were three counterfactual reasoning scores: 0, 1, and 2. Therefore, two comparisons were made. A score of 2 was chosen as the reference category because children with full marks were likely to consistently apply counterfactual reasoning. This study intended to examine what abilities might help children become consistent counterfactual reasoners rather than basic conditional reasoners or inconsistent counterfactual reasoners (with a score of 0 or 1). For the two comparisons, the first compared the outcome of getting a score of 2 to the outcome of getting a score of 1 (i.e., a score of 2 versus 1). The second compared the outcome of getting a score of 2 to the outcome of getting a score of 0 (i.e., a score of 2 versus 0). If a continuous variable was positively associated with a score of 2 versus 1, this means that as this variable increased, children were more likely to get a score of 2 rather than 1 on the location change task. Similarly, if a continuous variable was positively associated with a score of 2 versus 0, this means that as this variable increased, children were more likely to get a score of 2 rather than 0 on the location change task. How these comparisons were interpreted will be explained later in the Discussion. The independent variables will be specified later in the Results section.

Results

Preliminary analysis

Responses to the "why" question

In the counterfactual reasoning task in Phase 1, after each counterfactual question, a "Why" question was asked to shed light on why children made reasoning errors. For all responses to the "Why" question following an incorrect answer to the counterfactual question, 65.71% responses simply repeated the premise (e.g., Because the child did not hurt himself/Because there was no emergency). When an emergency did not occur, it was possible for the character to be at the initial location (e.g., park in the doctor story). Hence, such responses did not reveal why children did not choose the initial location as the correct answer. Also, 20% responses were related to the person's occupation (e.g., Because he is a doctor), showing that children's reliance on general knowledge might be one source of counterfactual reasoning errors. Finally, 8.57% responses were silence or "I don't know" and 5.71% responses were idiosyncratic responses hardly interpretable. Overall, the "Why" question

Table 3. Descriptive statistics for all cognitive tasks.

Variable	M	SD	Minimum	Maximum
Counterfactual Reasoning	1.65	0.66	0	2
Inhibitory Control	85.83	13.01	23.53	100.00
Working Memory	78.69	10.57	40.63	100.00
Receptive Language	63.87	19.11	21	99

Table 4. Correlations between study Variables.

	Age	Gender	Mother Education	Phase	Counterfactual Reasoning	Inhibitory Control	Working Memory
Gender	0.02						
Mother Education	0.03	0.03					
Phase	0.20***	−0.02	−0.05				
Counterfactual Reasoning	0.23***	0.05	0.04	0.01			
Inhibitory Control	0.18***	0.08	0.10	−0.03	0.11*		
Working Memory	0.24***	0.09	0.04	−0.01	0.21***	0.18***	
Receptive Language	0.42***	0.01	0.17***	0.09	0.16**	0.19***	0.28***

Note. The reference category for gender is males. The reference category for Phase is Phase 1.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

did not seem to be efficient for elucidating why children made counterfactual reasoning errors, and thus was dropped in Phase 2 to reduce task length.

Descriptive statistics

Descriptive Statistics for all cognitive tasks are presented in Table 3.

There was one outlier for inhibitory control and one outlier for working memory ($Z < -3.29$). Outliers were excluded from data analysis.

Correlations between study variables are presented in Table 4.

As shown in Table 4, counterfactual reasoning was significantly and positively correlated with age, inhibitory control, working memory, and receptive language.

Counterfactual reasoning accuracy across ages

Counterfactual reasoning accuracy was calculated by dividing the number of counterfactual questions that were answered correctly by the total number of counterfactual questions. Counterfactual reasoning accuracy in 4-, 5-, and 6-year-olds is presented in Figure 2. The proportion of children who answered zero, one, or two counterfactual questions correctly in each age group is presented in Figure 3.

Figure 2 showed that children's accuracy was similar for the doctor story and the policeman story with no significant difference between the two stories in 4-, 5-, and 6-year-olds ($p > 0.05$). Figure 2 also showed that counterfactual reasoning accuracy increased with age. Counterfactual reasoning accuracy was 67.42% in 4-year-olds, 85.65% in 5-year-olds, and 90.54% in 6-year-olds. In addition, Figure 3 showed that the proportion of children who answered both counterfactual questions correctly increased with age. The proportion of children who answered both questions correctly was 54.55% in 4-year-olds, 80.00% in 5-year-olds, and 86.49% in 6-year-olds. Combining both stories, counterfactual reasoning scores were 1.35 in 4-year-olds, 1.71 in 5-year-olds, and 1.81 in 6-year-olds. A Kruskal–Wallis test showed that counterfactual reasoning scores significantly changed with age, H

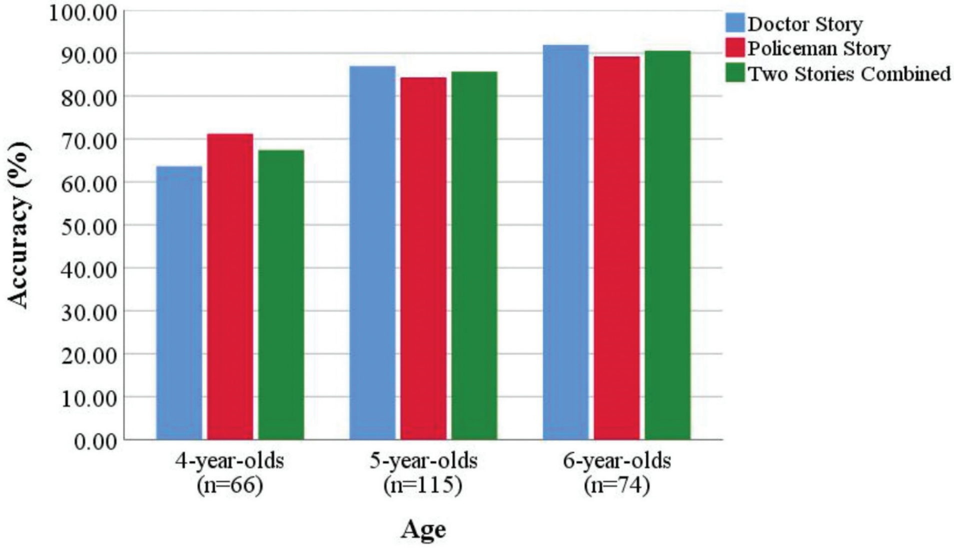

Figure 2. Counterfactual reasoning accuracy across Ages.

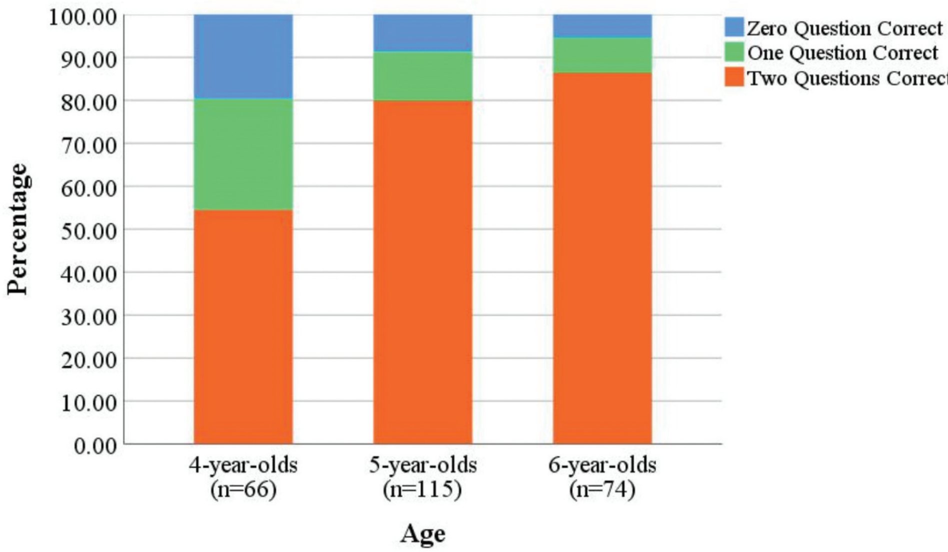

Figure 3. Proportion of children answering zero, one, or two counterfactual questions correctly across Ages.

(2) = 21.27, $p < 0.001$. Pairwise comparisons with adjusted p values found that counterfactual reasoning scores significantly differed between 4- and 5-year-olds ($p < 0.001$, $r = .28$) and between 4- and 6-year-olds ($p < 0.001$, $r = .37$) but not between 5- and 6-year-olds ($p = 0.93$, $r = .07$).

Table 5. Comparisons between location change task scores and the chance score.

Age (year;month)	Z	p	r
4;2-4;11	3.29	0.001	0.40
5;0-5;11	8.12	<0.001	0.76
6;0-6;6	7.28	<0.001	0.85

To compare children's counterfactual reasoning performance against chance, the chance level (i.e., the probability of answering a counterfactual question correctly by chance) needed to be specified first. The chance level of 50% was chosen because in the location change task, there were three locations. Previous studies have shown that on counterfactual reasoning tasks solvable by general knowledge, 3-year-olds mostly selected the non-real option in response to a counterfactual question (Chen et al., 2007; Harris et al., 1996). Therefore, 4- to 6-year-olds in the present study were likely to recognize that the correct answer to a counterfactual question differed from reality. In this case, if children randomly chose one of the two locations that were not the character's real location, the probability of getting correct would be 50%. As the chance level was 50%, and as there were two counterfactual questions in the location change task, if children reasoned by chance, the expected number of counterfactual questions that would be answered correctly was one. Hence, the chance score or the counterfactual reasoning score that children would get if reasoning by chance was 1. Children's counterfactual reasoning scores (0–2) were compared against 1 by the one-sample Wilcoxon signed-rank test, the results of which are presented in Table 5.

As Table 5 shows, in each of 4-, 5-, and 6-years-olds, scores on the location change task were significantly above the chance score with medium-to-large effects.

Association between counterfactual reasoning and executive functions

Before data collection, it was expected that the regression model would contain five predictors: age, gender, inhibitory control, working memory, and receptive language. Three changes were made. First, gender was excluded from the regression model because females ($M = 1.69$) and males ($M = 1.61$) performed similarly on the location change task ($Z = 0.83$, $p = 0.41$ on the Mann-Whitney U-test). Dropping gender did not significantly deteriorate model fit. Second, as data collection took place in two phases, to control for inter-phase procedural differences, phase was included as a predictor. Third, data on mother education were available in this study, dropping which affected the associations between counterfactual reasoning and other cognitive abilities. Hence, mother education was retained as a predictor. The final predictors were as follows: age, phase, mother education, receptive language, inhibitory control, and working memory. The results of the regression analysis are presented in Table 6. Continuous variables were standardized for effect size comparability.

As shown in Table 6, inhibitory control did not account for unique variance in counterfactual reasoning. Working memory significantly predicted whether a child got a counterfactual reasoning score of 2 or 1. Children with higher working memory capacity were more likely to have a score of 2 rather than 1 on the counterfactual reasoning task.

Table 6. Associations between counterfactual reasoning and other cognitive abilities.

					95% Confidence Interval	
Variable	*B* (SE)	Chi-Square	*p*	Odds Ratio	Lower	Upper
A Score of 2 versus 1						
Age	0.27 (0.26)	1.11	0.29	1.32	0.79	2.19
Phase	0.12 (0.43)	0.08	0.77	1.13	0.49	2.63
ME: University and Beyond	0.72 (0.49)	2.14	0.14	2.06	0.78	5.40
ME: Junior College	0.61 (0.50)	1.47	0.23	1.83	0.69	4.87
Receptive Language	0.44 (0.26)	2.90	0.09	1.55	0.94	2.56
Inhibitory Control	0.04 (0.20)	0.04	0.84	1.04	0.71	1.53
Working Memory	0.88 (0.23)***	14.55	<0.001	2.42	1.54	3.80
A Score of 2 versus 0						
Age	1.04 (0.30)***	12.02	0.001	2.82	1.57	5.07
Phase	−0.70 (0.47)	2.20	0.14	0.50	0.20	1.25
ME: University and Beyond	−0.13 (0.52)	0.06	0.81	0.88	0.32	2.46
ME: Junior College	0.03 (0.57)	0.003	0.96	1.03	0.34	3.15
Receptive Language	−0.40 (0.30)	1.79	0.18	0.67	0.38	1.20
Inhibitory Control	−0.10 (0.22)	0.21	0.64	0.90	0.58	1.39
Working Memory	0.24 (0.25)	0.95	0.33	1.27	0.78	2.06

Note. Model chi-square (14) = 56.72, $p < 0.001$. The reference category for Gender is males. The reference category for Phase is Phase 1. ME = Mother Education, the reference category for which is high school or below.
***$p < 0.001$.

Also, age significantly predicted whether a child got a counterfactual reasoning score of 2 or 0. Older children were more likely to have a score of 2 rather than 0 on the counterfactual reasoning task.

Discussion

This study investigated counterfactual reasoning performance and the associations between counterfactual reasoning and executive functions in Chinese preschool children. The results showed that counterfactual reasoning performance was above-chance in each of 4-, 5-, and 6-year-olds. Counterfactual reasoning scores significantly increased from 4 to 5 years old but not from 5 to 6 years old. Also, while age was significantly associated with getting a score of 2 rather than 0 on the counterfactual reasoning task, working memory was significantly associated with getting a score of 2 rather than 1 on the counterfactual reasoning task. These results are discussed in relation to current theories and future research directions.

Counterfactual reasoning performance in Chinese preschool children

Current theories on counterfactual reasoning development draw a conceptual distinction between basic conditional reasoning and counterfactual reasoning (Rafetseder & Perner, 2014). While basic conditional reasoners reason solely by general regularities, counterfactual reasoners follow the nearest possible world constraint to create a counterfactual world closest to reality (Leahy et al., 2014). Studies in English and German-speaking children have shown that children passed counterfactual reasoning tasks insolvable by basic conditional reasoning after they passed tasks solvable by basic conditional reasoning (Rafetseder et al., 2010; 2013; Rafetseder & Perner, 2010, 2018).

A similar developmental pattern emerges in Chinese children. Just like the present study, Chen et al. (2007) and Kang (2017) studied Mandarin-speaking children in kindergartens in mainland China. They showed that scores on tasks solvable by basic conditional reasoning significantly increased between 3 and 4 years old but not between 4 and 5 years old with most 4-year-olds passing these tasks. In contrast, in the present study with a counterfactual reasoning task insolvable by basic conditional reasoning, scores significantly increased between 4 and 5 years old but not between 5 and 6 years old with high accuracy already present in 5-year-olds. In the Introduction, it was noted that studies have shown that English- and German-speaking children shift from basic conditional reasoning to counterfactual reasoning between 3 and 6 years of age. This pattern is likely to be true for Chinese children as well based on the results of past studies and the present study. In Chinese children, the shift from basic conditional reasoning to counterfactual reasoning might take place earlier, with above-chance counterfactual reasoning performance in 4-year-olds.

To further probe into this earlier shift from basic conditional reasoning to counterfactual reasoning in Chinese children, it is helpful to compare the findings of this study to the findings of past studies with the same location change task. Hitherto, there are three studies applying the location change task. Rafetseder and Perner (2010) devised the location change task in German. They also translated the German task into English. Francis and Gibson (2023) applied the English version of the location change task. The present study translated a subset of the English task (stories in the 2-Atypical condition) into Chinese. Hence, apart from minor adaptations (e.g., a supermarket/shopping mall in the Chinese/English version), stories in the 2-Atypical condition were the same in all three studies.

For stories in the 2-Atypical condition, children's accuracy was <58% in the UK and Austrian 4- to 5-year-olds ($M \leq 65$ months) (Francis & Gibson, 2023; Rafetseder & Perner, 2010) and 79.01% in Chinese 4- to 5-year-olds (M = 62 months) in the present study. Also, the accuracy was 66.7% in Austrian 6-year-olds (M = 81 months) (Rafetseder & Perner, 2010) compared to 90.54% in Chinese 6-year-olds (M = 75 months). It is worth noting that the above comparisons are made with four stories in Francis and Gibson (2023) and Rafetseder and Perner (2010) and two stories in the present study. In the present study, Phase 1 data showed that the mean accuracy was higher across four rather than two stories, as children's accuracy was similar for the doctor, policeman, and teacher stories while much higher for the firefighter story. Hence, excluding the firefighter story and the teacher story in the present study is likely to lower the overall task accuracy. An implication is that if comparing children's performance for the two stories that were under focus in the present study (the doctor story and the policeman story), the difference between Chinese children and English- and German-speaking children may become even larger. In short, although this study did not compare performance between multiple language groups, thus could not draw a strong conclusion on cross-linguistic differences, its findings do at least suggest the possibility that Chinese children might show counterfactual reasoning abilities at an earlier age than children speaking English or German.

To explain why children's performance on the location change task was better in this study compared to previous studies, one possibility is that different counterfactual reasoning performance is due to age differences. However, as explained above, higher counterfactual reasoning accuracy in Chinese children was revealed despite the fact that the mean sample age was lower in the present study compared to previous studies in English- and

German-speaking children (noted above). Also, past studies included more iterations of the stories than the present study, as the present study only adopted stories in the 2-Atypical condition. Hence, a shorter location change task might introduce less cognitive fatigue, enhancing task performance in the present study. However, while the location change task seemed to be administered at the start of a session in Francis and Gibson (2023) and Rafetseder and Perner (2010), it followed two other cognitive tasks in the present study. Therefore, it is uncertain whether cognitive fatigue was more of an issue in past studies than in the present study.

Another explanation for the better counterfactual reasoning performance in Chinese children in the present study relative to English- and German-speaking children in previous studies is that past studies have found that Chinese children outperformed their Western counterparts on executive functions (Ellefson et al., 2017; Lan et al., 2011; Sabbagh et al., 2006). As executive functions are likely to be involved in counterfactual reasoning, the higher counterfactual reasoning accuracy in Chinese children might be due to more advanced executive functions. This explanation is unlikely. First, the findings regarding Chinese children outperforming Western children on executive function tasks are not consistent, as studies such as Schirmbeck et al. (2021) did not find any cross-cultural differences in executive function performance. Second, as Lan et al. (2011) observed, relative to US teachers, Chinese teachers more frequently asked children to act properly and avoid inappropriate behaviors. This emphasis on impulse control promotes the development of inhibitory control. Accordingly, Lan et al. (2011) found that while Chinese and US children showed comparable working memory performance, Chinese preschool children showed superior inhibitory control performance. As the present study showed that inhibitory control did not account for unique variance in counterfactual reasoning, differences in inhibitory control cannot account for the better counterfactual reasoning performance in Chinese children in the present study than English- and German-speaking children in past studies.

A more likely explanation for the better counterfactual reasoning performance in Chinese children is linguistic differences. For all versions of the location change task (in Chinese/English/German) as used in past studies and the present study, stories were acted out by props and the final story scene remained in sight when counterfactual questions were asked (e.g., the doctor attended to the hurt child at the swimming pool). Hence, visual cues related to the past emergency were available in the location change task in all languages, which provided clear contextual information on what really happened. In Chinese, premise counterfactuality can be inferred from pragmatic contexts. In the location change task, as there was clear contextual information related to the past emergency as noted above, Chinese children could readily deduce that the question on where the character would be if the emergency had not taken place was counterfactual. In other words, when pragmatic inference of counterfactuality was efficient, Chinese children did not need to understand any complex linguistic structures to interpret counterfactual questions. In contrast, in English and German, subjunctive forms as counterfactual markers are key for expressing counterfactuality. Also, subjunctive forms are part of the predicate as a key constituent of counterfactual sentences. Hence, when English- and German-speaking children do not know subjunctive forms, they might not be able to understand counterfactual questions despite the presence of visual cues. The comprehension of subjunctive forms can pose a linguistic challenge for preschool children (Badger et al., 2020), hindering them from passing the location change task. In short, the reason for why Chinese children were able to interpret and reason from counterfactual premises at an earlier age than their English- and German-speaking counterparts might be due to the fact that they did not need to understand subjunctive forms when pragmatic contexts clearly showed what happened.

In summary, when counterfactuality is readily inferable from pragmatic contexts, Chinese children might pass counterfactual reasoning tasks at an early age. Future studies may examine Chinese children's counterfactual reasoning performance on tasks where counterfactuality is not readily inferable from pragmatic contexts. For example, when counterfactual questions focus on repetitive actions, which happened in the past and can happen in the future, it might be ambiguous whether counterfactual premises refer to past or future events. For example, in the rainy season, it rained yesterday and is likely to rain tomorrow according to the weather forecast. When asked a question on what would have happened if it had not rained, Chinese children might not readily judge that this question referred to what could have happened in the past rather than what might happen in the future. In this condition, counterfactual premise interpretation might be more efficient when languages have counterfactual markers to directly signal counterfactuality. While this study raises the possibility that counterfactual reasoning development might vary across languages, future studies with multiple language groups and multiple counterfactual reasoning tasks may directly test whether and how languages affect counterfactual reasoning development.

Associations between counterfactual reasoning and executive functions

This study found that while age was significantly associated with getting a score of 2 rather than 0 on the counterfactual reasoning task, working memory was significantly associated with getting a score of 2 rather than 1 on the counterfactual reasoning task. This pattern of results might be puzzling. Nevertheless, as both Phase 1 and 2 data showed this pattern of results, it might be a consistent finding that is worth engaging with.

First, this section discusses the finding that age but not working memory/inhibitory control was associated with getting a score of 2 rather than 0 on the counterfactual reasoning task. To interpret this finding, it is helpful to clarify what a score of 0, 1, or 2 on the counterfactual reasoning task reflected. In this study, except for one child (excluding this child did not change the pattern of results), all counterfactual reasoning errors were the character's typical workplace (hospital/police station in the doctor/policeman story). Hence, children with a score of 0 are likely to reason according to general knowledge (i.e., basic conditional reasoning). Children with a score of 2 are likely to be able to consistently apply counterfactual reasoning. Children with a score of 1 might apply a mixture of basic conditional reasoning and counterfactual reasoning. Thus, the finding that age was significantly associated with getting a score of 2 rather than 0 on the counterfactual reasoning task means that with age increases, children were more likely to consistently apply counterfactual reasoning rather than consistently apply basic conditional reasoning. Hence, abilities other than executive functions and receptive language, which improved with age, might explain the difference between consistent basic conditional reasoners and consistent counterfactual reasoners.

Two such abilities that have been proposed by researchers are temporal reasoning and the understanding of the nearest possible world constraint (Beck, 2020; Rafetseder et al., 2013). Temporal reasoning, or the ability to represent and reason about time as one dimension of the world (Hoerl & McCormack, 2019), is important for children to generate counterfactual alternatives that could have occurred at a specific past time (Perner, 2000). Understanding the nearest possible world constraint is fundamental to counterfactual reasoning. If children are unable to represent events in time or do not understand the nearest possible world constraint, they are prone to relying on atemporal general knowledge regardless of receptive language and executive function capabilities. Hence, children's growth in temporal reasoning and the understanding of the nearest possible world constraint with age might explain why age significantly predicted getting a score of 2 (consistent counterfactual reasoning) rather than 0 (consistent basic conditional reasoning) on the counterfactual reasoning task.

The acquisition of the nearest possible world constraint and temporal reasoning abilities might be the first stage in counterfactual reasoning development, prior to which children may consistently apply basic conditional reasoning as noted above. After this stage, children are aware that they need to ground counterfactual inferences in reality. However, they may not be able to consistently apply counterfactual reasoning because to store past event sequences and integrate them with counterfactual antecedents is cognitively demanding, which exerts working memory. In counterfactual premise interpretation, working memory is needed to retain contextual information on what happened in the past (e.g., the story narrative on the child getting hurt) and contrast it with the premise (e.g., "if the child had not got hurt") to infer that the premise is counterfactual. In counterfactual outcome deduction, working memory is involved for storing and combining past story events (e.g., the doctor read in the park before receiving the emergency call) and the counterfactual premise to construct the most plausible counterfactual scenario (e.g., the doctor would continue reading in the park). Hence, to solve counterfactual reasoning tasks in Chinese heavily taps working memory. Accordingly, studies have shown that Chinese children's counterfactual reasoning accuracy was higher when stories had shorter causal chains with fewer events to recall and process in working memory (Chen et al., 2007; Zhu, 2014).

When children have limited working memory, and when facing multiple cognitive tasks that exert working memory resources, they may sometimes lapse into less effortful basic conditional reasoning despite that they understand the nearest possible world constraint. In the present study, children with a score of 1 on the counterfactual reasoning task might fall into this category. The significant association between working memory and getting a score of 2 rather than 1 on the counterfactual reasoning task means that with increases in working memory capacity, children were more likely to consistently apply counterfactual reasoning rather than inconsistently apply counterfactual reasoning. In other words, working memory might assist children to grow from inconsistent to consistent counterfactual reasoners.

Hence, the findings that age was associated with getting a score of 2 rather than 0 (consistent counterfactual reasoning versus consistent basic conditional reasoning) while working memory was associated with getting a score of 2 rather than 1 (consistent versus inconsistent counterfactual reasoning) might reflect that different abilities contributed to different stages of counterfactual reasoning development. In the first stage, the acquisition of the nearest possible world constraint and temporal reasoning abilities prevented children from relying entirely on basic conditional reasoning. In the second stage, children knew that the counterfactual world is maximally similar to reality. However, limited working memory meant that in cognitively demanding situations, they occasionally resorted to less complex basic conditional reasoning. Working memory growth further enhanced their ability to consider past events in making counterfactual inferences, helping children become consistent counterfactual reasoners. These finding interpretations are relatively speculative. Future studies may further examine these possibilities.

For the association between working memory and counterfactual reasoning, it is worth noting that different from this study, Beck et al. (2009) did not find that working memory was significantly associated with counterfactual reasoning after controlling for inhibitory control and receptive language. This might be because Beck et al. (2009) applied counterfactual reasoning tasks solvable by basic conditional reasoning. To tackle such tasks, children only needed to represent the alternative possibility and did not need to relate it to reality (Robinson & Beck, 2000), reducing the working memory demands. Francis and Gibson (2023) used a task insolvable by basic conditional reasoning. They did not find an association between working memory and counterfactual reasoning. This finding might be because the working memory task had validity concerns. The working memory task was negatively correlated with another executive function task that involved working memory. This unexpected negative association pointed out that the working memory task used by Francis and Gibson (2023) might have validity concerns. In contrast, the working memory task in the present study has been well validated (Willoughby et al., 2010, 2012). In the preliminary analysis, this task was significantly loaded onto the working memory factor. Overall, this study is the first one to reveal a significant association between working memory and a counterfactual reasoning task insolvable by basic conditional reasoning.

Unlike working memory, inhibitory control was not found to account for unique variance in counterfactual reasoning. There are multiple explanations. First, with no emergency, a doctor/policeman can be at places other than a hospital/police station (e.g., the policeman can go on patrol outside the police station). With no strict rules such as "If there is no emergency, a doctor/policeman *must* work in the hospital/police station," the typical workplace might not be highly prepotent. Also, Bloom (1981) wrote that with no counterfactual markers, Chinese children combine their knowledge of the past and the counterfactual premise to derive counterfactuality. Past event recall in premise interpretation might lessen the tendency to rely on habitual responses in outcome deduction. For example, to interpret the premise "if the child had not got hurt" as counterfactual, children might think of past events related to the child getting hurt, such as the doctor receiving the emergency call (about the child getting hurt) in the park. This consideration of past events decreased the saliency of "hospital" while increasing the saliency of "park," reducing the difficulty of general knowledge inhibition. On the other hand, if children did not have sufficient working memory capacity to recall past events, inhibitory control was unlikely to be strongly activated, as children were unaware that general knowledge misaligned with specific past contexts.

In summary, in Chinese preschool children, age was significantly associated with getting a score of 2 (consistent counterfactual reasoning) rather than 0 (consistent basic conditional reasoning) while working memory was significantly associated with getting a score of 2 (consistent counterfactual reasoning) rather than 1 (inconsistent counterfactual reasoning) on the location change task. One possibility is that to become counterfactual reasoners, firstly, children need to understand the nearest possible world constraint and acquire temporal reasoning abilities, after which working memory growth further increases the consistency of counterfactual reasoning performance. As this interpretation is tentative, longitudinal studies may assess whether counterfactual reasoning is related to different abilities at different developmental stages. Also, the current findings were revealed for a counterfactual reasoning task about people's social actions. As counterfactual reasoning development might differ by subject matters (Beck, 2020), different findings might emerge for counterfactual reasoning tasks in the physical domain. For example, a machine manipulation always leads to an outcome. If the manipulation had not been made, would the outcome have resulted? – "No" based on general rules while "Yes" when two manipulations were made, each of which could lead to the outcome (McCormack et al., 2018). With a strong link between the machine manipulation and the outcome, inhibitory control is important for suppressing general rules such as "if a machine manipulation is absent, an outcome will not follow". Studies with such tasks may complement this study for understanding whether and how the associations between counterfactual reasoning and executive functions change with the type of causal relations.

Strengths and limitations

To summarize study strengths and limitations, a major strength of this study is that while previous studies on counterfactual reasoning development mostly focused on children in Western societies, this study focused on children speaking Chinese (a language without subjunctive forms). In this population, it found that 4-year-olds achieved above-chance performance on the location change task, suggesting that Chinese children may show counterfactual reasoning abilities earlier than English- and German-speaking children. Hence, an interesting direction for studying counterfactual reasoning development in the future is to investigate cross-linguistic developmental differences. Also, while working memory has been proposed to aid counterfactual reasoning (Beck & Riggs, 2014), this study is the first one to demonstrate an association between working memory and a counterfactual reasoning task insolvable by basic conditional reasoning.

A major limitation of this study is that while the location change task showed good variability in past studies, in Chinese 4- to 6-year-olds, it had limited variability. This issue occurred because Chinese children showed better counterfactual reasoning performance than English- and German-speaking children in past studies. Future studies may adopt counterfactual reasoning tasks with higher difficulty to capture the full range of counterfactual reasoning abilities in Chinese preschool children. Also, while the present study focused only on Chinese children, future studies may include both Chinese-speaking children and English-/German-speaking children to directly test cross-linguistic differences in counterfactual reasoning development.

Conclusion

This study investigates the performance of Chinese preschool children on a counterfactual reasoning task insolvable by basic conditional reasoning and the associations between counterfactual reasoning, inhibitory control, and working memory. Four-year-olds achieved above-chance counterfactual reasoning performance, and working memory was significantly associated with counterfactual reasoning. Hence, when counterfactuality is readily inferable from pragmatic contexts, without the linguistic barrier of understanding complex subjunctive forms, Chinese children might pass counterfactual reasoning tasks at an early age. Also, working memory might assist counterfactual reasoning in Chinese. These findings may enrich the understanding of counterfactual reasoning development across languages.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Yanwen Wu http://orcid.org/0000-0001-8964-992X

Data availability statement

The data necessary to reproduce the findings presented in this paper are not publicly accessible due to a lack of full parental consent for data sharing.

Ethics approval

This research was approved by Faculty of Education, University of Cambridge.

References

Amsel, E., & Smalley, J. D. (2000). Beyond really and truly: Children's counterfactual thinking about pretend and possible worlds. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 121–147). Psychology Press/Taylor & Francis (UK).

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/s0079-7421(08)60452-1

Badger, J. R., Pascoe, E., Jagmetti, L., Baldwin, E., & Mellanby, J. (2020). The developmental trajectory of English conditional grammar in 4- to 11-year-old children. *Journal of Child Language Acquisition and Development*, *8*(1), 1–21.

Baggetta, P., & Alexander, P. A. (2016). Conceptualization and operationalization of executive function: Executive function. *Mind Brain & Education*, *10*(1), 10–33. https://doi.org/10.1111/mbe.12100

Beck, S. R. (2020). The subject of children's counterfactual thoughts. *Psychology of Consciousness: Theory, Research, & Practice*, *7*(4), 340–350. https://doi.org/10.1037/cns0000235

Beck, S. R., Carroll, D. J., Brunsdon, V. E. A., & Gryg, C. K. (2011). Supporting children's counterfactual thinking with alternative modes of responding. *Journal of Experimental Child Psychology*, *108*(1), 190–202. https://doi.org/10.1016/j.jecp.2010.07.009

Beck, S. R., & Riggs, K. J. (2014). Developing thoughts about what might have been. *Child Development Perspectives*, *8*(3), 175–179. https://doi.org/10.1111/cdep.12082

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2009). Relating developments in children's counterfactual thinking and executive functions. *Thinking & Reasoning*, *15*(4), 337–354. https://doi.org/10.1080/13546780903135904

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2010). The effect of causal chain length on counterfactual conditional reasoning. *British Journal of Developmental Psychology*, *28*(3), 505–521. https://doi.org/10.1348/026151009X450836

Bloom, A. (1981). *The linguistic shaping of thought: A study in the impact of language on thinking in China and the West*. Lawrence Erlbaum Associates.

Carroll, D. J., Blakey, E., & Simpson, A. (2021). Can we boost preschoolers' inhibitory performance just by changing the way they respond? *Child Development*, *92*(6), 2205–2212. https://doi.org/10.1111/cdev.13617

Chen, Y., Xiao, X., Wang, J., & Zhang, L. (2007). The development of young children's counterfactual reasoning and its influencing factors. *Psychological Science*, *30*(4), 774–777. https://doi.org/10.16719/j.cnki.1671-6981.2007.04.002

Cragg, L., & Nation, K. (2007). Self-ordered pointing as a test of working memory in typically developing children. *Memory*, *15*(5), 526–535. https://doi.org/10.1080/09658210701390750

Department for Education. (2014, July 16). *National curriculum in England: English programmes of study*. GOV.UK. https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study/national-curriculum-in-england-english-programmes-of-study#key-stage-1—year-1

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

Drayton, S., Turley-Ames, K. J., & Guajardo, N. R. (2011). Counterfactual thinking and false belief: The role of executive function. *Journal of Experimental Child Psychology*, *108*(3), 532–548. https://doi.org/10.1016/j.jecp.2010.09.007

Ellefson, M. R., Ng, F. F.-Y., Wang, Q., & Hughes, C. (2017). Efficiency of executive function: A two-generation cross-cultural comparison of samples from Hong Kong and the United Kingdom. *Psychological Science*, *28*(5), 555–566. https://doi.org/10.1177/0956797616687812

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, *12*(2), 168–192. https://doi.org/10.1177/1088868308316091

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

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. SAGE Publications.

Francis, G. A. (2019). *Pretence and counterfactual reasoning: What is the nature of the relationship?* [Doctoral dissertation]. University of Cambridge. Apollo. https://doi.org/10.17863/CAM.48585

Francis, G. A., & Gibson, J. L. (2023). A plausible role of imagination in pretend play, counterfactual reasoning, and executive functions. *British Journal of Psychology*, *114*(3), 749–770. https://doi.org/10.1111/bjop.12650

Frazier, M. S. (2011). Receptive language. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of child behavior and development* (pp. 1128–1129). Springer. https://doi.org/10.1007/978-0-387-79061-9_2357

Gautam, S., Suddendorf, T., Henry, J. D., & Redshaw, J. (2019). A taxonomy of mental time travel and counterfactual thought: Insights from cognitive development. *Behavioural Brain Research*, *374*, Article 112108. https://doi.org/10.1016/j.bbr.2019.112108

German, T. P., & Nichols, S. (2003). Children's counterfactual inferences about long and short causal chains. *Developmental Science*, *6*(5), 514–523. https://doi.org/10.1111/1467-7687.00309

Gong, Z., & Guo, D. (1984). An intelligence screening test for preschool and primary school children - picture vocabulary test. *Acta Psychologica Sinica*, *1984*(4), 392–401.

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233–259. https://doi.org/10.1016/S0010-0277(96)00715-9

Harris, P. L., Kavanaugh, R. D., Wellman, H. M., & Hickling, A. K. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development*, *58*(1), i–107. https://doi.org/10.2307/1166074

Hoerl, C., & McCormack, T. (2019). Thinking in and about time: A dual systems perspective on temporal cognition. *Behavioral and Brain Sciences*, *42*, Article e244. https://doi.org/10.1017/S0140525X18002157

Hu, X. (2011). *The research on the impact of children's inhibitory control on the counterfactual reasoning ability* [Master's thesis]. Southwest University. CNKI.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47–70. https://doi.org/10.1037/0096-3445.132.1.47

Kang, Y. (2017). Features of young children's counterfactual reasoning under the perspective of domain knowledge. *Journal of Campus Life & Mental Health*, *15*(5), 339–341. https://doi.org/10.19521/j.cnki.1673-1662.2017.05.005
Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160
Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*(3), 677–692. https://doi.org/10.1016/j.jecp.2010.11.001
Leahy, B., Rafetseder, E., & Perner, J. (2014). Basic conditional reasoning: How children mimic counterfactual reasoning. *Studia Logica*, *102*(4), 793–810. https://doi.org/10.1007/s11225-013-9510-7
Lerner, M. D., & Lonigan, C. J. (2014). Executive function among preschool children: Unitary versus distinct abilities. *Journal of Psychopathology & Behavioral Assessment*, *36*(4), 626–639. https://doi.org/10.1007/s10862-014-9424-3
McCormack, T., Ho, M., Gribben, C., O'Connor, E., & Hoerl, C. (2018). The development of counterfactual reasoning about doubly-determined events. *Cognitive Development*, *45*, 1–9. https://doi.org/10.1016/j.cogdev.2017.10.001
Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, *13*, 395–423. http://dx.doi.org/10.1080/15248372.2011.585478
Müller, U., Miller, M. R., Michalczyk, K., & Karapinka, A. (2007). False belief understanding: The influence of person, grammatical mood, counterfactual reasoning and working memory. *British Journal of Developmental Psychology*, *25*(4), 615–632. https://doi.org/10.1348/026151007X182962
Nyhout, A., & Ganea, P. A. (2019). The development of the counterfactual imagination. *Child Development Perspectives*, *13*(4), 254–259. https://doi.org/10.1111/cdep.12348
Perner, J. (2000). About + belief + counterfactual. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 367–401). Psychology Press/Taylor & Francis (UK).
Perner, J., & Rafetseder, E. (2011). Counterfactual and other forms of conditional reasoning: Children lost in the nearest possible world. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology* (pp. 91–109). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199590698.001.0001
Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of "nearest possible world". *Child Development*, *81*(1), 376–389. https://doi.org/10.1111/j.1467-8624.2009.01401.x
Rafetseder, E., & Perner, J. (2010). Is reasoning from counterfactual antecedents evidence for counterfactual reasoning? *Thinking & Reasoning*, *16*(2), 131–155. https://doi.org/10.1080/13546783.2010.488074
Rafetseder, E., & Perner, J. (2014). Counterfactual reasoning: Sharpening conceptual distinctions in developmental studies. *Child Development Perspectives*, *8*(1), 54–58. https://doi.org/10.1111/cdep.12061
Rafetseder, E., & Perner, J. (2018). Belief and counterfactuality: A teleological theory of belief attribution. *Zeitschrift für Psychologie*, *226*(2), 110–121. https://doi.org/10.1027/2151-2604/a000327
Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, *114*(3), 389–404. https://doi.org/10.1016/j.jecp.2012.10.010
Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, *13*(1), 73–90. https://doi.org/10.1016/S0885-2014(98)90021-1

Robinson, E. J., & Beck, S. (2000). What is difficult about counterfactual reasoning? In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 101–119). Psychology Press/Taylor & Francis (UK).

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science*, *17*(1), 74–81. https://doi.org/10.1111/j.1467-9280.2005.01667.x

Schirmbeck, K., Rao, N., Wang, R., Richards, B., Chan, S. W. Y., & Maehler, C. (2021). Contrasting executive function development among primary school children from Hong Kong and Germany. *European Journal of Psychology of Education*, *36*(4), 923–943. https://doi.org/10.1007/s10212-020-00519-9

Tiego, J., Testa, R., Bellgrove, M. A., Pantelis, C., & Whittle, S. (2018). A hierarchical model of inhibitory control. *Frontiers in Psychology*, *9*, Article 1339. https://doi.org/10.3389/fpsyg.2018.01339

Wang, J., Xhang, Y., Zheng, C., Lu, Y., & Qin, J. (2019). Study on the development of executive function subcomponents in 3-6-year-old children. *Psychological Development & Education*, *35*(1), 1–10. https://doi.org/10.16187/j.cnki.issn1001-4918.2019.01.01

Wang, Y. (2012). *The ingredients of counterfactuality in mandarin Chinese* [Doctoral dissertation]. The Hong Kong Polytechnic University. Pao Yue-kong Library. https://theses.lib.polyu.edu.hk/handle/200/6733

Willoughby, M. T., Blair, C. B., Wirth, R. J., Greenberg, M., & The Family Life Project Investigators. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment*, *22*(2), 306–317.

Willoughby, M. T., Blair, C. B., Wirth, R. J., Greenberg, M., & The Family Life Project Investigators. (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment*, *24*(1), 226–239.

Zhu, C. (2014). *Impact of causal chain length and situational causality on the counterfactual reasoning of 3-5 years old children* [Master's thesis]. Southwest University. CNKI.