COPiL CAMBRIDGE OCCASIONAL PAPERS *in* LINGUISTICS
Volume **18**, Article 6: pp. 123–151, 2026 | ISSN: 2050-5949

The Relationship between Theory of Mind and Evaluative Language Use in Narratives of Cantonese-speaking Adults*

Esther Yip
University of Cambridge

Abstract The current study aimed to add to the wider literature by examining whether there is a relationship between theory of mind and evaluative language use in narrative production in two narrative genres —fictional and personal narratives —elicited by Cantonese-speaking adults. 30 Cantonese-speaking participants (24 females and 6 males) aged between 20 and 35 years (M_{age} = *25.47*; SD_{age} = *3.60*), who are either bilingual or multilingual, completed the Strange Stories task as the theory of mind task and two narrative elicitation tasks, one corresponding to each narrative genre. Results showed that theory of mind was not associated with evaluative language use in both fictional and personal narratives produced by Cantonese-speaking adults, suggesting that producing evaluative language in Cantonese does not necessarily relate to the process of understanding the mental states of others in adulthood. The current findings contribute to the inconsistent literature surrounding theory of mind and evaluative language in narrative production. With respect to theory of mind, it appears that the cognitive and linguistic demands of narrating might interfere with the use of theory of mind when producing evaluative language. Evaluations might also be kept at an implicit level by the narrators, resulting in a lack of relationship.

1 Introduction

Narrative production, also known as storytelling, refers to the process of constructing a cohesive discourse with a sequential, temporal and causal structure (Labov & Waletzky 1997). Being competent in producing a cohesive and compelling narrative has been demonstrated to contribute significantly to academic performance and school readiness (Bishop & Edmundson 1987, O'Neill, Pearce & Pick 2004, Reese, Suggate, Long & Schaughency 2010), including reading comprehension (Griffin, Hemphill, Camp & Wolf 2004, Suggate, Schaughency, McAnally & Reese 2018) and even mathematical ability (Chow & Jacobs 2016, O'Neill et al. 2004), thereby, making it one of the important linguistic skills to develop since childhood. To understand the individual differences observed in narrative skills, it is important to examine what its underlying factors may be.

One of the contributing factors was generally recognised to be theory of mind, the ability to attribute mental states to others. This socio-cognitive ability may

* Thank you to Dr Hale Ögel-Balaban for supervising this work and providing invaluable feedback to the manuscript, and the designs and analysis reported in this paper. Thank you to the COPiL Team for their help and support throughout the editing process.

facilitate narrators to create comprehensive stories by using evaluative language, expressions referring to characters' mental states and narrators' interpretation of the events, in turn enabling them to better align their narratives with the audience's perspectives. However, previous studies have reported mixed findings when conducted in varying age groups and languages. For example, a relationship has been found in young children speaking Indo-European languages, such as English and Italian (Guajardo & Watson 2002, Pelletier & Astington 2004), but not in older children and above and when conducted in other languages (Ögel-Balaban & Hohenberger 2019).

Given the theoretical significance of the topic, as well as the importance of understanding this language development, the current study aimed to examine the relationship between theory of mind and evaluative language use in both fictional and personal narratives produced by Cantonese-speaking adults. It would be one of the first studies to examine this across both narrative genres, as well as looking into the developmental stage of adulthood, notably in a Cantonese sample.

2 Literature Review

2.1 The role of evaluative language in narrative production

Narrative production requires the ability to integrate different types of knowledge in relation to the story structure, such as setting, goal, attempt and outcome, as well as linguistic devices, such as utterances, lexicon, and morphosyntax (Aksu-Koç & Aktan-Erciyes 2018). These two aspects of narrative can be primarily identified as microstructure and macrostructure (Gagarina, Klop, Kunnari, Tantele, Välimaa, Balčiūnienė & Walters 2012, Gagarina, Klop, Kunnari, Tantele, Välimaa, Bohnacker & Walters 2019). Microstructure refers to the language level that communicates the events at both utterance and word levels (Gagarina et al. 2019), with measures including lexical and grammatical diversity, such as the number of different words and the mean length of C-unit, a spoken language measure in which a C-Unit consists of one main clause and any additional subordinate clauses attached to it (Lucero 2015, Otwinowska, Opacki, Mieszkowska, Białecka-Pikul, Wodniecka & Haman 2022). On the other hand, macrostructure reflects the cognitive abilities which involve organising and ordering related events into a story structure (Berman 2008, Heilmann, Miller, Nockerts & Dunaway 2010). The two most common analytical frameworks of macrostructure include the sociolinguistic approach known as high-point analysis (Labov 1972, Labov & Waletzky 1967, Peterson & McCabe 1983) and the cognitive approach referred to as the story grammar analysis (Stein & Glenn 1975), whereby the former places a greater emphasis on the evaluative or emotional aspect of a narrative in addition to the fundamental elements of a story structure.

Indeed, expressions relating to emotions, thoughts and intentions of either characters in narratives or narrators themselves are one of the important macrostructural components in narratives (Berman, Slobin, Aksu-Koç, Bamberg, Dasinger, Marchman, Neeman, Rodkin & Sebastián 1994, Peterson & McCabe 1983). These

expressions are often referred to as evaluative language (Labov & Waletzky 1967, Labov 1972). Evaluative language use in narrative production is an important component embedded within Labov's narrative structure theory (Labov 1972). Labov & Waletzky (1967) first established the foundational framework of evaluative language by identifying two functions of narratives —the referential function and the evaluative function —as essential components in forming coherent narratives that make sense to audiences. While referential function relates to factual information about the characters and events of a narrative, evaluative function refers to expressions of the narrator's emotions, attitude and interpretation of events. In other words, any linguistic features that convey narrators' emotions, thoughts, and assessments about the events or story characters are considered as evaluative language, contributing to the evaluative function of narratives. According to Labov (1972), a coherent narrative should provide the audience with not only ample background information, including the context leading up to the event, the characters involved, and when and where the event took place, but also the "point", meaning or significance of events inferred from the narrators' point of view (Labov & Waletzky 1967, Peterson & McCabe 1983). Evaluative language is not restricted to linguistic clauses. It can also take forms as simple as words and phrases (Peterson & McCabe 1983), such as adjectives, modal verbs, or even sentence-ending particles in the case of Cantonese (Ho 2007).

The types of evaluative categories have changed significantly over the years. According to Labov (1972), evaluative language was first identified into four categories: intensifiers (emphasising particular aspects of narratives that are either important or emotionally important such as really), comparators (comparing the event to what did not happen such as but), extensives (extended actions and use of compound adjectives such as left-handed man), and explanations (causal and contrastive statements such as because, though). Since then, these categories have been adapted and modified accordingly. For instance, the classic study of Bamberg & Damrad-Frye (1991) modified their coding criteria for fictional narratives to comprise five evaluative categories, including frames of mind (features related to affect and mental states such as happy, believed), hedges (features related to uncertainty such as maybe, probably), negative qualifiers (features indicating the discrepancy between the narrators' expectations and what happened in the story such as could not find), causal connectors (features of inferred causes of events and contrastive events such as because and but), and character speech (including direct, indirect and free speech), which has since been applied by later research (Eaton, Collis & Lewis 1999, Leung 2001, Ukrainetz, Justice, Kaderavek, Eisenberg, Gillam & Harm 2005). Many subsequent studies have also adopted and identified additional categories which might be more appropriate strategies in different languages and cultures, for example, adding onomatopoeia in the Japanese and Turkish study of Küntay & Nakamura (2004). Küntay & Nakamura (2004) explained that onomatopoeia is considered a common strategy used in Japanese narratives to convey mental states and intensify actions, such as *punpun*, which means furious in Japanese, or *botcchan*, which means splash in Japanese. Since the function of evaluative language is also to capture the audience's attention, onomatopoeia serves precisely this purpose.

Other studies have further refined Labov's categories, as they claimed Labov's concept has limitations concerning perspective-building (Astington 2014, Chafe 1994, Shiro 2003), for example, Shiro (2003) included 'perception' (referring to anything that is perceived through the senses) and 'physical state' (referring to a character's internal state that is not emotional, such as tired).

Overall, the use of evaluative language adds to the objective elements of the event, thereby constructing a coherent and interactive narrative (Labov 1972). Therefore, converging studies have examined evaluative language across all narrative genres as an important component of narrative production (Bamberg & Damrad-Frye 1991, Chang & McCabe 2013, Küntay & Nakamura 2004, Longobardi, Spataro, Renna & Rossi-Arnaud 2014, Ögel-Balaban & Hohenberger 2019, Peterson & McCabe 1983, Shiro 2003, Westerveld, Nelson, Claessen & Westby 2023).

2.2 Theory of mind and evaluative language use

Navigating through social scenarios, such as narrative production, requires sociocognitive skills (Selcuk, Gonultas & Ekerim-Akbulut 2023). While theory of mind has been defined and used interchangeably across the literature, including perspective-taking (Galinsky, Ku & Wang 2005), mentalising (Frith & Frith 2012) and emotion recognition (Baron-Cohen, Wheelwright, Hill, Raste & Plumb 2001), theory of mind is identified as the process of attributing mental states, such as feelings, thoughts, desires and motivations, to others in the present study (Gallese & Sinigaglia 2011, Wellman 2011).

Since narrators with stronger theory of mind capacity are more likely to be adept at shifting their perspectives to characters in narratives, a greater level of evaluative language use will be associated with stronger theory of mind. Moreover, being more competent in attributing mental states to others would also imply that narrators are more able to shift their perspectives to the audience's, thereby understanding that an overuse of evaluative language might start to reduce the flow and coherence of their narratives. As a result, theory of mind is commonly regarded as being associated with the use of evaluative language in both narrative genres — fictional and personal (Drijbooms, Groen & Verhoeven 2017, Fernández 2013, Longobardi et al. 2014), the most common forms of narrative produced in childhood (Allen, Kertoy, Sherblom & Pettit 1994, McCabe & Rollins 1994).

2.3 Relationship in fictional narratives

Fictional narratives are stories produced with imaginary characters and events (Engel 2005, Pesco & Gagné 2017). To produce evaluative language in fictional narratives, narrators must adopt the perspectives of the story characters, distinct from themselves. This genre of narratives is comparatively challenging relative to other genres, as it requires a coherent story structure temporally and sequentially ordered simultaneously with evaluative language (Licandro 2016, McCabe, Bliss, Barra & Bennett 2008). Converging evidence has demonstrated a bidirectional association between theory of mind and evaluative language use (Guajardo

& Watson 2002, Tompkins, Farrar & Montgomery 2019). Guajardo & Watson (2002) found that English-speaking 3 to 4-year-olds, who received training on engaging mental states of story characters, performed significantly better in theory of mind measures than those who did not. This finding confirms the narrative practice hypothesis that practice in producing narratives, specifically evaluative language, facilitates children in learning and understanding goal-directed behaviours of others (Gallagher 2014, Hutto 2008). On the other hand, other empirical studies have shown that theory of mind predicts evaluative language use in fictional narratives. For example, in a cross-sectional study, Pelletier & Astington (2004) found that English-speaking preschoolers who performed better in theory of mind tasks produced more evaluative language in their fictional narratives than those children who scored lower in theory of mind tasks. Similarly, Fernández (2013) found that performance on second-order theory of mind tasks was associated with the use of internal state terms in Spanish-speaking 5 to 7-year-olds' fictional narratives. Ultimately, it is evident that narrators who are better at attributing mental states to others are likely to produce more evaluative language in fictional narratives (Apperly, Warren, Andrews, Grant & Todd 2011).

In contrast, some other studies have not found such a relationship between evaluative language and theory of mind (Longobardi et al. 2014, Meins, Fernyhough, Johnson & Lidstone 2006, Ögel-Balaban & Hohenberger 2019). For example, (Longobardi et al. 2014) found that theory of mind task performance was not associated with the frequency of mental state language in fictional narratives in Italian-speaking 7 to 12-year-olds. Ögel-Balaban & Hohenberger (2019) further observed that the level of evaluative language was low across all Turkish-speaking 4 to 11-year-olds, with no relationship found with theory of mind. Overall, these mixed findings demonstrated that the link between evaluative language use and theory of mind varies by age, linguistic and cultural differences (Berman et al. 1994, Ögel-Balaban & Hohenberger 2019). In light of these contrasting findings, existing evidence seems to suggest potential cross-linguistic and cross-cultural differences in the use of evaluative language in fictional narratives concerning theory of mind. In addition, whether this relationship persists in personal narratives remains largely under-researched, despite the fact that personal narratives are acquired at a very young age and represent the most prevalent type of narratives produced in childhood (McCabe & Rollins 1994, Preece 1987).

2.4 Relationship in personal narratives

Personal narrative involves recounts of past experiences about a one-off event (Hudson & Shapiro 1991). Early structures of this form of discourse have been observed at the age of 2 and 3 years in the form of sharing interesting experiences (Nicolopoulou & Ünlütabak 2017) with caregivers, friends, or teachers (Eisenberg 1985, Fivush, Gray & Fromhoff 1987).

Unlike fictional narratives, the relationship between theory of mind and evaluative language use in personal narratives have not been as widely researched in the narrative literature. Given the limited number of studies investigating personal

narratives in general, it remains unclear whether evaluative language might be associated with theory of mind when recounting previous experiences.

2.5 Comparing the relationship between fictional and personal narratives

Personal narratives emerge as early as the age of two and three years as forms of sharing interesting occurrences to caregivers, teachers, and friends (Eisenberg 1985, McCabe & Rollins 1994), whereas fictional narratives were found to develop at a slower rate and only improve when children start receiving formal education or practices of forming imaginary or made-up events (Bamberg & Damrad-Frye 1991). This difference in developmental progression may stem from fictional narratives being fundamentally more challenging and from the natural occurrence of daily caregiver-child discussions around real-life experiences from early childhood facilitating personal narratives (McCabe & Rollins 1994).

Although a number of studies have researched the use of evaluative language in both fictional and personal narratives (Bamberg & Damrad-Frye 1991, Ukrainetz et al. 2005), only limited studies have investigated the two genres simultaneously (Longobardi et al. 2014, Shiro 2003). Research comparing evaluative language use in fictional and personal narratives has produced inconsistent findings, with some studies reporting greater evaluative language use in fictional narratives and others finding the opposite pattern (Fox 1991, Longobardi et al. 2014, Losh & Capps 2003, Shiro 2003). For example, Losh & Capps (2003) found that emotional terms in evaluative language were produced twice as often in personal narratives compared to fictional narratives in English-speaking children aged 8 and 14 years. Longobardi et al. (2014) were also able to replicate this result in Italian-speaking 9, 10, and 11-year-olds, where they further found that volition and cognitive terms occurred more frequently in fictional than personal narratives. Indeed, the latter finding regarding volition and cognitive terms was also replicated in Venezuelan-speaking 7 and 10-year-olds (Shiro 2003), with the overall frequency of evaluative language being significantly higher in fictional narratives than in personal narratives. However, this overall higher frequency in fictional narratives was inconsistent with the findings of Fox (1991) and Longobardi et al. (2014). Although the nature and direction of differences in evaluative language use between fictional and personal narratives remain unclear, it is evident that evaluative language is used differently across genres (Fox 1991, Longobardi et al. 2014, Shiro 2003). Hence, the findings across fictional and personal narratives seem to indicate that theory of mind may influence evaluative language use differently.

However, the existing literature has yet to explore whether this relationship varies across genres. To produce evaluative language in fictional narratives, narrators must adopt the perspectives of the story characters distinct from themselves. Theory of mind might be more involved in understanding the mental states, such as emotions, thoughts, and intentions, of fictional characters in the story based on existing social knowledge compared to personal narratives, as the narrators' perspectives are displaced to the perspectives of story characters. On the contrary, the process of recounting past experiences only displaces the narrator's consciousness

from the present to the past in the spatio-temporal dimension (Chafe 1994). In other words, narrators primarily reflect on their own mental states when recounting previous personal experiences. Nonetheless, there may be other individuals involved in the past events who the narrators feel the need to mention during personal narratives. Inferring mental states of real people and fictional story characters engage with the same sociocognitive network (Jacobs 2015). Therefore, drawing on both the conceptual distinctions and empirical evidence outlined above, it is plausible that theory of mind interacts differently with evaluative language use in fictional narratives versus personal narratives.

2.6 Present study

Existing literature has yet to explore whether the relationship between evaluative language use and theory of mind differs across fictional and personal narratives, despite the reported differences in evaluative language use (Fox 1991, Longobardi et al. 2014, Shiro 2003). Hence, it is not yet fully understood whether differences in evaluative language use between fictional and personal narratives reflect the underlying variations in the process of theory of mind. Understanding these relationships may reveal the processes involved when producing narratives, thereby explaining the individual differences observed.

With existing literature in evaluative language use primarily focusing on children and adolescents, there remains a large body of unanswered questions about adult development. Examining adults may reveal how evaluative language use continues to evolve throughout adulthood, and how different factors may influence individual differences in the use of evaluative language.

Similarly, there has been a lack of attention on different genres of narratives, for example, personal narratives. Within the context of narrative production, discourses produced can be classified into different narrative genres, such as fictional, personal, procedural and exposition narratives (Heath 1986, Whitworth, Claessen, Leitão & Webster 2015). Not only does the genre determine what the narrators should produce at topic, syntactic, and lexical levels, but it also provides an understanding for the narrators of the extent to which emotions and perspectives of the characters and themselves should be expressed (Wolf, Moreton & Camp 1994). Therefore, identifying how evaluative language is used in corresponding narrative genres is essential in understanding the development of narrative production. Considering that differences in evaluative language use have been observed between fictional and personal narratives, it appears to be a critical question to examine whether theory of mind relates differently to evaluative language used in fictional versus personal narratives. This may provide a more comprehensive understanding of narrative production as a whole.

Given that mixed findings have been reported when conducted in different languages, examining evaluative language use in cultures and languages beyond those previously studied may offer further insights into whether cultural and linguistic differences are involved in evaluative language use. An additional contribution of this study is its unique population of Cantonese speakers. Although there are many

different languages worth investigating, the universality of the evaluative language framework and its processes in Cantonese speakers remains ambiguous. While findings may generalise to Cantonese speakers based on limited research (Ho 2007, Leung 2001), different language systems and families may shape how narrators represent language distinctly in the auditory and visual domains. As opposed to the alphabetic language system commonly researched in the current literature, Cantonese belongs to the Sino-Tibetan language family and uses a logographic writing system composed of characters. It also consists of very different grammatical features from Indo-European languages where the SVO structure, tense inflections on verbs, and grammatical singulars and plurals are absent (Luo, Li & Mok 2020). Moreover, the demographic sample in Hong Kong is unique, given that the majority of Cantonese speakers are either bilingual or multilingual, with their second language mostly being English (Census and Statistics Department 2022). English has been taught in Hong Kong as part of the preschool curriculum, whereas Mandarin is introduced in Year 1. Thus, over 60% of the Hong Kong population has been reported to be multilingual, according to the 2021 Population Consensus (Census and Statistics Department 2022). Considering these interesting characteristics of Cantonese speakers from Hong Kong, this study will address the research gap existing in the Cantonese population.

All in all, the current study aimed to test the following research questions:

i. Does theory of mind relate to the use of evaluative language in both fictional and personal narratives in Cantonese-speaking adults?

ii. Does the relationship between theory of mind and the evaluative language of narratives differ in fictional versus personal narrative?

Open Science Practices were also engaged, having pre-registered the study on OSF (`osf.io`) and uploaded materials and data analysis scripts to OSF (`osf.io/yr3pz`) at the end of the project.

3 Method

3.1 Participants

30 participants (six males and 24 females) aged between 20 and 35 years (M = *25.2*; SD = *3.39*) completed the present study, which were recruited via convenience and snowball sampling. Of this sample, 37% of the participants received more than 16 years of Cantonese Education, with 23% having received 13 to 15 years. All participants reported to be bilinguals, with 87% of participants being trilingual or multilingual. In this sample, the two most commonly spoken languages alongside Cantonese were English and Mandarin. Table 1 presents the participants' characteristics. The study was reviewed and approved by the Faculty of Education Research Ethics Committee at the University of Cambridge.

	Total ($N = 30$)	**Male ($N = 6$)**	Female ($N = 24$)
Age, *M* (*SD*)	25.5 (3.60)	28.0 (5.33)	24.8 (2.84)
Education, n (%)			
High School/Secondary School	5 (17)	2 (33)	3 (12)
Bachelor's Degree	16 (53)	4 (67)	12 (50)
Master's Degree	9 (30)	0 (0)	9 (38)
Country of residence, n (%)			
From Hong Kong	13 (43)	5 (83)	8 (33)
From Elsewhere	17 (57)	1 (17)	16 (67)
Years of Cantonese education, n (%)*			
$<$ 4 years	3 (10)	0 (0)	3 (12)
4–9 years	3 (10)	1 (17)	2 (8)
10–12 years	6 (20)	1 (17)	5 (21)
13–15 years	7 (23)	0 (0)	7 (29)
$>$ 16 years	11 (37)	4 (67)	7 (29)
Bilingual, n (%)	30 (100)	6 (100)	24 (100)
Languages on top of Cantonese, n (%)			
Mandarin	1 (3)	1 (17)	0 (0)
English	3 (10)	1 (7)	2 (8)
English & Mandarin	24 (80)	4 (67)	20 (83)
English, Mandarin, & Others	2 (7)	0 (0)	2 (8)

Table 1 Participants' Characteristics

Note. *Categories of 'Years of Cantonese education' were determined by the years of each educational stage. < 4 years refers to the years of early years. 4-9 years refers to the years of primary school plus early years. 10-12 years refers to the years of the first few grades of secondary school plus primary school and early years. Participants in this category either left Hong Kong to study abroad in English-speaking countries or switched educational institutions to international schools, which is rather common in Hong Kong. 13-15 years refers to the years of the whole of secondary school plus primary school and early years. > 16 years refers to the years of any further education conducted after secondary school including secondary school, primary school and early years.

3.2 Materials

3.2.1 Questionnaires

Participants first completed a demographic questionnaire assessing their age, gender, level of education, country of residence, years of Chinese education, and any other languages they speak.

3.2.2 Tasks

Strange Stories task

Strange Stories task (Happé 1994, White, Hill, Happé & Frith 2009) is an advanced theory of mind task measuring mentalistic understanding. The study implemented a Cantonese version of the task translated in a Master's thesis from the University of Hong Kong. Through a set of 12 story vignettes, participants were required to explain the story character's behaviour in relation to the non-literate mental states embedded in the 12 social situations, such as lies, white lies, misunderstanding, sarcasm, persuasion, contrary emotions, pretend jokes, figure of speech, double bluff, reality and forget. Participants were first presented with a set of texts and illustrations for each story, with the researcher reading the text aloud simultaneously. This will be followed by a comprehension question 'Was it true what X said' and a mental state question 'Why did X say that?'. Only the mental state question was scored following standard task procedures. Using the 3-point scale, responses that only describe the physical state of events or characters were scored as 0; responses that are partially related to the corresponding mental state were credited as 1, where responses that address the appropriate mental state with dispositional words such as 'want' and 'afraid' were scored as a full score of 2. The total score of the whole task is 24, and the percentage score of the task was calculated by adding up the scores of all stories and dividing it by 24. 25% of participants were further coded and scored by a second-rater, resulting in an inter-rater reliability of 96.9% ($K = 0.943$, $p < .001$).

Fictional Narrative Elicitation task

To elicit fictional narratives, participants were presented with a six-picture wordless story adapted from the Cantonese version of the Multilingual Assessment Instrument of Narrative (MAIN; Gargarina et al., 2019). The six pictures depicted a goat family's day in the woods two at a time, where participants were expected to generate a story based on the pictures presented. In the first two pictures, a goat was shown to save a baby goat which fell into a pond; the second set of pictures depicted a fox's attempt to catch the baby goat; the last two pictures illustrated a bird saving the baby goat from the fox. The pictures remained in view when the participants were narrating.

Personal Narrative Elicitation task

To elicit personal narratives, a verbal prompt was used and chosen from the Cantonese version of the Global TALES Protocol (Westerveld et al., 2023). Participants were required to iterate a personal experience about a time when they had experienced a problem and how they fixed it. Further prompts, such as 'can you tell me more?' and 'I would like to know more about your story. Is there anything else you can tell me?' were asked when participants only responded with one or two sentences.

3.3 Coding criteria

To calculate the evaluative language use, each narrative was first separated into individual clauses.

In this study, C-Units (clausal units) were used to determine the segmentation of a clause (Hughes, McGillivray & Schmidek 1997, Loban 1976). C-Units are main clauses with modifiers, meaning that one C-Unit includes one main clause and all subordinate clauses preceding or following it. In other words, it can be either considered as a simple sentence (a single main clause) or a complex sentence (a main clause coupled with additional subordinate clauses). Subordinate clauses were not considered as a C-Unit and would be transcribed to its attached main clauses, as they are dependent on the main clauses to have essential meaning.

An example of a main clause:

(1) 隻 羊 媽媽 即刻 去 救 隻 羊 *bb* 啦。
CL sheep mother immediately go save CL sheep baby SFP

'The mama goat immediately went to save the baby goat.'

An example of a main clause, including a subordinate clause:

(2) 啱啱 係 到 飲緊 水 嘅 羊 依家 係 河邊 食 草。
just.now be there drink-ProM water LP sheep now be riverside eat grass

'The goat which was previously drinking water is now eating grass by the river.'

However, there are certain exceptions. Utterances containing coordinating conjunctions, such as *and*, *but*, and *so*, could be separated into two main clauses.

An example of coordinating conjunctions:

(3) 個 隻 羊 跳 左 落 水 跟著 就 開始 游水 啦。
CL CL sheep jump PerM fall water then then start swim SFP

'The goat jumped into the water and (it) started swimming.'

C-unit segmentation:

C-unit 1: 'The goat jumped into the water,'
C-unit 2: 'and (it) started swimming.'

Another exception when segmenting C-units includes sentence fragments, which describe utterances that do not necessarily fulfil the criteria of a main clause but indicate a complete thought based on intonation.

An example of sentence fragments:

(4) 個 三 隻 羊 就 一家團聚 。 好 開心 咁 。
CL three CL sheep then reunite very happy SFP

'The three goats then had a family reunion. Very happily.'

C-unit segmentation:

C-unit 1: 'The three goats then had a family reunion.'
C-unit 2: 'Very happily.'

Another notable rule concerns pauses and intonation. C-units should always be segmented based on clause structure (main vs. subordinate), even when there are significant pauses.

An example of pauses and intonation:

(5) 隻 狐狸 就 跑 走 左 <> 因為 隻 烏鴉 追住 佢 。
CL fox then run away PerM because CL crow chase-RP him

'The fox ran away <> because the crow was chasing him.'

Further information about how C-units are determined in Cantonese can be found in Wong & Chen (2009).

Overall, the ratio score of evaluative language was derived by the number of evaluative clauses divided by the total number of clauses. 25% of the participants in both fictional and personal narratives were further coded by a second rater, with the interrater reliability found to be 98.2% ($K = 0.96$, $p < .001$) for the fictional narrative elicitation task and 100% ($K = 1$, $p < .001$) for the personal narrative elicitation task.

3.4 Procedure

The study was built on Qualtrics (https://www.qualtrics.com/uk/), an online survey platform, where participants were required to first arrange a time with the researcher to schedule a synchronous study session via Zoom (https://zoom.us/). Participants were also informed about the requirement of having their responses audio-recorded via their electronic devices and being situated at a quiet place with stable internet connection when they were invited.

Once the participants joined the Zoom meeting, they were greeted with the information sheet shown on the computer screen using the 'share screen' function of Zoom. Ample time was then given to the participants to read the information sheet and give consent using the 'remote access' function of Zoom to tick the corresponding consent statement boxes. Participants were then required to complete a few questions about themselves and language background at their own pace.

Having filled out the questionnaire, 'remote access' was turned off in preventing any misclicks from the participants' side. Participants were then presented with the text and illustration of the first story from the Strange Stories task (Fung 2012). They were advised to start recording their responses once they felt ready to do

Category	Definition	Examples
1) Mental state terms	Internal states of the narrator or characters	
a) Emotional states	Emotive states	生氣了 'get angry', 開心 'happy', 傷心了 'become sad', 害怕了 'be scared of', 喜歡 'like'
b) Affect expression	Expression of affect	笑 'smile', 擁抱 'hug'
c) Motivation and ability	Expression of intention, desire, and ability	想 'want', 嘗試 'try', 可以 'be able, can'
d) Cognitive states	Cognitive mental states	諗 'think', 決定 'decide', 知道 'know'
2) Hedges	Expression of uncertainty of the narrator for the content of his/her statement	應該 'probably', 可能 'maybe', 我以為 'I assume'
3) Negative qualifiers	Any direct negation indicating the discrepancy between narrator's expectations and what happened in the story or event	搵唔到 'could not find'
4) Character speech	Direct and indirect statements of the utterances of the story characters or characters in the personal experience	我個鄰居就話:「我醒左啦。」'My neighbor replied, "I am awake now."'
5) Enrichment expressions	Adverbial phrases for unexpected/inferred nature of an action	再 'again', repetitions, intensifiers such as 非常 'very', 每次 'every'
6) Evaluative remarks	Expression of the subjective point of view of the narrator	「當然，那隻狗正躺在它的主人床邊。」'Of course, the dog is lying together with his owner in bed.'
7) Causal expressions	Expression of the inferred causes of events and actions	因為 'because', 為左 'for'
8) Contrastive expressions	Expression of unexpected/contrastive events	但係 'but'

Table 2 Categories of Evaluative Devices Adapted from Ögel-Balaban & Hohenberger (2019)

so. Upon completion of the first task, participants were encouraged to take a break before proceeding to the Personal Narrative Elicitation task (Westerveld et al. 2023),

followed by the MAIN (Gagarina et al. 2019). Participants were then advised to stop their recording and send it to the researcher immediately. At the end of the study, participants were thanked and debriefed.

The whole study took approximately 30 to 40 minutes. The tasks were presented in a fixed order considering the non-experimental nature of the current study, which would not benefit from a randomised task order.

3.5 Statistical Analyses

Narrative recordings were transcribed using CLAN (Computerised Language Analysis) and all statistical analyses were conducted with R version 4.3.2 (2023-10-31) (R Core Team, 2023) and RStudio using the *tidyverse* package (Wickham, Averick, Bryan, Chang, McGowan, François & Yutani 2019). Preliminary analyses were first conducted to check data distribution using the Shapiro-Wilk test. All data was winsorised to include all data within three standard deviations. Given there were no missing data, no further multiple imputation was conducted.

Pearson correlations were conducted between the ratio of evaluative language use in both narrative elicitation tasks and the percentage score of the Strange Stories task to investigate the relationship between theory of mind and evaluative language use in both genres of narratives. In case the correlations were significant, further regression would be conducted, with the percentage score of the Strange Stories task entered as predictor variable and the ratio score of evaluative language use from respective genre entered as outcome variable. To examine whether a difference could be found between fictional narratives' relationship with theory of mind and that of personal narratives, the Fisher r-to-z transformation would be conducted to analyse whether the two correlation coefficients were significantly different. All task scores were scored by a further second rater to calculate the inter-rater reliability

4 Result

4.1 Preliminary analyses

By examining the box plots, two data points were identified as outliers (any values beyond 1.5 x IQR were considered outliers in the study) in the Strange Stories task, with one outlier also found in the fictional narrative elicitation task. No outliers were found in the personal narrative elicitation task. All reported outliers were then winsorised. In other words, the outliers were replaced with the nearest acceptable value, that is, the upper boundary (Q3 + 1.5 x IQR) and lower boundary (Q1 - 1.5 x IQR). This approach retains all observations in the dataset while limiting the influence of extreme values, which is particularly important given the small sample size of the present study.

All variables were found to be normally distributed ($p > .05$), having run Shapiro-Wilk tests and examined histograms and QQ plots. The criteria of skewness ($< \pm 2$) and kurtosis ($< \pm 7$) were also met (Hair, Black, Babin & Anderson 2010), indicat-

ing that all variables were not skewed or kurtotic. All descriptive statistics and preliminary analysis of the three variables are listed in Table 3.

			Shapiro-Wilk test			
Variables	*M*	*SD*	**Statistic**	*p*	**Skewness**	**Kurtosis**
Percentage score of the Strange Stories task	74.2	9.81	0.948	.145	0.241	2.27
Ratio of evaluative language use in the fictional elicitation task	0.710	0.148	0.956	.248	-0.406	2.27
Ratio of evaluative language use in the personal elicitation task	0.712	0.160	0.950	.167	-0.491	2.45

Table 3 Descriptive Statistics and Normality Statistics of Study Variables

4.2 Relationship between theory of mind and evaluative language use

Pearson's correlation was first conducted between the percentage score of Strange Stories task and the ratio score of evaluative language use in fictional narratives, which showed a negative yet non-significant correlation, *r(28) = -0.235, p = .212*.

For personal narratives, there was a positive yet non-significant correlation between the percentage scores of Strange Stories task and the ratio score of evaluative language use, *r(28) = 0.235, p = .210*.

The two correlations are displayed as scatterplots in Figure 1. Correlation across all variables are presented in Table 4.

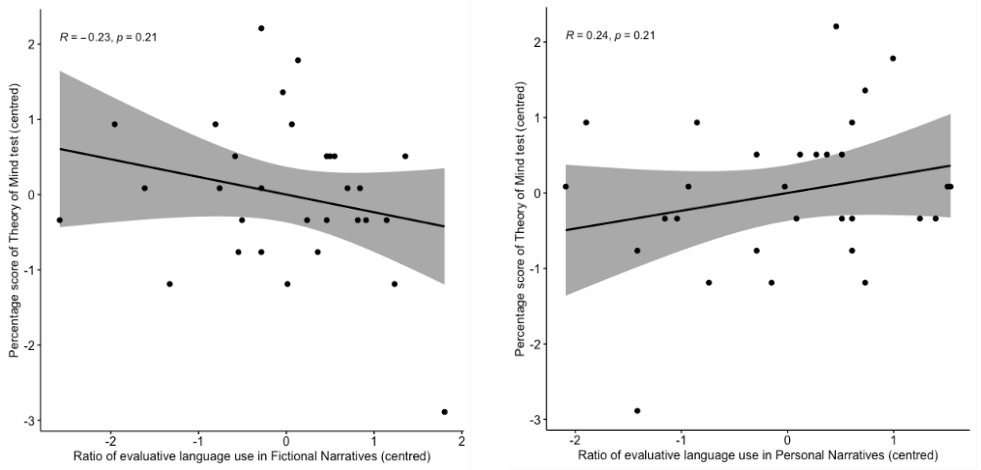

Figure 1 Scatter-plots showing the relationship between Theory of Mind scores and the ratio of evaluative language use in (left) fictional narratives and (right) personal narratives.

Variables	1	2	3
1. Percentage score of the Strange Stories task	–	-.235	.236
2. Ratio of evaluative language use in the fictional elicitation task	-.235	–	-.009
3. Ratio of evaluative language use in the personal elicitation task	.236	-.009	–

Table 4 Correlations of Study Variables

Since both correlations between theory of mind and evaluative language use were non-significant, no further regression analyses were run.

4.3 Genres differences in the relationship between theory of mind and evaluative language use

Given no significant correlations were found between theory of mind and evaluative language use in both narrative genres, the proposed Fisher's r-to-z transformation was not carried out to test whether the two correlation coefficients of fictional and personal narratives were significantly different.

5 Discussion

The present study examined the relationship between theory of mind and evaluative language use in both fictional and personal narratives produced by Cantonese-speaking adults. In addition, the study also proposed to explore whether this relationship differs between fictional versus personal narratives. The non-significant findings suggest that theory of mind is not related to evaluative language use in both narrative genres —fictional and personal narratives —in Cantonese-speaking adults. However, since no significant correlations were found between theory of mind and evaluative language use, it was unclear whether this relationship would differ significantly between fictional and personal narratives.

5.1 Theory of mind and evaluative language use in fictional narratives

Previous studies have shown that theory of mind is associated with evaluative language use in fictional narratives, notably produced by Indo-European children and adolescents (Bamberg & Damrad-Frye 1991, Fernández 2013, Riggio & Cassidy 2009). Theory of mind was conceptualised to facilitate evaluative language in inferring feelings, thoughts and intentions of story characters and events, in turn, constructing more engaging and coherent narratives (Drijbooms et al. 2017, Van Beijsterveldt & Van Hell 2009). Evaluative expressions were shown to increase rapidly over the first few years in schools (Bamberg & Damrad-Frye 1991, Drijbooms et al. 2017, Ukrainetz et al. 2005), similar to development of theory of mind

(Drijbooms, Groen & Verhoeven 2015, Wimmer & Perner 1983). Existing literature in evaluative language use focuses primarily on the development of childhood and adolescence, thereby leaving a large body of unanswered questions about adult development. Findings from the current study seem to demonstrate that the relationship between theory of mind and evaluative language use in fictional narratives does not necessarily persist into adulthood. Although replication is needed to establish the reliability of this finding, the present results are consistent with prior research reporting no significant association between theory of mind and evaluative language use in older children and adolescents (Longobardi et al. 2014, Meins et al. 2006, Ögel-Balaban & Hohenberger 2019). For example, Longobardi et al. (2014) did not find a significant correlation between theory of mind and the frequency of evaluative language use in fictional narratives in Italian-speaking 8 to 12-year-olds. This was also observed in English-speaking 7 to 9-year-olds, where task performance of theory of mind was not found to be associated with evaluative language use in Meins et al. (2006). Taking these into account, it appears the relationship between theory of mind and evaluative language use may only be salient in early childhood, a developmental stage where cognitive, sociocognitive and linguistic abilities just emerge. The relationship may potentially start to decline once it reaches a certain developmental age. With limited research conducted in adults, this warrants further research to verify the current findings.

Another implication from the current finding is that theory of mind may not necessarily relate to evaluative language use when producing fictional narratives in Cantonese specifically. Cross-cultural and cross-linguistic differences have been observed extensively in the existing narrative literature (Berman et al. 1994, Küntay & Nakamura 2004). Although the limited studies on Cantonese narratives (Ho 2007, Leung 2001) have found a similar increase in evaluative language use from 5-year-olds to adults as that observed in English samples (Bamberg & Damrad-Frye 1991), there seems to be discrepancies when examining the independent categories of evaluative language used. For example, hedges were found to be used more significantly in older English-speaking age groups (9-year-olds and adults) than 5-year-olds (Bamberg & Damrad-Frye 1991). Using the same coding criteria, Leung (2001) found that usage of causal connectors and negative qualifiers increased the most out of the five evaluative language categories when comparing 5-year-olds to 9-year-olds instead. Although the current study did not investigate the individual categories per se, results also showed that causal connectors and negative qualifiers (causal expressions, negative qualifiers and contrastive expressions in this study) occurred more frequently when compared to other categories in descriptive terms. One probable factor may be attributed to the common use of sentence-final particles in Cantonese (Cheung, Hsuan-Chih, Creed, Ng, Ping Wang & Mo 2004). Sentence-final particle is a morphosyntactic characteristic absent in English but commonly used in spoken Cantonese (Cheung et al. 2004). For example, for the sentence of 'Nick didn't take the apple', one of the ways to express this in Cantonese could be 'Nick didn't take the apple *wo3*'. *Wo3* acts as the sentence-final particle here, which implies that the speaker might find the information noteworthy, or contradictory to the previous speaker's statement. If the same word but in a different tone wo5 is

used instead, the narrator is relaying this information obtained from someone else. On other occasions, it can convey that the speaker does not necessarily believe this information is correct. Therefore, the usage of sentence-final particles might provide a subtle context to the underlying disposition of the narrators themselves in Cantonese. While this might influence Cantonese narrators to feel the need to explain themselves further in certain occasions, thus using more causal connectors, it might also result in fewer evaluative clauses observed in Cantonese narratives. Hence, a lack of relationship was found between theory of mind and evaluative language use in fictional narratives in the present study.

Indeed, it is widely acknowledged in the translation literature that one language cannot be fully equivalent to another (Nida 1969, Snell-Hornby 1988), and the cross-linguistic study by Berman et al. (1994) demonstrated exactly how each language facilitates narrating. Having examined the fictional narratives in English-, German-, Spanish-, Hebrew-, Turkish-speaking 9-year-olds, they noticed differences in rhetorical style and how the narrators position themselves in the narratives. For example, the English-speaking and German-speaking children engaged in complicated descriptions about the motion of one event, since English and German have a wider range of locative particles such as hurled and tipped. The Spanish-, Hebrew-, and Turkish-speaking children, on the other hand, depicted a simple change of state and location using verbs such as stop, and throw. Although this example is not essentially related to evaluative language use, it demonstrates how one language itself may predispose narrators in certain narrators' perspectives. Perhaps, having more locative particles facilitates narrators to engage more from the characters' perspectives, in turn, leading to greater evaluative language use. Therefore, different language systems might influence not only how often evaluative language is being applied, but also how it is being used.

Additionally, converging evidence has suggested that differences found across varying languages may be explained in regards to cross-cultural differences (Küntay & Nakamura 2004, Ögel-Balaban & Hohenberger 2019). In Hong Kong, expressing one's emotions freely is less encouraged since childhood (Soto, Perez, Kim, Lee & Minnick 2011), given how it may imply weakness and vulnerability (Yeung, Mak & Cheung 2015). On the other hand, American and most European cultures are more encouraging with expressing emotions (Soto et al. 2011). Therefore, less evaluative opinions might be expressed in the Cantonese narratives due to this inherent cultural difference, resulting in an overall reduced use of evaluative language. Some prior studies also considered cross-cultural differences in explaining the inconsistent findings in the evaluative language use literature. In Küntay & Nakamura (2004), Japanese-speaking children were found to rely more heavily on onomatopoeia than the Turkish-speaking children and English-speaking counterparts in Bamberg & Damrad-Frye (1991). Similarly, Ögel-Balaban & Hohenberger (2019) explained that emotional language is often expressed at an implicit level if the event is a common occurrence. As it might be considered to be too obvious for their audience Mesquita (2001), Turkish-speaking children were found to use their evaluative language differently when compared to the Japanese-speaking and English-speaking children (Küntay & Nakamura 2004). In view of this, it proves to be vital

to examine whether evaluative language use differs across different languages and cultures. Taken together, it appears that both cross-linguistic and cross-cultural differences might contribute to the non-significant relationship between theory of mind and evaluative language use in Cantonese fictional narratives in the present study.

5.2 Theory of mind and evaluative language use in personal narratives

There was no correlation found between the theory of mind task performance and the evaluative language use ratio in personal narratives in Cantonese speakers, suggesting that the socio-cognitive skill to infer mental states of oneself and others is not related to evaluative language use in personal narratives.

There has been a lack of research conducted to examine whether theory of mind relates to evaluative language use in terms of personal narratives in general. Given this relationship was not found in both fictional and personal narratives, another possible explanation might be attributed to the fundamental issue with the theory of mind task. Having implemented the Strange Stories task (Happé 1994) similarly as their measure of theory of mind, Meins et al. (2006) also did not find a relationship between theory of mind and evaluative language use. They proposed that this finding may arise from the fact that the theory of mind in evaluative language production is actually independent of the ability of theory of mind itself. In other words, expressing evaluative language might only indicate the frequency with which theory of mind is engaged rather than the competency of theory of mind. Moreover, forming a narrative is a complex process that requires a large amount of cognitive effort and linguistic demand (Reilly, Losh, Bellugi & Wulfeck 2004). Arranging different story events in a temporal and sequential manner involves higher-level perception-action skills (Akimoto 2018) and a range of cognitive skills, such as abstraction, non-verbal memory, and memory organisation (Akimoto 2019). Even if theory of mind underlies the ability to use evaluative language, the cognitive load of producing a narrative might prevent narrators from applying their theory of mind capacity in expressing evaluative language (Aksu-Koç & Aktan-Erciyes 2018). Thus, evaluative language production does not fully capture theory of mind ability itself, resulting in the lack of relationship found between theory of mind and evaluative language use (Meins et al. 2006, Ögel-Balaban & Hohenberger 2019).

5.3 Comparing the relationship between fictional and personal narratives: implications

Considering that non-significant results were found in both fictional and personal narratives, no further analysis was conducted. The overall findings seem to suggest that the relationship was similarly non-significant in both genres. Current literature has yet to investigate whether the relationship between theory of mind and evaluative language use would differ across genres, but the development of evaluative language has shown to be different between fictional and personal narratives (Ho 2007, Longobardi et al. 2014, Shiro 2003). Both Shiro (2003) and Ho

(2007) found that evaluative language use increased significantly across the years when producing fictional narratives whereas a lack of development was shown in personal narratives. These findings seem to suggest that evaluative language use is developed and applied differently across genres, in turn, resulting in a different relationship with theory of mind as well.

Reflecting on the present findings, although the correlation was non-significant, a negatively correlated relationship could be observed between theory of mind and evaluative language use in fictional narratives while a positive correlation was found in personal narratives. If there is a greater statistical power, these patterns might be attributable to the varying cognitive load between genres. Fictional narratives require greater cognitive effort, as it involves making up or organising an imaginary event into a goal-directed story (Ukrainetz et al. 2005). This leads to fewer cognitive resources to be used to elicit evaluative language, thereby resulting in a negative correlation with theory of mind. On the other hand, personal narratives are easier to elicit, since narrators are familiar with the events and characters (Hudson & Shapiro 1991). Therefore, more cognitive load is available to produce evaluative expression, thus, a positive correlation with theory of mind. Given the lack of power in the current study might be due to the small sample size, further research and replication with a large sample size would be essential to establish this proposition on the relationship between theory of mind and evaluative language use across narrative genres.

Converging evidence has further suggested that there may be other underlying factors of narrative discourse, which undermines the relationship between theory of mind and evaluative language use. For example, socio-economic status has been shown to influence the development of evaluative language (McConnell 2011, Shiro 2003). Shiro (2003) found that evaluative language use in fictional narratives only increased significantly from the age of 7 and 10 years in upper-middle class, but not in children from low-income families. Other studies examining socioeconomic status with respect to general language development also found poorer task performance in language assessments in children from lower socioeconomic backgrounds compared to peers from higher socioeconomic families (Ellwood-Lowe, Whitfield-Gabrieli & Bunge 2021, Noble, McCandliss & Farah 2007). Since children from high socioeconomic context have greater access to learning materials, opportunities, and social interactions with caregivers, they have more exposure to narrative practices, leading to better narrative discourse skills (Daneri, Blair, Kuhn, Vernon-Feagans, Greenberg & Mills-Koonce 2019, Hoff 2013, Hart & Risley 1995). This is contrary to children coming from lower socioeconomic backgrounds who might not only have no access to such resources and interactions, but also have to experience day-to-day stressors (Roseberry-McKibbin 2010). Therefore, the relationship between theory of mind and evaluative language use may be underscored by socioeconomic status.

Additionally, there is a general consensus that narrating events requires a set of cognitive abilities known as executive functions (Mozeiko, Le, Coelho, Krueger & Grafman 2011). Executive functions have been used as an umbrella term which encompasses a set of cognitive skills that regulates and underpins cognition and goal-

directed behaviour (Baggetta & Alexander 2016, Miller & Cohen 2001, Zelazo, Blair & Willoughby 2016). The most widely implemented framework proposes working memory, inhibition, and cognitive flexibility as the core components of executive functions (Diamond 2013, Miyake, Friedman, Emerson, Witzki, Howerter & Wager 2000). While there is a lack of studies investigating its relationship with evaluative language specifically, executive functions may serve to facilitate evaluative language production. For instance, it is necessary to hold and manipulate different pieces of information to infer feelings, thoughts, and emotions of story characters in fictional narratives and narrators themselves in personal narratives, which involves the skill of working memory. Regularly changing perspectives and monitoring communicative flow through alternating between referential and evaluative language also relies on the fundamental skill of cognitive flexibility. Inhibition may also support the use of evaluative language by neglecting irrelevant information and lexical representations to understand the feelings and thoughts behind the characters' actions in both fictional and personal narratives. Existing literature has mostly examined other aspects of narrative production such as plot, another macrostructural aspect (Balioussis, Pascual-Leone & Johnson 2012, Khan 2013), and syntax, the microstructural aspect (Drijbooms et al. 2017, Tonér & Gerholm 2021). Given the limited research, further investigation on the relationship between executive function and evaluative language use in adults would be valuable.

Monolingualism versus bilingualism might be another contributing factor (Schroeder 2018). The number of bilingual individuals has been growing rapidly over the past few decades (Tan, Marchman & Frank 2024). Although there is a limited body of studies that examined evaluative language use in narratives produced by monolinguals versus bilinguals, existing evidence demonstrates there are differences in the overall narrative structure depending on whether the narrators are monolinguals or bilinguals (Chen & Yan 2011, Fiestas & Peña 2004). For instance, Fiestas & Peña (2004) found a higher number of 'attempts' and 'initiating events' in the Spanish narratives compared to the English narratives produced by Spanish-English speaking 4 to 6-year-olds bilinguals. In contrast, they found a greater use of 'consequences' in their English narratives as opposed to the Spanish narratives. Looking at evaluative language use more closely, Chen & Yan (2011) found that Mandarin-English speaking bilinguals produced more evaluative language than both Mandarin and English-speaking monolinguals. Additionally, bilingual individuals often engage in practices of code-switching, the process of switching lexical elements from one linguistic code to another (Chan 1999), which leads to grammatical convergence (Cacoullos & Travis 2018). In the current study, narratives elicited indeed consisted of code-switching and grammatical structures that did not fully match with Cantonese but rather English. Taking these into account, bilingualism is another vital construct to address when researching language development. Overall, these factors seem to be integral to narrative development, hence, it might be useful to consider these contributing factors when understanding evaluative language use.

6 Limitations

The main limitation of the present study is its underpowered sample size. Due to time constraints, a large effect size was entered in the a-priori power analysis to determine the sample size. Larger sample sizes in future studies could clarify whether the present non-significant correlations reflect true null effects or low statistical power.

In addition, another limitation of the study is the lack of linguistic proficiency measures. Linguistic proficiency has been shown to be highly associated with evaluative language in narrative production (Berman 2008, Nippold 2004). Given evaluative language production requires complex language skills, such as acquisition and organisation of lexicon, syntax, semantics and pragmatics (Drijbooms et al. 2017), the current study could have assessed linguistic proficiency in understanding the underlying process of evaluative language. Due to time constraints, the current sample was recruited only via convenience and snowball sampling. This results in a narrowed sampled demographics. With over half of the present sample residing outside of Hong Kong and the majority completing their education in English from an early age, English might have been a more commonly spoken language for most of the participants in the present study. As a result, the linguistic proficiency in Cantonese might be lower than a sample who lives in Hong Kong, thereby undermining the relationship between theory of mind and evaluative language use in both genres of narratives.

7 Conclusion

Narrative production has been shown to hold significant educational implications, including literacy skills and reading comprehension. Therefore, identifying the underlying factors that influence specific aspects of narrative discourse, particularly evaluative language production, is crucial. The present study is one of the first few studies examining the relationship between theory of mind and evaluative language use across both narrative genres in Cantonese-speaking adults. No significant relationships were found between theory of mind and evaluative language use in both fictional and personal narratives, suggesting that theory of mind may not be related to evaluative language use in narratives elicited by Cantonese-speaking adults. Although it remains unclear whether this relationship will persist given a larger sample size, the current study provided potential evidence of cross-linguistic, cross-cultural and age differences. Further research is needed to shed light on the underlying processes of narrative production, which may ultimately inform pedagogical practices in the long run.

References

Akimoto, Taisuke. 2018. Emergentist view on generative narrative cognition: Considering principles of the self-organization of mental stories. *Advances in Human-Computer Interaction* 2018(1). 6780564.

Akimoto, Taisuke. 2019. Key issues for generative narrative cognition in a cognitive system: Association and blending of stories. In *Story-enabled intelligence, aaai 2019 spring symposium*, .

Aksu-Koç, Ayhan & Aslı Aktan-Erciyes. 2018. Narrative discourse: Developmental perspectives. In *Handbook of communication disorders*, 329–356.

Allen, Melissa S., Mary K. Kertoy, John C. Sherblom & Jill M. Pettit. 1994. Children's narrative productions: A comparison of personal event and fictional stories. *Applied Psycholinguistics* 15(2). 149–176.

Apperly, Ian A., Frances Warren, Ben J. Andrews, Joanne Grant & Sarah Todd. 2011. Developmental continuity in theory of mind: Speed and accuracy of belief–desire reasoning in children and adults. *Child Development* 82(5). 1691–1703.

Astington, Janet Wilde. 2014. Narrative and the child's theory of mind. In *Narrative thought and narrative language*, 151–171. Psychology Press.

Baggetta, Peter & Patricia A. Alexander. 2016. Conceptualization and operationalization of executive function. *Mind, Brain, and Education* 10(1). 10–33.

Balioussis, Christos, Juan Pascual-Leone & Janice Johnson. 2012. Fluency and complexity in children's writing: The role of mental attention and executive function. *Rivista di psicolinguistica applicata* 12(3). 33–46.

Bamberg, Michael & Rachel Damrad-Frye. 1991. On the ability to provide evaluative comments: Further explorations of children's narrative competencies. *Journal of Child Language* 18(3). 689–710.

Baron-Cohen, Simon, Sally Wheelwright, Jacqueline Hill, Yogini Raste & Ian Plumb. 2001. The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 42(2). 241–251.

Berman, Ruth A. 2008. Narrative development in multilingual contexts: A cross-linguistic perspective. In *Narrative development in a multilingual context*, 419–428. John Benjamins.

Berman, Ruth A., Dan I. Slobin, Ayhan A. Aksu-Koç, Michael Bamberg, Lisa Dasinger, Virginia Marchman, Yaffa Neeman, Philip C. Rodkin & Eugenia Sebastián. 1994. *Relating events in narrative: A crosslinguistic developmental study*. Lawrence Erlbaum Associates.

Bishop, Dorothy V. M. & Angela Edmundson. 1987. Specific language impairment as a maturational lag: Evidence from longitudinal data on language and motor development. *Developmental Medicine & Child Neurology* 29(4). 442–459.

Cacoullos, Rena Torres & Catherine E. Travis. 2018. *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge University Press.

Census and Statistics Department. 2022. 2021 population census. Hong Kong SAR Government.

Chafe, Wallace. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. University of Chicago Press.

Chan, Brian H. S. 1999. *Aspects of the Syntax, the Pragmatics and the Production of Code-switching: With Special Reference to Cantonese and English*. Peter Lang.

Chang, Chien-Ju & Allyssa McCabe. 2013. Evaluation in Mandarin Chinese children's personal narratives. In Allyssa McCabe & Chien-Ju Chang (eds.), *Chinese*

language narration: Culture, cognition, and emotion, 33–56. John Benjamins. doi:10.1075/sin.19.04cha.

Chen, Lieven & Rebecca Yan. 2011. Development and use of English evaluative expressions in narratives of Chinese–English bilinguals. *Bilingualism: Language and Cognition* 14(4). 570–578.

Cheung, Hoi, Chen Hsuan-Chih, Nicolette Creed, Lisa Ng, Shirley Ping Wang & Lei Mo. 2004. Relative roles of general and complementation language in theory-of-mind development: Evidence from Cantonese and English. *Child Development* 75(4). 1155–1170.

Chow, James C. & Melody Jacobs. 2016. The role of language in fraction performance: A synthesis of literature. *Learning and Individual Differences* 47. 252–257.

Daneri, M. P., Clancy Blair, Lisa J. Kuhn, Lynne Vernon-Feagans, Mark Greenberg & W. Roger Mills-Koonce. 2019. Maternal language and child vocabulary mediate relations between socioeconomic status and executive function during early childhood. *Child Development* 90(6). 2001–2018.

Diamond, Adele. 2013. Executive functions. *Annual Review of Psychology* 64(1). 135–168.

Drijbooms, Eva, Margriet A. Groen & Ludo Verhoeven. 2015. The contribution of executive functions to narrative writing in fourth grade children. *Reading and Writing* 28. 989–1011.

Drijbooms, Eva, Margriet A. Groen & Ludo Verhoeven. 2017. How executive functions predict development in syntactic complexity of narrative writing in the upper elementary grades. *Reading and Writing* 30(1). 209–231.

Eaton, Jane H., Gillian M. Collis & Vicky A. Lewis. 1999. Evaluative explanations in children's narratives of a video sequence without dialogue. *Journal of Child Language* 26(3). 699–720.

Eisenberg, Ann R. 1985. Learning to describe past experiences in conversation. *Discourse Processes* 8(2). 177–204.

Ellwood-Lowe, M. E., Susan Whitfield-Gabrieli & Silvia A. Bunge. 2021. Brain network coupling associated with cognitive performance varies as a function of a child's environment in the ABCD study. *Nature Communications* 12(1). 7183.

Engel, Susan. 2005. The narrative worlds of what is and what if. *Cognitive Development* 20(4). 514–525.

Fernández, Cristina. 2013. Mindful storytellers: Emerging pragmatics and theory of mind development. *First Language* 33(1). 20–46.

Fiestas, Christine E. & Elizabeth D. Peña. 2004. Narrative discourse in bilingual children.

Fivush, Robyn, Janet T. Gray & Faith A. Fromhoff. 1987. Two-year-old talk about the past. *Cognitive Development* 2(4). 393–409.

Fox, Richard. 1991. Developing awareness of mind reflected in children's narrative writing. *British Journal of Developmental Psychology* 9(2). 281–298.

Frith, Chris D. & Uta Frith. 2012. Mechanisms of social cognition. *Annual Review of Psychology* 63(1). 287–313.

Fung, E. 2012. *Performance of 5-to 8-year-old typically developing children and children with autism spectrum disorder in a Chinese version of the Happés' strange stories*. University of Hong Kong MA thesis. HKU Theses Online (HKUTO).

Gagarina, Natalia, Daleen Klop, Sari Kunnari, Konstadina Tantele, Taina Välimaa, Ute Bohnacker & Joel Walters. 2019. MAIN: Multilingual assessment instrument for narratives—revised. *ZAS Papers in Linguistics* 63. 20–20.

Gagarina, Natalia V., Daleen Klop, Sari Kunnari, Konstadina Tantele, Taina Välimaa, Ingrida Balčiūnienė & Joel Walters. 2012. MAIN: Multilingual assessment instrument for narratives. *ZAS Papers in Linguistics* 56. 155–155.

Galinsky, Adam D., Gillian Ku & Cynthia S. Wang. 2005. Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes & Intergroup Relations* 8(2). 109–124.

Gallagher, Shaun. 2014. Self and narrative. In *The routledge companion to hermeneutics*, 403–414. Routledge.

Gallese, Vittorio & Corrado Sinigaglia. 2011. What is so special about embodied simulation? *Trends in Cognitive Sciences* 15(11). 512–519.

Griffin, Timothy M., Linda Hemphill, Lynne Camp & Dennie P. Wolf. 2004. Oral discourse in the preschool years and later literacy skills. *First Language* 24(2). 123–147.

Guajardo, N. R. & A. C. Watson. 2002. Narrative discourse and theory of mind development. *The Journal of Genetic Psychology* 163(3). 305–325.

Hair, J. F., W. C. Black, B. J. Babin & R. E. Anderson. 2010. *Multivariate data analysis: A global perspective*. Prentice Hall.

Happé, Francesca G. 1994. An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders* 24(2). 129–154.

Hart, Betty & Todd R. Risley. 1995. *Meaningful differences in the everyday experience of young american children*. Baltimore, MD: Brookes Publishing.

Heath, Shirley Brice. 1986. Taking a cross-cultural look at narratives. *Topics in Language Disorders* 7(1). 84–94.

Heilmann, John, Jon F. Miller, Ann Nockerts & Carrie Dunaway. 2010. Properties of the narrative scoring scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology* 19(2). 154–166.

Ho, K. Y. 2007. *Children's use of evaluative devices in telling fictional and personal narratives*. University of Hong Kong MA thesis. HKU Theses Online (HKUTO).

Hoff, Erika. 2013. Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology* 49(1). 4.

Hudson, Judith A. & Lauren R. Shapiro. 1991. From knowing to telling: The development of children's scripts, stories, and personal narratives. In *Knowing and remembering in young children*, .

Hughes, D. L., L. McGillivray & M. Schmidek. 1997. *Guide to narrative language: Procedures for assessment*.

Hutto, Daniel D. 2008. Limited engagements and narrative extensions. *International Journal of Philosophical Studies* 16(3). 419–444.

Jacobs, Arthur M. 2015. Towards a neurocognitive poetics model of literary reading. In *Cognitive neuroscience of natural language use*, 135–159.

Khan, Amjad. 2013. *Pakistan and the narratives of extremism.* United States Institute of Peace.

Küntay, Aylin C. & Keiko Nakamura. 2004. Linguistic Strategies Serving Evaluative Functions: A Comparison between Japanese and Turkish Narratives. In *Relating events in narrative, volume 2*, 329–358. Psychology Press.

Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* 1(1). 97–120.

Labov, William & Joshua Waletzky. 1967. Narrative analysis: Oral versions of personal experiences. *Journal of Narrative and Life Experiences* 7(1–4). 3–38.

Labov, William & Joshua Waletzky. 1997. Narrative analysis: Oral versions of personal experience.

Leung, W. F. 2001. *Evaluative comments in narratives of Cantonese speaking children.* University of Hong Kong MA thesis. HKU Theses Online (HKUTO).

Licandro, Ute. 2016. Fictional narratives in preschool-aged DLLs–collection, analysis, and current state of research. In *Narrative skills of dual language learners: Acquisition and peer-assisted support in early childhood education and care*, 25–52. Wiesbaden: Springer Fachmedien Wiesbaden.

Loban, Walter. 1976. *Language development: Kindergarten through grade twelve.* NCTE Committee on Research. Report No. 18.

Longobardi, Emiddia, Paola Spataro, Monica Renna & Clelia Rossi-Arnaud. 2014. Comparing fictional, personal, and hypothetical narratives in primary school: Story grammar and mental state language. *European Journal of Psychology of Education* 29. 257–275.

Losh, Molly & Lisa Capps. 2003. Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of Autism and Developmental Disorders* 33(3). 239–251.

Lucero, Audrey. 2015. Cross-linguistic lexical, grammatical, and discourse performance on oral narrative retells among young Spanish speakers. *Child Development* 86(5). 1419–1433.

Luo, Jing, Victor G. Li & Peggy P. K. Mok. 2020. The perception of Cantonese vowel length contrast by Mandarin speakers. *Language and Speech* 63(3). 635–659.

McCabe, Allyssa, Lynn Bliss, Graciela Barra & Margaret Bennett. 2008. Comparison of personal versus fictional narratives of children with language impairment.

McCabe, Allyssa & Pamela R. Rollins. 1994. Assessment of preschool narrative skills. *American Journal of Speech-Language Pathology* 3(1). 45–56.

McConnell, G. E. 2011. *Story presentation effects on the narratives of preschool children from low and middle socioeconomic homes*: University of Kansas dissertation. Doctoral dissertation.

Meins, Elizabeth, Charles Fernyhough, Fiona Johnson & Janette Lidstone. 2006. Mind-mindedness in children: Individual differences in internal-state talk in middle childhood. *British Journal of Developmental Psychology* 24(1). 181–196.

Mesquita, Batja. 2001. Emotions in collectivist and individualist contexts. *Journal of Personality and Social Psychology* 80(1). 68.

Miller, Earl K. & Jonathan D. Cohen. 2001. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 24(1). 167–202.

Miyake, Akira, Naomi P. Friedman, Michael J. Emerson, Alexander H. Witzki, Amy Howerter & Tor D. Wager. 2000. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology* 41(1). 49–100.

Mozeiko, Jennifer, K. Le, Carl Coelho, Frank Krueger & Jordan Grafman. 2011. The relationship of story grammar and executive function following TBI. *Aphasiology* 25(6-7). 826–835.

Nicolopoulou, Ageliki & B. Ünlütabak. 2017. Narrativity and mindreading revisited: Children's understanding of theory of mind in a storybook and in standard false belief tasks. In *Social environment and cognition in language development: Studies in honor of ayhan aksu-koç*, 151–166. John Benjamins Publishing Company.

Nida, Eugene A. 1969. Science of translation. *Language* 483–498.

Nippold, Marilyn A. 2004. Research on later language development: International perspectives. *Language Development Across Childhood and Adolescence* 3(1). 1–8.

Noble, Kimberly G., Bruce D. McCandliss & Martha J. Farah. 2007. Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science* 10(4). 464–480.

Ögel-Balaban, H. & Annette Hohenberger. 2019. The development of narrative skills in Turkish-speaking children: A complexity approach. *PloS One* 15(5). e0232579.

O'Neill, Daniela K., Michael J. Pearce & Jessica L. Pick. 2004. Preschool children's narratives and performance on the peabody individualized achievement test–revised: Evidence of a relation between early narrative and later mathematical ability. *First Language* 24(2). 149–183.

Otwinowska, Agnieszka, Michał Opacki, Katarzyna Mieszkowska, Marta Białecka-Pikul, Zofia Wodniecka & Ewa Haman. 2022. Polish–English bilingual children overuse referential markers: MLU inflation in Polish-language narratives. *First Language* 42(2). 191–215.

Pelletier, Janette & Janet Astington. 2004. Action, consciousness and theory of mind: Children's ability to coordinate story characters' actions and thoughts. *Early Education and Development* 15(1). 5–22.

Pesco, Diane & Annie Gagné. 2017. Scaffolding narrative skills: A meta-analysis of instruction in early childhood settings. *Early Education and Development* 28(7). 773–793.

Peterson, Carole & Allyssa McCabe. 1983. *Developmental psycholinguistics: Three ways of looking at a child's narrative*. Springer Science & Business Media.

Preece, Alison. 1987. The range of narrative forms conversationally produced by young children. *Journal of Child Language* 14(2). 353–373.

Reese, Elaine, Sebastian Suggate, Jan Long & Elizabeth Schaughency. 2010. Children's oral narrative and reading skills in the first 3 years of reading instruction.

Reading and Writing 23. 627–644.

Reilly, Judy, Molly Losh, Ursula Bellugi & Beverly Wulfeck. 2004. "frog, where are you?" narratives in children with specific language impairment, early focal brain injury, and williams syndrome. *Brain and Language* 88(2). 229–247.

Riggio, M. M. & Kimberly W. Cassidy. 2009. Preschoolers' processing of false beliefs within the context of picture book reading. *Early Education and Development* 20(6). 992–1015.

Roseberry-McKibbin, Celeste. 2010. *Increasing language skills of students from low-income backgrounds: Practical strategies for professionals.* Plural Publishing.

Schroeder, Scott R. 2018. Do bilinguals have an advantage in theory of mind? a meta-analysis. *Frontiers in Communication* 3. 36.

Selcuk, Bilge, S. Gonultas & M. Ekerim-Akbulut. 2023. Development and use of theory of mind in social and cultural context. *Child Development Perspectives* 17(1). 39–45.

Shiro, Martha. 2003. Genre and evaluation in narrative development. *Journal of Child Language* 30(1). 165–195.

Snell-Hornby, Mary. 1988. *Translation studies: An integrated approach.* John Benjamins Publishing.

Soto, Jose A., Christopher R. Perez, Y. H. Kim, E. A. Lee & M. R. Minnick. 2011. Is expressive suppression always associated with poorer psychological functioning? a cross-cultural comparison between european americans and hong kong chinese. *Emotion* 11(6). 1450.

Stein, Nancy L. & C. G. Glenn. 1975. An analysis of story comprehension in elementary school children: A test of a schema.

Suggate, Sebastian, Elizabeth Schaughency, Heather McAnally & Elaine Reese. 2018. From infancy to adolescence: The longitudinal links between vocabulary, early literacy skills, oral narrative, and reading comprehension. *Cognitive Development* 47. 82–95.

Tan, A. W., Virginia A. Marchman & Michael C. Frank. 2024. The role of translation equivalents in bilingual word learning. *Developmental Science* 27(4). e13476.

Tompkins, Virginia, Michael J. Farrar & D. E. Montgomery. 2019. Speaking your mind: Language and narrative in young children's theory of mind development. In *Advances in child development and behavior*, vol. 56, 109–140.

Tonér, Sara & Tomas N. Gerholm. 2021. Links between language and executive functions in swedish preschool children: A pilot study. *Applied Psycholinguistics* 42(1). 207–241.

Ukrainetz, Teresa A., Laura M. Justice, Joan N. Kaderavek, Sarita L. Eisenberg, Ronald B. Gillam & Heidi M. Harm. 2005. The development of expressive elaboration in fictional narratives. *Journal of Speech, Language, and Hearing Research* 48(6). 1363–1377.

Van Beijsterveldt, L. M. & Janet G. Van Hell. 2009. Evaluative expression in deaf children's written narratives. *International Journal of Language & Communication Disorders* 44(5). 675–692.

Wellman, Henry M. 2011. Developing a theory of mind. In *The wiley-blackwell handbook of childhood cognitive development*, 258–284. 2nd edn.

Westerveld, Marleen F., Nancy W. Nelson, M. Claessen & Carol Westby. 2023. Children's use of evaluative devices in response to the global TALES protocol. *Folia Phoniatrica et Logopaedica* 75(6). 480–494.

White, Sarah, Elisabeth Hill, Francesca Happé & Uta Frith. 2009. Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development* 80(4). 1097–1117.

Whitworth, Anne, M. Claessen, Suze Leitão & Joan Webster. 2015. Beyond narrative: Is there an implicit structure to the way in which adults organise their discourse? *Clinical Linguistics & Phonetics* 29(6). 455–481.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D. A. McGowan, Romain François & Hiroaki Yutani. 2019. Welcome to the tidyverse. *Journal of Open Source Software* 4(43). 1686.

Wimmer, Heinz & Josef Perner. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13(1). 103–128.

Wolf, D., J. Moreton & L. Camp. 1994. Children's acquisition of different kinds of narrative discourse: Genres and lines of talk. In *Handbook of research in language development using childes*, 286–323.

Wong, A. W. K. & H. C. Chen. 2009. What are effective phonological units in Cantonese spoken word planning? *Psychonomic Bulletin & Review* 16(5). 888–892.

Yeung, N. C., W. W. Mak & L. K. Cheung. 2015. Conformity to the emotional-control masculine norm and psychological well-being among Chinese men in Hong Kong: The mediating role of stress appraisal for expressing tender emotions. *Psychology of Men & Masculinity* 16(3). 304.

Zelazo, Philip D., Clancy B. Blair & Michael T. Willoughby. 2016. Executive function: Implications for education.

Esther Yip
The University of Cambridge
ewcy2@cam.ac.uk