

Research Article

Exploring Personal Narrative Coherence in 10-Year-Old Children: A Global Study Using the Global TALES Protocol

Marleen F. Westerveld,^a Stephanie Malone,^b Carol Westby,^c Cibelle Amato,^d Josephine Ohenewa Bampoe,^e Angel Chan,^f Kai-Mei Chen,^g Maria Chrysostomou,^h Mary Claessen,ⁱ Jóhanna Einarsdóttir,^j Sara Ferman,^k Fernanda Dreux Fernandes,^l Pamela Filiatrault-Veilleux,^m Mateja Gabaj,ⁿ Eleftheria Geronikou,^o Vani Gupta,^p Kristine Jensen de López,^q Khaloob Kavar,^r Jelena Kuvač Kraljević,^s HaeJi Lee,^t Jisun Lee,^u Rena Lyons,^v İlknur Maviş,^w Nadia Nørgaard,^q Mary-Pat O'Malley Keighran,^x Laura Quintanilla,^y Semra Selvi Balı,^z Elena Theodorou,^{aa} Tze Peng Wong,^{bb} Anita Mei-Yin Wong,^{cc} Esra Yasar Gunduz,^{dd} Wanlin Zhu,^f and Nickola Nelson^{ee}

^aGriffith Institute for Educational Research, Griffith University, Southport, Queensland, Australia ^bAutism Centre of Excellence, Griffith University, Brisbane, Queensland, Australia ^cBilingual Multicultural Services, Albuquerque, NM ^dUniversidade Presbiteriana Mackenzie, São Paulo, Brazil ^eDepartment of Audiology, Speech and Language Therapy, University of Ghana, Accra ^fThe Hong Kong Polytechnic University, Hong Kong SAR, China ^gChung Shan Medical University, Taichung, Taiwan ^hCyprus University of Technology, Limassol ⁱSpeech Pathology Australia, Melbourne, Victoria, Australia ^jUniversity of Iceland, Reykjavik ^kDepartment of Communication Disorders, Faculty of Medicine, Tel Aviv University, Israel ^lSchool of Medicine, University of São Paulo, Brazil ^mDepartment of Communication Sciences and Disorders, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, Canada ⁿPostgraduate doctoral study "Speech, Language and Hearing Disorders," University of Zagreb, Faculty of Education and Rehabilitation Sciences, Croatia ^oDepartment of Speech and Language Therapy, University of Patras, Greece ^pSchool of Allied Health Sciences, Sport and Social Work, Griffith University, Gold Coast, Queensland, Australia ^qInstitute of Communication & Psychology, Aalborg University, Denmark ^rBeit Berl College, Kfar Saba, Israel ^sUniversity of Zagreb Faculty of Education and Rehabilitation Sciences, Croatia ^tUniversity of Minnesota, Twin Cities, Minneapolis ^uMore Than Words, Albuquerque, NM ^vSchool of Health Sciences, College of Medicine, Nursing & Health Sciences, University of Galway, Ireland ^wAssociation for Turkish SLPs (DKTD), Eskişehir, Türkiye ^xUniversity of Galway, Ireland ^yUniversidad Nacional de Educación a Distancia (UNED), Madrid, Spain ^zDepartment of Speech and Language Therapy, Anadolu University, Eskişehir, Türkiye ^{aa}Department of Rehabilitation Sciences, Cyprus University of Technology, Limassol ^{bb}School of Education, University of Nottingham Malaysia, Semenyih ^{cc}The University of Sydney, New South Wales, Australia ^{dd}Department of Speech and Language Therapy, School of Health Sciences, Cappadocia University, Nevşehir, Türkiye ^{ee}Western Michigan University, Kalamazoo

ARTICLE INFO

Article History:

Received August 1, 2025

Revision received December 12, 2025

Accepted March 5, 2026

Editor-in-Chief: Karla N. Washington

Editor: Natalia Gagarina

https://doi.org/10.1044/2026_LSHSS-25-00143

ABSTRACT

Purpose: The overall aim of this exploratory study was to evaluate personal narrative coherence in groups of children representing a range of cultural and/or linguistic backgrounds. Based on previous research, we expected to find effects based on culture/language participant group and the emotional valence of the prompt (positive, negative, neutral). We also anticipated potential interactions between culture/language participant group and prompt type.

Method: Four hundred seventeen 10-year-old children from 21 culture/language participant groups took part in this cross-sectional study. The children were from 19 countries or regions, speaking 18 different languages or dialects. All children were seen individually, either face-to-face or online, and asked to produce personal narratives in response to the Global Talking About Lived Experiences in Stories (TALES) protocol. Three of their narratives, each representing a different emotional valence—happy, angry, and a problem situation—were included in the current study. All narratives were coded for coherence using the multidimensional

Correspondence to Marleen F. Westerveld: m.westerveld@griffith.edu.au. **Publisher Note:** This article is part of the Forum: Global Assessment of Narratives. **Disclosure:** Marleen F. Westerveld declares that she has a consultancy agreement with SALT Software, LLC, for which she receives compensation. However, SALT Software provided no direct funding for the current project and had no direct influence on this work. All other authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

Narrative Coherence Coding Scheme (NaCCS), which yields a total coherence score and three coherence dimension subscores: context, chronology, and theme. Linear mixed models were used to identify differences in narrative coherence between culture/language participant groups and narrative prompt types.

Results: We found significant main effects for both participant group and narrative prompt type on total coherence scores, as well as a significant interaction effect. Overall, the problem prompt elicited more coherent narratives than the angry prompt. Performance varied notably across culture/language participant groups, with no consistent pattern of strength linked to prompt type. Additionally, some culture/language participant groups scored significantly lower on total coherence than others. Significant participant group effects were also observed across all three coherence dimension subscores, with the theme dimension showing the greatest number of between-groups differences.

Conclusions: In an increasingly multicultural and plurilingual world, the need for ecologically valid tools to assess children's spoken language skills is more critical than ever. Findings from this exploratory study offer further support for the use of the Global TALES protocol for eliciting personal narratives in children from diverse cultural and/or linguistic backgrounds. Furthermore, results indicated that the NaCCS is a viable method for describing differences and similarities in personal narrative coherence among these diverse groups of children. Our next steps involve refining the coding framework and expanding our participant pool to include larger, more age- and ability-representative samples—with the ultimate aim of integrating personal narrative assessment and analysis into routine clinical practice worldwide.

Supplemental Material: <https://doi.org/10.23641/asha.32653314>

Personal narrative coherence relates to an individual's ability to produce well-structured stories that make sense to a naive listener about personally experienced past events (Reese et al., 2011). Producing coherent personal narratives plays a prominent role in everyday communication (Preece, 1987), at home, at school, and in the community. Children share personal narratives for a variety of communicative purposes, including connecting with friends and family; advocating for themselves; and describing, interpreting, and processing past experiences and traumatic events (Michelson, 2023; Mossige et al., 2005; Peterson & McCabe, 1983; Vanaken et al., 2022).

Production of *coherent* personal narratives draws on a range of language and cognitive skills (Gabaj et al., 2025a). It involves retrieving a specific experience from memory, formulating sentences grammatically, and using appropriate vocabulary to provide enough detail about when and where the event happened, with sentences ordered into a logical sequence so the story is understandable to the listener, while conveying thematically what the experience meant to the individual. Our focus on coherence in the current article arises from research linking the importance of personal narrative coherence to positive developmental outcomes, including mental well-being and maintaining peer relationships, from early childhood into adolescence and beyond (Chen et al., 2012; Davidson et al., 2017; Mitchell et al., 2020; Vanaken & Hermans, 2021). Considering the importance of personal narrative

coherence to everyday functioning and well-being, personal narrative discourse offers an ecologically valid sampling context when appraising children's spontaneous language skills, which may be useful for speech-language pathology assessment and progress-monitoring purposes.

The growing number of culturally and/or linguistically diverse children and adolescents, associated with rising global mobility (Piccardo, 2018), presents a challenge for speech-language pathologists globally. Difficulty distinguishing between language differences (due to cultural and/or linguistic variations) and language disorders (due to a neurodevelopmental disorder learning [any] language) can result in either over- or underidentification of children or adolescents with developmental language disorder (e.g., see Bedore & Peña, 2008; Buac & Jarzynski, 2022). Misidentification, in turn, can result in inappropriate intervention for children who do not need it or failure to provide intervention in a timely fashion for those who do need it, with significant long-term negative socio-emotional and academic implications (McGregor, 2020). This complexity underscores the need for ecologically valid, evidence-based assessment tools that are appropriate for children from culturally and/or linguistically diverse backgrounds and capable of distinguishing difference from disorder.

To address this need, the Global TALES (Talking About Lived Experiences in Stories) Network, an international team of child language researchers, set out to develop

the Global TALES protocol. It is a standard personal narrative elicitation protocol aimed at promoting cross-cultural and cross-linguistic understanding of children's personal narrative abilities. The Global TALES protocol includes six emotion- and situation-based prompts. Children are asked to "tell me a story about a time when ...": you felt happy or excited, you felt worried or confused, you felt annoyed or angry, you felt proud of yourself, you had a problem and had to fix it, and something happened that was very important to you. To avoid cultural bias, the child is not provided with pictures or a model story. For further details, please visit <https://osf.io/ztqg6/>.

Evidence of initial feasibility of the Global TALES protocol was reported in a pilot study, conducted with 249 children from 10 different countries, speaking eight different languages (Westerveld et al., 2022). We limited the age range to 9- to 10-year-old children in the initial feasibility study because this age group matches children participating in the global Progress in International Literacy Study (Mullis & Martin, 2019), and we wished to control the number of variables when piloting the tool across multiple culture/language groups. That study showed that the protocol could be successfully translated into different languages and was effective in eliciting comparable discourse samples from all participating children. Following the pilot investigation, the protocol has been successfully used with 5- to 16-year-old children from 15 different countries (Westerveld & Nelson, 2023), demonstrating that it is useful in capturing developmental advances from 7 to 13 years of age (Kuvač Kraljević et al., 2023), can be administered face-to-face or via an online platform (Ferman & Kavar, 2023), and differentiates between 9- and 11-year-old children with and without developmental language disorder (Gabaj et al., 2025b). The current study builds on that body of work by including groups of children from a wider range of language and/or cultural backgrounds and investigating group-level similarities and differences by measuring three dimensions of coherence of personal narratives produced in response to the Global TALES protocol.

In prior research, evaluating personal narrative coherence has been conducted in a range of ways. Of the 211 studies identified in a recent systematic review (Gupta et al., in press), 29 studies evaluated "coherence." The most frequently used method (13 peer-reviewed studies) was the theoretically motivated Narrative Coherence Coding Scheme (NaCCS), developed by Reese et al. (2011; e.g., Bauer et al., 2016; Chen et al., 2012), or an adapted version of it (e.g., Foldager et al., 2024). In the NaCCS, Reese et al. (2011) conceptualized analysis of personal narrative coherence as encompassing three key dimensions: context, chronology, and theme. *Context* helps the listener interpret where and when the event took place. *Chronology* refers to the temporal ordering of the events in the narrative, often indicated

through linguistic markers. *Theme* refers to the degree to which the narrative stays on topic, contains enough development for the emotional point to be clear, and includes a resolution accompanied by affective and evaluative information.

In validating the coding scheme, Reese et al. (2011) analyzed personal narratives produced by 498 English-speaking participants aged 3 years to adulthood. Their results showed that the NaCCS could be used with narratives produced by children from preschool into adolescence and also revealed developmental changes in narrative coherence across all three coherence dimensions (Reese et al., 2011). The study showed further that *chronological* coherence develops from preschool into young adulthood, with maximum scores observed in young adulthood, whereas *contextual* coherence develops more slowly, peaking a little later than chronological coherence, and, finally, that *thematic* coherence develops the slowest, showing continued development throughout the lifespan.

Cultural and/or Linguistic Variations in Personal Narrative Ability

A broad goal of the Global TALES research effort has been to develop and test a common elicitation protocol to investigate cultural/linguistic variation in personal narrative production. It is well known that sharing of personal narratives is universal (Wang, 2016) and that personal narrative coherence is valued across cultures (Bliss & McCabe, 2008; McCabe & Chang, 2013). Degrees and mechanisms of cultural variation are less well known, however.

In line with social constructivist theories (Vygotsky, 1978), children's personal narrative development is influenced across cultures by how their parents and other community members engage in narrative conversations with them. For example, research comparing cultures that are known for valuing independence (e.g., Australia, the United States) to cultures that are known for valuing interdependence (e.g., East Asia, China) has shown that mothers from cultures valuing independence used a high-elaborative, independence-oriented style, whereas mothers from cultures valuing interdependence used a low-elaborative style that focused on social hierarchy and social norms (e.g., Wang & Fivush, 2005). It is hypothesized that these cultural differences in maternal discourse can explain associated variations in how much children talk (Minami & McCabe, 1991; Westerveld et al., 2022) and in the way personal narratives are structured (Kupersmitt, 2015; Lai, 2020), including what children choose to talk about (Srivastava et al., 2023; Westby et al., 2023). Consistent with this line of research, better understanding of how children's cultural background may influence their personal narrative coherence is important to ensure that personal narrative *differences* (associated with cultural expectations and norms)

are not mistaken for personal narrative *difficulties* (associated with language disorders).

Apart from cultural influences, language-specific linguistic conventions may influence which linguistic devices and structures children rely on for conveying personal narratives (Berman & Slobin, 1994) and how they might affect dimensions of coherence. Examples include language differences in grammatical morphemes, interclausal connectives, and syntactic structures, along with lexical items to denote temporality, manner, and causation, all of which could influence chronological coherence. For example, research on written narratives showed that language-specific grammatical conventions may influence narrative production for expressing temporality (Kupersmitt, 2015). Another example, which could affect the dimension of context and counts of “complete” utterances, is that some languages (e.g., Spanish, Italian, Japanese) consistently allow the subject pronoun to be omitted (pro-drop) because the verb morphology encodes the subject (compare Spanish “voy al mercado” vs. English “I go to the market”). This results in shorter sentences, potentially affecting the classification of utterances as “complete and intelligible,” which could influence the overall utterance count as well as coding of the chronology dimension of the NaCCS. Another example of language-specific influences on word/morpheme counts is that, in English, verbs undergo morphological changes through verbal inflections to mark tense (past, present, or future), whereas Chinese languages such as Mandarin and Cantonese do not grammaticalize tense through verbal morphology. Instead, temporal information in Chinese, also relevant to the dimension of chronology, is conveyed through aspect markers, time adverbials, and contextual cues, adding to word counts, as well as potentially affecting total coherence scores. Languages also may differ in their use of vocabulary, reflecting unique concepts or cultural habits with words that cannot be directly translated into other languages (e.g., *hygge* in Danish or *gezellig* in Dutch). Since the dimension of theme for the NaCCS (Reese et al., 2011), as applied in the current research, involves observations of causal linkages, elaborations of actions, and personal evaluations, it stands to reason that language-specific and cultural-specific features could influence children’s performance on any or all of the three dimensions of coherence coding and the total coherence score.

Differences in performance across languages and/or cultures also may arise from lexical variations in the prompts that are used to elicit children’s personal narratives. To illustrate, research into children’s understanding of emotion words (e.g., *ashamed*, *guilty*, *disgusted*, *angry*) showed significant differences between Chinese and American 6-year-old children, when asked to “tell me one time when you felt really” The results showed that only 55% of American children understood the word *ashamed*

compared to 96% of Chinese children (Wang & Leichtman, 2000), whereas 69% of American children understood the word *guilty* compared to only 15% of the Chinese children. Interestingly, no effect of culture was found on the amount of talking (narrative volume). In our previous research, we found some language-specific challenges (Westerveld et al., 2022) when translating verbal prompts. For example, in Brazil, the term for *annoyed* was replaced with *angry*; in Iceland, children did not readily understand the word *worried*. In Korean, there was a word that was similar to *annoyed*, but one would be annoyed with a situation, not a person—therefore, the suggested subprompt (included in the Global TALES protocol) for *annoyed* was not appropriate for Korean students (i.e., a time you were annoyed with a sibling).

Studies investigating the effects of emotional valence on children’s personal narrative performance have highlighted how the valence of the prompt type (i.e., *positive*, *negative*, or *neutral*) might influence children’s personal narrative coherence. Fivush et al. (2008) argued that negative events may create “a problem to be solved” (p. 590), providing a structure to the story and prompting the child to more clearly search for meaning and thereby potentially enhancing the coherence theme score in particular. This may explain why children with typical development have been found to produce better structured, more coherent stories (as scored with the NaCCS) in response to negative prompts (e.g., *scary*, *angry*) than to positive prompts (e.g., *happy*; Fivush et al., 2008; Kuvač Kraljević et al., 2023). Kuvač Kraljević et al. (2023) also found that personal narratives in response to negative and neutral prompts were chronologically better ordered and thematically more elaborated than narratives in response to positive prompts. To the authors’ knowledge, whether these effects extend to children from a range of cultural and/or language backgrounds has not yet been investigated.

Personal Narrative Coherence: Cross-Cultural and/or Cross-Linguistic Research

Single culture/language studies of personal narrative development are fairly common. A recent systematic review identified 211 peer-reviewed papers and theses (not limited to date of publication) focused on personal narratives of children aged between 4 and 18 years (Gupta et al., in press). The majority of studies (~70%) reported on English-speaking participants (although the language used was not explicitly reported in ~35% of studies); 8% reported on Spanish-speaking children, and 4% involved Mandarin-speaking children. Other languages (< 4 papers per language) included Arabic, Croatian, Danish, Dutch, Estonian, Filipino, French, German, Greek, Hebrew, Hindi, Icelandic, Indonesian, Italian, Japanese, Norwegian, Russian, Swedish,

Thai, Turkish, Ukrainian, and Urdu. Although 35% of studies reported recruiting a “culturally diverse” group to constitute their sample, only 4% (nine published studies) reported recruiting explicitly different cultural groups for comparison purposes. As reported by Gupta et al., major difficulties in comparing results across studies include the wide range of personal narrative elicitation and analysis methods used in previous research.

Most of the 13 studies identified by Gupta et al. (in press) as evaluating the “coherence” of personal narrative with the NaCCS rating scale involved English-speaking children. However, one study recruited 7- to 14-year-old children speaking Danish (Foldager et al., 2024) and reported using a heavily adapted version of the NaCCS. Another study involved 7- to 13-year-old Croatian-speaking participants with typical development (Kuvač Kraljević et al., 2023), whereas a more recent Croatian study (not included in the systematic review) involved 10-year-old children with developmental language disorder (Gabaj, et al., 2025b). These Croatian studies showed improving performance with age on the NaCCS, consistent with previous research (Reese et al., 2011), and they showed sensitivity of the NaCCS to language difficulty on the dimensions of chronology and theme (Gabaj et al., 2025b), with the typically developing children demonstrating better performance on these dimensions. Although these studies showed that a translated version of the NaCCS can be used effectively with some non-English-speaking children, these results need to be confirmed with children from other non-English-speaking backgrounds.

Of the studies included in the systematic review (Gupta et al., in press) as specifically assessing personal narrative coherence, only one study recruited children from multiple cultural backgrounds (Reese et al., 2017). Reese et al. investigated the links between narrative identity, personality traits, and well-being in 263 adolescents (ages 12–21 years) who were living as part of communities with one of three cultural backgrounds in New Zealand: Māori, Chinese, or European. All participants were asked to provide turning point narratives (an event that changed their lives or the kind of person they are) in English. Narratives were coded for causal coherence (reference to a change in personality) and thematic coherence (based on the theme dimension of the NaCCS). Although thematic coherence showed a positive association with well-being for Māori adolescents (only), there were no culture-based differences for thematic coherence. These results indicate that the thematic element of the NaCCS may be useful when appraising personal narrative coherence in English-speaking children from different cultural backgrounds. Taken together, there is currently no unified picture of the extent to which, or the ways in which, personal narrative coherence depends on language-specific properties and/or cultural practices. In this large-scale,

exploratory study, we make a first attempt to address this gap and advance our understanding of this important issue.

The Current Study

In this exploratory study, we used the Global TALEs protocol to elicit personal narratives from a large number of children from a range of different countries and evaluated coherence using the NaCCS (Reese et al., 2011). A total of 417 children (from 21 culture/language participant groups, some of which were described as participants in previous studies) produced personal narratives in response to the six Global TALEs protocol prompts (either face-to-face or online). We deliberately recruited “typically developing” 9- to 11-year-old children who spoke their native language > 80% of the time at home, as per our previous investigation (Westerveld et al., 2022). Responses to three of the prompts (happy/excited, angry/worried, and problem situation) were chosen to be analyzed for coherence because they could reflect three different emotional valences (positive, negative, and neutral), based on previous research (Kuvač Kraljević et al., 2023). We selected the NaCCS for analysis of coherence, as this was the most frequently used coherence coding scheme in previous research, and it was based on a clear theoretical foundation for considering three coherence dimensions of context, chronology, and theme (Reese et al., 2011). Thus, this coding system choice also supported our second research aim of investigating culture/language participant group effects on the three coherence dimensions in the coding scheme. Our specific research questions were as follows:

1. What are the effects of participant group (defined by cultural/language background) and narrative prompt type (*happy*, *angry*, *problem situation*) on the total coherence scores of children’s personal narratives, and how do these factors interact?
2. Do coherence subscores for the three dimensions (context, chronology, and theme) vary across participant groups (defined by cultural/language background) when examined collectively across the three narrative prompt types (*happy*, *angry*, *problem situation*)?

Because of the number of culture/language participant groups ($N = 21$) and the risk of overinterpreting our results based on relatively small individual participant group sizes, we specify broad, directional hypotheses focused on overall effects and their interaction rather than group-specific contrasts. For Research Question 1, we hypothesized that total coherence scores would differ across culture/language participant groups, based on previous research outlining cultural influences on children’s coherence of (written) personal narratives (Kupersmitt, 2015). We further expected that total coherence scores would

differ across narrative prompt types (Fivush et al., 2008; Gabaj et al., 2025c), with the negative prompt (*angry/annoyed*) eliciting more coherent personal narratives than the positive prompt (*happy/excited*). We also hypothesized interactions between culture/language group and prompt type, such that the magnitude of prompt-type differences would vary across culture/language participant groups (e.g., Wang & Leichtman, 2000), due to the possible influence of cultural differences in parental engagement in narrative conversations that could affect children's personal narrative performance.

For Research Question 2, we expected coherence dimension subscores to differ across culture/language participant groups, based on prior research suggesting children's cultural/language background influences how they may structure their narratives (e.g., Lai, 2020; Reese et al., 2017). Finally, we anticipated that the effect of narrative prompt type on coherence dimension subscores might vary across culture/language participant groups. However, given the exploratory nature of the study and the limited prior research, we did not specify an a priori directional hypothesis for this interaction.

Method

Participants

Study participants were recruited through the authors' research and personal networks. All authors named on this report (with the exception of the project's statistical consultant [S.M.]) were members of the Global TALES Network. There were 417 participants from 21 culture/language participant groups, representing 19 different countries (two distinct culture/language groups each from Canada [Canada-E for English and Canada-F for French] and Israel [Israel_A for Arabic and Israel_H for Hebrew]) or regions (Hong Kong is considered a special administrative region of China) and speaking 18 different languages (i.e., three groups spoke English) or dialects (Cypriot Greek and Ghanaian English were considered dialects). Participant numbers ranged from 18 to 20 children per cultural/language participant group, with an overall mean age of 10;2 (years;months; $SD = 5.9$ months; range: 89–142 months).

Twelve of the 21 culture/language participant groups in the current study were reported on in previous studies (Einarsdóttir & Þráinsdóttir, 2023; Ferman & Kavar, 2023; Kuvač Kraljević et al., 2023; Lyons et al., 2023; Theodorou et al., 2023; Westby et al., 2023; Westerveld et al., 2022; Yaşar-Gündüz, 2024). An additional participant group (Ghanaian children) was part of a larger PhD study (J.B.). All children met the following inclusion criteria, as confirmed through written parent report in the form of a brief demographic questionnaire (with exceptions noted for Cyprus and Ghana): (a) ages between 9;6 and

10;11 (except children in Cyprus ranging from age 7;4 to 11;9 and in Ghana ranging from age 8;10 to 10;11), (b) no history of speech and language difficulties, (c) currently not receiving specialist services (e.g., speech-language therapy), and (d) native speakers of the majority language of the country or region in which the protocol was administered (except children in Ghana speaking "Ghanaian English" but using other languages at home).

To control for potential socioeconomic variations, we aimed to invite children who attended schools located in a middle-income area. To ensure roughly equal sample sizes, we aimed for 20 children per participant group. If there were more children in a group, we randomly selected 20 children while maintaining an equal number of male and female participants in each group where possible. Table 1 provides an overview of participant details, including the child's country or region, language spoken, chronological age, gender, relative income area of the child's school (low, middle, high), (first) parent's highest level of education, and family's reported income relative to the average income in their country.

Procedure

Ethics permission was obtained through the Human Research Ethics Committee at Griffith University (No. 2018/273 and No. 2020/942) in Australia. All other research groups obtained ethics approval from their universities or governing bodies, when required, often using a prior approval process. Parents provided written consent for their children to participate in the study, and children provided written and/or verbal assent prior to the session. Across the different culture/language participant groups, data collection occurred between 2019 and 2025. In the face-to-face condition, children were seen individually in a quiet location, often at the child's school, in their home, or in a community venue. In the online condition, children were often at home, occasionally with a parent present. All examiners were qualified speech-language pathologists, psychologists, linguists, or students in speech-language pathology, psychology, or linguistics under supervision of qualified professionals. The exception was one teacher pursuing a Master of Arts degree in special and inclusive education (Malaysia). All examiners were native speakers of the language in which the task was elicited, except in Malaysia where the examiner, although not a native speaker, was fluent in Malay. Prior to commencing the data collection, all examiners read the administration manual, watched the demonstration video, and practiced administering the protocol with at least one child who did not participate in the study.

Personal Narrative Task

All children were asked to "tell a story" in response to the six Global TALES prompts (available for download in translated versions at <https://osf.io/ztqg6>). As per the

Table 1. Participant demographics.

Group	n	F/M	Mode	Age (months)	Parent education (first parent)						SES				Relative income			
				M (SD)	PS	HS	TQ	B	PG	NR	Low	Mid	H	NR	L	Av	H	NR
AUS	20	10/10	F2F	124.00 (4.26)		20%	10%	50%	20%	5%		70%	30%			35%	45%	10%
BRA	20	11/9	F2F	124.70 (1.72)	30%	15%		50%		5%	40%	50%	5%	5%		100%		
CAN_EN	20	11/9	OL	119.75 (6.72)		5%	10%	40%	40%	5%		65%	30%	5%	5%	35%	45%	5%
CAN_FR	20	11/9	OL	116.25 (7.69)			10%	40%	50%			65%	35%		5%	35%	60%	
CRO	20	10/10	F2F	123.20 (1.88)		9.5%		10.5%						100%	5%	55%	25%	15%
CYP	20	11/9	F2F	121.55 (10.67)		30%		65%	5%					100%				100%
DK	20	14/6	F2F	124.50 (3.47)			20%	35%	65%			100%			5%	65%	20%	5%
GHA	18	9/11	F2F	117.89 (9.35)						100%	39%	33%	28%		39%	33%	28%	
GRE	20	10/10	F2F	122.35 (7.03)		45%		25%	30%					100%				100%
HK_SAR	20	12/8	OL	126.20 (4.15)		25%		40%	35%		5%	75%	10%	10%	10%	60%	30%	
ICE	20	10/10	OL	121.95 (3.56)		10%	25%	25%	40%		5%	90%		5%		75%	25%	
IRE	19	13/6	OL	125.37 (3.73)			21%	26%	47%	5%		100%				90%	10%	
ISR_A	20	10/10	OL	125.90 (3.85)		30%		65%	5%			30%	70%		5%	65%	30%	
ISR_H	20	10/10	F2F	122.90 (3.85)		20%	10%	65%	5%					100%	5%	65%	30%	
MAL	20	10/10	OL	124.65 (5.93)	10%	5%	10%	35%	25%	15%		75%	20%	5%	5%	55%	35%	5%
ML_CHINA	20	9/11	F2F	117.85 (2.18)	5%	45%	30%	20%			10%	70%	5%	15%	10%	70%	5%	15%
S_KOR	20	10/10	OL	119.90 (3.16)			15%	70%	15%			95%	5%		5%	65%	30%	
SP	20	10/10	F2F	128.35 (4.93)		20%			10%	70%		35%		65%		25%		75%
TAIW	20	9/11	F2F	123.25 (3.08)		20%	10%	40%	30%			90%	10%			90%	10%	
TUR	20	10/10	OL	124.25 (4.14)	25%	35%		25%	15%		5%	90%	5%		10%	70%	20%	
USA	20	10/10	F2F	123.10 (4.40)		20%	10%	25%	25%	20%	10%	45%	25%	20%	15%	40%	25%	20%

Note. F/M indicates the number of female and male participants. Mode refers to the elicitation method. F2F = face-to-face; OL = online. Age is presented in months, with mean and standard deviation. Parent education (first parent) includes the highest level of education attained by the first parent: PS = primary school; HS = high school; TQ = trade qualification; B = bachelor's degree; PG = postgraduate degree; and NR = not reported. SES (socioeconomic status) categories include Low, Mid, H = High, and NR = not reported. Relative income refers to household income levels relative to national standards: L = very low/low; Av = average; H = high; and NR = not reported. The groups listed in the table refer to the following countries or regions: AUS = Australia; BRA = Brazil; CAN_EN = English-speaking children from Canada; CAN_FR = French-speaking children from Canada; CRO = Croatia; CYP = Cyprus; DK = Denmark; GHA = Ghana; GRE = Greece; HK_SAR = Hong Kong Special Administrative Region; ICE = Iceland; IRE = Ireland; ISR_A = Arabic-speaking children in Israel; ISR_H = Hebrew-speaking children in Israel; MAL = Malaysia; ML_CHINA = Mainland China; S_KOR = South Korea; SP = Spain; TAIW = Taiwan; TUR = Turkey; and USA = United States of America.

protocol, if the child could not think of a story, the examiner provided one or two scripted follow-up prompts. To encourage the child to continue their story, neutral encouragements were allowed, such as “Tell me more” or “And then?” All protocol prompts were asked in a set order and were presented simultaneously in written format on laminated cards or on the computer screen, except in Mainland China where the prompts were provided only verbally. The protocol was translated, when required, by one of the researchers who was a native speaker of the language.

Transcription and Analysis

The sessions were audio-recorded for transcription and analysis purposes. The three prompts that were analyzed for the current study, *happy/excited*, *angry/annoyed*, and *problem situation*, were chosen to reflect three different emotional valences (positive, negative, and neutral). Children’s personal narrative samples were transcribed verbatim in their native language by native language speakers experienced in transcribing language samples. As per standard Systematic Analysis of Language Transcripts (SALT) conventions (Miller, Andriacchi, et al., 2020), all reformulations, repetitions, false starts, and filler words (e.g., *uhm*, *like*) were put in brackets and not included as part of the analysis. Utterance segmentation was based on communication units (C-units), defined as an independent clause with its modifiers. Only complete and intelligible utterances were counted as C-units; unfinished and interrupted utterances were not included.

Personal narratives were defined as a narrative containing at least two time-bound (called “restricted”) narrative clauses (Labov & Waletzky, 1997). Restricted narrative clauses are thus clauses that refer to past personally experienced events that can be put on a timeline (e.g., *I went to the park*; *My brother asked me if I wanted to come*). Evaluative statements were not considered narrative clauses (e.g., *That was so much fun*; *We really loved going there*; *We never know how to solve these problems*). Children’s responses to the three prompts that were relevant to the current study were thus categorized as follows: No PN = not a personal narrative, if it did not contain at least two restricted narrative clauses; PN = personal narrative, if it contained at least two restricted narrative clauses; and NR = no response to either the original prompt or follow-up prompts. Children’s verbal productivity, which was measured only for attempts that met the PN criteria, involved counting the total number of words (or characters, for transcripts from Hong Kong, Mainland China, and Taiwan) per personal narrative. Word (or character) counts were calculated automatically using SALT (Miller, Gillon, & Westerveld, 2020), manually, or using a language-specific program (<https://malgreinir.herokuapp.com/>). Table 2 shows the distribution of the types of responses (No PN, PN, NR) to each personal narrative prompt

(*happy*, *angry*, *problem situation*) by participant group. Supplemental Material S1 shows these same data as stacked bar graphs. Children’s verbal productivity was measured as the total number of words or characters per PN. These measures are reported for descriptive purposes only in Table 3.

Coding Personal Narrative Coherence

The NaCCS, as described by Reese et al. (2011), was used to assess personal narrative coherence, with minor refinements and added descriptors to facilitate interrater reliability (Gabaj et al., 2025b). Each narrative was scored on a scale from 0 to 3 on the three dimensions: context, chronology, and theme (for a maximum total score of 9). The *context* dimension evaluates the degree to which the child orients the narrative in time and place. The *chronology* dimension evaluates whether the child relates narrative components along a timeline. The *theme* dimension evaluates how well the child maintains the topic and elaborates it by using causal statements, personal interpretations, and evaluations, with links to autobiographical experiences. The adapted coding scheme is included in the Appendix.

Reliability

Establishing reliability of coherence coding within and between culture/language participant groups that vary in language and/or cultural background is a challenging task. To maximize reliability coding for this global study, we used the following five-step process: Step 1: Two native, or highly proficient, speakers from each culture/language participant group viewed the coding scheme and the exemplars, focusing on the criteria for identifying narrative clauses (determining PN status) and for awarding scores of 1–3 for *theme* (as the most challenging of the three judgments). Step 2: One native (or highly proficient) speaker from each participant group attended at least one of the 10 scheduled drop-in sessions or arranged an individual meeting with the first author (M.F.W.) to discuss elements that were difficult to code. Step 3: One native speaker coded all 20 transcripts in their own language. Step 4: English translations were created for a representative subgroup of samples using Google Translate (if needed), which were sent to the first author (M.F.W.), who, acknowledging that these were translated stories, reviewed them and provided feedback regarding consistency with NaCCS standards and coding by other researcher groups. Step 5: Following any adjustments, the second native coder reviewed the coding of the first native coder and marked disagreements. At that point, the percent of agreement was calculated for each element (context, chronology, theme). Then, the two native speaker coders met and discussed discrepancies to reach 100% agreement on scores to be used for analysis. If a score discrepancy could not be resolved, the narrative was discussed at a drop-in meeting with the first author (M.F.W.)

Table 2. Percentage of responses by participant group for each prompt.

Group	Happy PN			Angry PN			Problem PN		
	No PN	PN	NR	No PN	PN	NR	No PN	PN	NR
Australia	5.0%	95.0%	0.0%	20.0%	80.0%	0.0%	0.0%	100.0%	0.0%
Brazil	10.0%	90.0%	0.0%	10.0%	90.0%	0.0%	25.0%	75.0%	0.0%
Canada_EN	5.0%	95.0%	0.0%	30.0%	70.0%	0.0%	5.0%	80.0%	15.0%
Canada_FR	0.0%	100%	0.0%	20.0%	75.0%	5.0%	0.0%	90.0%	10.0%
Croatia	5.0%	95.0%	0.0%	10.0%	90.0%	0.0%	5.0%	90.0%	5.0%
Cyprus	10.0%	90.0%	0.0%	25.0%	70.0%	5.0%	20.0%	60.0%	20.0%
Denmark	15.0%	85.0%	0.0%	20.0%	75.0%	5.0%	0.0%	90.0%	10.0%
Ghana	27.8%	72.2%	0.0%	5.6%	88.9%	5.6%	5.6%	77.8%	16.7%
Greece	10.0%	90.0%	0.0%	30.0%	70.0%	0.0%	10.0%	90.0%	0.0%
HK SAR	25.0%	75.0%	0.0%	5.0%	95.0%	0.0%	15.0%	85.0%	0.0%
Iceland	25.0%	75.0%	0.0%	50.0%	50.0%	0.0%	0.0%	100%	0.0%
Ireland	15.8%	84.2%	0.0%	26.3%	68.4%	5.3%	10.5%	73.7%	15.8%
Israel_A	30.0%	70.0%	0.0%	30.0%	70.0%	0.0%	10.0%	90.0%	0.0%
Israel_H	10.0%	90.0%	0.0%	0.0%	90.0%	10%	15.0%	85.0%	0.0%
Malaysia	15.0%	85.0%	0.0%	10.0%	85.0%	5.0%	15.0%	75.0%	10.0%
ML China	15.0%	85.0%	0.0%	10.0%	85.0%	5.0%	15.0%	85.0%	0.0%
S. Korea	10.5%	89.5%	0.0%	10.0%	90.0%	0.0%	5.0%	95.0%	0.0%
Spain	5.0%	95.0%	0.0%	5.0%	95.0%	0.0%	5.0%	95.0%	0.0%
Taiwan	0.0%	100%	0.0%	0.0%	100%	0.0%	0.0%	100%	0.0%
Turkey	30.0%	70.0%	0.0%	15.0%	85.0%	0.0%	10.0%	90.0%	0.0%
USA	75.0%	25.0%	0.0%	65.0%	35.0%	0.0%	10.0%	90.0%	0.0%

Note. PN = personal narrative; NR = no response; Canada_EN = English-speaking children from Canada; Canada_FR = French-speaking children from Canada; HK SAR = Hong Kong Special Administrative Region; Israel_A = Arabic-speaking children in Israel; Israel_H = Hebrew-speaking children in Israel; ML China = Mainland China; S. Korea = South Korea; USA = United States of America.

to reach consensus. Throughout this process, M.F.W. consulted Professor Elaine Reese (the first author of the article that introduced the NaCCS), who provided feedback on her agreement with the accuracy of the coding of 24 representative narrative samples (some in English and some translated to English). These were added to the list of exemplars. Supplemental Material S2 provides an overview of the reliability process and data. Supplemental Material S3 provides examples of three coded personal narratives.

Data Screening and Analysis Approach

As shown in Table 2, not all children produced responses to all three of the prompts used in this study that qualified as personal narratives, with 17.5% of children not producing a *happy* (or *excited*) narrative, 19.9% not producing an *angry* (or *annoyed*) narrative, and 13.2% of children not producing a *problem* narrative. To potentially assist in our choice of statistical analysis method, Little's Missing Completely At Random (MCAR) test was conducted to assess whether the missing data were missing completely at random. The test was not statistically significant, $\chi^2(9) = 15.63$, $p = .075$, indicating that the data are consistent with the assumption of MCAR. Regardless, to adequately deal

with unequal sample sizes, model complexity, and repeated measures, we selected to run linear mixed models (LMMs).

Approximately half of the personal narratives were elicited in face-to-face mode (594; 56.9%), with the remaining narratives elicited online. To determine if elicitation mode significantly influenced narrative coherence (without complicating the overall model), a linear mixed-effects model was conducted to examine the effect of mode on coherence scores, with random effects accounted for as appropriate. The fixed effect of mode was not statistically significant, $F(1, 408.56) = 1.37$, $p = .243$, partial $\eta^2 = .003$, indicating that mode did not significantly predict coherence scores. Therefore, elicitation mode was not considered in the subsequent analyses.

To answer the two research questions, LMMs were estimated using restricted maximum likelihood. For the first research question, an LMM was fitted to examine the effects of narrative prompt type, participant group, and their interaction on total coherence scores. The model included random intercepts for participants and specified an autoregressive (AR(1)) covariance structure matrix for repeated measures across narrative prompt types. To answer the second research question, three separate LMMs were conducted to examine potential differences in the three dependent coherence variables, which were scores for context, chronology,

Table 3. Length of personal narrative by prompt and participant group in number of words/characters.^a

Group	Happy prompt	Angry prompt	Problem prompt
	<i>M (SD); range</i>	<i>M (SD); range</i>	<i>M (SD); range</i>
Australia	108.6 (40.7); 54.0–239.0	108.6 (54.5); 29.0–236.0	126.5 (64.8); 22.0–246.0
Brazil	76.7 (67.8); 22.0–315.0	92.6 (89.3); 21.0–370.0	91.2 (72.4); 22.0–287.0
Canada_EN	86.1 (43.6); 31.0–190.0	92.7 (33.7); 14.0–137.0	133.1 (48.5); 69.0–225.0
Canada_FR	94.1 (70.2); 14.0–269.0	115.1 (92.0); 27.0–332.0	147.4 (115.1); 42.0–448.0
Croatia	110.1 (129.9); 36.0–623.0	78.5 (39.2); 29.0–168.0	85.8 (48.7); 35.0–201.0
Cyprus	32.3 (17.3); 9.0–84.0	45.7 (55.2); 10.0–218.0	66.5 (50.5); 22.0–190.0
Denmark	91.3 (37.8); 48.0–185.0	92.8 (35.1); 50.0–146.0	120.4 (72.0); 21.0–354.0
Ghana	69.8 (67.6); 13.0–283.0	67.3 (48.3); 16.0–220.0	84.6 (68.0); 27.0–294.0
Greece	48.6 (23.2); 13.0–94.0	53.0 (23.2); 14.0–89.0	68.8 (45.0); 16.0–197.0
HK SAR ^a	199.1 (151.3); 62.0–576.0	165.6 (86.0); 36.0–310.0	163.3 (100.0); 46.0–367.0
Iceland	78.9 (37.3); 38.0–173.0	90.7 (40.1); 52.0–185.0	78.0 (45.7); 38.0–245.0
Ireland	68.6 (24.3); 34.0–114.0	55.5 (22.8); 17.0–97.0	69.6 (32.9); 26.0–156.0
Israel_A	32.6 (14.7); 20.0–75.0	38.3 (21.2); 20.0–94.0	36.9 (15.6); 16.0–74.0
Israel_H	72.3 (45.9); 14.0–171.0	57.2 (28.8); 27.0–117.0	81.6 (48.8); 25.0–220.0
Mainland China ^a	156.3 (72.4); 73.0–361.0	179.2 (84.3); 90.0–428.0	130.3 (54.8); 57.0–212.0
Malaysia	30.4 (18.7); 11.0–73.0	49.4 (38.6); 15.0–152.0	40.4 (24.7); 12.0–88.0
South Korea	56.1 (57.4); 17.0–270.0	54.1 (36.1); 18.0–137.0	55.3 (36.3); 11.0–152.0
Spain	37.7 (17.3); 10.0–78.0	36.3 (18.0); 12.0–76.0	73.0 (48.0); 24.0–202.0
Taiwan ^a	162.9 (197.7); 24.0–840.0	127.7 (165.4); 29.0–747.0	113.1 (83.8); 30.0–312.0
Turkey	55.4 (41.1); 12.0–138.0	78.0 (77.0); 11.0–354.0	80.4 (101.7); 24.0–469.0
USA	74.4 (38.1); 45.0–140.0	86.3 (52.1); 29.0–181.0	104.9 (69.9); 33.0–344.0

Note. EN = English-speaking; FR = French-speaking; HK = Hong Kong; SAR = Special Administrative Region; A = Arabic-speaking; H = Hebrew-speaking; Malaysia = Malay-speaking children only; USA = United States of America.

^aLength measured in the number of characters.

and theme, across the three narrative prompt types (*happy/excited*, *angry/annoyed*, and *problem situation*). The model included random intercepts for each participant to account for repeated measures, and an unstructured covariance structure was used because it makes no assumptions about the pattern of covariance between measures. Group was entered as a fixed factor to allow for the establishment of group differences.

Data screening indicated most assumptions were met, including residual normality (normal histogram and Q–Q plot of residuals). Although a potential violation of the homoscedasticity assumption was observed (e.g., diagonal patterns in residual scatter plot of predicted values by residuals), the linear mixed-effects model was retained due to its robustness to moderate heteroscedasticity.

Results

Effects of Participant Group and Narrative Prompt Type on Total Coherence Scores

Question 1 investigated the effects of culture/language participant group and narrative prompt type on the total

coherence scores of children's personal narratives, as well as the interaction. An LMM with an AR(1) covariance structure revealed meaningful variability between participants; therefore, the random intercept of participant was retained (variance = 0.39, standard error [*SE*] = 0.20, intraclass correlation coefficient [ICC] = .17). Significant main effects of narrative prompt type (*happy*, *angry*, *problem*), $F(2, 537.88) = 4.38$, $p = .013$ (partial eta squared [η^2_p] = .016), and group (culture/language group), $F(20, 386.86) = 8.82$, $p < .001$, $\eta^2_p = .31$, were found on total coherence scores, which could range from 0 to 9 (based on a maximum score of 3 for each of the three dimensions, context, chronology, and theme). Furthermore, a significant Narrative Prompt Type $\times$ Group interaction was found for total coherence scores, $F(40, 523.37) = 1.83$, $p = .002$, indicating that the effect of narrative prompt type was not consistent across culture/language groups.

Because there was a significant main effect for narrative prompt type, post hoc pairwise comparisons using Bonferroni correction were conducted to examine differences between narrative prompt type across groups. The *problem* narratives had the highest estimated marginal mean coherence score ($M = 5.10$, $SE = 0.08$), followed by *happy* narratives ($M = 4.93$, $SE = 0.09$), and *angry*

narratives ($M = 4.79$, $SE = 0.09$). Pairwise comparisons showed a statistically significant difference between *problem* and *angry* ($p = .010$) but no significant differences between *happy* and *angry* ($p = .552$) or *problem* and *happy* ($p = .413$). To investigate the effect of narrative prompt type on total coherence scores within culture/language participant groups, simple effects analyses were conducted. Pairwise comparisons of total coherence scores by narrative prompt type within each participant group (adjusted using Bonferroni correction) revealed significant differences in coherence scores for the following participant groups only: Taiwan and Spain (*angry* < *problem*) and Mainland China (*happy* > *problem*). Figure 1 shows the estimated marginal means by narrative prompt type for each culture/language group, with SE indications.

Because there was a significant main effect for group (culture/language participant group), we investigated the effect of participant group on total coherence scores by conducting simple effects analyses. Pairwise comparisons of total coherence scores by group (using Bonferroni correction) revealed significant differences in coherence by group. Supplemental Material S4 shows the pairwise comparisons table, illustrating the mean difference in coherence score, SE , significance, and 95% confidence intervals. The differences are graphically displayed as a heatmap in Figure 2. Each cell represents the mean difference from Group 1 (rows) to

Group 2 (columns). The color gradient indicates the direction and magnitude of the differences, with annotations showing the exact values. As shown in Figure 2, for example, participants in Australia obtained significantly higher (red shadings) coherence scores than three other participant groups: Iceland (performed lower by an estimated mean of 1.43), Spain (1.67 lower), and Malaysia (1.75 lower). In comparison, Brazil scored significantly lower (blue shadings) than seven other participant groups, including Canada_EN (1.58 higher than Brazil). As shown, the performance of children in the United States and Taiwan did not differ significantly from any other countries or regions.

Interaction Effects for Prompt Type by Participant Group on Total Coherence Scores

The significant interaction found for narrative prompt type by culture/language participant group was investigated further. An exploratory k -means cluster analysis was conducted on estimated marginal means of coherence scores across three narrative prompt types (*happy*, *angry*, and *problem*) for all 21 culture/language participant groups. The goal was to identify patterns in narrative coherence across the culture/language participant groups. The elbow plot of the within-cluster sums of squares revealed a two-cluster solution. Cluster 1 includes Australia, Canada_EN, Canada_FR, Croatia, Denmark, Greece, Hong Kong SAR, Ireland, Israel_A, Israel_H, Mainland China, Malaysia, South Korea, Spain, Taiwan, Turkey, and USA.

Figure 1. Estimated marginal means of total scores by narrative prompt type and group. Canada_EN = English-speaking children from Canada; Canada_FR = French-speaking children from Canada; SAR = Special Administrative Region; Israel_A = Arabic-speaking children in Israel; Israel_H = Hebrew-speaking children in Israel; USA = United States of America.

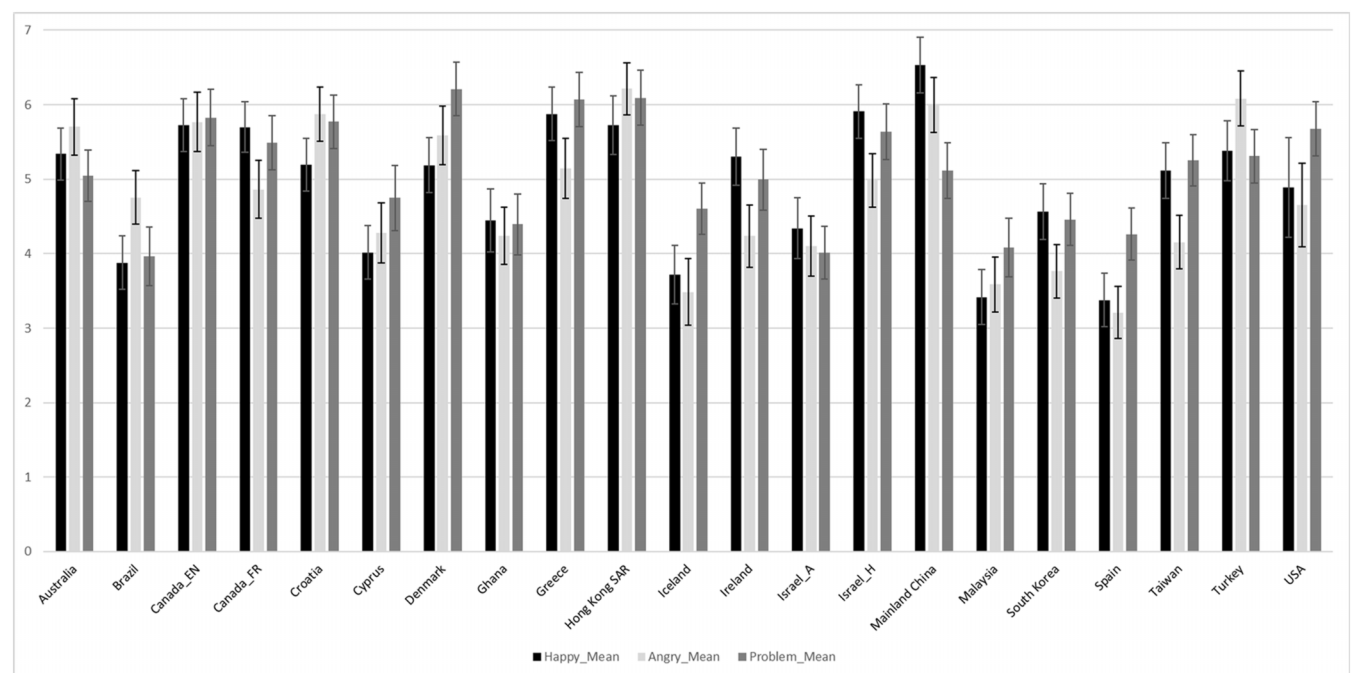
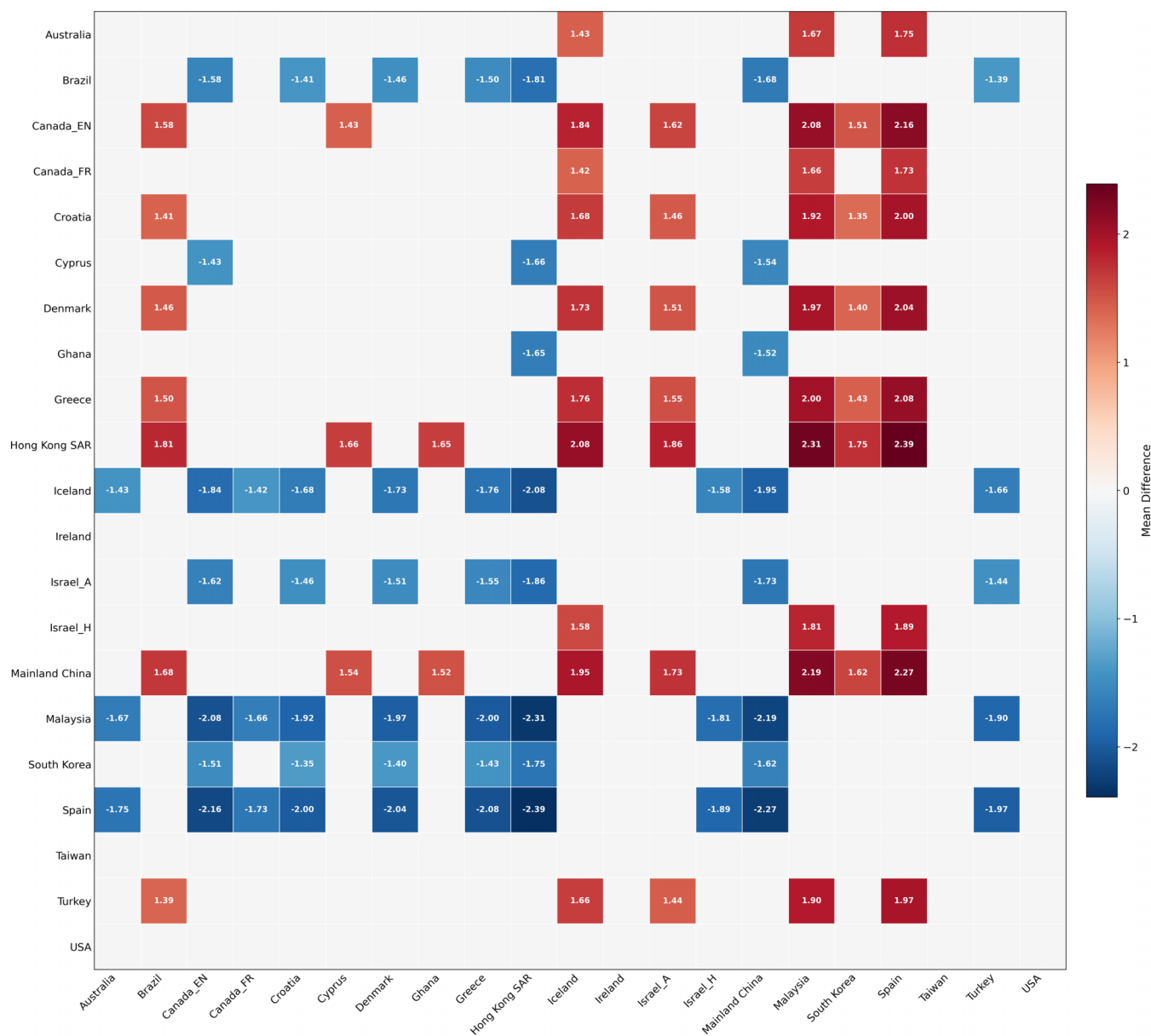


Figure 2. Heatmap showing significant pairwise comparisons of total coherence scores between culture/language participant groups, with numbers indicating difference scores. Canada_EN = English-speaking children from Canada; Canada_FR = French-speaking children from Canada; SAR = Special Administrative Region; Israel_A = Arabic-speaking children in Israel; Israel_H = Hebrew-speaking children in Israel; USA = United States of America.



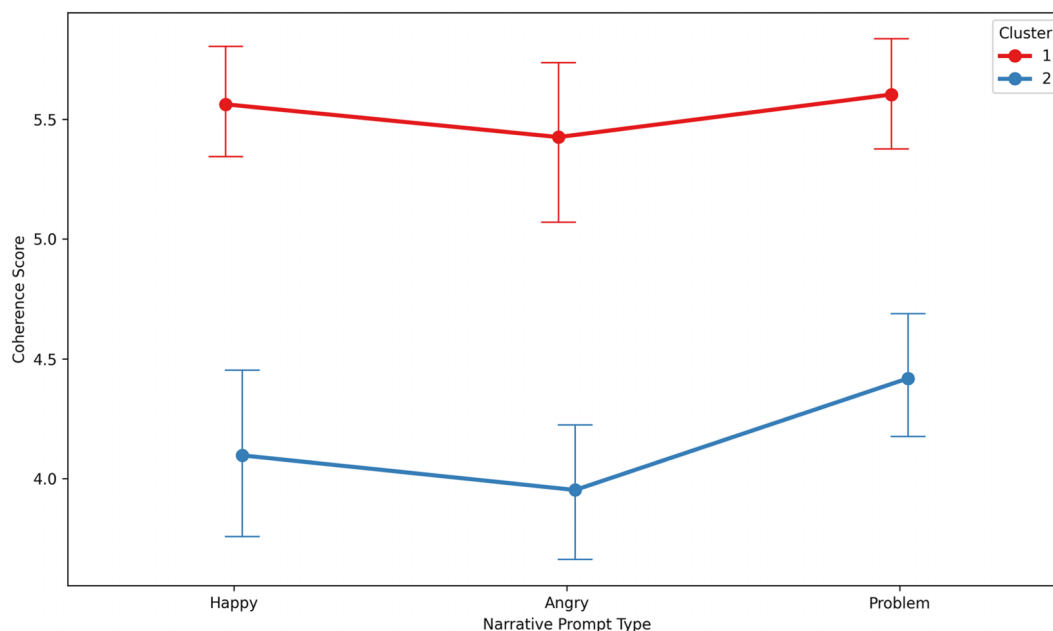
Israel_H, Mainland China, Turkey, and the United States. As shown in Figure 3, Cluster 1 shows consistently higher scores across all narrative prompt types, with relatively balanced performance across the three prompt types (emotional valence). Cluster 2 includes Brazil, Cyprus, Ghana, Iceland, Israel_A, Malaysia, South Korea, Spain, and Taiwan. This cluster displays lower coherence scores overall, particularly for the *angry* and *happy* narratives. The steeper slope in the interaction plot suggests that the

narrative prompt type has a stronger influence on coherence in this cluster.

Group Differences in Coherence Dimensions Across Prompt Types

The second research question asked whether there were coherence dimension (context, chronology, theme) differences across narrative prompt types (considered

Figure 3. Interaction plot: total coherence score by narrative prompt type for each cluster (based on estimated marginal means).



collectively) and culture/language participant groups. All three of the LMMs provided an intercept of participant value that indicated meaningful variability between participants (context: variance = 0.10, $SE = 0.028$, $ICC = .14$; chronology: variance = 0.08, $SE = 0.027$, $ICC = .12$; theme: variance = 0.08, $SE = 0.013$, $ICC = .28$); therefore, the random intercept of participant was retained in each model. For each of the models, the fixed effect of group was significant: context: $F(20, 382.26) = 5.35$, $p < .001$; chronology: $F(20, 364.92) = 5.51$, $p < .001$; and theme: $F(20, 382.14) = 8.75$, $p < .001$. Subsequent Bonferroni post hoc analyses revealed the following number of significant group differences for each of the three dimensions (with maximum number of group differences = 210): 25 for context (see Table 4), 21 for chronology (see Table 5), and 56 for theme (see Table 6). Figures 4–6 visualize the estimated marginal means, ordered from high to low, for context, chronology, and theme, based on the LMM, with dots indicating the point estimate for each culture/language group and error bars representing 95% confidence intervals. All participant groups that showed statistically significant differences with at least one other participant group (as presented in the Tables 4–6) are shown in red. Performance of the culture/language groups in blue does not differ from any of the other participant groups.

Discussion

The collective members of the Global TALES Network were interested in continuing to develop a global tool that

would be useful for evaluating the personal narrative language skills of individuals across a range of languages and cultures. The current study contributes evidence for measuring coherence in personal narratives, which, pending further research, could help differentiate between students who struggle to produce coherent personal narratives due to language learning difficulties and those whose personal narrative coherence may be influenced by their cultural and/or linguistic background.

To that end, the current study investigated similarities and differences in performance in personal narrative coherence—defined as the ability to produce well-structured stories that make sense to a naive listener about personally experienced past events—of typically developing children from 21 culture/language participant groups ($N = 417$; from 19 countries or regions; speaking 18 languages or dialects), whose ages ranged from 7;4 to 11;8 ($M = 10;2$, $SD = 6$ months). All children produced personal narratives in their own languages in response to the Global TALES protocol, either face-to-face or online. The current study involved analysis of children's personal narratives produced orally in response to three of the six original prompts, each representing a different emotional valence: *happy/excited* (positive), *angry/annoyed* (negative), and *problem situation* (neutral).

Effects of Group and Narrative Prompt Type on Children's Total Coherence Scores

In answering the first question regarding the effects of participant culture/language group and narrative prompt

Table 4. Significant pairwise group differences for context score across narrative prompt types.

Group I	Group II	Mean difference (I–II)	<i>p</i> value
Brazil	Canada_EN	–0.705	.048
Brazil	Mainland China	–0.797	.006
Canada_EN	Ghana	0.777	.019
Canada_EN	Iceland	0.773	.017
Canada_EN	Malaysia	0.878	.001
Canada_EN	South Korea	0.698	.046
Canada_EN	Spain	0.950	< .001
Denmark	Spain	0.764	.010
Ghana	Greece	–0.739	.040
Ghana	Mainland China	–0.869	.002
Greece	Iceland	0.736	.036
Greece	Malaysia	0.841	.003
Greece	Spain	0.913	< .001
Hong Kong SAR	Iceland	0.714	.049
Hong Kong SAR	Malaysia	0.819	.004
Hong Kong SAR	Spain	0.891	< .001
Iceland	Mainland China	–0.865	.002
Israel_H	Spain	0.744	.013
Mainland China	Malaysia	0.971	< .001
Mainland China	South Korea	0.790	.005
Mainland China	Spain	1.042	< .001
Malaysia	Taiwan	–0.707	.037
Malaysia	Turkey	–0.742	.024
Spain	Taiwan	–0.779	.005
Spain	Turkey	–0.814	.003

Note. Canada_EN = English-speaking children from Canada; HK SAR = Hong Kong Special Administrative Region; Israel_H = Hebrew-speaking children in Israel.

type (with positive, negative, or neutral valence) on the coherence of children’s personal narratives (coded for context, chronology, and theme), including any interaction effects, we found, as hypothesized, a significant main effect (with a small effect size) of narrative prompt type on total coherence scores (Wang & Leichtman, 2000). However, contrary to our expectations, we found that, across culture/language participant groups, the *problem* prompt (considered neutral emotional valence; Kuvač Kraljević et al., 2023) yielded the highest coherence scores and, furthermore, that significant differences were found between *problem* narratives and *angry* narratives in favor of the *problem* narratives. Previous research using the NaCCS to rate children’s narrative coherence had indicated that negative prompts, such as our *angry/annoyed* prompt, are likely to elicit more coherent personal narratives in children in this age group (Fivush et al., 2008; Gabaj et al., 2025c; Kuvač Kraljević et al., 2023). One possibility for the difference in this study is that narratives produced in response to the problem prompt might trend more negative than positive and, thus, are not uniformly neutral. Another possibility, as suggested by Fivush et al. (2008), is that the “problem” prompt in the current study was the most likely to provide

a structure to the story, even more so than the “angry” prompt, thereby enhancing the coherence theme score to the “problem” stories in particular. Further analysis of the children’s responses would be needed to determine whether children chose to narrate positive, negative, or in fact neutral experiences when prompted, “Tell me a story about a time when you had a problem and you had to fix it. Tell me all about what happened and what you had to do to fix it.”

Another possibility is that variations in methodology could help explain the differences. In their study, Fivush et al. (2008) enrolled 51 English-speaking children who had asthma and prompted them to tell a story about a scary asthma episode; in the current study, we used a more open-ended prompt to invite the children to tell a story about a time that they were *angry* or *annoyed*. Perhaps the more directed prompt in the Fivush study helped their participants structure their asthma narratives chronologically and produce a more coherent story. This may also help explain why children in the current study produced more coherent narratives in the *problem* condition. As noted, this prompt also provided more support for structuring the narrative. Although Gabaj et al. (2025c) and Kuvač Kraljević et al.

Table 5. Significant pairwise group differences for chronology score across narrative prompt types.

Group I	Group II	Mean difference (I–II)	<i>p</i> value
Australia	Israel_A	0.818	.002
Canada_EN	Israel_A	0.826	.002
Canada_FR	Israel_A	0.709	.023
Croatia	Israel_A	0.870	< .001
Denmark	Israel_A	0.996	< .001
Denmark	Malaysia	0.768	.007
Denmark	Spain	0.769	.004
Ghana	Israel_A	0.725	.033
Greece	Israel_A	1.064	< .001
Greece	Malaysia	0.836	.002
Greece	Spain	0.837	.001
Hong Kong SAR	Israel_A	1.068	< .001
Hong Kong SAR	Malaysia	0.840	.001
Hong Kong SAR	Spain	0.841	.001
Ireland	Israel_A	0.902	.001
Israel_A	Israel_H	–0.788	.004
Israel_A	Mainland China	–0.792	.004
Israel_A	Turkey	–1.016	< .001
Israel_A	USA	–0.803	.015
Malaysia	Turkey	–0.788	.005
Spain	Turkey	–0.789	.003

Note. Israel_A = Arabic-speaking children in Israel; Canada_EN = English-speaking children from Canada; Canada_FR = French-speaking children from Canada; HK SAR = Hong Kong Special Administrative Region; Israel_H = Hebrew-speaking children in Israel; USA = United States of America.

(2023) used the Global TALES protocol in their studies of the effects of emotional valence on coherence, they combined children's responses across several prompts to evaluate the effect of emotional valence on coherence (using the *angry* and *worried* prompts for negative valence, *happy* and *proud* prompts for positive valence, and *problem* situation and *important* prompts for neutral valence). Neither Gabaj et al. (2025c) nor Fivush et al. (2008) reported overall coherence scores. Rather, Gabaj et al. reported significant effects of emotional valence only on performance for chronology (not for context; and theme results were not reported). Similarly, Fivush et al. reported higher levels of coherence for their *scary* narrative compared to the *happy* one, but only for the coherence dimensions of chronology and theme.

Interestingly, significant within-group differences were only observed for three culture/language participant groups. In Taiwan and Spain, the *problem* (neutral) prompt elicited higher total coherence scores than the *angry* (negative) prompt, whereas in Mainland China, the *happy* prompt (positive) resulted in higher total coherence scores than the *problem* (neutral) prompt. As shown in Figure 1, there is clear variability in performance across the participant groups, with no single pattern of strength in narrative coherence by narrative prompt type. Thus, our findings indicate

no generalizable assumptions can be made about the effects the type of prompt (emotional valence) may have on narrative coherence in school-age children from different cultural and/or language backgrounds.

When examining the effect of culture/language participant group on narrative coherence across the three prompt types, a significant effect was found. However, results from the post hoc comparisons are difficult to interpret. As shown in the heatmap (see Figure 2), five culture/language groups scored consistently lower than other participant groups. These included Brazil, Iceland, Israel_A, Malaysia, and Spain. In contrast, Canada_EN, Denmark, Greece, Hong Kong SAR, and Mainland China scored consistently higher than other participant groups. These findings rule out relatively straightforward interpretations such as differentiating between Western and Eastern cultures described as valuing independence (e.g., Canada, Iceland) versus interdependence (e.g., Mainland China; see Westby et al., 2023). Other explanations for our findings include potential narrative structure differences that might relate to cultural traditions (e.g., Iceland; Einarsdóttir & Þráinsdóttir, 2023) or culture-specific use of evaluative clauses to engage the listener (Lai, 2020). Some cultures may favor storytelling structures that do not adhere to a more traditional appreciation of “coherence”

Table 6. Significant pairwise group differences for theme score across narrative prompt types.

Group I	Group II	Mean difference (I–II)	<i>p</i> value
Australia	Cyprus	0.509	.023
Australia	Ghana	0.570	.004
Australia	Ireland	0.637	< .001
Australia	Israel_A	0.569	.003
Australia	Malaysia	0.619	< .001
Australia	South Korea	0.541	.005
Australia	Spain	0.588	.001
Australia	Taiwan	0.497	.020
Brazil	Croatia	–0.656	< .001
Brazil	Hong Kong SAR	–0.508	.019
Canada_EN	Ghana	0.550	.010
Canada_EN	Ireland	0.616	.001
Canada_EN	Israel_A	0.548	.007
Canada_EN	Malaysia	0.598	.001
Canada_EN	South Korea	0.520	.012
Canada_EN	Spain	0.567	.002
Canada_EN	Taiwan	0.476	.046
Canada_FR	Ghana	0.539	.012
Canada_FR	Ireland	0.606	.001
Canada_FR	Israel_A	0.538	.008
Canada_FR	Malaysia	0.588	.001
Canada_FR	South Korea	0.509	.014
Canada_FR	Spain	0.557	.003
Croatia	Cyprus	0.737	< .001
Croatia	Ghana	0.798	< .001
Croatia	Greece	0.475	.050
Croatia	Iceland	0.656	< .001
Croatia	Ireland	0.865	< .001
Croatia	Israel_A	0.797	< .001
Croatia	Malaysia	0.847	< .001
Croatia	South Korea	0.769	< .001
Croatia	Spain	0.816	< .001
Croatia	Taiwan	0.725	< .001
Cyprus	Hong Kong SAR	–0.589	.002
Cyprus	Mainland China	–0.502	.032
Denmark	Ireland	0.539	.013
Denmark	Malaysia	0.521	.015
Denmark	Spain	0.490	.031
Ghana	Hong Kong SAR	–0.650	< .001
Ghana	Mainland China	–0.563	.006
Hong Kong SAR	Iceland	0.508	.029
Hong Kong SAR	Ireland	0.717	< .001
Hong Kong SAR	Israel_A	0.649	< .001
Hong Kong SAR	Malaysia	0.699	< .001
Hong Kong SAR	South Korea	0.620	< .001
Hong Kong SAR	Spain	0.668	< .001
Hong Kong SAR	Taiwan	0.577	.002
Ireland	Israel_H	–0.538	.012
Ireland	Mainland China	–0.630	.001
Israel_A	Mainland China	–0.562	.004
Israel_H	Malaysia	0.520	.013

(table continues)

Table 6. (Continued).

Group I	Group II	Mean difference (I–II)	<i>p</i> value
Israel_H	Spain	0.489	.028
Mainland China	Malaysia	0.612	.001
Mainland China	South Korea	0.534	.007
Mainland China	Spain	0.581	.001
Mainland China	Taiwan	0.490	.029

Note. Israel_A = Arabic-speaking children in Israel; HK SAR = Hong Kong Special Administrative Region; Israel_H = Hebrew-speaking children in Israel.

as captured by the NaCCS. Alternatively, language-specific linguistic conventions, such as those described by Berman and Slobin (1994), could influence narrative coherence.

In response to further investigation of the significant Group $\times$ Narrative Prompt Type interaction effects, the exploratory *k*-means cluster analysis revealed two clusters. Cluster 1 (Australia, Canada [English and French], Croatia, Denmark, Greece, Hong Kong SAR, Ireland, Israel [Hebrew-speaking], Mainland China, Turkey, and the United States) showed consistently higher scores across all three narrative prompt types, with relatively balanced performance across the three. The second cluster (Brazil, Cyprus, Ghana, Iceland, Israel [Arabic-speaking], Malaysia, South Korea, Spain, Taiwan) showed lower coherence scores overall, particularly for the *angry* and *happy* narratives. Although a detailed interpretation of these preliminary findings goes beyond the scope of the current study and further statistical analysis is not possible due to the small size of the individual participant group samples, the differentiating factor between the two clusters likely stems from a combination of cultural, educational, and linguistic influences. For example, most countries or regions included in Cluster 1 share Western, individualistic cultural values that may emphasize more elaborate storytelling, with clear personal expression, potentially fostering more coherent narratives. However, as Mainland China and Hong Kong SAR were included in Cluster 1, simple explanations based on Western versus Eastern cultural views do not work, and perhaps, educational experience plays a role as well. Examining the influence of educational experiences across countries can be investigated in future studies. Cluster 2 might represent regions where emotional expression is treated differently in storytelling or where children are less familiar with the kinds of narrative structures emphasized in the task, leading to lower coherence, particularly in emotional personal narratives. It is important to note that the two-cluster result could also be a result of the relatively low sample size per participant group. Based on our results, it is evident that future studies evaluating personal narrative coherence in more detail, perhaps involving fewer culture/language participant groups, but with larger sample sizes,

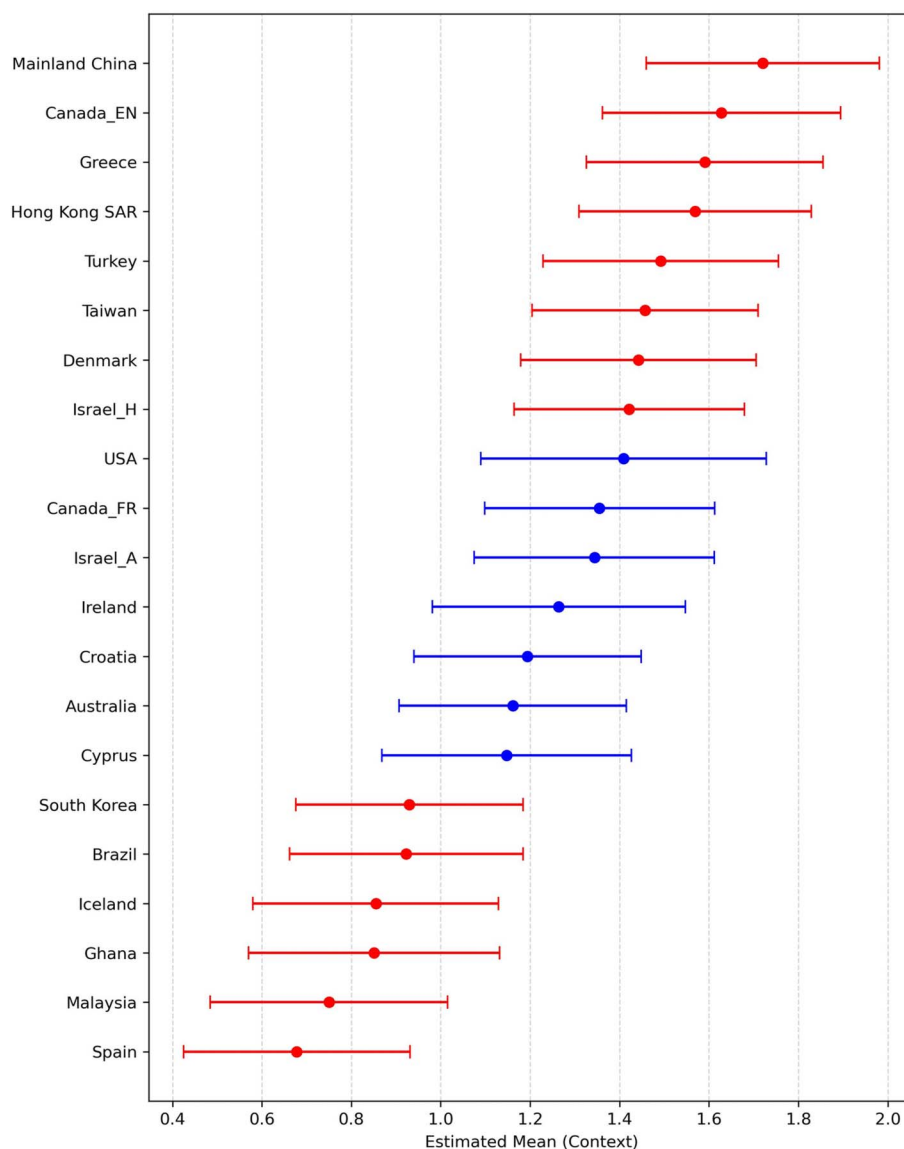
are needed to better disentangle the intriguing possibilities of cultural, educational, and language-specific linguistic influences on children's personal narrative coherence.

Closer Inspection of Context, Chronology, and Theme, Narrative Prompt Types Combined

When answering Research Question 2, the results showed significant culture/language group effects for each of the three individual coherence dimensions: context, chronology, and theme (for all three prompt types combined). This means that children's inclusion of context information (when and where), their chronological sequencing of events (chronology), and their ability to maintain and elaborate on the topic (theme dimension) all might be influenced by their cultural and/or language background (i.e., participant group differences). Figures 4, 5, and 6 show the estimated mean scores by group for each of the dimensions. Although estimated mean scores are not the same as raw scores, they are particularly useful for comparing groups fairly by accounting for repeated measures and individual differences. They can still be interpreted as raw scores as we did not transform our outcome variables.

As shown in Figure 4, despite considerable overlap in estimated mean scores between culture/language groups for context, some participant groups scored significantly higher than others, with children from Mainland China including significantly more context information (specifying time and place) than children from Brazil, Ghana, Iceland, Malaysia, South Korea, and Spain (see Table 4). The estimated mean scores ranged between 0.7 (Spain) and 1.7 (Mainland China), which is roughly in line with previous research on English-speaking children from the United States and New Zealand reporting a mean score for context of 1.1 for their 11- and 12-year age groups (Reese et al., 2011). For chronology (see Figure 5), the estimated mean scores ranged between 1.7 (Israel_A) and 2.7 (Hong Kong SAR), which is similar to the results reported by Reese et al. (2011) for their English-speaking school-age cohorts (2.25–2.4) and to those reported by Kuvač Kraljević et al. (2023)

Figure 4. Dot plot of mean estimates (context) with 95% confidence intervals. Canada_EN = English-speaking children from Canada; SAR = Special Administrative Region; Israel_H = Hebrew-speaking children in Israel; USA = United States of America; Canada_FR = French-speaking children from Canada; Israel_A = Arabic-speaking children in Israel.

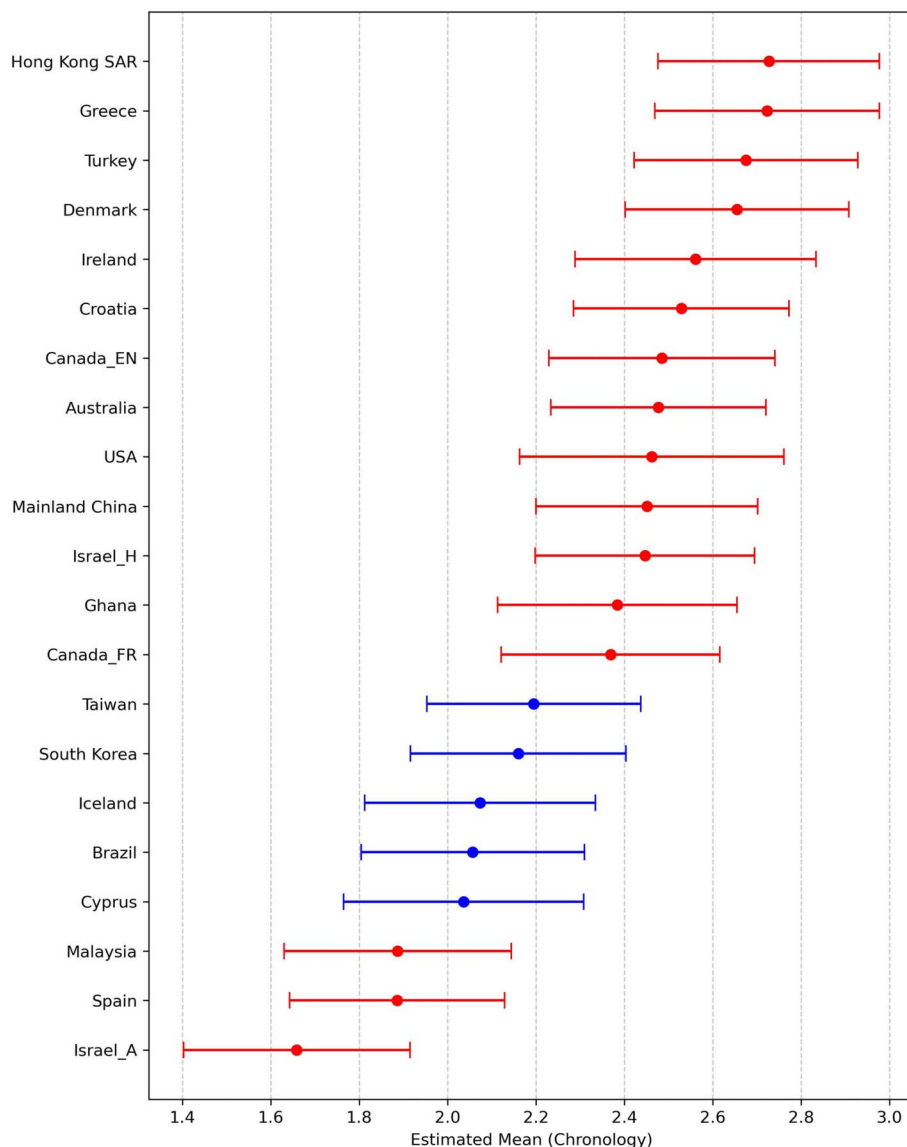


and Gabaj et al. (2025c) for their Croatian-speaking 9- to 11-year-olds (1.88 and 2.23, respectively).

As shown in Figure 6, performance on theme ranged between 1.05 (Ireland) and 1.9 (Croatia). In contrast to context and chronology, the results for some participant groups for theme seem somewhat lower than those reported by Reese et al. (2011) for their 11- and 12-year-old groups (1.9 and 1.7, respectively) and the mean of 1.83 (range: 0.5–3.0) reported by Kuvač Kraljević et al. (2023). Future research, with larger individual group samples, combined with a more detailed coding scheme with coded exemplars

across all languages, may help elucidate whether these relatively lower scores by some culture/language participant groups in the current study are truly due to cultural or linguistic differences among children or perhaps among researchers coding their samples. Based on these findings, it seems that the lower overall coherence scores of the Cluster 2 participant groups (Brazil, Cyprus, Ghana, Iceland, Israel_A, Malaysia, South Korea, Spain, and Taiwan) were mainly driven by their relatively lower scores for context and theme, except for the Arabic-speaking children from Israel who performed in the mid-range for context but obtained the lowest scores in chronology.

Figure 5. Dot plot of mean estimates (chronology) with 95% confidence intervals. SAR = Special Administrative Region; Canada_EN = English-speaking children from Canada; USA = United States of America; Israel_H = Hebrew-speaking children in Israel; Canada_FR = French-speaking children from Canada; Israel_A = Arabic-speaking children in Israel.

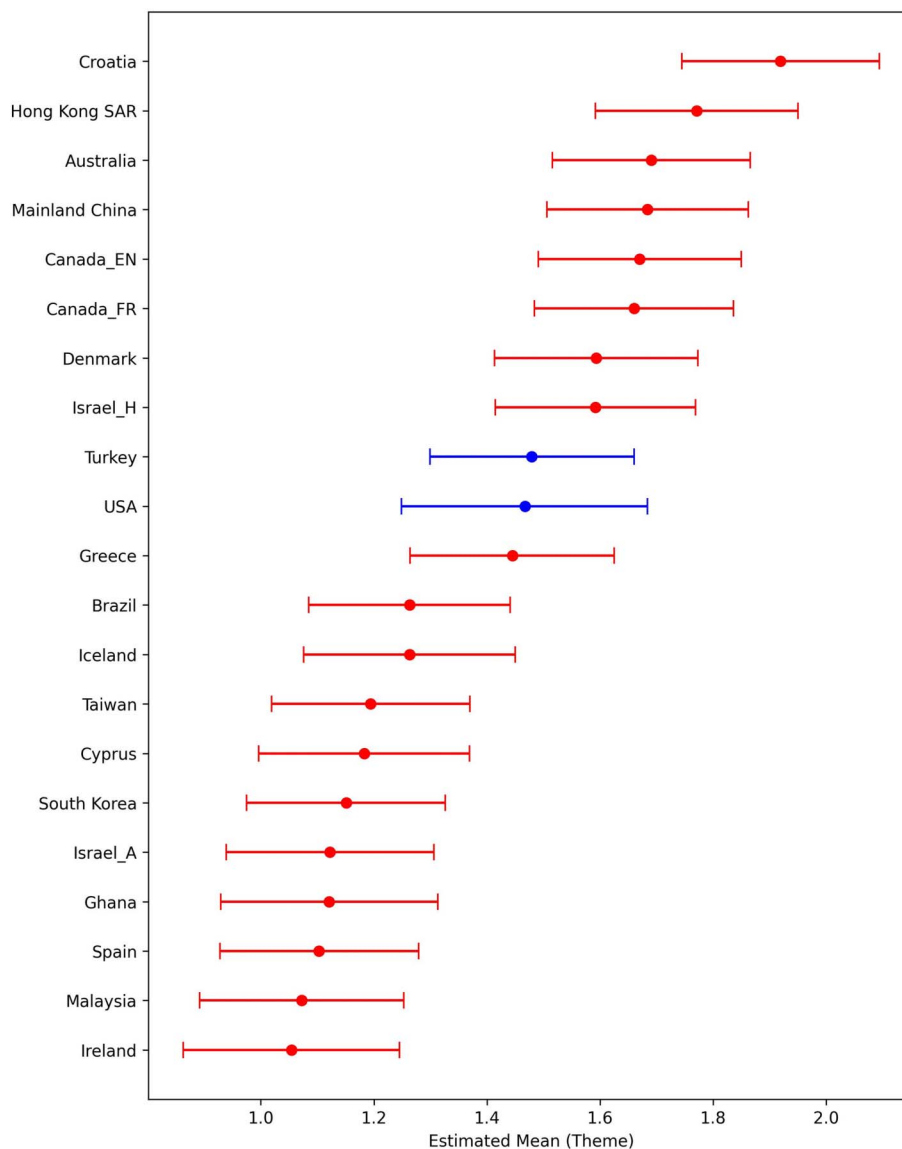


Strengths and Limitations

This exploratory study has yielded intriguing findings about similarities and differences in personal narrative coherence among children from a wide range of countries or regions, speaking a range of different languages or dialects. A strength of the study was the global participation of researchers from 21 culture/language groups who all used the Global TALES elicitation protocol (adapted/translated appropriately) to gather the personal narrative samples. Beyond answering the specific research questions about coherence, this study adds to existing evidence regarding

the viability of the Global TALES protocol with diverse groups of participants (Westerveld et al., 2022). It also adds evidence that, similar to previous research investigating the effect of assessment mode (face-to-face vs. online) on children's language productivity, syntactic complexity, and semantic diversity (Ferman & Kavar, 2023), we found no significant main effects for assessment mode on personal narrative coherence scores. Furthermore, across culture/language participant groups, at least 75% of children produced a personal narrative in response to at least one of the three prompt types, regardless of assessment mode. Taken together, these findings provide further support for

Figure 6. Dot plot of mean estimates (theme) with 95% confidence intervals. SAR = Special Administrative Region; Canada_EN = English-speaking children from Canada; Canada_FR = French-speaking children from Canada; Israel_H = Hebrew-speaking children in Israel; USA = United States of America; Israel_A = Arabic-speaking children in Israel.



the use of the Global TALES protocol with children from a wide range of cultural and/or language backgrounds in both face-to-face and online situations.

As in any study, but especially for a study with global scope, we acknowledge its limitations. Although the total sample size was over 400 and the numbers of children per participant group were similar to previous well-described cross-cultural and cross-linguistic investigations (Berman & Slobin, 1994; Peterson & McCabe, 1983), our findings of group differences are based on a relatively low number of children from each participant group, and most of them were recruited through the authors' personal

as well as research networks, some during the COVID pandemic. To better understand the cultural, educational, and linguistic influences on children's personal narrative coherence, future research is needed involving fewer participant groups but with larger, more representative sample sizes.

We acknowledge that differences in the degree of shared knowledge of past events of participants and interviewers with prior connections could influence the amount of contextual information in particular that a participant conveys explicitly in the narratives. As shown in Supplemental Material S1, our results indicate that the protocol is successful in eliciting personal narrative samples from

children representing a relatively wide range of cultures and speaking 18 different languages or dialects. The only outlier seems to be the United States, where we observed a relatively high number of responses that did not meet our criteria of two narrative clauses to the *happy* and *angry* prompts. Some of these were “stories” about particular situations that made the narrators always feel happy or situations in which siblings did habitually annoying things. One potential explanation is that the examiner was known to many of the participants, which could have resulted in shared background knowledge and elicited more descriptive evaluation clauses as opposed to narrative clauses that expressed specific one-time events using past tense. Future research studies should more carefully collect information regarding the elicitation context, including examiner familiarity, the presence of parents during the assessment, and the location (e.g., school or home).

In this study, we opted to use the NaCCS (Reese et al., 2011) to appraise narrative coherence, based on promising results from previous studies demonstrating that this tool potentially can be used in different languages (Foldager et al., 2024; Gabaj et al., 2025c) and across different cultures (Reese et al., 2017). It is a strength that the use of this coding system made comparison with prior studies possible; however, the researchers were aware of challenges in using this coding system (or perhaps any coding of qualitative features across cultures) reliably. Although we carefully checked reliability of coding within and between culture/language participant groups, maximum scores for the theme dimension (score of 3) relied on the researchers’ own culturally influenced expectations of adult performance, in line with the lifespan model of coherence as captured by the NaCCS. Due to a lack of empirical data involving non-English adult speakers, this may have resulted in making assumptions about adult performance in non-English-speaking participant groups from differing cultural backgrounds, which could have inadvertently influenced the coding. Future research including adult speakers from a range of different countries or regions is clearly needed to determine the validity of this lifespan model of personal narrative coherence suggested by Reese et al. (2011).

Consistent with this concern, many of the authors commented on the limitations of the scoring scale used for the theme dimension, in particular. As per the NaCCS coding scheme, theme coding involves evaluating the child’s ability to maintain and elaborate on the topic, using causal linkages, personal evaluations or reactions, elaboration of actions, and interpretations. Anecdotal observations of children’s narratives that were awarded a score of 2 showed large variability in narrative “quality,” often resulting in informal differentiation between low-2

and high-2 (although analysis was based on the NaCCS coding process of a single 2 value). Even though the NaCCS was successful in highlighting similarities and differences between the culture/language participant groups, future research may consider using a more refined coding scheme, especially when studying children across a narrow age range.

Future Directions

This study should be considered to contribute to initial steps in finding a global, culturally sensitive tool for describing and/or evaluating personal narrative coherence. In cross-cultural studies like this, one cannot make broad assumptions about cultural and/or linguistic influences and their effects on personal narrative coherence. To help move the field forward and better understand the cultural and/or linguistic influences on children’s personal narrative production, future studies are strongly advised to employ existing elicitation and analysis methods to enable meaningful comparisons (see also Gupta et al., in press). To date, we have focused on children speaking the dominant languages of their countries, which they have been learning since birth, and most studies have enrolled children from a specific age group (predominantly 10-year-olds). Future research should aim to recruit participants speaking more than one language, representing a range of socioeconomic backgrounds, and ranging in age from preschool into adulthood, as well as children with known communication difficulties and responding in written as well as spoken modalities.

At this point, the Global TALES protocol has been used successfully with more than 700 children, including children from a wide range of cultural and/or linguistic backgrounds (see Westerveld & Nelson, 2023) and children with neurodevelopmental conditions, including autism (Westerveld, 2024), developmental language disorder (Gabaj et al., 2025b), and dyslexia (Van Vreckem et al., 2023). One concern may be that the protocol may not elicit the most coherent personal narratives from children who are not accustomed to sharing emotional content with unfamiliar adults or from children who are more familiar with co-constructing stories. However, using a standardized protocol to collect data across the lifespan and from representatives of a range of languages and dialects and cultures will enable researchers from various disciplines (linguistics, speech-language pathology, psychology) to better understand this universal discourse genre, its development with age or years of schooling, and the ways in which culture and language may influence performance. Adhering to findable, accessible, interoperable, and reusable guiding principles for scientific data management and stewardship (Wilkinson et al., 2016) will support knowledge discovery and innovation and maximize

the reuse of existing resources, thus accelerating knowledge creation while capitalizing on the finite funding resources available.

Clinical Implications

In a world that is becoming increasingly multicultural and plurilingual, finding ecologically valid assessment tools for evaluating children's spoken language skills is of utmost importance. In an ideal scenario, norms of "typical" performance would be available across the lifespan, across different cultures, and across different languages. However, considering that personal narrative coherence is influenced by a myriad factors, including culture, education, socioeconomic status, and linguistics, developing such benchmarks will be expensive and possibly unrealistic to achieve in the near future.

To mitigate the barriers to using personal narrative assessment in clinical practice, the Global TALES Network set out to develop freely available global assessment and analysis tools. Since its development in 2018, the Global TALES protocol has been translated into 20 languages and can be downloaded from the Open Science Framework site (<https://osf.io/ztqg6/files/osfstorage>). By not providing the child with pictures, model stories, or specific topics, the protocol elicits children's spontaneous personal narrative skills in a time-efficient manner. In this article, we have demonstrated the potential usefulness of the NaCCS in appraising the personal narrative skills of children from a wide range of cultural and/or language backgrounds. Describing a child's performance across the three coherence dimensions of context, chronology, and theme and comparing it to their typically developing peers are important first steps in differentiating between a language disorder and a language difference, regardless of a child's cultural or linguistic background.

Conclusions

This is the first study to investigate the personal narrative skills of school-aged children from 21 culture/language participant groups (i.e., representing 19 countries and regions and 18 different languages or dialects). The results from this exploratory study provide further evidence for the feasibility of using the Global TALES protocol to elicit personal narrative language samples from children across diverse cultural and linguistic backgrounds. Furthermore, the NaCCS shows promise as a tool for evaluating a child's personal narrative coherence, regardless of cultural background or language spoken. We now turn to refining the coding scheme and recruiting larger, more representative samples to further explore some of the larger trends observed, with the ultimate goal of promoting the

global adoption of personal narrative assessment and analysis as a routine component of clinical practice.

Data Availability Statement

The Excel spreadsheet containing the NaCCS scores is available from the corresponding author upon reasonable request.

Acknowledgments

This work was supported by the Hong Kong Polytechnic University (1-WZ9D; to Angel Chan); the University of Iceland (to Jóhanna Einarsdóttir); the Faculty of Rehabilitation Medicine New Investigator Grant, University of Alberta (to Pamela Filiatrault-Veilleux); the Department of Language Science and Technology, Hong Kong Polytechnic University (to Wanlin Zhu); and the Australian Government Research Training Program (to Josephine Ohenewa Bampoe). We are thankful to the children who participated in our study and shared their stories and to their families or teachers who supported their participation. We extend our heartfelt thanks to Professor Elaine Reese for her assistance with the coherence coding. We also thank our colleagues who helped with the data collection and/or reliability checks, including Zhuowen Spring Li, Sidus Ang, Cassandra Baragar, Justine Gillert, Roxanne Bedard, Mairi Cade-Hemphill, Kristy Fyfe, Jade Eckl, Sophie Barbotin, Maeve Molloy, Jessica Spelman, Yarminee Gonasagan, Chien Chien Wong, Alex Pak-Wah Tse, Chris Pak-Lam Lee, Chana Sverdllov, Hrafnhildur Halldórsdóttir, and Tzu Wei Wang.

References

- Bauer, P. J., Hättenschwiler, N., & Larkina, M. (2016). "Own-ing" the personal past: Adolescents' and adults' autobiographical narratives and ratings of memories of recent and distant events. *Memory*, 24(2), 165–183. <https://doi.org/10.1080/09658211.2014.995673>
- Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11(1), 1–29. <https://doi.org/10.2167/beb392.0>
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Erlbaum.
- Bliss, L. S., & McCabe, A. (2008). Personal narratives—Cultural differences and clinical implications. *Topics in Language Disorders*, 28(2), 162–177. <https://doi.org/10.1097/01.TLD.0000318936.31677.2d>
- Buac, M., & Jarzynski, R. (2022). Providing culturally and linguistically responsive language assessment services for multi-lingual children with developmental language disorder: A scoping review. *Current Developmental Disorders Reports*, 9(4), 204–212. <https://doi.org/10.1007/s40474-022-00260-6>

- Chen, Y., McAnally, H. M., Wang, Q., & Reese, E. (2012). The coherence of critical event narratives and adolescents' psychological functioning. *Memory*, 20(7), 667–681. <https://doi.org/10.1080/09658211.2012.693934>
- Davidson, A. J., Walton, M. D., Kansal, B., & Cohen, R. (2017). Narrative skills predict peer adjustment across elementary school years. *Social Development*, 26(4), 891–906. <https://doi.org/10.1111/sode.12219>
- Einarsdóttir, J. T., & Þráinsdóttir, E. (2023). Personal narratives of 10-year-old children in Iceland: Verbal productivity, word diversity, and topics. *Folia Phoniatrica et Logopaedica*, 75(6), 372–381. <https://doi.org/10.1159/000533134>
- Ferman, S., & Kavar, K. (2023). Tele-assessment of oral personal narratives in Arabic- and Hebrew-speaking children using the Global TALES protocol. *Folia Phoniatrica et Logopaedica*, 75(6), 456–469. <https://doi.org/10.1159/000533833>
- Fivush, R., McDermott Sales, J., & Bohanek, J. G. (2008). Meaning making in mothers' and children's narratives of emotional events. *Memory*, 16(6), 579–594. <https://doi.org/10.1080/09658210802150681>
- Foldager, M., Simonsen, E., Lassen, J., Petersen, L. S., Oranje, B., Aggnaes, B., & Vestergaard, M. (2024). Narrative coherence and mentalizing complexity are associated in fictive storytelling and autobiographical memories in typically developing children and adolescents. *Cognitive Development*, 71, Article 101484. <https://doi.org/10.1016/j.cogdev.2024.101484>
- Gabaj, M., Kuvač Kraljević, J., & Westerveld, M. F. (2025a). Contribution of memory mechanisms and socio-emotional functioning to the production of personal narratives in children with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research*, 68(10), 4845–4864. https://doi.org/10.1044/2025_JSLHR-25-00047
- Gabaj, M., Kuvač Kraljević, J., & Westerveld, M. F. (2025b). Linguistic organisation and coherence in personal narratives of 10-year-old children with developmental language disorder. *First Language*, 45(2), 128–151. <https://doi.org/10.1177/01427237241298163>
- Gabaj, M., Kuvač Kraljević, J., & Westerveld, M. F. (2025c). The personal narratives of Croatian-speaking children with developmental language disorder: Investigating the influence of emotional valence at linguistic, propositional, and macrostructure levels. *Language, Speech, and Hearing Services in Schools*, 56(3), 617–632. https://doi.org/10.1044/2025_LSHSS-24-00111
- Gupta, V., Malone, S., Dipper, L., & Westerveld, M. F. (in press). Assessment of personal narrative skills in children aged 4 to 18: A systematic review. *Language, Speech, and Hearing Services in Schools*.
- Kupersmitt, J. R. (2015). Moving on and off the timeline in narratives of personal experience written in three languages: A form-function developmental study. *Linguistics*, 53(6), 1303–1339. <https://doi.org/10.1515/ling-2015-0033>
- Kuvač Kraljević, J., Matic Škorić, A., & Gabaj, M. (2023). Personal narratives of school-age children: A cross-sectional developmental study. *Folia Phoniatrica et Logopaedica*, 75(6), 412–430. <https://doi.org/10.1159/000533399>
- Labov, W., & Waletzky, J. (1997). Narrative analysis: Oral versions of personal experience. *Journal of Narrative and Life History*, 7(1–4), 3–38. <https://doi.org/10.1075/jnlh.7.02nar>
- Lai, W.-F. (2020). Narrative structure and evaluative language of personal narratives shared by urban aboriginal children in Taiwan. *Early Child Development and Care*, 190(14), 2214–2228. <https://doi.org/10.1080/03004430.2019.1566229>
- Lyons, R., Antonijevic-Elliott, S., Barbotin, S., Molloy, M., O'Malley-Keighran, M. P., Spelman, J., & Westerveld, M. F. (2023). Feasibility of using the Global TALES protocol to elicit personal narratives in 10 year-old children in Ireland. *Folia Phoniatrica et Logopaedica*, 75(6), 393–400. <https://doi.org/10.1159/000533140>
- McCabe, A., & Chang, C.-j. (2013). *Chinese language narration: Culture, cognition, and emotion*. John Benjamins. <https://doi.org/10.1075/sin.19>
- McGregor, K. K. (2020). How we fail children with developmental language disorder. *Language, Speech, and Hearing Services in Schools*, 51(4), 981–992. https://doi.org/10.1044/2020_LSHSS-20-00003
- Michelson, S. (2023). Children's agency when experiencing family-related adversities: The negotiation of closeness and distance in children's personal narratives. *Qualitative Social Work*, 22(2), 286–301. <https://doi.org/10.1177/14733250211066823>
- Miller, J. F., Andriacchi, K., Nockerts, A., Westerveld, M., & Gillon, G. (2020). *Assessing language production using SALT software: A clinician's guide to language sample analysis* (New Zealand - Australia version, 3rd ed.). SALT Software.
- Miller, J. F., Gillon, G., & Westerveld, M. (2020). *Systematic Analysis of Language Transcripts (SALT)* (New Zealand/Australia Version 20) [Computer software]. SALT Software.
- Minami, M., & McCabe, A. (1991). Haiku as a discourse regulation device: A stanza analysis of Japanese children's personal narratives. *Language in Society*, 20(4), 577–599. <https://doi.org/10.1017/S0047404500016730>
- Mitchell, C., Reese, E., Salmon, K., & Jose, P. (2020). Narrative coherence, psychopathology, and wellbeing: Concurrent and longitudinal findings in a mid-adolescent sample. *Journal of Adolescence*, 79(1), 16–25. <https://doi.org/10.1016/j.adolescence.2019.12.003>
- Mossige, S., Jensen, T. K., Gulbrandsen, W., Reichelt, S., & Tjersland, O. A. (2005). Children's narratives of sexual abuse: What characterizes them and how do they contribute to meaning-making? *Narrative Inquiry*, 15(2), 377–404. <https://doi.org/10.1075/ni.15.2.09mos>
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2019). *PIRLS 2021 assessment frameworks*. International Association for the Evaluation of Educational Achievement; TIMSS & PIRLS International Study Centre. https://pirls2021.org/wp-content/uploads/sites/2/2019/04/P21_Frameworks.pdf [PDF]
- Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. Plenum.
- Piccardo, E. (2018). Plurilingualism: Vision, conceptualization, and practices. In P. Trifonas & T. Aravossitas (Eds.), *Handbook of research and practice in heritage language education* (pp. 207–225). Springer. https://doi.org/10.1007/978-3-319-44694-3_47
- Preece, A. (1987). The range of narrative forms conversationally produced by young children. *Journal of Child Language*, 14(2), 353–373. <https://doi.org/10.1017/S0305000900012976>
- Reese, E., Haden, C. A., Baker-Ward, L., Bauer, P., Fivush, R., & Ornstein, P. A. (2011). Coherence of personal narratives across the lifespan: A multidimensional model and coding method. *Journal of Cognition and Development*, 12(4), 424–462. <https://doi.org/10.1080/15248372.2011.587854>
- Reese, E., Myftari, E., McAnally, H. M., Chen, Y., Neha, T., Wang, Q., Jack, F., & Robertson, S. J. (2017). Telling the tale and living well: Adolescent narrative identity, personality traits, and well-being across cultures. *Child Development*, 88(2), 612–628. <https://doi.org/10.1111/cdev.12618>
- Srivastava, V., Chan, A., & Westerveld, M. F. (2023). What do children in India talk about? Personal narratives of typically developing Hindi-speaking children. *Folia Phoniatrica et Logopaedica*, 75(6), 447–455. <https://doi.org/10.1159/000534298>

- Theodorou, E., Vogindroukas, I., Giannakou, E., Tsouti, L., & Phiniketos, I. (2023). Telling personal narratives: Comparing stories told by 10-year-old speakers of Cypriot Greek Dialect and of Standard Modern Greek. *Folia Phoniatrica et Logopaedica*, 75(6), 401–411. <https://doi.org/10.1159/000533400>
- Vanaken, L., Bijttebier, P., Fivush, R., & Hermans, D. (2022). Narrative coherence predicts emotional well-being during the COVID-19 pandemic: A two-year longitudinal study. *Cognition and Emotion*, 36(1), 70–81. <https://doi.org/10.1080/02699931.2021.1902283>
- Vanaken, L., & Hermans, D. (2021). Be coherent and become heard: The multidimensional impact of narrative coherence on listeners' social responses. *Memory & Cognition*, 49(2), 276–292. <https://doi.org/10.3758/s13421-020-01092-8>
- Van Vreckem, C., Desoete, A., Simoens, D., Van de Vyver, A., Pauwels, J., Van Laethem, C., & Van Lierde, K. (2023). Vocabulary diversity in personal narratives produced in response to the Global TALEs protocol in Dutch-speaking students with and without dyslexia. *Folia Phoniatrica et Logopaedica*, 75(6), 470–479. <https://doi.org/10.1159/000532024>
- Vygotsky, L. S. (1978). *Mind and society: The development of higher psychological processes*. Harvard.
- Wang, Q. (2016). Remembering the self in cultural contexts: A cultural dynamic theory of autobiographical memory. *Memory Studies*, 9(3), 295–304. <https://doi.org/10.1177/1750698016645238>
- Wang, Q., & Fivush, R. (2005). Mother–child conversations of emotionally salient events: Exploring the functions of emotional reminiscing in European-American and Chinese families. *Social Development*, 14(3), 473–495. <https://doi.org/10.1111/j.1467-9507.2005.00312.x>
- Wang, Q., & Leichtman, M. D. (2000). Same beginnings, different stories: A comparison of American and Chinese children's narratives. *Child Development*, 71(5), 1329–1346. <https://doi.org/10.1111/1467-8624.00231>
- Westby, C. E., Chen, K.-M., Lee, J. P., & Wong, A. M.-Y. (2023). Topic and content of personal narratives of children from three East Asian cultures and three English-speaking cultures: A collaborative qualitative analysis. *Folia Phoniatrica et Logopaedica*, 75(6), 431–446. <https://doi.org/10.1159/000533559>
- Westerveld, M. F. (2024, November 25–28). *Personal narrative coherence of school-age autistic children in response to the Global TALEs protocol* [Paper presentation]. Early Childhood Voices 2024, Online.
- Westerveld, M. F., Lyons, R., Nelson, N. W., Chen, K. M., Claessen, M., Ferman, S., Fernandes, F. D. M., Gillon, G. T., Kavar, K., Kuvač Kraljević, J., Petinou, K., Theodorou, E., Tumanova, T., Vogandroukas, I., & Westby, C., on behalf the Global TALEs Consortium. (2022). Global TALEs feasibility study: Personal narratives in 10-year-old children around the world. *PLOS ONE*, 17(8), Article e0273114. <https://doi.org/10.1371/journal.pone.0273114>
- Westerveld, M. F., & Nelson, N. (2023). Evaluating children's personal narrative skills using the Global TALEs protocol: Implications for practice. *Folia Phoniatrica et Logopaedica*, 75(6), 367–371. <https://doi.org/10.1159/000534510>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3(1), Article 160018. <https://doi.org/10.1038/sdata.2016.18>
- Yaşar-Gündüz, E. (2024). *Evaluating the personal narrative skills of typically developing 7–10-year-old monolingual Turkish-speaking children through Global TALEs-TR* [Unpublished doctoral thesis]. Anadolu University.

Appendix

Narrative Coherence Coding Scheme Coding Criteria (summarized from Reese et al., 2011)

Only responses that contain a least two restricted narrative clauses, defined as clauses that refer to past personally experienced events that can be put on a timeline (Labov & Waletzky, 1997), were considered personal narratives and coded for coherence.

Criteria	Context	Chronology	Theme
0	No information	2 clauses (or more) that could be ordered chronologically, but no information is provided about temporal order	The story constantly deviates from the topic and/or is characterized by numerous digressions that make it difficult to define the topic. No attempt at self-correction.
1	Mention of time or place	2 clauses in correct order. If the story has 3 or more clauses, at least 50% are in order.	The topic can be defined, and most utterances are consistent with the topic. However, elaboration of the topic through references to causes or personal judgments or elaboration of actions is minimal.
2	Time and place are mentioned, but only one is specific.	3 clauses: all in order. More than 3 clauses: > 75% are in order (not all).	The story substantially develops the topic, and the story contains at least two of the following (one of each type): causal relationship (explicitly expressed), personal evaluation/interpretation or reaction, or elaboration of action, or resolution (that does not have to provide new information).
3	Both time and place are mentioned and specific.	More than 3 clauses—almost all in order, with minimal deviation.	The story substantially develops the topic and includes all of the above plus a resolution to the story or a connection to autobiographical experiences, including future events, self-concept, or identity. The story as a whole must be on topic and developed for the emotional point to be clear.