



City Research Online

City St George's, University of London

Citation: Harvey, A., Spicer-Cain, H., Botting, N., Ryan, G. & Henry, L. (2023). Assessing 'coherence' in the spoken narrative accounts of autistic people: A systematic scoping review. *Research in Autism Spectrum Disorders*, 102, 102108. doi: 10.1016/j.rasd.2023.102108

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/29651/>

Link to published version: <https://doi.org/10.1016/j.rasd.2023.102108>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Research in Autism Spectrum Disorders

journal homepage: www.elsevier.com/locate/rasd

Assessing ‘coherence’ in the spoken narrative accounts of autistic people: A systematic scoping review

Anna Harvey^{*}, Helen Spicer-Cain, Nicola Botting, Gemma Ryan, Lucy Henry

Department of Language and Communication Science, City, University of London, Northampton Square, London EC1V 0HB, United Kingdom

ARTICLE INFO

Keywords:

Autism
Narrative
Coherence
Macrostructure
Story grammar

ABSTRACT

Background: The ability to produce a well-structured, coherent narrative account is essential for successful everyday communication. Research suggests that autistic people may find this challenging, and that narrative assessment can reveal pragmatic difficulties in this population that are missed on sentence-level tasks. Previous studies have used different methodologies to assess spoken narrative skills in autism. This review systematically examined these approaches and considered their utility for assessing narrative coherence.

Method: Keyword database searches were conducted, with records screened by two independent reviewers. Eligible studies (n = 59) included specified frameworks for evaluating structure/coherence in spoken narrative accounts by autistic participants of any age. Studies were categorised according to the type of narrative scoring scheme used, and strengths and limitations were considered.

Results: Over 80% of included articles reported observational cross-sectional studies, with participants generally matched on age and cognitive ability with non-autistic comparison groups. The most common approaches involved coding key elements of narrative structure (‘story grammar’) or scoring the inclusion of pre-determined ‘main events’. Alternative frameworks included ‘holistic’ rating scales and subjective quality judgements by listeners. Some studies focused specifically on ‘coherence’, measuring diverse aspects such as causal connectedness and incongruence. Scoring criteria varied for each type of framework.

Conclusions: Findings indicated that solely assessing story structure ignores important features contributing to the coherence of spoken narrative accounts. Recommendations are that future research consider the following elements: (1) context, (2) chronology, (3) causality, (4) congruence, (5) characters (cognition/emotion), and (6) cohesion; and scoring methods should include rating scales to obtain sufficiently detailed information about narrative quality.

1. Introduction

Oral narrative is a complex form of discourse in which the speaker recounts a series of events to an audience, drawing upon a range of linguistic, cognitive and pragmatic abilities (Kunnari et al., 2016; Norbury et al., 2014). Narrative skills are fundamental to human communication, social development and learning across numerous contexts, and it is well established that they are essential for both

^{*} Corresponding author.

E-mail addresses: anna.harvey@city.ac.uk (A. Harvey), helen.spicer-cain.2@city.ac.uk (H. Spicer-Cain), nicola.botting.1@city.ac.uk (N. Botting), gemma.ryan@city.ac.uk (G. Ryan), lucy.henry.1@city.ac.uk (L. Henry).

<https://doi.org/10.1016/j.rasd.2023.102108>

Received 25 July 2022; Received in revised form 1 December 2022; Accepted 10 January 2023

Available online 24 January 2023

1750-9467/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Table 1
Study characteristics.

First author	Year	Country	Study design	Groups	Matching criteria	Type of narrative task	Total sample size [no. autistic participants]	Age group	M:F autistic participants	Analytical approach
Banney	2015	Australia	Between-groups observational study	ASD; TD	Age; non-verbal ability; receptive and expressive language; visual and auditory attention	Storybook task	28 [11]	Child; Adolescent	M>F	Story grammar
Beytollahi	2020	Iran	Between-groups observational study	ASD; TD	None	Narrative retell	57 [7]	Child	All male	Story grammar
Birri	2016	USA	Single subject multiple baseline intervention study	N/A	N/A	Personal narrative	4 [4]	Young adult	All male	Story grammar
Brown	2007	USA	Between-groups observational study	Asperger Syndrome; TD	Age	Personal narrative	57 [30]	Child; Adolescent	All male	Coherence (context, chronology, theme)
Canfield	2016a	USA	Between-groups observational study	ASD, TD	Age; sex, full-scale IQ; receptive language	Storybook task	30 [15]	Adolescent	M>F	Story grammar / Subjective story quality ratings
Canfield	2016b	USA	Between-groups observational study	ASD; ASD-"optimal outcomes"; TD	Age, full-scale IQ	Storybook task	45 [15]	Child; Adolescent	M>F	Story grammar / Subjective story quality ratings
Carlsson	2020	Sweden	Between-groups observational study	ASD; TD (age matched); TD (language-matched)	Age (age-matched group); receptive and expressive language (language-matched group)	Narrative retell	119 [45]	Child	M>F	Holistic macrostructure
Colozzo	2015	Canada	Between-groups observational study	ASD; Specific Language Impairment (SLI); TD	Age; maternal education	Picture narration task	36 [12]	Child	M>F	Story grammar
Conlon	2019	Canada	Between-groups observational study	ASD (girls); ASD (boys)	Age; non-verbal IQ; receptive and expressive language	Storybook task	26 [13]	Child	M=F	Main events
de Marchena	2010	USA	Between-groups observational study	ASD; TD	Age; sex; full-scale IQ	Storybook task	30 [15]	Adolescent	M>F	Subjective story quality ratings
Diehl	2006	USA	Between-groups observational study	ASD; TD	Age; sex; receptive and expressive language; verbal reasoning	Narrative retell	34 [17]	Child; Adolescent	M>F	Coherence (causal connectedness)
Engberg-Pedersen	2017	Denmark	Between-groups observational study	ASD; Language Impairment (LI); TD	Age; non-verbal IQ; (ASD /TD groups also matched on receptive and expressive language and sentence repetition)	Storybook task	69 [27]	Child; Adolescent	M>F	Main events
Engel	2018	USA	Matched pairs intervention study (random assignment)	Intervention; Control	Age; word identification ability	Narrative retell	20 [10]	Child	M>F	Story grammar
Estigarribia	2011	USA	Between-groups observational study	Fragile X without ASD (FX-O); Fragile X with ASD (FX-ASD); Down's Syndrome (DS); TD	Mental age	Narrative retell	129 [28]	Child; Adolescent	All male	Story grammar

(continued on next page)

Table 1 (continued)

First author	Year	Country	Study design	Groups	Matching criteria	Type of narrative task	Total sample size [no. autistic participants]	Age group	M:F autistic participants	Analytical approach
Favot	2018	Australia	Multiple baseline across participants intervention study	N/A	N/A	Narrative retell	4 [4]	Child	M=F	Story grammar/ Subjective story quality ratings
Favot	2019a	Australia	Single subject (AB) intervention study	N/A	N/A	Picture narration	1 [1]	Child	All male	Story grammar
Favot	2019b	Australia	Multiple baseline across participants intervention study	N/A	N/A	Personal narrative	3 [3]	Child	M>F	Story grammar/ Subjective story quality ratings
Favot	2021a	Australia	Multiple probe across participants intervention study	N/A	N/A	Picture narration	4 [4]	Child	All male	Story grammar/ Subjective story quality ratings
Favot	2021b	Australia	Multiple baseline with probe across participants intervention study	N/A	N/A	Personal narrative	4 [4]	Child	M=F	Story grammar/ Subjective story quality ratings
Ferretti	2018	Italy	Between-groups observational study	ASD; TD	Age; non-verbal ability	Picture narration	132 [66]	Child	M>F	Coherence (causal connectedness/ incongruence)
Geelhand	2020	Belgium	Between-groups observational study	ASD; TD	Age and sex (pairwise); full-scale IQ, verbal IQ; perceptual IQ (groups)	Storybook task	36 [18]	Adolescent; Adult	M>F	Main events
Gillam	2015	USA	Multiple baseline across participants intervention study	N/A	N/A	Picture narration	5 [5]	Child; Adolescent	M>F	Story grammar
Goldman	2008	USA	Between-groups observational study	ASD; Developmental Language Disorder (DLD); TD	Age; non-verbal ability	Personal narrative	38 [14]	Child; Adolescent	M>F	Story grammar
Henry	2020	UK	Between-groups observational study	ASD; TD	Age; IQ; receptive language	Witness narrative	104 [52]	Child	M>F	Story grammar
Hilvert	2016	USA	Between-groups observational study	ASD; TD	Non-verbal reasoning, receptive language	Narrative retell	45 [19]	Child; Adolescent	M>F	Holistic macrostructure
Hogan-Brown	2013	USA	Between-groups observational study	ASD; FX-ASD; FX-O; DS; TD	Receptive and expressive language; non-verbal mental age	Storybook task	94 [43]	Child; Adolescent	All male	Main events
Iandolo	2020	Spain	Between-groups observational study	ASD; TD	Verbal age (based on IQ); sex; SES; level of narrative cohesion (Bears Family Projective Test)	Story generation (with props)	50 [25]	Child; Adolescent	All male	Story grammar
Johnels	2013	Sweden	Cross-sectional observational study	N/A	N/A	Narrative retell	55 [21]	Child	M>F	Main events
Kauschke	2016	Germany	Between-groups observational study	ASD (girls); ASD (boys); TD (girls)	Age; verbal IQ, performance IQ, full-scale IQ	Storybook task	33 [11]	Child; Adolescent	M=F	Coherence (global orientation)
Kenan	2019	Israel	Between-groups observational study	ASD; TD	Age; verbal and non-verbal ability	Storybook task	48 [24]	Child	All male	Main events
King	2014	UK	Between-groups observational study	ASD; TD (age-matched); TD	Age; non-verbal ability (age-matched group) / expressive language and	Picture narration	81 [27]	Adolescent	Not reported	Holistic macrostructure

(continued on next page)

Table 1 (continued)

First author	Year	Country	Study design	Groups	Matching criteria	Type of narrative task	Total sample size [no. autistic participants]	Age group	M:F autistic participants	Analytical approach
				(language-matched)	non-verbal ability (language-matched group)					
King	2018	UK	Between-groups observational study	ASD; TD	Age; non-verbal ability	Narrative retell; Picture narration	28 [14]	Adolescent	All male	Holistic macrostructure
Lee	2018	USA	Between-groups observational study	ASD; TD	None	Picture narration	33 [19]	Adolescent; Adult	M>F	Story grammar
Lind	2014	UK	Between-groups observational study	ASD; TD	Age; verbal and non-verbal IQ ability	Storybook task	56 [27]	Adult	M>F	Story grammar/Main events
Losh	2003	USA	Between-groups observational study	ASD; TD	Age; verbal IQ	Storybook task; Personal narrative	50 [28]	Child; Adolescent	Not reported	Story grammar/Main events
Loveland	1990	USA	Between-groups observational study	ASD; DS	Age and composite verbal age equivalent	Witness narrative	32 [16]	Child; Adolescent; Adult	M>F	Main events
Mäkinen	2014	Finland	Between-groups observational study	ASD; TD	Age	Storybook task	32 [16]	Child	M>F	Main events
Manolitsi	2011	Greece	Between-groups observational study	ASD; SLI	Age; sex; non-verbal IQ	Narrative retell	26 [13]	Child; Adolescent	M>F	Holistic macrostructure
Marini	2019	Italy	Between-groups observational study	ASD; TD	Age; non-verbal ability; level of formal education	Picture narration	154 [77]	Child	M>F	Holistic macrostructure
Marini	2020	Italy	Between-groups observational study	ASD; TD	Age; non-verbal ability; level of formal education	Storybook task	74 [24]	Child	Not reported	Coherence (incongruence)
McCabe	2013	USA	Between-groups observational study	ASD; TD	Age; sex; reading ability	Personal narrative	34 [16]	Young adult	M>F	High-point analysis
McCabe	2017	USA	Randomised controlled trial (pilot)	N/A	None	Personal narrative	10 [10]	Adolescent; Adult	M>F	High-point analysis
Miniscalco	2007	Sweden	Cross-sectional observational study	N/A	N/A	Narrative retell	21 [5]	Child	M>F	Main events
Norbury	2014	USA/UK	Between-groups observational study	ASD; SLI; TD	Age; non-verbal ability; (TD and ASD groups also matched on receptive and expressive language)	Storybook task	89 [25]	Child; Adolescent	M>F	Main events/ Holistic macrostructure
Norbury	2003	UK	Between-groups observational study	ASD; SLI, Pragmatic Language Impairment (PLI); TD	Age; non-verbal ability; (ASD, SLI and PLI groups also matched on receptive and expressive language)	Storybook task	68 [12]	Child	Not reported	Main events/ Story grammar
Pearlman-Avnion	2002	Israel	Between-groups observational study	ASD; Williams Syndrome (WS); TD (two age groups)	ASD and WS groups were matched on verbal comprehension (Comprehension, Vocabulary, and Similarities scores, WISC-R)	Narrative retell	52 [13]	Child; Adolescent; Young Adult	M>F	Main events/ Story grammar

(continued on next page)

Table 1 (continued)

First author	Year	Country	Study design	Groups	Matching criteria	Type of narrative task	Total sample size [no. autistic participants]	Age group	M:F autistic participants	Analytical approach
Peristeri	2017	Greece	Between-groups observational study	ASD (higher range language skills - HL); ASD (lower range language skills - LL); TD	Age; performance IQ; (TD and ASD-HL also matched on expressive vocabulary and verbal IQ)	Narrative retell	45 [30]	Child; Adolescent	All male	Story grammar
Peristeri	2020	Greece	Between-groups observational study	ASD (monolingual); ASD (bilingual); TD (monolingual); TD (bilingual)	Age	Narrative retell	80 [20]	Child; Adolescent	All male	Story grammar
Persson	2006	Sweden	Cross-sectional observational study	N/A	N/A	Narrative retell	19 [5]	Child	M>F	Main events
Petersen	2014	USA	Single subject multiple baseline intervention study	N/A	N/A	Personal narrative	3 [3]	Child	All male	Story grammar
Rollins	2014	USA	Within-subjects observational study	N/A	N/A	Storybook task; Personal narrative	10 [10]	Young adult	M>F	Holistic macrostructure
Rumpf	2012	Germany	Between-groups observational study	Aspergers Syndrome, ADHD, TD	Age; IQ; mean length of utterance	Storybook task	31 [11]	Child; Adolescent	All male	Coherence (global orientation)
Sah	2015	Taiwan	Between-groups observational study	ASD; TD	Full-scale IQ; receptive/ expressive language	Storybook task	36 [18]	Child	All male	Coherence (causal connectedness)
Tager-Flusberg	1995	USA	Between-groups observational study	ASD; TD; intellectual disability (ID)	Verbal mental age.; (ASD and ID groups also matched on chronological age)	Storybook task	30 [10]	Child; Adolescent	M>F	Story grammar
Veneziano	2019	France	Between-groups observational study	ASD; TD	Age; sex; verbal age	Storybook task	26 [13]	Child	M>F	Main events
Volden	2017	Canada	Cross-sectional observational study	N/A	N/A	Narrative retell	74 [74]	Child	M>F	Main events
Westerveld	2017	Australia	Cross-sectional observational study	N/A	N/A	Narrative retell	29 [29]	Preschooler	Not reported	Main events
Young	2005	USA	Between-groups observational study	ASD; TD	Age; sex; language; Verbal IQ	Storybook task	34 [17]	Child; Adolescent	M>F	Story grammar
Zielinski Deane	2016	USA	Randomised controlled trial	N/A	Age; IQ; sex; ADOS scores	Storybook task	73 [73]	Child; Adolescent	M>F	Main events

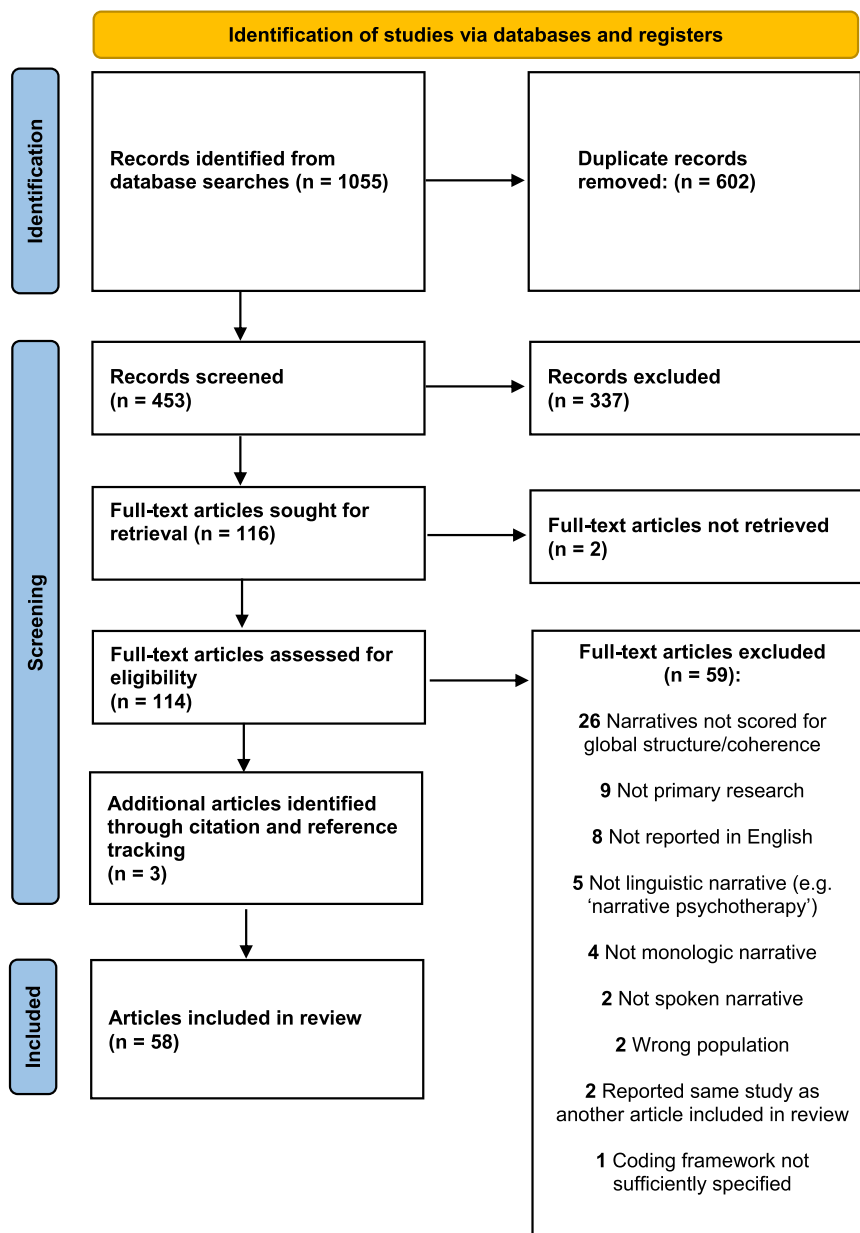


Fig. 1. PRISMA flow diagram.

social and academic success (Petersen et al., 2008). However, the pragmatic use of language in social contexts can be an area of challenge for autistic individuals, and this is reflected in current diagnostic criteria for autism (American Psychiatric Association, 2013). Although early language difficulties associated with autism typically resolve or lessen over time, difficulties with pragmatics tend to be more pervasive (Rapin & Dunn, 2003). Standardised language assessments, which often target sentence-level skills, may therefore not provide a full picture of functional communicative abilities in this population. Previous research in this area has demonstrated that narrative tasks can reveal higher order difficulties with extended discourse that may not be apparent when traditional language assessments are used (Botting, 2002; King & Palikara, 2018; Manolitsi & Botting, 2011; Volden et al., 2017). Spoken narrative generation is therefore considered to be both a more sensitive and ecologically valid form of assessment for autistic people who demonstrate language abilities in the typical range (Volden et al., 2017).

Research into the spoken language skills of autistic people has investigated a range of narrative features, with structural abilities typically examined at two levels: microstructure and macrostructure. Microstructural analysis tends to focus on linguistic form and content at the level of individual utterances, whereas macrostructural analysis considers the overall organisation of the narrative (Heilmann et al., 2010). To produce a coherent account, a narrator must “temporally and causally organize a narrative into a sequence

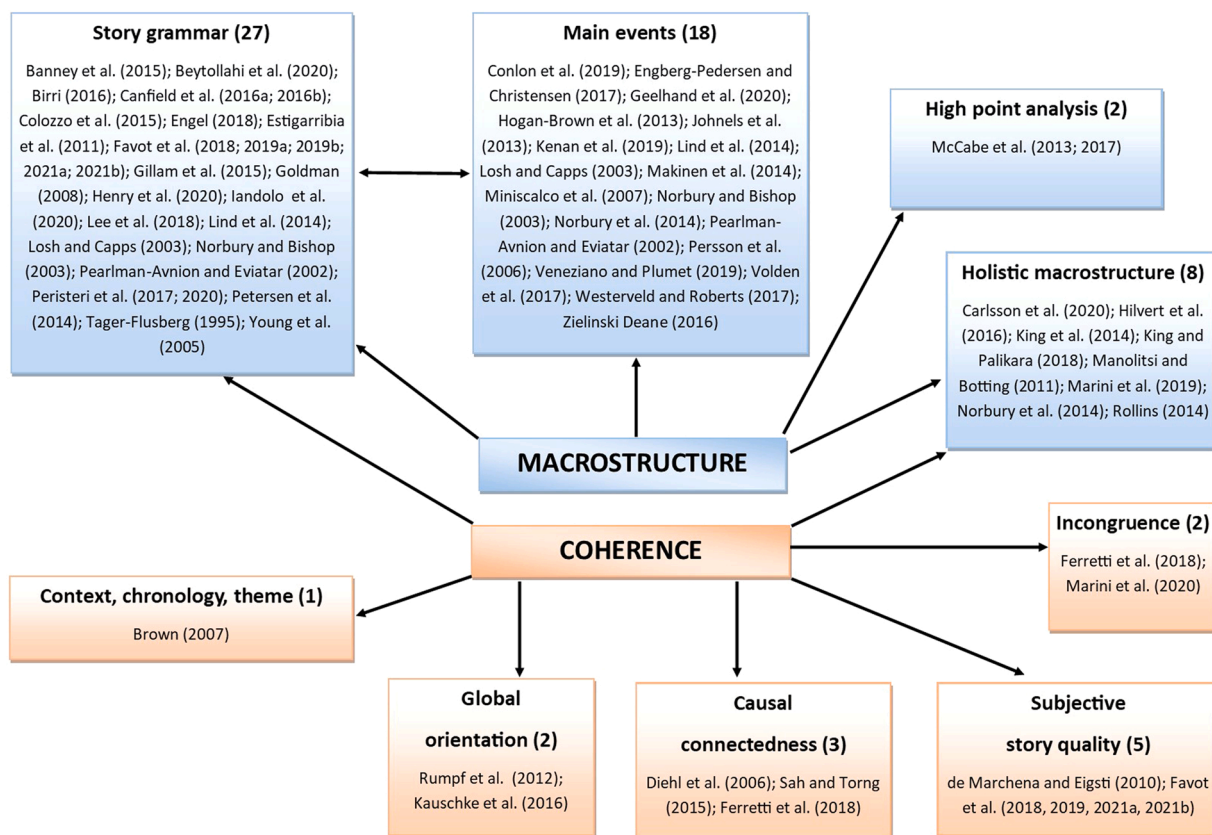


Fig. 2. Map of approaches to analysing narrative structure and coherence.

that is meaningful to themselves and their listeners. That is, all the parts of the story must be structured so that the entire sequence of events is interrelated in a meaningful way” (Shapiro & Hudson, 1991, p. 960). Macrostructure is therefore closely linked with the concept of narrative coherence.

Previous studies in the autism literature have produced a complex set of findings, with conflicting evidence about various aspects of narrative skill across a range of experimental tasks (Baixauli et al., 2016). Research has mostly focused on narrative production by autistic children and adolescents, although there are some studies with adult samples (e.g., McCabe et al., 2013; Rollins, 2014; Lee et al., 2018; Geelhand et al., 2020). In terms of microstructure, findings have included increased morphosyntactic errors in autistic children’s narratives when compared to non-autistic peers (Kuijper et al., 2017); reduced use of complex syntax (Banney et al., 2015; Capps et al., 2000; Losh & Capps, 2003; Norbury & Bishop, 2003); difficulties with referential language and pronouns (Banney et al., 2015; Manolitsi & Botting, 2011; Novogrodsky, 2013; Suh et al., 2014) and reduced vocabulary (King & Palikara, 2018). With reference to macrostructure, several studies have found that autistic people tend to produce shorter (King et al., 2013; Rumpf et al., 2012; Siller et al., 2014) and more simplistic narrative accounts (Banney et al., 2015; Kuijper et al., 2017; Rumpf et al., 2012). Autistic groups have frequently been found to use less causal language to link story events (Diehl et al., 2006; Hilvert et al., 2016; Kuijper et al., 2017; Losh & Capps, 2003; Sah & Torng, 2015); fewer cohesive devices (Hilvert et al., 2016; Kuijper et al., 2017) and fewer instances of internal state language (Kauschke et al., 2016; King & Palikara, 2018; Pearlman-Avniion & Eviatar, 2002; Rumpf et al., 2012). There is also evidence for difficulties with higher-order narrative structure (e.g., Goldman, 2008; Manolitsi & Botting, 2011; Suh et al., 2014; Volden et al., 2017; Kenan et al., 2019).

In contrast, research groups have also found minimal or no differences between autistic and non-autistic groups in terms of story length (Capps et al., 2000; Kauschke et al., 2016; Losh & Capps, 2003; Norbury & Bishop, 2003; Sah & Torng, 2019); morphosyntactic errors (Capps et al., 2000; Suh et al., 2014), syntactic complexity (Diehl et al., 2006; Rumpf et al., 2016; Peristeri et al., 2017; Sah & Torng, 2019); cohesion (Kauschke et al., 2016); lexical diversity (Kauschke et al., 2016; Suh et al., 2014), causal language (Capps et al., 2000; King et al., 2013; Suh et al., 2014); use of evaluative devices, use of internal state language (Canfield, Eigsti, de Marchena, & Fein, 2016; Capps et al., 2000; Suh et al., 2014); ability to structure an account around key story elements (Henry et al., 2020; Norbury & Bishop, 2003) and overall coherence (Kauschke et al., 2016).

Many of these previous studies had small sample sizes, and the inconsistency in results is not yet reconciled. Differences in both the choice of comparison group, and how these study participants are matched to the autistic group, further complicate the picture. However, existing evidence appears to indicate that although autistic individuals can perform similarly to non-autistic participants on many microstructural measures, they tend to have more difficulties with macrostructure, resulting in less coherent spoken narrative

Table 2
Story grammar elements included across studies.

First author and year	Formal opening	Characters/ persons /'who' or 'who with'	Setting (physical /location / time)	Initiating event/ problem/ obstacle	Internal response/ emotional state	Plan/ goal	Action/ attempt	Consequence/ outcome (direct or long-term)	Main story events/ episodes/ 'what happened'	Reaction	Ending emotion	Ending/ resolution/ conclusion	Formal ending
Banney (2015)		✓	✓	✓	✓	✓	✓	✓					
Beytollahi (2020)		✓	✓	✓	✓	✓	✓	✓					
Birri (2016)			✓	✓	✓		✓	✓			✓		
Canfield (2016a)				✓			✓	✓					
Canfield (2016b)				✓			✓	✓					
Colozzo (2015)			✓	✓	✓		✓	✓		✓			
Engel (2018)		✓	✓	✓	✓		✓		✓			✓	
Estigarribia (2011)		✓	✓	✓	✓		✓					✓	
Favot (2018)		✓	✓	✓	✓		✓	✓				✓	
Favot (2019a)		✓	✓	✓	✓							✓	
Favot (2019b)		✓	✓	✓	✓				✓			✓	
Favot (2021a)		✓	✓	✓	✓				✓			✓	
Favot (2021b)		✓	✓	✓	✓				✓			✓	
Gillam (2015)		✓	✓	✓	✓	✓	✓	✓					
Goldman (2008)	✓	✓	✓	✓			✓					✓	
Henry (2020)			✓	✓	✓	✓	✓	✓				✓	
Iandolo (2020)	✓	✓	✓	✓	✓			✓				✓	
Lee (2018)		✓	✓	✓	✓		✓	✓		✓			
Lind (2018)				✓					✓			✓	
Losh (2003)			✓	✓					✓			✓	
Norbury (2003)				✓					✓			✓	
Pearlman-Avniion (2002)			✓	✓	✓		✓	✓		✓			
Peristeri (2017)						✓	✓	✓					
Peristeri (2020)						✓	✓	✓					
Petersen (2014)			✓	✓	✓	✓	✓	✓			✓		
Tager-Flusberg (1995)	✓	✓	✓						✓			✓	✓
Young (2005)			✓	✓	✓	✓	✓	✓		✓			
Total no. of studies:	3	14	21	20	16	8	17	16	7	4	2	12	1
% of studies:	11%	52%	78%	74%	59%	30%	63%	59%	26%	15%	7%	44%	4%

accounts (Baixauli et al., 2016; Conlon et al., 2019; Hewitt, 2019; Volden et al., 2017). For this reason, the current review focused on measures relating to macrostructure.

Researchers have used a range of approaches to evaluate macrostructure in narratives produced by autistic people. These include well-established analytical frameworks such as 'story grammar' (Stein & Glenn, 1979), in which narratives are scored for the presence of key story elements, and 'high point analysis' (Labov, 1972), which considers how closely narrative accounts adhere to a prototypical story structure. However, many research groups investigating language skills in autistic individuals have proposed alternative frameworks for assessing structure and coherence in spoken narrative accounts. Despite a sizeable body of research evidence in this area, to our knowledge, there has been no attempt to systematically compare narrative assessment methodologies and consider their suitability for judging the coherence of accounts by autistic narrators. Preliminary database searches found no existing reviews focusing on narrative assessment methods used with autistic people, supporting the rationale for the present work. One meta-analysis investigating the narrative skills of autistic children and adolescents was retrieved (Baixauli et al., 2016). This paper discussed narrative macrostructure as a variable; however, since the focus of the article was on analysing research findings, the authors did not provide detailed information on the scoring frameworks used in different studies. Therefore, the present review provides a novel contribution to the literature on narrative skills in autism.

1.1. The current study

The key objectives of this review were:

1. To identify relevant frameworks used within the autism research literature to assess and score narrative macrostructure and/or coherence.
2. To categorise the identified narrative coding frameworks according to their overall theoretical approach, key features, and the type of scoring system used.
3. To consider the advantages and limitations of these different approaches for evaluating the coherence of spoken narrative accounts by autistic people (this objective was added following pre-registration).

2. Methods

A scoping review methodology was selected as the most suitable approach for mapping the breadth and variety of existing research evidence on this topic. Arksey and O'Malley (2005) suggest that scoping reviews can also be an effective means of clarifying working definitions and the conceptual boundaries of a topic; in this case, the concept of 'narrative coherence'. The present scoping review was approached systematically, so that relevant methodological concerns could be addressed in detail and the quality of evidence relating to narrative assessment in autism could be considered. The review was carried out following guidance proposed by the Joanna Briggs Institute's Manual for Evidence Synthesis (Aromataris and Munn, 2020) and reported according to the Preferred Reported Items for Systematic Reviews and Meta-Analyses - Extension for Scoping Reviews (Tricco et al., 2018). The review protocol was pre-registered on the Open Science Framework (<https://osf.io/d9jrf/>).

2.1. Eligibility criteria

This review included published and unpublished studies reported in English that investigated spoken narrative skills in individuals of any age with a diagnosis of autism spectrum disorder or Asperger's Syndrome. To ensure that all previous relevant research literature was included in this review, no restrictions were placed on the date of publication. Studies needed to focus on the analysis of monologic spoken narrative accounts (i.e., conversation/discourse analyses were excluded). Importantly, included papers also needed to include a specified framework for evaluating overall macrostructure and/or coherence. Papers that exclusively analysed micro-structural elements of narrative were excluded. Studies that scored narratives for individual elements of macrostructure, such as internal state language or causal connectives, were excluded if they did not consider the overall structure or coherence of the account.

2.2. Search strategy

Preliminary searches using a range of keywords relating to 'autism' with the terms 'narrative' and 'coherence' were carried out on 04/10/2021, across the following databases: CINAHL, Communication Source, Medline and PsychInfo (via EBSCOHost); and Amed, Embase, Cochrane databases and Ovid Emcare (via Ovid Online). A small sample of relevant articles was identified from these searches. The titles and abstracts of these papers were then analysed for additional keywords and synonyms that were incorporated into the proposed search terms for this review. A complete search including all the identified keywords was then carried out. 'Grey literature' sources, such as doctoral theses, were included.

Search terms.

1. Narrative* N5 (production or abilit* or skill* or performance or competence or quality or macrostructure or "macro structure" or structure or coheren* or element* or component*) or stor* N5 (production or abilit* or skill* or performance or competence or quality or macrostructure or "macro structure" or structure or coheren* or element* or component*) or storytelling or "story telling" or "high point" or "story grammar"

2. Autis* or “autis* spectrum disorder” or asd or “autis* spectrum condition” or asc or asperger*

The first author searched the following databases on 15/10/21: CINAHL, Communication Source, Medline and PsychInfo (via EBSCOHost); and Amed, Embase, Cochrane Databases and Ovid Emcare (via Ovid Online). These sources were selected with the aim of covering all key publications within the healthcare and psychology literature.

2.3. Selection process

Results from each database search were uploaded to RefWorks reference management software, with duplicates removed automatically at this stage. Article references with accompanying abstracts were then imported to the Rayyan systematic review web app (Ouzzani et al., 2016).

The first and fourth authors (AH, GR) independently screened the titles and abstracts of the retrieved articles for inclusion against the eligibility criteria, with 84.5% agreement achieved at this stage (Cohen’s $\kappa = 0.62$). Discrepancies that arose in the screening process were resolved through discussion. As a result of these discussions, the screening guidelines were revised to clarify the following points:

1. Studies must be concerned with ‘narrative’ in the sense of linguistic discourse; any other usages of the term (e.g., ‘narrative psychotherapy’) should be excluded.
2. Studies must report empirical research carried out by the authors, thereby excluding literature reviews and articles describing narrative intervention programmes with no research findings (e.g., intervention manuals).

Full-text versions of the articles that met the inclusion criteria were retrieved and reviewed by the two authors, using the same blinded methodology as in the previous stage. Two articles were unretrievable, due to insufficient identifying information in the database records. Inter-rater agreement for the full-text screening stage was 88.2% ($\kappa = 0.76$).

Ten of the included articles were hand-selected for citation and reference tracking (Brown, 2007; Diehl et al., 2006; Ferretti et al., 2018; Henry et al., 2020; Hilvert et al., 2016; Kauschke et al., 2016; Marini et al., 2020; Rollins, 2014; Rumpf et al., 2012; Sah & Torng, 2015). These studies were chosen as they aimed to evaluate narrative coherence specifically and were therefore more directly relevant to the research questions. The first author used SCOPUS to track citations and references for these articles. Three further papers were identified by this method (Johnels et al., 2013; Pearlman-Avni & Eviatar, 2002; Tager-Flusberg, 1995). The second reviewer was consulted and approved the inclusion of these studies in the review.

2.4. Data collection

Each included paper was individually appraised by the first author, and a data extraction form was completed for each study in Microsoft Excel. Missing information was entered as ‘not reported’.

The data items were agreed in advance by the research team. These included:

- General information (authors; title; year of publication; journal/DOI; country of study; language used)
- Participant characteristics (age; sex; diagnostic information)
- Methodology (experimental design; recruitment strategy; sample size; number of autistic participants; number of participants in each group; inclusion/exclusion criteria; diagnostic confirmation; comparison groups; matching criteria; type of narrative task used; reliability procedures and reporting; blinding procedures)
- Scoring framework (overall analytical approach; method for scoring narrative structure; method for scoring narrative coherence; any other details scored)
- Key study findings

2.5. Quality assessment

The quality of the included studies was assessed, to determine the strength of the existing evidence base in relation to the narrative abilities of autistic people. An adapted version of the Joanna Briggs Institute’s Critical Appraisal Checklist for Analytical Cross-Sectional Studies (Moola et al., 2017) was used. This checklist was modified to make it more relevant to the focus of the review, with all changes discussed and agreed in advance by the research team. Some of the questions were removed or altered to ensure that the checklist was appropriate for assessing all the included papers, and specific items were added relating to how autistic participants’ diagnostic status was verified, how participants’ narratives were scored, and whether adequate inter-rater reliability measures were used. The first and second reviewers independently appraised all included studies using the adapted checklist and achieved 84.7% inter-rater agreement overall, with levels ranging from 64.4%–100% across the nine checklist items. Quality appraisal ratings were added to a colour-coded table (see Appendix for a full version of the modified checklist with quality ratings).

2.6. Synthesis

A table summarising the key characteristics of the included studies was created (Table 1). Studies were then categorised according

to the type of narrative scoring scheme used, and a diagram was produced to map out the range of approaches used in the literature.

3. Results

3.1. Selection process

1055 records were identified and included in the selection process. Following the removal of duplicates and title and abstract screening, 114 full-text articles were retrieved and assessed for eligibility, with 55 articles meeting the inclusion criteria.¹ A further three articles were identified through citation and reference tracking, resulting in a total of 59 studies included in the final review (note that the paper by [Canfield et al., 2016](#), reports two separate studies carried out by the authors – these were designated [Canfield et al., 2016a](#); [2016b](#) for the purposes of this review). See PRISMA flow diagram for details of the selection process ([Fig. 1](#)).

3.2. Study Characteristics

The included studies covered 31 years of research (1990–2021) across 16 countries (Australia: 7, Belgium: 1, Canada: 3, Denmark: 1, Finland: 1, France: 1, Germany: 2, Greece: 3, Iran: 1, Israel: 2, Italy: 3, Spain: 1, Sweden: 4, Taiwan: 1, UK: 5 and USA: 23). Included in these figures is one study that was a UK-USA collaboration ([Norbury et al., 2014](#)). Most of the studies (64%) were carried out with English-speaking participants. There were four unpublished doctoral theses ([Birri, 2019](#); [Brown, 2007](#); [Engel, 2018](#); [Zielinski Deane, 2017](#)) and one book chapter ([Veneziano & Plumet, 2019](#)); the remaining 54 were peer-reviewed journal articles.

The 59 studies reported on a total of 1218 autistic individuals. The average sample size of autistic participants was $M(SD) = 20.64 (17.56)$, ranging from 1 to 77 participants across the studies. 19% of the studies had fewer than 10 autistic participants; 47% included a sample of between 10 and 20 autistic participants; and 22% included between 20 and 40 autistic participants. Only 12% of studies had a sample of more than 40 autistic participants. Studies involved participants of all ages, although children aged 5–11 years were represented most in the research (48 studies, 81%), while 31 studies (53%) included adolescent participants aged 12–17 years, and 9 studies (15%) included young adults aged 18–30 years. There were fewer studies at either end of the age range, with 6 papers (10%) reporting on under-fives, and only 3 studies (5%) reporting on adults over 30. 25% of studies included in this review comprised entirely male autistic samples. The sex of autistic participants was not reported in five studies (8%). Male autistic participants outnumbered female autistic participants in all but four of the remaining 54 papers (85%), at an average ratio of 5.2: 1 across these studies. This figure surpasses the 4.2: 1 male-female ratio reported in the wider autistic population ([Maenner et al., 2021](#)) and likely reflects the systematic underrepresentation of women within autism research ([D’Mello et al., 2022](#)). There were four studies with equal numbers of male and female participants, but no studies in which there were more female participants.

Most of the included articles (43/59, 73%) did not mention any co-occurring diagnoses in the exclusion criteria or the description of study participants. Only four studies (7%) stated clearly that autistic participants with additional diagnoses were excluded from the sample. In twelve studies (20%), autistic individuals with additional diagnoses were included in the sample (language disorder: 9; other learning or psychiatric conditions: 5). Detailed information about individual participants demonstrating the overlap between multiple diagnoses was provided in only three of the studies (5%).

81% of the papers reported observational cross-sectional studies. Of these, 43 out of 48 were between-groups designs, typically including either two comparison groups (26 studies, 60%) or three groups (11 studies, 26%). Five studies (12%) compared four or more groups. The most frequent comparison group across the studies comprised non-autistic individuals with ‘typical development’, appearing in 91% of the between-groups observational studies. Autistic participants were also compared with other clinical groups in several of the studies (Specific Language Impairment/Developmental Language Disorder: 6; Down’s Syndrome: 3; Fragile X Syndrome (with/without ASD): 2; Williams’ Syndrome: 1; Pragmatic Language Impairment: 1; ADHD: 1; Intellectual Disability: 1). One study compared autistic participants to children who were previously diagnosed with autism but no longer met diagnostic criteria (described in the article as “optimal outcome” - [Canfield et al., 2016b](#)).

Matching criteria for these studies were varied, although the majority matched participants on their chronological age (78%) and cognitive abilities (76%), with 29 studies (49%) using both of these criteria in their group-matching strategy. Participants were typically matched on measures of non-verbal ability (39%), although full-scale IQ scores (20%) and verbal IQ measures (17%) were also used. A small number of studies (7%) matched groups on verbal age or mental age scores. Language skills were another key matching criterion, with receptive skills matched for in 41% of studies and expressive skills in 28% of studies. The sex of participants was balanced between comparison groups in a minority of studies (22%). Other matching criteria used across the studies included: reading ability (4%); visual/auditory attention (2%); sentence repetition ability (2%); scores on ADOS (2%); level of formal education (4%); level of maternal education (2%); socio-economic status (2%). There were three between-groups studies whose authors did not attempt to match their comparison groups on key variables ([Beytollahi, Soleymani, & Jalaie, 2020](#); [Lee et al., 2018](#); [McCabe et al., 2017](#)).

Intervention studies were reported by 19% of the papers. These included eight studies with single case experimental designs (1–5

¹ Two records were excluded for reporting the same study as another article included in the review. One of these was an unpublished PhD thesis ([Goldman, 2002](#)), the findings of which were later written up as a published journal article ([Goldman, 2008](#)). The other was a descriptive article written to promote a narrative intervention programme ([Gillam et al., 2016](#)), which reproduced some of the findings from an earlier research article by the authors ([Gillam et al., 2015](#)).

participants), five of which were carried out by the same research group (Favot and colleagues). There was one study with a matched pairs design and two randomised control trials, one of which was a pilot study.

Across all the included studies, the most common narrative elicitation task was the 'storybook' paradigm (39%), in which the participant is asked to look at a wordless picture book and tell the examiner the story. The other narrative tasks used were story retellings (29%); personal narratives (17%); picture narration tasks (15%); eyewitness recall narratives (3%) and story generation with props (2%). Three studies used more than one of these approaches: two research groups elicited both 'storybook' and personal narratives from their participants (Losh & Capps, 2003; Rollins, 2014) and one study used a story retelling task, personal narratives (describing 'general' versus 'specific' events) and a picture narration task (King & Palikara, 2018).

Most of the 23 studies that used 'storybook' tasks elicited narratives with stimuli commonly used for the purpose of autism diagnostic assessments, such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2012). Seven research groups used the book 'Frog, where are you?' (Mayer, 1969), while two used another book in the series by the same author ('A boy, a dog and a frog'; Mayer, 1967). Six studies used the story 'Tuesday' (Wiesner, 1991), and three used the 'Monkey cartoon' task from the ADOS (Lord et al., 2012). The remaining studies used either children's picture books selected by the research team or a sequence of picture cards, with participants asked to narrate the story. The eight research teams that used 'picture narration' tasks took a similar approach but tended to use a single picture stimulus as a narrative prompt, often depicting an everyday 'problem'. Story retellings were often elicited using single or multiple picture stimuli to accompany the initial presentation of the story and to support the participants' retellings, although in some cases, these were removed for the retellings (e.g., Beytollahi et al., 2020; Diehl et al., 2006). Studies that investigated personal narratives most frequently used a 'conversational' elicitation procedure (6 of the 8 studies), although two intervention studies with primary aged children instead used a stimulus photograph of the participant engaged in an activity (Favot, Carter, & Stephenson, 2021b). In both studies that used eyewitness recall tasks, participants viewed either a live or videoed sequence of events, involving puppets or actors (Henry et al., 2020; Loveland et al., 1990).

3.3. Approaches to analysing narrative structure and coherence

The following section outlines the major analytical frameworks identified in the literature and describes the different scoring methods used for each approach. It is important to note that several research groups used more than one of these frameworks (see Table 1) and many of the studies also scored narratives for additional variables not reported in this review (e.g., microstructure, use of gesture). Across all the studies, there were no obvious links between participants' characteristics (e.g., age, cognitive ability), the type of elicitation task that was chosen, and the analytical frameworks that were used to assess participants' narratives.

3.4. Story grammar

The most frequently used approach to assessing narrative macrostructure involved story grammar frameworks (27 of 59 studies, 46%), although studies differed as to which story elements were scored and how these were labelled. For example, some of the intervention studies with young children carried out by Favot et al. (2021b) renamed the typical story grammar elements using more accessible language for the participants (e.g., 'where', 'who with', 'what happened', 'feelings/personal reaction'). Table 2 shows a breakdown of story grammar elements coded across these studies.

78% of the story grammar frameworks included the element of 'setting' (physical, location or time) and 74% included an 'initiating event' or 'problem'. 52% of studies included the introduction of 'characters' as a story element (typically coded as 'who' or 'who with' in frameworks assessing personal/autobiographical narratives). 63% of frameworks scored 'actions' or 'attempts' by the characters, with 59% including either the direct or long-term 'consequences' of these actions. The 'internal response' or 'emotional state' of characters was also scored in 59% of studies, with two research groups additionally scoring for an 'ending emotion'. 15% of studies coded characters' 'reactions' as a separate element to their internal responses/emotions, while 30% considered the protagonists' 'plan' or 'goal' in the story. 44% of studies scored for some form of 'ending', 'resolution', or 'conclusion' to the narrative. 11% of studies included a 'formal opening' as a story element (e.g., 'Once upon a time...'), with one of these frameworks also including a 'formal ending'. There was also considerable variation in the number of story elements included, with some frameworks incorporating virtually all of the above elements, whilst others presented a pared-down version of story grammar including only 'goal', 'attempt' and 'outcome' for each story episode.

Nearly half of the story grammar frameworks (13/27 studies) scored each individual story element on a scale, with a predetermined maximum total score. The exact measure used varied between studies, but these typically awarded between 0 and 3 points to each story element, with a higher score indicating a more detailed, accurate, specific, relevant or elaborated usage of that element within the narrative (Banney et al., 2015; Beytollahi et al., 2020; Birri, 2016; Estigarribia et al., 2011; Favot et al., 2021a, 2021b; Gillam et al., 2015; Lind et al., 2014; Peristeri et al., 2017; Peristeri et al., 2020; Petersen et al., 2014). Some researchers (e.g., Estigarribia et al., 2011; Losh & Capps, 2003) weighted scores so that a varying number of potential points were available for different narrative elements. 15% of the story grammar frameworks used a binary coding scheme, in which each element was scored 0 or 1 depending on its presence or absence in the narrative (Engel, 2018; Goldman, 2008; Iandolo et al., 2020; Tager-Flusberg, 1995).

27% of the story grammar frameworks simply tallied the total number of instances of each story element in the narrative (Canfield et al., 2016; Colozzo et al., 2015; Henry et al., 2020; Lee et al., 2018; Pearlman-Avniion & Eviatar, 2002; Young et al., 2005) with these authors tending to produce both an overall summed score of elements and a breakdown of how these were distributed. Pearlman-Avniion and Eviatar (2002) created composite scores using these summed elements in different combinations to produce an index of 'emotional' language and an index of 'informational language', with both composite scores calculated as a proportion of the narrative).

3.5. 'Main events'

Eighteen of the 59 studies (31%) scored participants' narratives for the inclusion of the specific events particular to that story. In 13 of these studies (72%), the authors created bespoke lists of the key story events in their narrative stimuli, sometimes through carrying out prior pilot work with adults (e.g., Hogan-Brown et al., 2013; Mäkinen et al., 2014). However, five research groups (28%) used published assessments in which narrative retells are scored against a predetermined list of main story events. These were the Expression, Reception and Recall of Narrative Instrument (ERRNI; Bishop, 2004 - in Conlon et al., 2019 and Volden et al., 2017), and the Swedish translation of the Renfrew Bus Story Test (Renfrew, 1997; Svensson and Tuominen-Eriksson, 2002 - in Johnels et al., 2013, Miniscalco et al., 2007 and Persson et al., 2006).

It is interesting to note that there was a degree of overlap between 'main events' and story grammar analytical approaches. Some authors applied both frameworks to their data to generate separate scores (e.g., Canfield et al., 2016; Conlon et al., 2019; Losh & Capps, 2003; Norbury & Bishop, 2003). Other researchers used a predetermined list to identify the main events in the story, but then mapped these onto a story grammar framework by assigning the events to particular story grammar elements (e.g., Estigarribia et al., 2011; Geelhand et al., 2020; Pearlman-Avniot & Eviatar, 2002; Veneziano & Plumet, 2019).

Most research groups using a 'main events' approach simply calculated the total number of listed story events that were included in the participants' narratives (Geelhand et al., 2020; Hogan-Brown et al., 2013; Kenan et al., 2019; Loveland et al., 1990; Mäkinen et al., 2014; Veneziano & Plumet, 2019; Westerveld & Roberts, 2017; Zielinski Deane, 2016). However, another common method was to score each listed story event on a 0–2 scale, with one point awarded for its inclusion in the narrative, and a second point given for its completeness/accuracy (Engberg-Pedersen & Christensen, 2017; Lind et al., 2014; Norbury & Bishop, 2003). Studies that scored participants' narratives using published assessments also followed this method.

3.6. Holistic scoring schemes

Instead of coding story grammar elements or individual story events, 8 of the 59 studies (14%) rated how various aspects of macrostructure were realised across the narrative as a whole. These research groups tended to use either published narrative assessments, or analytical frameworks adopted from previously published literature. The categories that were scored differed between studies.

The 'Peter and the Cat' assessment (Leitão and Allan, 2003) was used in two of the studies (Hilvert et al., 2016; Manolitsi & Botting, 2011). Commonly used in clinical contexts, this narrative assessment tool includes both microstructure and macrostructure measures. The macrostructure score comprises 0–3 point ratings for the overall Structure and Content of the story. The 'Narrative Scoring Scheme' was used in three further studies (King & Palikara, 2018; King et al., 2014; Rollins, 2014). This framework was developed by the Madison Metropolitan School District SALT Working Group in 1998 and was published in Heilmann et al. (2010). It provides detailed guidance for rating narrative accounts on a scale of 1–5 points in the categories of Introduction, Character Development, Mental states, Referencing, Conflict Resolution, Cohesion, and Conclusion. Some of these categories were shared by the 'Narrative Assessment Profile' (Bliss et al., 1998) used by Carlsson et al. (2020), which includes Topic Maintenance, Event Sequencing, Explicitness, Referencing, Conjunctive Cohesion and Fluency. Each dimension is scored 1–3 (where 3 indicates appropriate usage, 1 indicates inappropriate usage, and 2 indicates variable usage).

Two research groups devised novel holistic frameworks for the purposes of their study. Norbury et al. (2014) rated participants' narratives in the following five domains: Setting, Referencing, Conflict resolution, Cohesion, and Conclusion, with each category scored 0–3 for presence, clarity and accuracy. Marini et al. (2019) were concerned with participants' ability to generate novel narratives by 'projecting' themselves in time from a stimulus story, so these authors created a scoring system to be used across conditions (i.e., Past, Middle and Future). A 0–2 point scale was used to score each narrative on two dimensions: the presence and elaboration of new story elements, and the inclusion of pertinent information that was causally connected to the stimulus story.

3.7. High point analysis

Two studies by McCabe and colleagues (2013; 2017) used high point analysis to assess how well participants' narrative accounts corresponded to a prototypical narrative structure. This method involved using a flowchart (adapted from previous work by the lead author) to categorise each narrative according to the number of story events included, whether these were related in accurate chronological order, whether the account featured a 'high point', and whether this was followed by a resolution. Each narrative was awarded a score of between 1 and 7, depending on its overall structure. A narrative featuring only one past-tense event would score 1 point, while a two-event narrative would score 2 points. A 'miscellaneous' narrative that featured more than two past-tense events but did not fit a logical pattern scored 3 points. A 'leapfrog' narrative (relating logically sequenced events but lacking accurate chronology) scored 4 points. A narrative with events correctly sequenced but without a high point ('chronological' narrative) scored 5 points. A narrative that included a high point, but not a resolution would score 6 points. To score the maximum 7 points, a narrative would need to include a series of logically sequenced events leading to a high point, followed by a resolution ('classic' narrative).

3.8. Subjective story quality ratings

Three studies used subjective ratings by a group of naive (blinded) listeners to determine the overall quality of participants' narrative accounts. de Marchena and Eigsti (2010) asked raters: "How well were you able to follow this story?" and "How engaged

Table 3
Study characteristics with quality assessment ratings.

First author	Year	1. Were the criteria for inclusion in the sample clearly defined?	2. Was the sample size appropriate for the study aims?	3. Were the study participants described in detail?	4. Was autism diagnosis confirmed using a valid measure?	5. Were confounding factors identified and managed?	6. Were the methods sufficiently described to enable them to be repeated?	7. Were standard criteria used for scoring participants' narratives?	8. Were adequate inter-rater reliability measures used?	9. Was appropriate statistical analysis used?
Banney	2015	Y	N	Y	Y	Y	Y	Y	Y	Y
Beytollahi	2020	Y	N	Y	Y	N	Y	Y	Y	Y
Birri	2016	Y	Y	Y	N	Y	Y	Y	Y	Y
Brown	2007	Y	Y	Y	Y	Y	Y	Y	Y	Y
Canfield	2016a	Y	N	Y	Y	Y	Y	Y	Y	Y
Canfield	2016b	Y	N	Y	Y	Y	Y	Y	Y	Y
Carlsson	2020	Y	Y	Y	U	Y	Y	Y	Y	U
Colozzo	2015	Y	N	Y	U	Y	Y	Y	Y	Y
Conlon	2019	Y	N	Y	U	Y	Y	Y	Y	Y
de Marchena	2010	Y	N	Y	Y	Y	Y	Y	Y	Y
Diehl	2006	Y	N	Y	U	Y	Y	Y	Y	Y
Engberg-Pedersen	2017	N	Y	Y	N	Y	Y	Y	Y	Y
Engel	2018	Y	N	Y	N	Y	Y	Y	Y	N
Estigarribia	2011	Y	Y	Y	Y	Y	Y	Y	Y	Y
Favot	2018	Y	Y	Y	Y	Y	Y	Y	Y	Y
Favot	2019a	Y	Y	Y	N	N	Y	Y	Y	U
Favot	2019b	Y	Y	Y	Y	N	Y	Y	Y	Y
Favot	2021a	Y	Y	Y	Y	Y	Y	Y	Y	Y
Favot	2021b	Y	Y	Y	Y	N	Y	Y	Y	Y
Ferretti	2018	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geelhand	2020	Y	N	Y	Y	Y	U	Y	Y	U
Gillam	2015	Y	Y	Y	N	Y	Y	Y	Y	Y
Goldman	2008	Y	N	Y	Y	Y	N	Y	Y	U
Henry	2020	Y	Y	Y	N	Y	Y	Y	Y	Y
Hilvert	2016	Y	N	Y	Y	Y	Y	Y	Y	Y
Hogan-Brown	2013	Y	N	Y	Y	Y	Y	Y	Y	U
Iandolo	2020	Y	Y	Y	N	N	Y	Y	Y	N
Johnels	2013	Y	Y	Y	N	Y	Y	N	Y	Y
Kauschke	2016	Y	N	Y	N	Y	Y	Y	Y	Y
Kenan	2019	Y	Y	Y	Y	Y	Y	Y	Y	Y
King	2014	N	Y	N	N	Y	N	Y	Y	Y

(continued on next page)

Table 3 (continued)

King	2018	N	N	Y	Y	Y	N	Y	Y	Y
Lee	2018	Y	N	Y	N	N	Y	Y	N	Y
Lind	2014	Y	Y	Y	Y	Y	Y	Y	Y	Y
Losh	2003	Y	Y	N	Y	Y	Y	Y	Y	Y
Loveland	1990	N	N	N	N	Y	Y	Y	Y	Y
Mäkinen	2014	Y	N	Y	N	Y	U	Y	Y	Y
Manolitsi	2011	Y	N	Y	N	Y	Y	Y	N	Y
Marini	2019	Y	Y	Y	Y	Y	Y	Y	Y	Y
Marini	2020	Y	Y	N	N	Y	Y	Y	Y	Y
McCabe	2013	Y	N	Y	Y	Y	Y	Y	Y	N
McCabe	2017	N	Y	Y	Y	Y	Y	Y	Y	N
Miniscalco	2007	Y	Y	Y	U	Y	Y	Y	Y	Y
Norbury	2014	Y	Y	Y	N	N	Y	Y	Y	Y
Norbury	2003	Y	N	N	Y	Y	Y	Y	N	Y
Pearlman-Avnion	2002	N	N	Y	N	N	N	Y	Y	Y
Peristeri	2017	Y	N	Y	Y	Y	Y	Y	Y	Y
Peristeri	2020	Y	Y	Y	Y	Y	Y	Y	Y	Y
Persson	2006	Y	N	Y	N	N	Y	Y	N	N
Petersen	2014	Y	Y	Y	N	Y	Y	Y	Y	Y
Rollins	2014	Y	Y	Y	Y	Y	U	Y	Y	Y
Rumpf	2012	Y	N	Y	N	Y	Y	Y	U	Y
Sah	2015	Y	N	Y	U	Y	N	Y	Y	U
Tager-Flusberg	1995	Y	N	Y	U	Y	Y	Y	Y	Y
Veneziano	2019	Y	N	Y	N	Y	Y	Y	N	N
Volden	2017	Y	Y	Y	N	Y	Y	Y	Y	Y
Westerveld	2017	Y	Y	N	N	Y	Y	Y	N	N
Young	2005	Y	N	Y	U	Y	Y	Y	Y	Y
Zielinski Deane	2016	Y	Y	Y	Y	Y	Y	Y	Y	Y

Key:

Y = Yes	U = Unclear	N = No
---------	-------------	--------

Inter-rater reliability levels achieved for each question: (Q1) 98.3%; (Q2) 86.4%; (Q3) 98.3%; (Q4) 64.4%; (Q5) 72.8%; (Q6) 83.1%; (Q7) 100%; (Q8) 84.7%; (Q9) 74.6%.

were you during this story?” with responses rated on a 7-point scale. In a similar vein, Canfield et al. (2016a; 2016b) asked raters the following questions: “How good a story is this?”; “How well were you able to follow this story?”; “How well does this story reflect the actual story in the [stimulus] cards?”; and “How odd/unusual did you find this story?”. Responses were rated on a 1–5 Likert scale. Four intervention studies conducted by (Favot, Carter, & Stephenson, 2018, 2021a,b) used a similar approach to obtain a “social validity” outcome measure. In these studies, a blinded observer was asked to rate pre- and post-intervention narratives produced by each participant.

3.9. Approaches to assessing ‘coherence’

Although most of the articles included in this review did not specifically address narrative coherence, several research groups suggested that macrostructure scores could be used to indicate the coherence of spoken accounts. For example, Henry et al. (2020) used a story grammar framework “to capture higher order hierarchical structure, organisation and coherence” (p.2) in the eyewitness accounts of autistic children, while Gillam et al. (2015) “measured the coherence among story elements” (p.931), using a subscale of selected story grammar items. Hilvert et al. (2016) used a composite of their ‘structure’ and ‘content’ scores to provide a measure of narrative coherence. Additionally, three of the included studies used the ‘Narrative Scoring Scheme’ (Heilmann et al., 2010), which was designed to measure children’s ability “to effectively tell a coherent and interesting story” (p.156).

Conversely, there were a small number of articles included in this review (7/59, 12%) whose authors did not rely on macrostructure as a proxy measure, but instead attempted to quantify narrative coherence more specifically. These studies presented a mixed assortment of analytical approaches, showcasing highly divergent views of what constitutes ‘coherence’ within spoken discourse. These are summarised below.

Brown (2007) used an early version of the ‘Narrative Coherence Coding Scheme’ developed by Reese and colleagues (later published in Reese et al., 2011). This framework separates narrative coherence into three dimensions: context, chronology and theme, with narratives scored on a 0–3 scale in each dimension, depending on the level of elaboration and specificity. In contrast, Rumpf et al. (2012) and Kauschke et al. (2016) attempted to judge coherence by assessing the narrator’s “global orientation”, that is, how they introduced characters, and how they referenced setting elements such as location and time. The number of references to these elements were counted and subdivided into implicit (general) versus explicit (specific) references. This scoring scheme also considered whether the main events of the story were verbalised coherently, by counting the number of core events included (0–2) and the number of propositions that were used to describe these.

Taking a completely different approach, Diehl et al. (2006) characterised coherence as the “causal connectedness” of a narrative. This was measured by identifying the total number of causal connections (“the direct causal relationship between two story events”, p.89), then dividing this by the total number of ‘communication-units’ (a verb and its arguments). A higher score indicated a greater level of connectedness, and therefore coherence, in the narrative. A similar framework was adopted by Sah and Torng (2015), who assessed narrative coherence by calculating the density of causal connections and causal-chain events (“a sequence of events that form the gist of a story”, p.191), dividing the total number of such elements by the total number of clauses in each narrative.

Ferretti et al. (2018) also argued that causality is a key aspect of coherent narratives. In this model, coherence was indicated by participants’ ability to introduce a high number of new elements that were causally linked (i.e., episodes of the story were connected and congruent). Conversely, a measure of ‘incoherence’ was obtained by calculating the number of new elements that were introduced without causal links (i.e., story episodes were tangential or conceptually incongruent). Similarly, Marini et al. (2020) analysed narrative accounts in terms of their lack of coherence. This approach involved calculating the overall percentage of ‘local coherence errors’ (i.e., shifts of topic or missing referents) and ‘global coherence errors’ (tangential, incongruent, repetitive and filler utterances).

It is important to note that several of the other studies in this review also included additional measures relating to ‘incongruence’ alongside their main scoring framework. Many authors noted aspects such as the number of off-topic, irrelevant, bizarre, or illogical remarks, the intrusion of extraneous/invented information, and the presence of vague, redundant, inappropriate or unintelligible utterances (Canfield et al., 2016; Diehl et al., 2006; Geelhand et al., 2020; Hogan-Brown et al., 2013; Kauschke et al., 2016; Lee et al., 2018; Losh & Capps, 2003; Loveland et al., 1990; Mäkinen et al., 2014; Norbury et al., 2014).

3.10. Quality assessment ratings

The papers included in this review were judged to have a low risk of bias across most of the critical appraisal checklist ratings – see Table 3 (Appendix) for details. 90% of articles clearly defined the inclusion criteria for their sample (Q1), and the same percentage described study participants in sufficient detail (i.e., by providing essential demographic information - Q3). 85% of papers identified and managed confounding factors (for example, by matching comparison groups on key variables - Q5), while 86% described their methods in enough detail to allow for replication of the procedure (Q6). 100% of the studies used standard criteria for scoring participants’ narratives (Q7 – N.B. this was a screening criterion for inclusion in this review). 86% of papers used adequate inter-rater reliability measures, such as using a second coder to check inter-rater reliability; calculating reliability based on at least 10% of the narratives; and achieving an adequate level of agreement (Q8). A slightly increased risk of bias was evident in relation to Q9, with only 78% of articles using appropriate statistical analysis when reporting study findings. Across the papers, a high risk of bias was identified in relation to sample size (only 51% had a sample size that was judged appropriate for the study aims – Q2), and confirmation of participants’ autism diagnosis (just 47% of studies used a valid measure for this purpose- Q4). Ten studies were rated as having a low risk of bias across all nine checklist items (Brown, 2007; Estigarribia et al., 2011; Favot et al., 2018, 2021a; Ferretti et al., 2018; Kenan et al., 2019; Lind et al., 2014; Marini et al., 2019; Peristeri et al., 2020; Zielinski Deane, 2016).

4. Discussion

This systematic scoping review aimed to investigate the analytical frameworks used in previous research into the spoken narrative accounts of autistic people, and to provide an overview and synthesis of these approaches. The 59 studies included in the review represented research from 16 countries over a period of 31 years. Despite the inclusion criteria specifying studies that were reported in English, more than a third of the studies were carried out in other languages (36%), indicating that the findings of this review spanned a range of different cultures and research settings. Most were observational cross-sectional studies (81%), with autistic participants usually compared to non-autistic comparison groups matched on key variables, such as age and cognitive ability. The remaining 19% were intervention studies. Study participants were predominantly children and adolescents. Most studies (66%) included 20 or fewer autistic participants, and autistic females were particularly underrepresented in the study samples. A range of narrative elicitation methods were used across studies, with wordless storybook narrations and story retellings being the most common tasks (39% and 29%, respectively). Most of the papers in this review presented a low risk of bias across areas such as reporting of inclusion criteria, participant demographics and methods, management of confounding factors and scoring and reliability measures. However, small sample sizes and failure to validate participants' diagnostic status were sources of bias in around half of the articles.

Across the included studies, story grammar frameworks were the most popular method for analysing narrative macrostructure (46% of papers). This approach aims to provide a clear overview of both the content and organisation of a narrative by scoring it for key story elements (Henry et al., 2020). However, it was apparent from the literature that there was variability in which narrative elements were considered essential by different research groups. Only six elements (characters; setting; initiating event; action/attempt; consequences; internal response) appeared in more than half of the story grammar frameworks. Some of this variation may be attributed to the diverse types of narrative accounts that were elicited by different research groups. For example, some of the scoring schemes included a 'formal opening' element such as "Once upon a time." (Goldman, 2008; Iandolo, 2020; Tager-Flusberg, 1995), which would clearly be inappropriate for assessing an eyewitness or personal narrative account. Another possibility is that some researchers focusing on children's narratives may have chosen to focus on a simplified set of story grammar elements because they did not expect young participants to produce a comprehensive narrative account at their current stage of linguistic development.

The second most common approach in the literature assessed how successfully narrators conveyed the gist of a narrative by relating the main story events in a retelling task (31% of studies). This approach is commonly used in published narrative assessments, and several of the included studies made use of existing formal assessments for this purpose (Conlon et al., 2019; Johnels et al., 2013; Miniscalco et al., 2007; Persson et al., 2006; Volden et al., 2017). Retelling tasks can be scored reliably for the inclusion of key story events, since the expected content of the narrative is known (Miniscalco et al., 2007). This type of framework also provides a measure of whether the most salient narrative information is transmitted to the listener, and therefore tells us something about how useful it is as a form of interpersonal communication. Several studies assigned their predetermined main story events to story grammar categories (i.e., explicitly mapping out the narrative structure of the story before coding the transcripts), thereby demonstrating the overlap between these two macrostructural approaches (Estigarribia et al., 2011; Geelhand et al., 2020; Pearlman-Avniot & Eviatar, 2002; Veneziano & Plumet, 2019).

The inclusion of key story components or events within a narrative is often interpreted as an indication of how coherent it will appear to a listener (e.g., Gillam et al., 2015; Henry et al., 2020). However, the mere presence or frequency of particular elements does not account for how effectively these are deployed in the narrative. For example, if story events are presented in a disordered sequence, this could result in a wholly incoherent narrative. Despite this, several of the studies included in this review specified that their scores did not reflect the chronology of story elements or how these were linked together (e.g., Carlsson et al., 2020; Favot et al., 2018). There is research evidence to suggest that autistic people may have difficulties recalling events in the correct order (e.g., Zalla et al., 2013), so this is a limitation of this approach for assessing coherence in this population. Similarly, although the episodic formula of story grammar somewhat accounts for causality (i.e., 'goal' or 'action/attempt' leading to 'outcome' or 'direct consequence'), the relationships between discrete story events are examined less stringently within 'main events' frameworks. Causality is a crucial aspect of narrative coherence (Diehl et al., 2006; Ferretti et al., 2018; Jolliffe & Baron-Cohen, 2000; Sah & Torng, 2015) and this has been identified as an area of difficulty for autistic individuals (Capps et al., 2000; Diehl et al., 2006; Hilvert et al., 2016; Jolliffe & Baron-Cohen, 2000; King et al., 2014; Kuijper et al., 2017; Lee et al., 2018; Losh & Capps, 2003; Losh & Gordon, 2014; McCabe et al., 2013).

Neither story grammar nor main events approaches account for referencing errors, which can significantly affect a listener's ability to follow the thread of a narrative. Accounting for referential accuracy is an important concern when investigating narrative skills in autism, as previous research demonstrates that many autistic people find this challenging (e.g., Manolitsi & Botting, 2011; Novogrodsky et al., 2013; Suh et al., 2014). In recognition of this, some of the studies discussed in this review chose to include a separate measure of cohesion/referencing alongside their principal macrostructural framework (e.g., Banney et al., 2015).

A further limitation specific to the 'main events' approach is that it is not possible to apply this type of framework to a spontaneous language sample (i.e., to assess a narrative account of events that are not known to the listener). This limits the usefulness of this method for evaluating real-life instances of narrative. One way to overcome this difficulty is by using a holistic macrostructure framework, as these can, in theory, be used to score any type of narrative account. This approach can also provide a much more comprehensive picture of narrative competence across a range of features, incorporating aspects of both macrostructure and coherence (Rollins, 2014). A further advantage of a holistic method is that it goes beyond cataloguing story elements and attempts to capture how the narrative is perceived as a whole, thus mirroring the way in which listeners experience narratives in real life. This is of particular significance when assessing narrative skill in the autistic population, as pragmatic difficulties may affect the quality of spoken discourse in very subtle ways that nonetheless result in a poorer overall impression on the part of the listener (e.g., Canfield et al.,

2016b). However, there is considerable variability in the level of detail and usefulness of macrostructure frameworks that have been used in the literature. Moreover, although some scoring schemes of this type include aspects relating to the coherence and cohesion of a narrative, such as accurate referencing, they do not tend to focus on evaluating coherence specifically.

High point analysis was only used in two of the included studies (3%), likely because this provides a broad categorisation of story structure rather than a detailed method of narrative evaluation. High point analysis can be useful when analysing specific kinds of stories, but not all types of occurrences recounted through narratives will necessarily conform to typical high point structure. For instance, there is evidence that narratives about negative events tend to include more high points (Fivush et al., 2003). Conversely, a personal narrative recounting a mundane day at work or school might lack such dramatic structure, whilst still being completely coherent. The usefulness of high point analysis is therefore dependent on the structure of the original event (Reese et al., 2011), and this may limit its applicability for evaluating coherence in diverse forms of narrative account.

The use of subjective story quality ratings in a handful of studies provides an alternative perspective that involves a less detailed level of analysis, and instead considers how communication partners perceive a narrative. As discussed above, pragmatic difficulties commonly experienced by autistic individuals may affect listeners' overall perception of a spoken account, even in the absence of obvious structural errors (Canfield et al., 2016b). This approach might therefore offer a more functionally useful measure of narrative coherence in the autistic population. It is interesting to note that, despite the subjective nature of the rating scales used, high levels of inter-rater agreement were achieved in the studies that used this methodology. For instance, de Marchena and Eigsti (2010) reported a Cronbach's α of 0.93 (collapsed across diagnosis) when raters were asked how well they could follow the story. Similarly, Canfield et al. (2016a) reported intra-class correlation coefficients (ICCs) between raters of at least 0.83 (i.e., excellent) across all their subjective story quality ratings, whereas the inter-rater reliability for subjective ratings in their second study (2016b) ranged from good to excellent (ICC = 0.66–0.87). These figures suggest that subjective perceptions of coherence may be accurate. This could therefore be a useful approach in conjunction with other frameworks to provide a measure of social validity, as in the studies carried out by Favot and colleagues (2018; 2019b; 2021a; 2021b).

A small number of studies considered coherence directly. Despite these research groups adopting very different approaches, some key shared features characterising 'coherence' did emerge. Elements that were included in at least three studies are listed below:

1. **Context** (Brown, 2007; Kauschke et al., 2016; Rumpf et al., 2012)
2. **Chronology** (Brown, 2007; Diehl et al., 2006; Sah & Torng, 2015)
3. **Causality** (Diehl et al., 2006; Ferretti et al., 2018; Sah & Torng, 2015)
4. **Incongruence** (Ferretti et al., 2018; Kauschke et al., 2016; Marini et al., 2020)

Some of these features reflected aspects of the more typical macrostructural approaches used in narrative research. For example, in relation to story grammar, 'context' corresponds closely to 'setting', with both categories scoring narratives for the inclusion of information relating to 'time' and 'place' or 'location'. This is an important dimension of narrative to consider when assessing autistic individuals, since there is evidence to suggest that 'source monitoring' difficulties are more commonly experienced in this population (Maras et al., 2020). These might include issues with accurately recalling 'when' or 'where' an event took place, which could affect autistic narrators' ability to include this information in their spoken accounts.

The coherence-focused frameworks also highlighted some features contributing to narrative coherence that were not adequately addressed by story grammar and 'main events' approaches, and these may be of particular significance for assessing narrative abilities in autism. As discussed earlier, presenting story events in a logical order (i.e., chronology) is important for the coherence of an account; yet this may present a challenge for autistic speakers. Similarly, pragmatic violations (i.e., perceived lack of congruence) can affect listeners' ability to make sense of a narrative and such incongruencies have frequently been reported as a feature of autistic discourse from a neurotypical perspective (Volden et al., 2009).

There are some further elements that can be seen as important for narrative coherence and yet are omitted by the small selection of coherence-focused frameworks discussed above. For example, using accurate chains of reference is crucial in producing a cohesive narrative, in which the listener easily comprehends who is doing what to whom. Referencing errors can lead to confusion and in extreme cases, may render spoken accounts incomprehensible. Like the story grammar and main events approaches, none of the included coherence frameworks considered referencing as a criterion (interestingly, however, this dimension was accounted for in many of the holistic macrostructure frameworks). Similarly, despite some of the coherence frameworks characterising the external context (i.e., setting) as a key feature, none of them included criteria for evaluating how the internal context of the characters' thoughts and motivations were presented in the narrative. For a listener trying to make sense of a story, understanding why characters act as they do may be just as valuable as awareness of the setting in which they appear. This might be considered especially relevant for assessing narrative skills in autism, since previous research has indicated that internal state/mental state terms may be used less frequently by autistic speakers (Kauschke et al., 2016; King & Palikara, 2018; Pearlman-Avniot & Eviatar, 2002; Rumpf et al., 2012). Again, it is important to highlight that some of the other frameworks discussed in this review did consider the internal states and motivations of characters, notably in most of the story grammar approaches ('internal response'/'emotional state') and in the 'mental states' category of the Narrative Scoring Scheme (Heilmann et al., 2010). These findings indicate that while some aspects of narrative coherence are captured by conventional analytical methods for macrostructure, other key features are omitted. In summary, none of the frameworks identified in this review appeared to account for all the dimensions of a narrative account that can be used to judge its coherence.

Another key aspect of this review was considering the different scoring methods used by researchers when assessing participants' narratives. Scoring schemes were found to vary across frameworks, including between studies that adopted the same general approach

to narrative analysis (e.g., story grammar). Several studies simply measured the amount of information delivered by the narrator, by counting the total number of key story events, the number of times a particular narrative feature appeared, or the number of propositions used to describe core story events. The drawback of this approach for assessing coherence is that if the content is poorly organised, a narrative containing lots of information could prove less coherent than a more succinct yet well-organised account. In the same way, noting the presence of certain story elements does not indicate how effectively these aspects of the story were conveyed. For this reason, the most commonly used scoring approach was to rate main story events/story grammar elements for their level of detail, specificity, relevance, elaboration or accuracy, typically using a 0–2 or 0–3 scale. Rating scales were also used for the holistic macrostructure frameworks and in some of the coherence-specific frameworks. The advantage of this scoring method is that it provides more detail about the quality of different aspects of the narrative than a simple tally count, while still producing a quantitative score. As evidenced by the quality ratings undertaken for this scoping review, if scoring frameworks are sufficiently detailed, it is also possible to achieve a high level of inter-rater reliability when using this approach.

A small number of studies opted for proportional scores to assess coherence. The ‘causal connectedness’ approach used by Diehl et al. (2006) and Sah and Torng (2015) calculated the proportion of the overall narrative output that was concerned with describing the causal relationships between story events, whilst Marini et al. (2020) scored narrative accounts by calculating the percentage of coherence errors. Proportional scores have the advantage of controlling for narrative length. However, this type of scoring method appears quite limited in application, as it is difficult to see how it could be used to measure other aspects of narrative coherence identified in this review, such as ‘context’ or ‘chronology’.

5. Implications

In previous studies that have focused on evaluating narrative coherence in autistic people’s narratives, the following elements have been regarded as important by multiple teams of researchers: **(1) context, (2) chronology, (3) causality, and (4) congruence** (see Brown, 2007; Diehl et al., 2006; Ferretti et al., 2018; Kauschke et al., 2016; Marini et al., 2020; Rumpf et al., 2012; Sah & Torng, 2015). Research evidence suggests that these aspects of narrative may present challenges for autistic speakers (e.g., Diehl et al., 2006; Hilvert et al., 2016; Maras et al., 2020; Volden et al., 2009; Zalla et al., 2013). Two further dimensions that merit consideration are: **(5) the internal states and motivations of characters (i.e., cognition/emotion)** and **(6) the use of accurate referencing (cohesion)**, as these can also contribute to the coherence of an account and have been identified as areas of potential difficulty for autistic narrators (Kauschke et al., 2016; King & Palikara, 2018; Manolitsi & Botting, 2011; Novogrodsky et al., 2013; Pearlman-Avni & Eviatar, 2002; Rumpf et al., 2012; Suh et al., 2014). We strongly recommend that researchers incorporate these six key indices when investigating narrative coherence in future studies. When compared to typical macrostructural analysis, this approach offers a more comprehensive understanding of how narrators construct spoken accounts that are clear and easy to follow, regardless of the type of narrative and the method of elicitation. In terms of scoring methodology, the use of a rating (i.e., scale) scoring system is recommended rather than a tally count, as this approach provides more detailed information about narrative quality and scores are less affected by total narrative length.

Limitations and areas for future research.

The search strategy for this review involved keyword searches with a range of synonyms relating to both participants (i.e., autism) and the topic of interest (i.e., frameworks for analysing narratives). Following a limited initial search, additional synonyms were selected by screening the abstracts of the retrieved articles for relevant terms used in the literature. This methodology was chosen because the keyword ‘narrative’ is so broadly used across different academic fields that searching for this term without further qualification resulted in an overwhelming number of irrelevant articles. Similar issues arose when conducting preliminary searches that included generic terms such as ‘analysis’ or ‘framework’. For this reason, using a selection of more specific search terms relating to narrative ability, narrative macrostructure and narrative coherence was preferred. However, by narrowing the parameters of the keyword search in this manner, it is possible that some relevant terms may have been omitted, so this is acknowledged as a potential limitation of the present review.

Following the database searches, citation and reference tracking was carried out for the retrieved articles that specifically aimed to score participants’ narratives for coherence ($n = 10$). The rationale for this selective approach was to efficiently identify the most relevant papers to inform our key research question. However, we recognise that this strategy might have led us to overlook articles that, despite being less relevant to narrative coherence, could have perhaps informed the discussion of macrostructural frameworks in the earlier part of the review.

Future research in this area could include a direct comparison of established macrostructural approaches, such as story grammar, with coherence-focused frameworks, to investigate whether these methods produce similar findings when applied to the same data. Another interesting avenue for future research might consider methods used for the analysis of spoken narrative discourse in other clinical populations. This could provide a broader perspective on how the issue of narrative coherence has been tackled by researchers working within different specialisms, and its relevance to clinical practice.

6. Summary

This scoping review systematically investigated methodologies used in previous research to evaluate spoken accounts by autistic people in terms of their overall narrative structure and coherence. The findings revealed a clear preference for ‘story grammar’ and ‘main events’ analytical frameworks when analysing narrative macrostructure in this population. However, approaches to evaluating narrative coherence varied significantly in the literature. As we have argued, scoring narratives solely for typical elements of

macrostructure or the inclusion of key events fails to consider several important features that contribute to the coherence of an account. Some frameworks have attempted to overcome this limitation by rating specific aspects of narrative related to coherence, such as chronology and the presence of incoherent content (incongruence). Despite this, no existing framework has fully captured all the features of narrative coherence that have been identified by this scoping review. We therefore recommend that researchers consider the dimensions of **context**, **chronology**, **causality**, **congruence**, **characters** (i.e., **cognition/emotion**), and **cohesion** when rating narratives for their overall coherence. We also recommend the use of rating scales for assessing the coherence of narrative accounts, as this method of scoring was found to provide more detailed information about narrative quality.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declarations of interest: None.

Data Availability

No data was used for the research described in the article.

Acknowledgements

City, University of London - School of Health and Psychological Sciences doctoral studentships awarded to the first (AH) and fourth (GR) authors.

Appendix

See [Table 3](#)

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Washington, D.C.: American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. <https://doi.org/10.1080/1364557032000119616>
- Aromataris, E., & Munn, Z. (Eds.). (2020). *Joanna Briggs Institute manual for evidence synthesis*. Joanna Briggs Institute. Available from <https://synthesismanual.jbi.global>.
- Baixaulli, I., Colomer, C., Roselló, B., & Miranda, A. (2016). Narratives of children with high-functioning autism spectrum disorder: A meta-analysis. *Research in Developmental Disabilities*, 59, 234–254. <https://doi.org/10.1016/j.ridd.2016.09.007>
- Banney, R. M., Harper-Hill, K., & Arnott, W. L. (2015). The Autism Diagnostic Observation Schedule and narrative assessment: Evidence for specific narrative impairments in autism spectrum disorders. *International Journal of Speech-Language Pathology*, 17(2), 159–171. <https://doi.org/10.3109/17549507.2014.977348>
- Beytollahi, S., Soleymani, Z., & Jalaie, S. (2020). The development of a new test for consecutive assessment of narrative skills in Iranian school-age children. *Iranian Journal of Medical Sciences*, 45(6), 425–433. <https://doi.org/10.30476/ijms.2019.81984>
- Birri, N. L. (2019). *A personal narrative intervention for adults with autism and intellectual disability: A single subject multiple baseline design*. PhD thesis. University of Cincinnati.
- Bliss, L. S., McCabe, A., & Miranda, A. E. (1998). Narrative assessment profile: Discourse analysis for school-age children. *Journal of Communication Disorders*, 31(4), 347–363. [https://doi.org/10.1016/S0021-9924\(98\)00009-4](https://doi.org/10.1016/S0021-9924(98)00009-4)
- Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching and Therapy*, 18(1), 1–21. <https://doi.org/10.1191/0265659002ct224oa>
- Brown, B.T. (2007). The content and structure of autobiographical memories in children with and without Asperger Syndrome. PhD thesis. North Carolina State University.
- Canfield, A. R., Eigsti, I. M., de Marchena, A., & Fein, D. (2016). Story goodness in adolescents with autism spectrum disorder (ASD) and in optimal outcomes from ASD. *Journal of Speech, Language, and Hearing Research: JSLHR*, 59(3), 533–545. <https://doi.org/10.1044/2015.JSLHR-L-15-0022>
- Capps, L., Losh, M., & Thurber, C. (2000). "The frog ate the bug and made his mouth sad": Narrative competence in children with autism. *Journal of Abnormal Child Psychology*, 28(2), 193–204. <https://doi.org/10.1023/a:1005126915631>
- Carlsson, E., Asberg Johnels, J., Gillberg, C., & Miniscalco, C. (2020). Narrative skills in primary school children with autism in relation to language and nonverbal temporal sequencing. *Journal of Psycholinguistic Research*, 49(3), 475–489. <https://doi.org/10.1007/s10936-020-09703-w>
- Colozzo, P., Morris, H., & Miranda, P. (2015). Narrative production in children with autism spectrum disorder and specific language impairment. [La production narrative chez des enfants ayant des troubles du spectre de l'autisme ou des troubles spécifiques du langage]. *Canadian Journal of Speech-Language Pathology and Audiology*, 39(4), 316–332.
- Conlon, O., Volden, J., Smith, I. M., Duku, E., Zwaigenbaum, L., Waddell, C., ... Ungar, W. J. (2019). Gender differences in pragmatic communication in school-aged children with autism spectrum disorder (ASD). *Journal of Autism & Developmental Disorders*, 49(5), 1937–1948. <https://doi.org/10.1007/s10803-018-03873-2>
- D'Mello, A. M., Frosch, I. R., Li, C. E., Cardinaux, A. L., & Gabrieli, J. D. E. (2022). Exclusion of females in autism research: Empirical evidence for a "leaky" recruitment-to-research pipeline. *Autism Research*, 15(10), 1929–1940. <https://doi.org/10.1002/aur.2795>
- Diehl, J. J., Bennetto, L., & Young, E. C. (2006). Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *Journal of Abnormal Child Psychology*, 34(1), 87–102. <https://doi.org/10.1007/s10802-005-9003-x>
- Engberg-Pedersen, E., & Christensen, R. V. (2017). Mental states and activities in Danish narratives: Children with autism and children with language impairment. *Journal of Child Language*, 44(5), 1192–1217. <http://doi.org/10.1017/S0305000916000507>
- Engel, K.S. (2018). Reading comprehension instruction for young students with high functioning autism: Forming contextual connections. PhD thesis.

- Estigarribia, B., Martin, G. E., Roberts, J. E., Spencer, A., Gucwa, A., & Sideris, J. (2011). Narrative skill in boys with fragile X syndrome with and without autism spectrum disorder. *Applied Psycholinguistics*, 32(2), 359–388. <https://doi.org/10.1017/S0142716410000445>
- Favot, K., Carter, M., & Stephenson, J. (2018). The effects of an oral narrative intervention on the fictional narrative retells of children with ASD and severe language impairment: A pilot study. *Journal of Developmental and Physical Disabilities*, 1–23. <https://doi.org/10.1007/s10882-018-9608-y>
- Favot, K., Carter, M., & Stephenson, J. (2019a). The effects of oral narrative intervention on the personal narratives of children with ASD and severe language impairment: A pilot study. *International Journal of Disability, Development and Education*, 66(5), 492–509. <https://doi.org/10.1080/1034912X.2018.1453049>
- Favot, K., Carter, M., & Stephenson, J. (2019b). Brief report: A pilot study into the efficacy of a brief intervention to teach original fictional narratives to a child with ASD and language disorder. *Australasian Journal of Special and Inclusive Education*, 43(2), 102–108. <https://doi.org/10.1017/jsi.2019.7>
- Favot, K., Carter, M., & Stephenson, J. (2021a). The effects of an oral narrative intervention on the fictional narratives of children with autism spectrum disorder and language disorder. *Journal of Behavioral Education*. <https://doi.org/10.1007/s10864-021-09430-9>
- Favot, K., Carter, M., & Stephenson, J. (2021b). The effects of oral narrative intervention on the personal narratives of children with ASD and severe language disorder. *Journal of Behavioral Education*, 30(1), 37–61. <https://doi.org/10.1007/s10864-019-09354-5>
- Ferretti, F., Adornetti, I., Chiera, A., Nicchiarelli, S., Valeri, G., Magni, R., ... Marini, A. (2018). Time and narrative: An investigation of storytelling abilities in children with autism spectrum disorder. *Frontiers in Psychology*, 9, 944. <https://doi.org/10.3389/fpsyg.2018.00944>
- Fivush, R., Hazzard, A., McDermott Sales, J., Sarfati, D., & Brown, T. (2003). Creating coherence out of chaos? Children's narratives of emotionally positive and negative events. *Applied Cognitive Psychology*, 17(1), 1–19. <https://doi.org/10.1002/acp.854>
- Geelhand, P., Papastamou, F., Deliens, G., & Kissine, M. (2020). Narrative production in autistic adults: A systematic analysis of the microstructure, macrostructure and internal state language. *Journal of Pragmatics*, 164, 57–81. <https://doi.org/10.1016/j.pragma.2020.04.014>
- Gillam, S. L., Hartzheim, D., Studenka, B., Simonsmeier, V., & Gillam, R. (2015). Narrative intervention for children with autism spectrum disorder (ASD). *Journal of Speech, Language, and Hearing Research: JSLHR*, 58(3), 920–933. https://doi.org/10.1044/2015_JSLHR-L14-0295
- Goldman, S. (2008). Brief report: Narratives of personal events in children with autism and developmental language disorders: Unshared memories. *Journal of Autism and Developmental Disorders*, 38(10), 1982–1988. <https://doi.org/10.1007/s10803-008-0588-0>
- Heilmann, J., Miller, J. F., Nockerts, A., & Dunaway, C. (2010). Properties of the narrative scoring scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology*, 19(2), 154–166. [https://doi.org/10.1044/1058-0360\(2009\)08-0024](https://doi.org/10.1044/1058-0360(2009)08-0024)
- Henry, L. A., Crane, L., Fesser, E., Harvey, A., Palmer, L., & Wilcock, R. (2020). The narrative coherence of witness transcripts in children on the autism spectrum. *Research in Developmental Disabilities*, 96, Article 103518. <https://doi.org/10.1016/j.ridd.2019.103518>
- Hewitt, L. E. (2019). Narrative as a critical context for advanced language development in autism spectrum disorder. *Perspectives of the ASHA Special Interest Groups*, 4(3), 430–437. https://doi.org/10.1044/2019_PERS-SIG1-2018-0021
- Hilvert, E., Davidson, D., & Gámez, P. B. (2016). Examination of script and non-script based narrative retellings in children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 29, 79–92. <https://doi.org/10.1016/j.rasd.2016.06.002>
- Hogan-Brown, A., Losh, M., Martin, G. E., & Mueffelmann, D. J. (2013). An investigation of narrative ability in boys with autism and fragile X syndrome. *American Journal on Intellectual & Developmental Disabilities*, 118(2), 77–94. <https://doi.org/10.1352/1944-7558-118.2.77>
- Iandolo, G., López-Florit, L., Venuti, P., Neoh, M. J. Y., Bornstein, M. H., & Esposito, G. (2020). Story contents and intensity of the anxious symptomatology in children and adolescents with autism spectrum disorder. *International Journal of Adolescence and Youth*, 25(1), 725–740. <https://doi.org/10.1080/02673843.2020.1737156>
- Johnels, J. A., Hagberg, B., Gillberg, C., & Miniscalco, C. (2013). Narrative retelling in children with neurodevelopmental disorders: Is there a role for nonverbal temporal-sequencing skills. *Scandinavian Journal of Psychology*, 54, 376–385. <https://doi.org/10.1111/sjop.12067>
- Jolliffe, T., & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or Asperger's syndrome. Is global coherence impaired? *Psychological Medicine*, 30(5), 1169–1187. <https://doi.org/10.1017/S003329179900241X>
- Kauschke, C., van der Beek, B., & Kamp-Becker, I. (2016). Narratives of girls and boys with autism spectrum disorders: Gender differences in narrative competence and internal state language. *Journal of Autism and Developmental Disorders*, 46(3), 840–852. <https://doi.org/10.1007/s10803-015-2620-5>
- Kenan, N., Zachor, D. A., Watson, L. R., & Ben-Itzhak, E. (2019). Semantic-pragmatic impairment in the narratives of children with autism spectrum disorders. *Frontiers in Psychology*, 10, 2756. <https://doi.org/10.3389/fpsyg.2019.02756>
- King, D., Dockrell, J., & Stuart, M. (2014). Constructing fictional stories: a study of story narratives by children with autistic spectrum disorder. *Research in Developmental Disabilities*, 35(10), 2438–2449. <https://doi.org/10.1016/j.ridd.2014.06.015>
- King, D., Dockrell, J. E., & Stuart, M. (2013). Event narratives in 11–14 year olds with autistic spectrum disorder. *International Journal of Language & Communication Disorders*, 48(5), 522–533. <https://doi.org/10.1111/1460-6984.12025>
- King, D., & Palikara, O. (2018). Assessing language skills in adolescents with autism spectrum disorder. *Child Language Teaching & Therapy*, 34(2), 101–113. <https://doi.org/10.1177/0265659018780968>
- Kuijper, S. J. M., Hartman, C. A., Bogaards-Hazenberg, S. T. M., & Hendriks, P. (2017). Narrative production in children with autism spectrum disorder (ASD) and children with attention-deficit/hyperactivity disorder (ADHD): Similarities and differences. *Journal of Abnormal Psychology*, 126(1), 63–75. <https://doi.org/10.1037/abn0000231>
- Kunnari, S., Valimäa, T., & Laukkanen-Nevala, P. (2016). Macrostructure in the narratives of monolingual Finnish and bilingual Finnish–Swedish children. *Applied Psycholinguistics*, 37(1), 123–144. <https://doi.org/10.1017/S0142716415000442>
- Labov, W. (1972). *Language in the inner city*. Philadelphia: University of Pennsylvania Press.
- Lee, M., Martin, G. E., Hogan, A., Hano, D., Gordon, P. C., & Losh, M. (2018). What's the story? A computational analysis of narrative competence in autism. *Autism*, 22(3), 335–344. <https://doi.org/10.1177/1362361316677957>
- Leitão, S., & Allan, L. (2003). *Peter and the cat*. Keighley: Blacksheep Press.
- Lind, S. E., Williams, D. M., Bowler, D. M., & Peel, A. (2014). Episodic memory and episodic future thinking impairments in high-functioning autism spectrum disorder: An underlying difficulty with scene construction or self-projection. *Neuropsychology*, 28(1), 55–67. <https://doi.org/10.1037/neu0000005>
- Lord, C., Luyster, R. J., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic Observation Schedule (Second ed. (ADOS-2))*. Torrance, CA: Western Psychological Services.
- Losh, M., & Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of Autism and Developmental Disorders*, 33(3), 239–251.
- Losh, M., & Gordon, P. (2014). Quantifying narrative ability in autism spectrum disorder: A computational linguistic analysis of narrative coherence. *Journal of Autism and Developmental Disorders*, 44(12), 3016–3025. <https://doi.org/10.1007/s10803-014-2158-y>
- Loveland, K. A., McEvoy, R. E., & Tunali, B. (1990). Narrative story telling in autism and Down's syndrome. *British Journal of Developmental Psychology*, 8(1), 9–23. <https://doi.org/10.1111/j.2044-835X.1990.tb00818.x>
- Maenner, M. J., Shaw, K. A., Bakian, A. V., Bilder, D. A., Durkin, M. S., Esler, A., ... Cogswell, M. E. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years - Autism and developmental disabilities monitoring network, 11 Sites, United States, 2018. *Morbidity and mortality weekly report. Surveillance summaries*, 70(11), 1–16. <https://doi.org/10.15585/mmwr.ss7011a1>
- Mäkinen, L., Loukusa, S., Leinonen, E., Moilanen, I., Ebeling, H., & Kunnari, S. (2014). Characteristics of narrative language in autism spectrum disorder: Evidence from the Finnish. *Research in Autism Spectrum Disorders*, 8(8), 987–996. <https://doi.org/10.1016/j.rasd.2014.05.001>
- Manolitsi, M., & Botting, N. (2011). Language abilities in children with autism and language impairment: Using narrative as an additional source of clinical information. *Child Language Teaching & Therapy*, 27(1), 39–55. <https://doi.org/10.1177/0265659010369991>
- Maras, K., Dando, C., Stephenson, H., Lambrechts, A., Anns, S., & Gaigg, S. (2020). The witness-aimed first account (Wafa): A new technique for interviewing autistic witnesses and victims. *Autism: The International Journal of Research and Practice*, 24(6), 1449–1467. <https://doi.org/10.1177/1362361320908986>
- de Marchena, A., & Eigsti, I. (2010). Conversational gestures in autism spectrum disorders: Asynchrony but not decreased frequency. *Autism Research: Official Journal of the International Society for Autism Research*, 3(6), 311–322. <https://doi.org/10.1002/aur.159>

- Marini, A., Ferretti, F., Chiera, A., Magni, R., Adornetti, L., Nicchiarelli, S., ... Valeri, G. (2019). Episodic future thinking and narrative discourse generation in children with autism spectrum disorders. *Journal of Neurolinguistics*, 49, 178–188. <https://doi.org/10.1016/j.jneuroling.2018.07.003>
- Marini, A., Ozbić, M., Magni, R., & Valeri, G. (2020). Toward a definition of the linguistic profile of children with autism spectrum disorder. *Frontiers in Psychology*, 11, 808. <https://doi.org/10.3389/fpsyg.2020.00808>
- Mayer, M. (1967). *A boy, a dog and a frog*. New York: Dial Press.
- Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.
- McCabe, A., Hillier, A., Dasilva, C., Queenan, A., & Tauras, M. (2017). Parental mediation in the improvement of narrative skills of high-functioning individuals with autism spectrum disorder. *Communication Disorders Quarterly*, 38(2), 112–118. <https://doi.org/10.1177/1525740116669114>
- McCabe, A., Hillier, A., & Shapiro, C. (2013). Brief report: Structure of personal narratives of adults with autism spectrum disorder. *Journal of Autism & Developmental Disorders*, 43(3), 733–738. <https://doi.org/10.1007/s10803-012-1585-x>
- Miniscalco, C., Hagberg, B., Kadesjo, B., Westerlund, M., & Gillberg, C. (2007). Narrative skills, cognitive profiles and neuropsychiatric disorders in 7-8-year-old children with late developing language. *International Journal of Language and Communication Disorders*, 42(6), 665–681. <https://doi.org/10.1080/13682820601084428>
- Moola, S., Munn, Z., Tufanaru, C., Aromataris, E., Sears, K., Sfetcu, R., ... Mu, P.-F. (2017). : Systematic reviews of etiology and risk. In E. Aromataris, & Z. Munn (Eds.), *Joanna Briggs Institute reviewer's manual*. The Joanna Briggs Institute (Available from <https://reviewersmanual.joannabriggs.org/>).
- Norbury, C. F., & Bishop, D. V. M. (2003). Narrative skills of children with communication impairments. *International Journal of Language and Communication Disorders*, 38(3), 287–313. <https://doi.org/10.1080/136820310000108133>
- Norbury, C. F., Gemmell, T., & Paul, R. (2014). Pragmatics abilities in narrative production: A cross-disorder comparison. *Journal of Child Language*, 41(3), 485–510. <https://doi.org/10.1017/S030500091300007X>
- Novogrodsky, R. (2013). Subject pronoun use by children with autism spectrum disorders (ASD). *Clinical Linguistics & Phonetics*, 27(2), 85–93. <https://doi.org/10.3109/02699206.2012.742567>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan — a web and mobile app for systematic reviews. *Systematic Reviews*, 5, 210. <https://doi.org/10.1186/s13643-016-0384-4>
- Pearlman-Avniot, S., & Eviatar, Z. (2002). Narrative analysis in developmental social and linguistic pathologies: Dissociation between emotional and informational language use. *Brain and Cognition*, 48(2–3), 494–499. <https://doi.org/10.1006/brcg.2001.1405>
- Peristeri, E., Andreou, M., & Tsimpli, I. M. (2017). Syntactic and story structure complexity in the narratives of high- and low- language ability children with autism spectrum disorder. *Frontiers in Psychology*, 8, 2027. <https://doi.org/10.3389/fpsyg.2017.02027>
- Peristeri, E., Baldimtsi, E., Andreou, M., & Tsimpli, I. M. (2020). The impact of bilingualism on the narrative ability and the executive functions of children with autism spectrum disorders. *Journal of Communication Disorders*, 85, Article 105999. <https://doi.org/10.1016/j.jcomdis.2020.105999>
- Persson, C., Niklasson, L., Öskarsdóttir, S., Johansson, S., Jönsson, R., & Söderpalm, E. (2006). Language skills in 5-8-year-old children with 22q11 deletion syndrome. *International Journal of Language & Communication Disorders*, 41(3), 313–333. <https://doi.org/10.1080/13682820500361497>
- Petersen, D. B., Brown, C. L., Ukrainetz, T. A., Wise, C., Spencer, T. D., & Zebre, J. (2014). Systematic individualized narrative language intervention on the personal narratives of children with autism. *Language, Speech, and Hearing Services in Schools*, 45(1), 67–86. https://doi.org/10.1044/2013_LSHSS-12-0099
- Petersen, D. B., Gillam, S. L., & Gillam, R. B. (2008). Emerging procedures in narrative assessment: The index of narrative complexity. *Topics in Language Disorders*, 28(2), 115–130. <https://doi.org/10.1097/01.TLD.0000318933.46925.86>
- Rapin, I., & Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain & Development*, 25(3), 166–172. [https://doi.org/10.1016/s0387-7604\(02\)00191-2](https://doi.org/10.1016/s0387-7604(02)00191-2)
- 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>
- Rollins, P. R. (2014). Narrative skills in young adults with high-functioning autism spectrum disorders. *Communication Disorders Quarterly*, 36(1), 21–28. <https://doi.org/10.1177/1525740114520962>
- Rumpf, A., Kamp-Becker, I., Becker, K., & Kauschke, C. (2012). Narrative competence and internal state language of children with Asperger Syndrome and ADHD. *Research in Developmental Disabilities*, 33(5), 1395–1407. <https://doi.org/10.1016/j.ridd.2012.03.007>
- Sah, W., & Torng, P. (2015). Narrative coherence of Mandarin-speaking children with high-functioning autism spectrum disorder: An investigation into causal relations. *First Language*, 35(3), 189–212. <https://doi.org/10.1177/0142723715584227>
- Sah, W., & Torng, P. (2019). Storybook narratives in Mandarin-speaking pre-adolescents with and without autism spectrum disorder: Internal state language and theory of mind abilities. *Taiwan Journal of Linguistics*, 17(2), 67–89. [https://doi.org/10.6519/TJL.201907.17\(2\).0003](https://doi.org/10.6519/TJL.201907.17(2).0003)
- Shapiro, L. R., & Hudson, J. A. (1991). Tell me a make-believe story: Coherence and cohesion in young children's picture-elicited narratives. *Developmental Psychology*, 27(6), 960–974. <https://doi.org/10.1037/0012-1649.27.6.960>
- Siller, M., Swanson, M. R., Serlin, G., & Teachworth, A. G. (2014). Internal state language in the storybook narratives of children with and without autism spectrum disorder: Investigating relations to theory of mind abilities. *Research in Autism Spectrum Disorders*, 8(5), 589–596. <https://doi.org/10.1016/j.rasd.2014.02.002>
- Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing* (pp. 53–120). New Jersey: Ablex Publishing Corporation.
- Suh, J., Eigsti, I., Naigles, L., Barton, M., Kelley, E., & Fein, D. (2014). Narrative performance of optimal outcome children and adolescents with a history of an autism spectrum disorder (ASD). *Journal of Autism and Developmental Disorders*, 44(7), 1681–1694. <https://doi.org/10.1007/s10803-014-2042-9>
- Tager-Flusberg, H. (1995). 'Once upon a ribbit': Stories narrated by autistic children. *British Journal of Developmental Psychology*, 13(1), 45–59. <https://doi.org/10.1111/j.2044-835X.1995.tb00663.x>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... Hempel, S. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- Veneziano, E., & Plumet, M. (2019). Promoting narratives through a short conversational intervention in typically developing and high-functioning children with ASD. In E. Veneziano, & A. Nicolopoulou (Eds.), *Narrative, Literacy and Other Skills: Studies in Intervention* (pp. 285–312). John Benjamins Publishing Company. <https://doi.org/10.1075/sin.25.14ven>
- Volden, J., Coolican, J., Garon, N., White, J., & Bryson, S. (2009). Brief report: pragmatic language in autism spectrum disorder: Relationships to measures of ability and disability. *Journal of Autism and Developmental Disorders*, 39(2), 388–393. <https://doi.org/10.1007/s10803-008-0618-y>
- Volden, J., Dodd, E., Engel, K., Smith, I. M., Szatmari, P., Fombonne, E., ... Duku, E. (2017). Beyond sentences: using the expression, reception, and recall of narratives instrument to assess communication in school-aged children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research: JSLHR*, 60(8), 2228–2240. https://doi.org/10.1044/2017_JSLHR-L-16-0168
- Westerveld, M. F., & Roberts, J. M. A. (2017). The oral narrative comprehension and production abilities of verbal preschoolers on the autism spectrum. *Language, Speech, and Hearing Services in Schools*, 48(4), 260–272. https://doi.org/10.1044/2017_LSHSS-17-0003
- Wiesner, D. (1991). *Tuesday*. New York: Clarion Books.
- Young, E. C., Diehl, J. J., Morris, D., Hyman, S. L., & Bennetto, L. (2005). The use of two language tests to identify pragmatic language problems in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools*, 36(1), 62–72. <https://doi.org/10.1044/0161-1461%282005/006%29>
- Zalla, T., Labruyère, N., & Georgieff, N. (2013). Perceiving goals and actions in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 43(10), 2353–2365. <https://doi.org/10.1007/s10803-013-1784-0>
- Zielinski Deane, K. (2017). *The remediation of episodic memory deficits in children with autism spectrum disorder: An examination of the efficacy of cognitive behavioral therapy*. PhD thesis. University of California, Los Angeles.