Pragmatics abilities in narrative production: a cross-disorder comparison

COURTENAY FRAZIER NORBURY, TRACEY GEMMELL and RHEA PAUL

Journal of Child Language / *FirstView* Article / September 2013, pp 1 - 26
DOI: 10.1017/S030500091300007X, Published online: 30 April 2013

J. Child Lang., Page 1 of 26. © Cambridge University Press 2013
doi:10.1017/S030500091300007X

Pragmatics abilities in narrative production: a cross-disorder comparison*

COURTENAY FRAZIER NORBURY
Royal Holloway, University of London

TRACEY GEMMELL
Winnebago County Special Education Cooperative

AND

RHEA PAUL
Sacred Heart University, Connecticut, USA

(*Received* 26 *August* 2011 – *Revised* 19 *March* 2012 – *Accepted* 3 *January* 2013)

ABSTRACT

We aimed to disentangle contributions of socio-pragmatic and structural language deficits to narrative competence by comparing the narratives of children with autism spectrum disorder (ASD; $n=25$), non-autistic children with language impairments (LI; $n=23$), and children with typical development (TD; $n=27$). Groups were matched for age (6½ to 15 years; mean: 10;6) and non-verbal ability; ASD and TD groups were matched on standardized language scores. Despite distinct clinical presentation, children with ASD and LI produced similarly simple narratives that lacked semantic richness and omitted important story elements, when compared to TD peers. Pragmatic errors were common across groups. Within the LI group, pragmatic errors were negatively correlated with story macrostructure scores and with an index of semantic–pragmatic relevance. For the group with ASD, pragmatic errors consisted of comments that, though extraneous, did not detract from the gist of the narrative. These findings underline the importance

[*] This work was supported by a Nuffield Foundation New Career Development Fellowship and an Experimental Psychology Society study visit grant to C. F. Norbury, and grants K24 HD045576 from U.S. National Institute of Deafness and Communication Disorders and P01-HD03008 from U.S. National Institute of Child Health and Development to R. Paul. The authors are indebted to the children and families who participated in this study. Address for Correspondence: Dr Courtenay Frazier Norbury, Department of Psychology, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, United Kingdom. tel: +44(0)1784 276 139; e-mail: courtenay.norbury@rhul.ac.uk

of both language and socio-pragmatic skill for producing coherent, appropriate narratives.

INTRODUCTION

Narrative discourse has long been seen as a critical aspect of human communication, appearing early in development (Berman, 2009) and serving important functions in organizing human experience (Bruner, 1991). Narrative development involves an extended trajectory that requires the integration of linguistic, cognitive, and social capacities. Despite their relative complexity, however, narratives are used widely across cultures and language groups (Kaplan, 1966; O'Connell, 1997), for a range of communicative purposes (Bruner, 1991; Goldstein, 2000; Heath, 1986). Thus the use of narratives as indices of language ability has considerable ecological validity (Cicourel, 1996).

It is not surprising, given narrative's central role in communication, as well as its multi-faceted nature drawing on a range of linguistic, social, and cognitive abilities, that narrative acquisition is vulnerable to a variety of developmental insults. Moreover, narrative skills have been shown to be significant predictors of academic achievement (Bishop & Edmundson, 1987; Boudreau, 2008; Tabors, Snow & Dickinson, 2001) and a prognostic indicator of persistent language impairment (Bishop & Edmundson, 1987; Paul, Hernandez, Taylor & Johnson, 1996).

Narrative is defined linguistically as a type of discourse, primarily monologic, involving the relation of a sequence of events enacted by an agent in which plans and goals play a role (Stein & Glen, 1979). Polyani (1989) highlights 'stories' as one type of narrative in which a specific past time world is represented, and the story is used to make a point about the world shared by speaker and listener(s). Stories have been seen as discourse contexts that involve particular cognitive demands, including mastery of a range of linguistic (lexical, syntactic, and pragmatic) skills, the ability to remember and sequentially organize a set of events, to establish and maintain perspectives of a range of characters, and to use information from both within and outside the text to construct a novel and creative unit of discourse. Thus story telling provides a context in which to examine speakers' abilities to engage in a complex cognitive–linguistic task, with social functions and parameters, that requires spatial and temporal knowledge, memory, the identification and description of internal states, perspective-taking, and the ability to coordinate, integrate, and encode a relatively large amount of information with little interactional support. For these reasons, many studies of language development and disorders have made use of tasks involving narrative as a means toward understanding language development beyond the acquisition of single words and sentence structures, and as a vehicle for

addressing more functional, complex, socially embedded, and ecologically valid aspects of language use.

Since narratives are seen as a key component of social communication, and are thus thought be important in furthering knowledge about the intersection of social cognition and language, two populations in which narratives have been of special interest are autism spectrum disorders (ASD) and developmental language impairments[1] (LI) not associated with autism.

The continuities and boundaries between these two disorders have been discussed frequently in the literature (e.g. Bishop, 2010; Cantwell, Baker & Rutter, 1978; Conti-Ramsden, Simpkin & Botting, 2006; Norbury, 2005; Paul, 2007; Tager-Flusberg & Joseph, 2003; Verhoeven *et al.*, 2012; Whitehouse, Barry & Bishop, 2008). Studies suggest a range of similarities in the phenotypic presentation of the communication aspects of these disorders, and overlaps occur in these areas of clinical presentation. That is, some individuals with ASD are reported to show language profiles similar to those seen in LI (e.g. Tager-Flusberg & Joseph, 2003) and some individuals with LI display autistic-like behaviors (e.g. Conti-Ramsden *et al.*, 2006). Despite these overlaps, however, the two disorders are considered diagnostically distinct, and have specified criteria in standard diagnostic references (e.g. DSM-IV 1994 and proposed DSM-V 2012; ICD-10 2010). In the present study, we aim to maximize the differentiation between diagnostic groups by excluding both participants with ASD who show LI-like profiles of language performance, as well as those with LI who show autistic symptoms.

For speakers with ASD, stories are seen as vehicles for exploring the pragmatic impairments universal in this syndrome by providing a setting that puts pressure on both their relatively (though not entirely) spared structural language skills as well as their weak pragmatic abilities. This 'pressure test' has been thought to be particularly important in the study of speakers with ASD because it has been so difficult to identify and quantify their communicative deficits using standardized assessments of linguistic ability (Reichow, Salamak, Paul, Volkmar, & Klin, 2008). In LI, on the other hand, narratives are a context in which to observe the relatively (though not entirely) spared pragmatic and social–cognitive skills of affected children, while highlighting the effects of their characteristic limitations in structural language (i.e. vocabulary and grammar) by providing a challenging and more ecologically valid linguistic situation in which to display these abilities.

[1] We use the term LI to refer to children with language impairments who do not meet diagnostic criteria for ASD. We acknowledge the controversies surrounding current terminology and diagnostic criteria and make no assumptions about non-verbal intelligence scores or discrepancies between verbal and non-verbal ability and therefore avoid the term 'specific' language impairment or SLI.

Experimental studies of narrative production in ASD have generally confirmed these predictions (see Eigsti, de Marchena, Schuh & Kelley, 2010 for overview), though there are numerous inconsistencies in the literature. For example, studies using wordless picture books such as *Frog Where are You*? (Mayer, 1969) have generally failed to find differences between groups with ASD and typically developing peers on measures such as story length, story macrostructure, or complex syntax (Diehl, Bennetto & Young, 2006; Losh & Capps, 2003; Kelley, Naigles & Fein, 2010). However, differences on these lexical and grammatical variables have emerged on more open-ended narrative tasks such as conveying personal experiences (Losh & Capps, 2003), and may be apparent even when children with ASD are matched to younger TD peers on language variables (cf. Capps, Losh & Thurber, 2000). Some studies report differences between ASD and TD groups in relaying the 'gist' of a story (Jolliffe & Baron-Cohen, 2000; Loveland & Tunali, 1994); while others do not (Diehl *et al.*, 2006; Norbury & Bishop, 2003); and while some studies (Capps *et al.*, 2000) find differences between ASD and TD comparison groups in the use of mental state language or other features thought to reflect perspective-taking ability, others find no significant differences (Beaumont & Newcombe, 2006; Capps *et al.*, 2000; Colle, Baron-Cohen, Wheelwright & van der Lely, 2008; Diehl *et al.*, 2006; Garcia-Perez, Hobson & Lee 2008; Norbury & Bishop, 2003; Tager-Flusberg & Sullivan, 1995). Differences between studies may reflect differences in participant selection and severity, as well as differences in narrative stimuli and task demands. Most groups report wide age ranges (typically between six years and adolescence), so it is unclear whether inconsistencies in results may also reflect developmental differences in narrative competence.

Across these studies, however, several common narrative features of children with ASD are observed. First, the ability to describe internal states or to shift perspectives between characters appears to be correlated with tested Theory of Mind ability or with measures of emotional understanding in speakers with ASD (Capps *et al.*, 2000; Losh & Capps, 2003; Tager-Flusberg & Sullivan, 1995). Second, there is a trend among these studies to find, not absent, but reduced use of causal language (Capps *et al.*, 2000; Diehl *et al.*, 2006; Losh & Capps, 2003; Tager-Flusberg & Sullivan, 1995). Third, speakers with ASD appear to have more difficulty using referential expressions, such as anaphoric pronouns, than speakers from other diagnostic groups (Colle *et al.*, 2008; Manolitsi & Botting, 2011; Norbury & Bishop, 2003). Fourth, the majority of studies of narrative in ASD have reported pragmatic errors, exemplified by increased use of 'bizarre' or irrelevant comments (Capps, Losh & Thurber, 2000; Diehl, Bennetto & Young, 2006; Loveland & Tunali, 1993) or misinterpretation of key events (Kelley *et al.*, 2010). Finally, the narratives of speakers with ASD tend to

focus more on minor details and descriptions of elements within visual stimuli than on the global sense of the story (Diehl *et al.*, 2006; Loveland, McEvoy & Tunali, 1990; Peng, 1988; Waterhouse & Fein, 1982).

For the most part, however, these findings result from comparisons with typically developing peers, rather than peers with other neurodevelopmental disorders. Only a handful of studies have explicitly compared the narratives of children with ASD and non-autistic children with language impairment (Norbury & Bishop, 2003; Goldman, 2008; Manolitsi & Botting, 2011). In these studies, few differences emerge between the two clinical groups on either structural language variables or more pragmatic variables such as evaluation and pragmatic error. For instance, Norbury and Bishop (2003) explicitly attempted to differentiate the narratives of children with ASD and children with specific LI on an index of irrelevant comments. Unlike previous reports, a quantitative measure was taken, such that propositions that were not on a prespecified list of story-relevant propositions were tallied; surprisingly, there were no differences between clinical groups or the TD comparison group on this quantitative measure. Norbury and Bishop (2003) observed that many children (including TD as well as those with LI or ASD) embellished their stories with information that could feasibly provide additional detail about the setting or events. Attempts to distinguish plausible off-script comments from those that were tangential or 'bizarre' were not successful due to low inter-rater reliability.

Another factor that may obscure clinically significant differences between groups is the wide range of language and pragmatic skills that exist within clinical groups. For instance, a large proportion of cognitively able children with ASD experience additional language impairments (Kjelgaard & Tager-Flusberg, 2001; Loucas *et al.*, 2008). On the other hand, non-autistic children with LI are not only characterized by structural language deficits, many also experience difficulties with pragmatic aspects of language including coherence (Norbury, Nash, Baird & Bishop, 2004). All studies explicitly comparing children with ASD and LI have matched the groups on language variables such as receptive vocabulary or grammar, suggesting that many of the children with ASD had concomitant language impairments. These overlapping phenotypes may serve to make narrative performance more similar. There is a clear need to disentangle aspects of narrative performance that may be more clearly associated with structural language skills, and those associated with pragmatic deficits that characterize children with ASD.

The present study aimed to add to the rich literature on narrative development in ASD, by including both TD and LI contrast groups who do not have significant pragmatic impairments or autistic symptoms, and comparing these children to peers with ASD who demonstrate no structural language deficits on standardized tests. Apart from attempting again to identify elements of pragmatic function that distinguish the stories of

speakers with ASD from contrast groups, the present study also aimed to identify dissociations between the pragmatic aspects of story telling and semantic/syntactic aspects across diagnostic groups. To achieve these aims we applied standard narrative measures, such as lexical diversity, utterance and text length, and appearance of story macrostructure elements (e.g. setting, resolution). In addition, we modified a schedule of pragmatic errors developed by Klin (2000) for use with the Social Attribution Task. This type of error analysis has not previously been applied to more standard narrative procedures, but we anticipated that coding errors in this way may enhance inter-rater reliability. In addition, we coded classes of enhancement of meaning (through the use of internal state language) and complexity within the narratives. By investigating a wider range of semantic and pragmatic function than have previous studies, and by contrasting pragmatic deficits with strengths in the enhancement of story content and form, we attempted to identify the elusive differences among clinical populations in narrative competence.

Our specific hypotheses were that on quantitative indices of narrative, such as total number of words and utterances, number of different words, and syntactic length and complexity, children with LI would have significantly lower scores than both ASD and TD peers. We predicted that children with ASD who do not have additional language impairments should perform like TD peers on quantitative measures. In contrast, despite structural language scores within the normal range, we anticipated that children with ASD would have pronounced difficulties relative to both LI and TD peers on qualitative indices of narrative, including the ability to convey semantic–pragmatic information, would use internal state language, produce more pragmatic errors, and have more difficulties with the macrostructure of stories. The wide age range in our study sample enabled us to explore the extent to which qualitative aspects of story telling changed with age.

METHODS

Participants

Eighty-nine participants were recruited from two participating sites: clinical referrals and research participants at the Yale Child Study Center in the USA, and research participants in a larger study of language processing in ASD and non-autistic LI in England. Four participants (3 with LI and 1 with ASD) were excluded due to non-verbal ability scores greater than -2 *SD* below the normative mean. A further ten children with ASD were excluded because they obtained scores on structural language assessments of more than $-1{\cdot}5$ *SD* below the normative mean, indicating a clinically significant language impairment. This left twenty-three children with language impairment and no evidence of ASD (LI: 20 boys; all recruited

from the UK); twenty-five children with autism spectrum disorder and no additional language impairment (ASD: 22 boys; 11 recruited from the UK), and twenty-seven typically developing children (TD: 22 boys; 13 recruited from the UK). Groups were matched for chronological age; however, the age range was wide in order to enable us to consider developmental trajectories within the clinical populations. The age range across the entire sample was 78 months to 189 months, and was similar to previous investigations of narrative in these populations (*cf.* Diehl *et al.*, 2006). Written, informed consent was obtained from all parents of participating children to allow the use of narrative samples for research purposes.

Participants with ASD recruited to the Yale site were seen for a full clinical assessment, and the narrative sample was taken as part of that diagnostic assessment. Diagnosis of autism spectrum disorder was made by a multidisciplinary team of experts specializing in the diagnosis of pervasive developmental disorders, including a child psychiatrist, psychologist, speech–language pathologist (third author), and psychiatric nurse practitioner, all with at least ten years' experience in the clinical diagnosis of ASD. The assessment battery included parental interview concerning social communication skills, standardized testing, and direct clinical observation (NB: although the *Autism Diagnostic Observation Schedule* (ADOS; Lord *et al.*, 2000) was not formally administered, the psychologist and nurse practitioner on the clinical team are both trained and certified ADOS administrators who incorporated aspects of the ADOS procedures within their clinical interviews with the participants). Best estimate consensus clinical diagnosis was based on a review of all standardized tests, observations and medical and family history, and extensive group discussion, considering DSM-IV guidelines.

Participants with clinical diagnoses recruited to the UK site were all in receipt of a Statement of Special Educational Need (SEN) and were receiving specialist support for ASD within school settings. Children with ASD had been diagnosed by a multidisciplinary team external to the research group according to DSM-IV/ICD-10 criteria and confirmed via parent report of existing diagnosis and the statement of SEN. Children with LI were in specialist language units or schools serving children with LI. They had been diagnosed by multidisciplinary teams including a speech–language pathologist (SLP) and had Total Language scores on the Clinical Evaluation of Language Fundamentals (CELF-4UK: Semel & Wiig, 2006) below 77 (approximately $-1{\cdot}5$ *SD* below the normative mean). They did not have a diagnosis of ASD and their current specialist teachers/SLPs reported that these children did not have pragmatic impairments of the kind associated with ASD. This was confirmed by asking parents to complete the Children's Communication Checklist-2 (CCC-2; Bishop, 2003). The CCC-2 provides a measure of pragmatic deficit relative to overall language ability

(Norbury *et al.*, 2004); positive scores indicate that structural deficits in grammar, phonology, or word knowledge are significantly greater than pragmatic deficits. All children included in the LI group had positive scores, providing further evidence that they did not have significant pragmatic language difficulties. All children in the US sample were assessed by RP and all children in the UK were assessed by CFN; both authors have considerable research and clinical experience in the differential diagnosis of LI and ASD.

Children with TD were recruited at the Yale site as part of a separate study through community advertisements and personal connections. These participants had no history of special educational placement, were all in the appropriate grade for age, were reported by parents to have no developmental concerns, and scored within the normal range on both the *Clinical Evaluation of Language Fundamentals*-4th ed. (Semel & Wiig, 2006) and the non-verbal portion of the *Wechsler Abbreviated Scales of Intelligence* (WASI: Wechsler, 1999). These participants, as part of the research study in which they were involved, engaged in a structured interview designed to detect pragmatic language difficulties. None were identified by the speech–language pathologist conducting the interviews as showing any significant pragmatic deficits (Gemmell, 2007). Children with TD in the UK sample were recruited through local schools. They had no history of SEN support and scored within the normal range on the language and cognitive measures administered.

All children with ASD recruited to the Yale site were clinical referrals and, as such, the assessment battery therefore reflected clinical need. Children recruited in the UK were part of a research study; assessment measures were predetermined by the research questions, but intended to overlap with the USA battery as far as possible. At Yale, non-verbal ability in the ASD group was assessed using either the WISC (Wechsler, 1991; 2004) or the Differential Abilities Scales (Elliot, 1990). At the UK site, non-verbal reasoning was assessed using the Matrix Reasoning subtest of the WASI. At both sites, structural language abilities were assessed using the *Clinical Evaluation of Language Fundamentals*-4th ed. (Semel & Wiig, 2006); this test is widely used in both the USA and UK for clinical diagnosis of language impairment. At Yale, verbal IQ was obtained from the WISC and in the UK verbal IQ was estimated using the *British Picture Vocabulary Scales* (BPVS: Dunn, Dunn, Whetton & Burley, 1997). We aimed to match the three groups on age and non-verbal ability. Although none of the children included in the final sample had non-verbal ability scores within the intellectual disabilities range (i.e. all non-verbal standard scores were greater than 70), children in the LI group had significantly lower scores than the TD group ($F(2, 72) = 3{\cdot}54$, $p = {\cdot}04$). The LI and ASD groups did not differ significantly in terms of non-verbal ability ($t = {\cdot}76$,

TABLE 1. *Descriptive statistics for age, non-verbal ability, and structural language scores*

	Group				
	LI ($n=22$)	ASD ($n=26$)	TD ($n=27$)	F	P
Variable					
Chron. age (mths)	129·15 (39·10)	134·13 (38·33)	118·09 (28·52)	1·41	·25
range:	79·00–184·68	78·00–189·24	81·00–182·00		
Non-verbal ability	99·78 (13·77)	102·68 (12·71)	108·85 (10·77)	3·54	·04
range:	79–125	81–123	90–124		
Verbal ability/vocabulary	88·87 (13·71)	105·20 (14·05)	105·59 (11·21)	13·28	·001
range:	70–118	81–131	90–130		
Clinical evaluation of language fundamentals	68·91 (5·47)	95·00 (7·38)	98·00 (11·22)	76·04	·001
range:	56–80	85–130	85–130		

NOTES: Non-verbal ability includes: WISC-III Performance; WASI-Matrix reasoning; or Differential Ability Scales. Verbal ability includes: WISC-III Verbal; British Picture Vocabulary Scales-II.

$p=·48$). In addition, the ASD and TD groups were well matched on non-verbal ability, verbal ability, and CELF scores (see Table 1).

There were some differences in recruitment and procedures between the two sites. For example, children with ASD from the UK were older than their US counterparts (UK mean = 164 months; US mean = 114 months; $t(23)=4·1$, $p=·001$). However, children with ASD from the two sites did not differ on non-verbal ability ($t=·25$, $p=·81$), vocabulary ($t=·66$, $p=·52$), or CELF structural language scores ($t=1·3$, $p=·20$).

Procedure

Narrative elicitation task. Participants were shown the wordless picture book *A Boy, a Dog and a Frog* (Mayer, 1969). After reviewing all of the pages in the book, the children were asked to return to the beginning and to 'tell the story' to the examiner, as they followed along in the book. Thus, both the children and the examiner could see the pictures in the book as the story unfolded. After the initial instruction to tell the story, only minimal prompts were provided if the participant did not continue narrating. These included, remarks such as 'Can you tell me more?' and 'Then what?' In this book, the boy spots a frog in a pond and, over a number of unsuccessful episodes, attempts to capture the frog in a net with the help of his pet dog. In frustration, the boy shouts at the frog and leaves the pond with his dog. The frog is then lonely and so follows the footprints of the boy and the dog back to the boy's house. He follows them right into the bathroom and jumps in the bathtub with them at the end of the story.

Story transcription and coding: quantitative indices. One of the authors (TG) transcribed and coded all of the participants' narratives using the Systematic Analysis of Language Transcripts (SALT: Miller, 2003) and following a detailed coding manual devised by RP and CFN (coding manual available in full at: http://www.pc.rhul.ac.uk/sites/lilac). The coder, who did not participate in diagnostic assessments, was blinded to participant diagnosis by removing identifying information from audio-recorded samples, which were labelled by number only. Transcription reliability was monitored by the third author, who providing training on transcription using SALT conventions and read through 10% of the transcriptions while listening to the story audio-recordings. Any discrepancies were resolved by consensus. However, because the basic text of the story was stable across all participants, there were very few errors of transcription detected.

Transcripts were divided into communication units (c-units), which were utterances containing a verb and its arguments, such as, 'the boy tried to catch the frog with his net'. We also measured syntactic complexity, by counting all c-units that included complex utterances and noun phrase elaborations. Complex utterances included all those with a complement, adverbial, or relative clause. Noun phrase elaboration other than relative clauses already counted included utterances with more than two modifiers preceding the noun (e.g. 'the *two big* dogs'), or participial clauses, including both gerunds and adjectival clauses ('The boy *sitting in the bathtub* was his friend') following the noun, as well as appositives following the noun ('Jester, *the frog*, sat on the log'). The number of complex utterances was divided by the total number of utterances to give an overall index of syntactic complexity.

Thirty percent of transcripts were randomly selected from each of the three participant groups for assessment of inter-rater coding reliability by CFN and RP. Coders achieved 94% reliability (range: 92–100%) on coding c-units and 95% reliability (range: 90–100%) on coding syntactic complexity; disagreements were resolved by discussion. SALT software provides for the automatic calculation of quantitative measures such as length (in words and c-units), mean length of utterance (in words), number of different words and type–token ratios (number of different words/number of total words).

We predicted that children with LI would find the quantitative aspects of narrative that draw on syntactic and lexical skills more challenging than TD and ASD peers, in the light of their low CELF scores, which measure morphology and syntax (see Table 1).

Story coding: qualitative indices. We computed a number of indices that tapped more qualitative aspects of story telling. These included a semantic enhancement index, a semantic–pragmatic relevance index, a pragmatic error index, and a story macrostructure index. Each index is described briefly below (and is available in the coding manual). Inter-rater reliability was computed for each index separately, on 30% of the transcripts from each

participant group. Percent agreement ranged from 82% to 96%. As in previous research (*cf*. Norbury & Bishop, 2003), agreement was highest for the semantic indices and lowest for the pragmatic error index. Overall levels of agreement above 80% are regarded as acceptable and in line with reliability indices for pragmatic codes in Klin (2000); thus reliability of coding was considered sufficient.

Internal state language. Codes were used to denote words or utterances which conveyed emotional (e.g. 'The frog was *lonely*') and mental states (e.g. 'He *thought* the frog was in the net'), as well as character intentions (e.g. 'He *tricked* the boy on the log'). All instances were summed and divided by the total number of utterances to give an overall index of use of internal state language.

Semantic–pragmatic relevance. We constructed a list of 33 propositions that would be relevant and convey information essential to understanding the story (see 'Appendix'). This could be construed as similar to the information index provided by standardized narrative measures such as the Renfrew Bus Story (Renfrew, 1969). Children were given two points for every utterance containing all relevant pieces of information. One point was given for utterances containing partial information. The maximum possible score on this index was 66. We also computed a ratio score of relevant propositions/total utterances. Thus, a score of 1 would indicate a story that was maximally relevant, a score of less than 1 would indicate a story which was under-informative, and scores greater than 1 would be indicative of stories in which a number of utterances were not central to the story themes.

Pragmatic errors. Following Klin (2000), we adapted codes to quantify pragmatic errors that altered the meaning of the story. These included misattributions, in which erroneous information was conveyed (e.g. 'Then another boy came to the pond'); irrelevant details (e.g. 'There were three lilypads'); inconstant reference, in which the way characters were referred to changed throughout the story and led to confusion (e.g. 'John had a pet dog. A boy saw the frog in the pond'); vagueness, in which the intended meaning was unclear (e.g. 'He followed them back', in which 'them' could refer to the boy and dog or the footprints); non-narrator speech, in which the child stepped out of narrator role to comment or question (e.g. 'How did the frog open the door?'). Pragmatic errors were summed and divided by the total number of utterances to provide a pragmatic error index.

Story macrostructure. We also rated each narrative on the child's ability to convey information about the setting, appropriate referencing, conflict resolution, cohesion, and an adequate conclusion, following Norbury and Bishop (2003). Each section was rated on a scale of 0–3, with 0 indicating that the child had not provided any information and 3 indicating the information was clear and correct. Scores were summed for a total of 15 points.

RESULTS

Due to the wide age ranges and cognitive ability scores involved in this study, our preliminary analysis investigated the relationship between age (in months), non-verbal ability, and the variables of interest across all three participant groups. For our quantitative measures of narrative length and complexity, age was significantly correlated with number of different words ($r(75) = \cdot 31$, $p < \cdot 01$) and mean length of c-unit ($r(75) = \cdot 32$, $p < \cdot 01$). Age was not significantly correlated with total number of words or c-units ($r = \cdot 11$, $p = \cdot 33$) or with syntactic complexity ($r = \cdot 18$, $p = \cdot 12$). However, age was significantly correlated with all of the qualitative indices (internal state language, $r = \cdot 32$, $p = \cdot 005$; semantic–pragmatic relevance, $r = \cdot 51$, $p = \cdot 001$; pragmatic error, $r = -\cdot 47$, $p = \cdot 001$; story macrostructure, $r = \cdot 49$, $p = \cdot 001$). There were no significant correlations between non-verbal ability and any story variable (total words: $r = -\cdot 06$, $p = \cdot 60$; number different words: $r = \cdot 02$, $p = \cdot 89$; mean length c-unit: $r = \cdot 27$, $p = \cdot 06$; syntactic complexity: $r = \cdot 16$, $p = \cdot 12$; internal state language: $r = \cdot 09$, $p = \cdot 42$; semantic–pragmatic relevance: $r = \cdot 13$, $p = \cdot 27$; pragmatic error: $r = -\cdot 06$, $p = \cdot 59$; story macrostructure: $r = \cdot 13$, $p = \cdot 26$). Results were therefore analyzed using age as a covariate; non-verbal ability was not used as a covariate because it was associated with diagnostic group (Dennis, Francis, Cirino, Schachar, Barnes & Fletcher, 2009; Miller & Chapman, 2001) and was not significantly correlated with any of the dependent variables.

Due to the large number of statistical tests undertaken, two preliminary MANOVAs were conducted. The first included quantitative measures of length and complexity (number of utterances, number of different words, mean length of c-unit, and syntactic complexity). Wilks' lambda revealed significant group differences ($F(8, 138) = 4 \cdot 12$, $p < \cdot 001$). The second analyzed qualitative measures and included internal state language, semantic–pragmatic relevance, pragmatic error, and story indices ($F(8, 138) = 3 \cdot 61$, $p = \cdot 001$). ANCOVAs, controlling for age, were therefore used to assess group differences on individual measures. For indices reported as proportion scores, arc sine transformations were performed to normalize the distribution (Sheskin, 2000). Raw proportion scores are presented in Tables 2 and 3. Where significant between-group differences existed, the magnitude of the difference between the clinical groups, or between a clinical group and the TD group, was examined using Cohen's *d*, an estimate of effect size. Values up to.20 are considered small effects, ·50 a medium-sized effect, and ·80 is considered a large effect of clinical significance (Cohen, 1988).

Quantitative indices of narrative length and syntactic complexity

Performance on quantitative indices for each group is reported in Table 2. Significant group differences were revealed on all quantitative indices.

TABLE 2. *Mean (SD) scores and range for performance on quantitative indices for each group. F-values from ANCOVA, with age as the covariate*

	Group					
	LI ($n=23$)	ASD ($n=25$)	TD ($n=27$)	*F*	*df*	*P*
Variable						
Number of c-units	$36{\cdot}17_{a,b}$ (12·05)	$30{\cdot}52_{a}$ (9·24)	$39{\cdot}15_{b}$ (16·43)	3·53	2, 71	·04
range:	18–68	280·65$_{a,b}$ 17–54	226·36$_{a}$ 10–87	338·78$_{b}$		
Number of words	(114·09)	(79·30)	(141·71)	8·45	2, 71	·001
range:	140–582	121–440	90–713			
Number of different words	$91{\cdot}65_{a,b}$ (29·23)	$84{\cdot}48_{a}$ (25·88)	$108{\cdot}19_{b}$ (37·08)	6·84	2, 71	·002
range:	56–145	43–162	45–202			
Mean length of c-unit (words)	$7{\cdot}69_{a}$ (1·31)	$7{\cdot}40_{a}$ (1·88)	$8{\cdot}78_{b}$ (1·49)	9·51	2, 71	·001
range:	4·94–10·07	3·36–10·94	6·54–12·94			
Syntactic complexity[1]	$0{\cdot}16_{a}$ (0·11)	$0{\cdot}22_{a,b}$ (0·14)	$0{\cdot}26_{b}$ (0·13)	4·19	2, 71	·02
range:	0·0–0·42	0·0–0·55	0·07–0·68			

NOTES: Values with different subscripts in each row differ significantly at $p<{\cdot}05$.
[1] Proportion of c-units containing complex (subordinate) clauses or noun-phrase elaboration. See text for definitions and examples.

TABLE 3. *Mean (SD) scores and range for performance on qualitative story indices for each group*

	Group					
	LI ($n=23$)	ASD ($n=25$)	TD ($n=27$)	*F*	*df*	*P*
Index						
Internal state language	$\cdot12_{a}$	$\cdot21_{b}$	$\cdot20_{b}$			
SD:	(·11)	(·14)	(·09)	5·92	2, 71	·004
range:	·00–·38	·00–·73	·06–·40			
Semantic–pragmatic relevance*	$29{\cdot}70_{a}$	$31{\cdot}80_{a,b}$	$36{\cdot}19_{b}$			
SD:	(10·89)	(5·62)	(8·07)	4·69	2, 71	·01
range:	07–49	21–42	19–52			
Pragmatic errors	·15	·17	·15			
SD:	(·14)	(·15)	(·13)	0·87	2, 71	·42
range:	·00–·57	·00–·53	·00–·43			
Story macrostructure+	$10{\cdot}39_{a}$	$10{\cdot}44_{a}$	$11{\cdot}41_{b}$			
SD:	(2·71)	(1·92)	(1·97)	5·18	2, 71	·008
range:	4–15	6–14	8–15			

NOTES: * Maximum score = 66; + Maximum score = 15. Log-transformations were used in the statistical tests as raw scores violated assumptions of homogeneity of variance. Values with different subscripts in the same row differ significantly at $p<{\cdot}05$.

Predictably, the LI group had shorter mean length of c-unit ($t = 2{\cdot}74$, $p = {\cdot}009$, $d = {\cdot}79$) and reduced syntactic complexity ($t = 3{\cdot}00$, $p = {\cdot}009$, $d = {\cdot}84$) relative to the TD comparison group. However, despite being matched to the TD group on standardized assessments of structural language skill, the ASD group obtained significantly lower values than TD peers on total number of utterances produced ($t = 2{\cdot}31$, $p = {\cdot}03$), number of different words produced (a measure of semantic diversity; $t = 3{\cdot}23$, $p = {\cdot}01$), and mean length of c-unit ($t = 2{\cdot}95$, $p = {\cdot}005$), with differences representing clinically significant effect sizes (Cohen's d values = ·65, ·76, ·83 respectively). The ASD and LI groups did not differ significantly on any of these quantitative measures (syntactic complexity, $t = 1{\cdot}16$, $p = {\cdot}16$, $d = {\cdot}48$; mean length c-unit, $t = {\cdot}60$, $p = {\cdot}55$, $d = {\cdot}18$; total utterances, $t = 1{\cdot}83$, $p = {\cdot}09$, $d = {\cdot}54$; number different words, $t = 1{\cdot}92$, $p = {\cdot}37$, $d = {\cdot}27$).

Qualitative indices of semantic and pragmatic competence

Internal state language. Mean proportions of internal state references are reported in Table 3. Inspection of the data revealed two outliers in the ASD group. These two boys were the two oldest children in the group and 43% and 73% of their respective utterances contained internal state language. Removing these participants from the analysis did not alter the main result, a significant group difference for internal state language use ($F(2, 71) = 5{\cdot}92$, $p = {\cdot}004$). Post-hoc comparisons indicated that the LI group produced fewer instances of internal state language than either the TD group ($t = 2{\cdot}89$, $p = {\cdot}02$, $d = {\cdot}82$) or the ASD group ($t = 2{\cdot}57$, $p = {\cdot}02$, $d = {\cdot}73$). The TD and ASD groups did not differ from one another ($t = {\cdot}19$, $p = {\cdot}69$, $d = {\cdot}08$).

Semantic–pragmatic relevance. Data regarding the number of relevant utterances produced violated assumptions of normality. Log transformations were applied in an attempt to normalize the data. There was a significant group difference, with the LI participants producing significantly fewer relevant utterances than the TD comparison group ($t = 2{\cdot}89$, $p = {\cdot}02$, $d = {\cdot}70$). The ASD group did not differ from either LI ($t = {\cdot}85$, $p = {\cdot}40$, $d = {\cdot}25$) or TD peers ($t = 2{\cdot}26$, $p = {\cdot}19$, $d = {\cdot}64$), though there was a clear trend for the ASD group to produce fewer relevant utterances than TD peers.

Pragmatic errors. Inspection of Figure 1 reveals considerable within-group variation for all groups in terms of pragmatic errors, with some children producing few if any such errors and others producing high rates of error. ANCOVA did not reveal any reliable group differences ($F(2, 71) = {\cdot}87$, $p = {\cdot}42$). However, it did appear that more participants in the ASD group were likely to have proportion scores outside the typical range of values. We therefore applied a cut-off of.20 (this proportion of error indicating an 'extreme value' in the distribution of TD values) and counted the percentage of children in each group falling above that cut-off (TD = 14·8%,

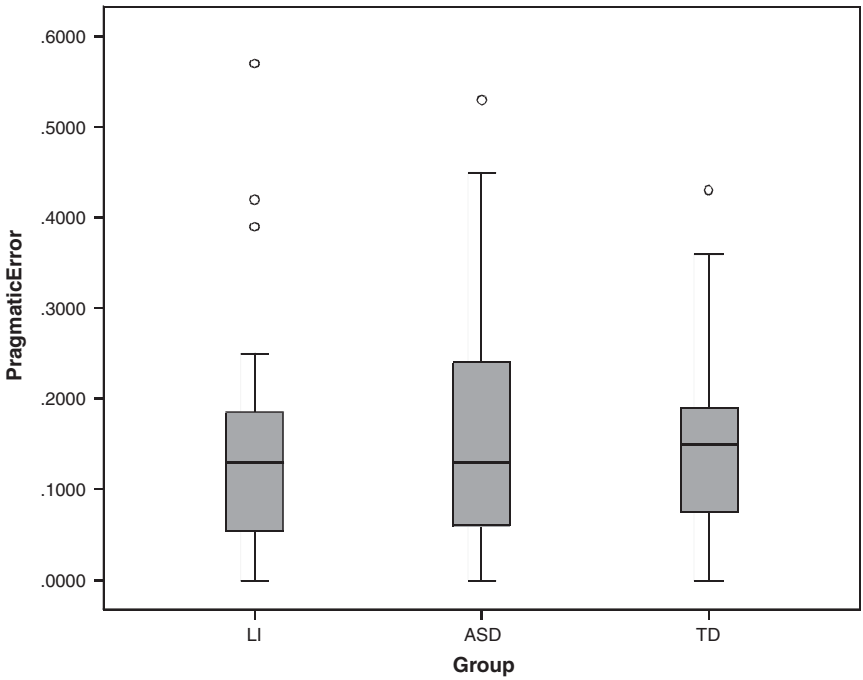

Fig. 1. Boxplot depicting proportion of pragmatic errors made in relation to total utterances for each participant group.

LI = 21·7%, ASD = 40%). Participants with ASD were significantly more likely to make pragmatic errors than TD peers ($\chi^2 = 4{\cdot}12$, $p = {\cdot}04$, odds ratio = 3·83). There was no significant difference in the number of pragmatic errors made by the ASD and LI groups ($\chi^2 = 1{\cdot}86$, $p = {\cdot}17$).

Story macrostructure. As reported in Table 3, there was a significant group effect for story macrostructure scores ($F(2, 71) = 5{\cdot}18$, $p = {\cdot}008$). The pattern of results indicated that both clinical groups had some difficulty organizing utterances into a cohesive narrative structure with a clear beginning, middle, and ending, and their mean scores were almost identical ($t(46) = {\cdot}07$, $p = {\cdot}94$). When controlling for age, both the LI and the ASD groups achieved significantly lower story macrostructure scores than TD peers ($p = {\cdot}01$, $d = {\cdot}45$ and $p = {\cdot}005$, $d = {\cdot}51$, respectively), with moderate effect sizes. We were interested to know if particular aspects of the story were differentially challenging to the clinical groups. We therefore conducted a 3 (group) × 5 (story component) repeated measures ANOVA (see Figure 2). This analysis revealed no main effect of story macrostructure component ($F = {\cdot}99$), but a significant group × story component interaction

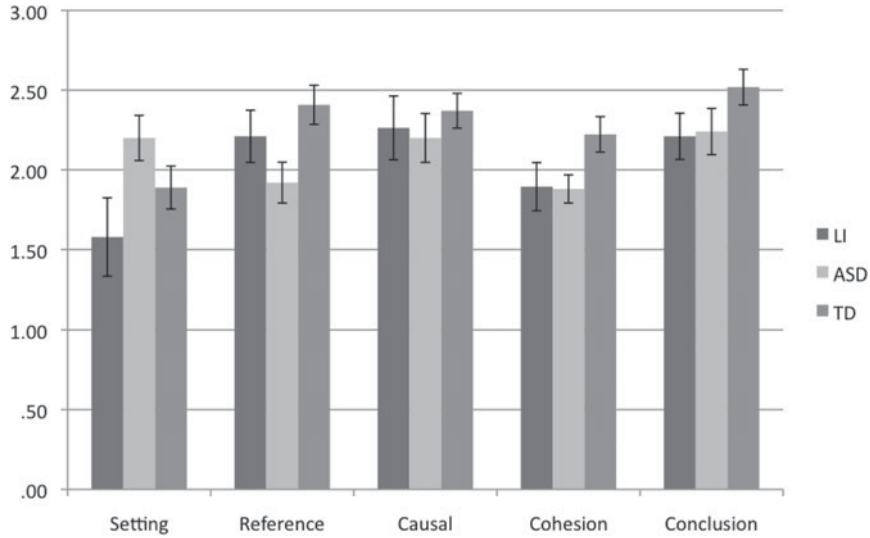

Fig. 2. Mean scores on story macrostructure components for each group. Error bars represent standard error.

($F(8, 284) = 2{\cdot}46$, $p = {\cdot}01$). For referencing, the difference between the TD and ASD groups was significant ($t = 2{\cdot}75$, $p = {\cdot}03$, $d = {\cdot}78$). No significant difference was found between the ASD and LI groups ($t = 1{\cdot}57$, $p = {\cdot}35$, $d = {\cdot}47$) but the moderate effect size suggests that the ASD group were the least able to provide adequate referencing. With regard to cohesion, no significant differences were found between the TD and the clinical groups (TD vs. ASD: $t = 2{\cdot}39$, $p = {\cdot}09$, $d = {\cdot}67$; TD vs. LI: $t = 1{\cdot}54$, $p = {\cdot}29$, $d = {\cdot}44$) although the moderate effect sizes indicate that the TD group tended to provide more cohesive stories.

Relationships amongst qualitative narrative indices, age, non-verbal ability, and language ability

Finally, we were interested in the extent to which pragmatic errors interfered with narrative performance. Within the clinical groups, pragmatic errors tended to decrease with age (LI: $r = -{\cdot}67$, $p = {\cdot}001$; ASD: $r = -{\cdot}65$, $p = {\cdot}001$), though this was not the case in the TD group ($r = {\cdot}13$, $p = {\cdot}51$). Language ability did not correlate with pragmatic error within the LI group ($r = -.14$, $p = {\cdot}52$), but was associated with pragmatic error within the ASD group ($r = -{\cdot}52$, $p = {\cdot}03$), such that increasing language competence was associated with fewer pragmatic errors. However, increasing language competence was also associated with fewer semantically–pragmatically relevant propositions in the ASD group only ($r = -{\cdot}56$, $p = {\cdot}03$). This would suggest that more

TABLE 4. *Regression analyses predicting story macrostructure from age, non-verbal ability, language ability (CELF scores), and pragmatic error scores*

Predictor	ΔR^2	B	SE B	β	t	p
Step 1						
(Constant)		7·52	2·93		2·56	·01
Age	·34	·02	·01	·30	2·31	·03
Step 2						
Non-verbal ability	·04	·02	·02	·08	·69	·49
Step 3						
Language ability	·001	·001	·02	·006	·06	·95
Step 4						
Pragmatic error	·13	−9·34	2·65	−·48	−3·53	·001

verbally able children with ASD may be verbose, but that their comments may not necessarily be relevant to the task at hand. Non-verbal ability was associated with pragmatic error in the LI group ($r = -·42$, $p = ·04$), but not the TD group ($r = -·16$, $p = ·42$). While the correlation was not statistically significant for the ASD group, the r value was moderate ($r = ·36$, $p = ·08$). It is notable that the direction of the relationship between non-verbal abilities and pragmatic error differs in the clinical groups. For children with LI, higher scores on measures of non-verbal ability are associated with fewer pragmatic errors; in contrast, within the ASD group, higher non-verbal scores are modestly associated with increasing pragmatic error. The reasons for this are unclear and warrant further investigation.

We therefore conducted a regression analysis with story macrostructure as the outcome variable, to investigate the unique contributions of each variable to story telling performance (see Table 4). Age, non-verbal ability, language ability (CELF scores), and pragmatic error scores were entered in stepwise fashion. The final model was highly significant, accounting for 72% of variance in story macrostructure. Age alone accounted for 34% of the variance in narrative performance, indicating again a positive change in narrative competence over time. Non-verbal ability contributed a small 3·7% of additional variance; this was not significant in the final model. Language ability scores also did not contribute significantly to story macrostructure, once age had been taken into account. Pragmatic errors, however, contributed a unique and significant 13·4% of variance. Thus, pragmatic errors may reflect a poorer understanding of the depicted events, resulting in less coherent and structured oral narratives.

DISCUSSION

In this study, we examined the narrative performance of children from two clinical populations: LI and ASD. There is considerable debate about the

degree to which these two disorders overlap (cf. Tomblin, 2011), but we selected our participant groups to be as contrastive as possible. We selected participants with LI who were rated by teachers and parents as having minimal social–pragmatic impairments. Similarly, we excluded children with ASD who had clinically significant language impairments on a standardized test of structural language competence typically used in the clinical diagnosis of language impairment. As such, we anticipated distinct narrative profiles, in which children with LI were more impaired on quantitative indices and measures of syntactic complexity, while peers with ASD would demonstrate significant impairments in the qualitative aspects of story telling that required the integration of linguistic and pragmatic skills. In reality, we found that two populations with different developmental trajectories converge on rather similar problems in narrative skill.

In line with previous research (e.g. Capps *et al.*, 2000), we found that participants with ASD differed significantly from age and non-verbal ability matched TD peers with regard to the number of utterances produced, despite being matched on standard tests of language. In addition, a salient finding from the current study is that these same participants produced utterances that were simpler than those produced by TD peers in terms of syntactic and semantic diversity. Similarly, their syntactic and semantic abilities in narrative did not differ significantly from the LI comparison group, despite significant differences between the two groups on standardized language assessment. Previous research has also failed to find differences between ASD and LI groups on structural and semantic aspects of narrative (Norbury & Bishop, 2003). These findings underline the importance of complementing standardized assessment with more ecologically valid measures of language use when assessing the communicative competence of children with ASD.

An equally striking finding was that on qualitative indices of internal state language use, semantic–pragmatic relevance, pragmatic error, and story macrostructure, non-autistic children with LI had as much, if not more, difficulty than peers with ASD. These findings suggest that the ability to talk about the mental and emotional lives of others depends crucially on having the vocabulary to do so (see also Norbury & Bishop, 2003). In this study, participants with ASD had sufficient vocabulary and social insight to recognize and label emotional and cognitive states; we would not wish to suggest that this necessarily means individuals with ASD experience those states in a similar manner to TD or LI peers. On the other hand, children with LI may recognize those internal states in others, but not have the vocabulary to discuss thoughts or feelings. We did not include independent measures of emotional or social cognition (*cf.* Losh & Capps, 2003), but it would be informative to explore the relationships between these cognitive capacities and narrative in diverse populations with

neurodevelopmental disorders, in order to better understand their role in narrative development.

Our measure of semantic–pragmatic relevance has much in common with standardized narrative measures such as the Information index of the Renfrew Bus Story (Renfrew, 1969). Our findings are consistent with a long-standing literature demonstrating that children with LI are less able to provide salient pieces of information in a narrative context (Gabig, 2008; Pankratz, Plante, Vance & Insalaco, 2007). We extend this finding to individuals with ASD. Given that the information index of the Bus Story has demonstrated prognostic value in identifying long-term language needs (Pankratz *et al.*, 2007; Stothard, Snowling, Bishop, Chipchase & Kaplan, 1998) and is a good predictor of both reading comprehension difficulties and responsiveness to language interventions to support reading comprehension (Bowyer-Crane *et al.*, 2008), we would predict that our participants with ASD might have additional difficulties with text comprehension. Clinically, it would seem prudent to monitor the wider language competencies and reading comprehension skills of children with ASD, especially those who are struggling with narrative.

Previous research has investigated 'bizarre' comments made by participants with ASD (Diehl *et al.*, 2006; Loveland, McEvoy, Tunali & Kelley, 1990), though attempts to quantify this have not always been successful (Norbury & Bishop, 2003) and levels of inter-rater reliability on 'bizarre' comments are rarely reported. Here, we attempted a novel method of coding a range of pragmatic errors, including misattributions and irrelevant comments, a method used successfully in assessing spontaneous attribution of animacy to ambiguous figures in adults with ASD (Klin, 2000). Consistent with the earlier work of Norbury and Bishop (2003), the group means did not differ on this pragmatic index and there was a large degree of within-group variation. However, when we looked at the numbers of children with very large error rates, we discovered that these children were more likely to have a diagnosis of ASD. Nevertheless, a significant proportion of children with LI made pragmatic errors. Regression analyses revealed that pragmatic errors were predictive of story macrostructure scores, even after accounting for age, non-verbal ability, and language competence. The more pragmatic errors a child made, the less coherent and well structured the story was. Thus pragmatic errors may reflect a fundamental problem with understanding the story that is crucial for making sense of depicted events and conveying them in a structured and interesting way.

We did find that children with ASD were specifically impaired in their ability to provide consistent and unambiguous references throughout the narrative. This finding replicates other studies indicating that referencing is particularly challenging for individuals with ASD (Diehl *et al.*, 2006; Norbury & Bishop, 2003). It could be argued that referential ambiguity may be an artefact of our task design; both the adult and the child could see

the pictures and so the need to mark reference was less crucial. However, it should be noted that the task was presented in the same manner to all three groups, yet the group with ASD had the most significant difficulties with referencing. In addition, the types of errors tended to be more striking; children with ASD tended not to start with an indefinite article plus noun ('a boy') and then refer to 'him' or 'the boy'. Instead, they may start with 'the boy' or 'he' and then revert to a proper noun ('Jim'), creating the impression that another character had entered the story. Alternatively, all of the characters (boy, dog, and frog) may be referred to as 'he', creating confusion, or given proper names without clearly establishing which name went with each character. The source of these referencing errors requires further investigation, but could reflect either a poor appreciation of listener need, or a lack of familiarity with narrative conventions.

Limitations and future directions

Our study required participants to tell the story while both the participant and the examiner could see the pictures in the book. We did this to minimize working memory demands that could have made the task impossible for younger children with LI. However, having the pictures available may have attenuated potential group differences. For instance, pragmatic errors may be reduced and use of internal state language enhanced for participants with ASD, because of the support provided by the pictures. Indeed, Diehl *et al.* (2006) investigated narrative production in a similar cohort of individuals with ASD, using a story retelling from memory task. Here children with ASD did make more bizarre comments than TD peers, though the absolute rate was rather low. However, in the Diehl *et al.* study, participants heard the story from a prerecorded audio sample, whereas in our study narratives were generated by the participants themselves. This might explain why we found that participants with ASD used shorter and less semantically and syntactically complex utterances relative to peers, while no such differences were reported by Diehl and colleagues. A systematic investigation of the influence of different narrative elicitation techniques on the quality of narrative production would be useful in identifying those techniques that are most sensitive to differential diagnosis, and have the potential to demonstrate change, both over time and in response to therapeutic intervention.

Our participant groups also included a wide age range and there were indications that the trajectories of narrative development, and the contributions of verbal abilities over time, may differ between the clinical groups and TD peers. However, as these are cross-sectional data, we cannot draw firm conclusions. Longitudinal studies charting narrative development in relation to linguistic and social–cognitive skills is a potentially rich arena for future research.

We pooled participants over two geographical sites and our testing protocols did not include a recognized measure of autistic symptomatology, such as the ADOS. It is possible therefore, that some of the overlap in clinical groups arises because of idiosyncrasies in diagnosis across the two countries. This is inevitable to some degree, even when standard diagnostic procedures are employed (*cf*. Lord *et al.*, 2012). We tried to minimize the impact of diagnostic differences by ensuring that participants with ASD did not differ across sites with respect to verbal or non-verbal ability. While there may be some differences in recruitment procedures between the two sites, pooling the data enabled us to include a greater number of participants than is usually the case in studies of narrative production in children with ASD. The study is further enhanced by including a non-ASD comparison group with language impairment, and comparing groups over a developmentally sensitive age range. Moreover, the authors made efforts to compare diagnostic procedures across sites in preparation for the study by joint work on videotaped case examples and by ensuring that coding of all transcripts was done by one individual who had not participated in diagnostic assessments and was blind to diagnostic status during the coding process. This added to the confidence that different recruitment practices or diagnostic differences would not skew outcomes.

CONCLUSIONS

Narrative continues to be an important line of investigation in clinical and academic research aimed at elucidating the pragmatic strengths and difficulties in a range of clinical populations. We attempted to identify aspects of narrative that could clearly differentiate ASD and LI, by including children with the most prototypical diagnostic phenotypes, and by extending our narrative analyses to include detailed indices of semantic diversity and pragmatic error. Instead of identifying significant differences between diagnostic groups on our measures as we predicted, we found that 'distinct' developmental trajectories converge on a similar problem space when it comes to using language in a cognitively demanding narrative task. How narrative develops in ASD and how it relates to other linguistic, cognitive, and social accomplishments is an important area for future investigation. Our findings demonstrate that cross-disorder comparisons will be necessary for determining narrative profiles that are 'specific' to ASD, and those that are common across neurodevelopmental disorders.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*, 4th ed. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2012). *Diagnostic and statistical manual of mental disorders: Proposed DSM-IV*, 5th ed. Washington, DC: American Psychiatric

Association. Available at <http://www.dsm5.org/ProposedRevision/Pages/NeurodevelopmentalDisorders.aspx>.
Beaumont, R. & Newcombe, P. (2006). Theory of mind and central coherence in adults with high-functiong autism and Asperger syndrome. *Autism* **10**, 365–82.
Berman, R. (2009). Language development in narrative contexts. In E. Bavin (ed.), *Cambridge handbook of child language*, 354–75. Cambridge: Cambridge University Press.
Bishop, D. V. M. (2003). *The children's communication checklist-2*. London: Psychological Corporation.
Bishop, D. V. M. (2010). Overlaps between autism and language impairment: phenomimicry or shared etiology? *Behavior Genetics* **40**, 618–29.
Bishop, D. & Edmundson, A. (1987). Language-impaired 4-year-olds: distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders* **52**, 156–73.
Boudreau, D. (2008). Narrative abilities: advances in research and implications for clinical practice. *Topics in Language Disorders* **28**(2), 99–114.
Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., Götz, K. & Hulme, C. (2008). Improving early language and literacy skills: differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology & Psychiatry* **49**, 422–32.
Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry* **18**, 1–21.
Cantwell, D., Baker, L. & Ruttter, M. (1978). A comparative study of infantile autism and specific developmental receptive language disorder – analysis of syntax and language function. *Journal of Child Psychology and Psychiatry* **19**, 351–62.
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.
Charman, T., Baird, G., Simonoff, E., Loucas, T., Chandler, S., Meldrum, D. & Pickles, A. (2007). Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *British Journal of Psychiatry* **191**, 554–59.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ: Lawrence Earlbaum Associates.
Cicourel, A. (1996). Ecological validity and 'white room effects': the interaction of cognitive and cultural models in the pragmatic analysis of elicited narratives from children. *Pragmatics & Cognition* **44**, 221–64.
Colle, L., Baron-Cohen, S., Wheelwright, S. & van der Lely, H. (2008). Narrative discourse in adults with high-functioning autism or Asperger syndrome. *Journal of Autism and Developmental Disorders* **38**, 28–40.
Conti-Ramsden, G., Simkin, Z. & Botting, N. (2006). The prevalence of autistic spectrum disorders in adolescents with a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry* **47**(6), 621–28.
Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A. & Fletcher, J. M. (2009). Why IQ is not a convariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society* **15**, 331–43.
Diehl, J., Bennetto, C. & Young, E. (2006). Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *Journal of Abnormal Child Psychology & Health* **34**, 87–102.
Dunn, L. M., Dunn, L. M., Whetton, C. & Burley, L. (1997). *British picture vocabulary scale*, 2nd ed. Windsor: NFER-Nelson.
Eigsti, I., de Marchena, A., Schuh, J. & Kelley, E. (2010). Language acquisition in autism spectrum disorders: a developmental review. *Research in Autism Spectrum Disorders* **5**(2), 681–91.
Elliot, C. (1990). *Differential Ability Scales*. Oxford: Harcourt Assessment.
Gabig, C. S. (2008). Verbal working memory and story retelling in school-age children with autism. *Language, Speech and Hearing Services in Schools* **39**, 498–511.
Garcia-Perez, R., Hobson, P. & Lee, A. (2008). Narrative role-taking in autism. *Journal of Autism and Developmental Disorders* **38**, 156–68.

Gemmell, T. (2007). Can conversational deficits in children and young adults with autism be more clearly identified using a semi-structured interview measure versus a standardized pragmatic protocol? Unpublished Master's thesis, Department of Communication Disorders, Southern Connecticut State University, New Haven, CT.
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), 1982–88.
Goldstein, B. (2000). *Cultural and linguistic diversity resource guide for speech-language pathology*. San Diego: Singular Publishing Group.
Heath, S. (1986). Talking a cross-cultural look at narratives. *Topics in Language Disorders* **7**(1), 84–94.
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), 1169–87.
Kaplan, R. (1966). Cultural thought patterns in intercultural education. *Language Learning* **16**, 1–2.
Kelley, E., Naigles, L. & Fein, D. (2010). An in-depth examination of optimal outcome in children with a history of autism spectrum disorders. *Research in Autism Spectrum Disorders* **4**, 526–38.
Kjelgaard, M. M. & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: implications for genetic subgroups. *Language and Cognitive Processes* **16**, 287–308.
Klin, A. (2000). Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: the Social Attribution Task. *Journal of Child Psychology and Psychiatry* **41**(7), 831–46.
Lord, C., Risi, S., Lambrecht, L., Cook, E. H. Jr., Leventhal, B. L., DiLavore, P. C., Pickles, A. & Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders* **30**, 205–23.
Lord, C., Petkova, E., Hus, V., Gan, W., Lu, F., Martin, D. M., … Risi, S. (2012). A multi-site study of the clinical diagnosis of different autism spectrum disorders. *Archives of General Psychiatry* **69**(3), 306–13.
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), 239–51.
Loucas, T., Charman, T., Pickles, A., Simonoff, E., Chandler, S., Meldrum, D. & Baird, G. (2008). Autistic symptomatology and language ability in autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry* **49**(11), 1184–92.
Loveland, K., McEvoy, R., Tunali, B. & Kelley, M. (1990). Narrative story telling in autism and Down's syndrome. *British Journal of Developmental Psychology* **8**, 9–23.
Loveland, K. & Tunali, B. (1993). Narrative language in autism and the Theory of Mind hypothesis: a wider perspective. In Baron-Cohen, S., Tager-Flusberg, H. *et al*. (eds.), *Understanding other minds: perspectives from autism*, 247–66. Oxford: Oxford University Press.
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 and Therapy* **27**(1), 39–55.
Mayer, M. (1969). *A boy, a dog and his frog*. New York: Dial Books for Young Readers.
Miller, G. & Chapman, J. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology* **110**, 40–48.
Miller, J. (2003). *Systematic analysis of language transcripts*. Madison: University of Wisconsin.
Norbury, C. (2005). The relationship between Theory of Mind and metaphor: evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology* **23**, 383–99.
Norbury, C. F. & Bishop, D. V. M. (2003). Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders* **38**(3), 287–313.

Norbury, C. F., Nash, M., Baird, G. & Bishop, D. V. M. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the Children's Communication Checklist–2. *International Journal of Communication Disorders* **39**, 345–64.

O'Connell, P. (1997). *Speech, language, and hearing problems in schools: a guide for students and practitioners*. Gaithersburg, MD: Aspen Publishers.

Pankratz, M., Plante, E., Vance, R. & Insalaco, D. (2007). The diagnostic and predictive validity of the Renfrew Bus Story. *Language, Speech, & Hearing Services in Schools* **38**, 390–99.

Paul, R. (2007). *Language disorders from infancy through adolescence*, 3rd ed. St Louis: Mosby.

Paul, R., Hernandez, R., Taylor, L. & Johnson, K. (1996). Narrative development in later talkers: early school age. *Journal of Speech and Hearing Research* **39**, 99–107.

Peng, F. (1988). On the acquisition of discourse among autistic children. *Language Sciences* **10**, 193–224.

Polyani, L. (1989). *Telling the American story: a structural and cultural analysis of conversational storytelling*. Cambridge, MA: MIT Press.

Reichow, B., Salamak, S., Paul, R., Volkmar, F. & Klin, A. (2008). Pragmatic assessment in autism spectrum disorders: a comparison of a standard measure with parent report. *Communication Disorders Quarterly* **29**, 169–76.

Renfrew, C. (1969). *The bus story*. Oxford: Author.

Semel, E. & Wiig, E. (2006) *Clinical Evaluation of Language Fundamentals*, 4th ed. London, UK: Psychological Corporation.

Sheskin, D. J. (2000) *Handbook of parametric and nonparametric statistical procedures*, 2nd edn. Boca Raton, FL: Chapman & Hall / CRC.

Stein, N. & Glen, C. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (ed.), *New directions in discourse processing*, Vol. 2, 53–120. Norwood, NJ: Ablex.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B. & Kaplan, C. A. (1998). Language-impaired preschoolers: a follow-up into adolescence. *Journal of Speech, Language & Hearing Research* **41**, 407–18.

Tabors, P., Snow, C. & Dickinson, D. (2001). Home and schools together: supporting language and literacy development. In D. Dickinson & P. Tabors (eds.), *Beginning literacy with language: young children learning at home and school*, 313–34. Baltimore, MD: Paul H. Brookes Publishing.

Tager-Flusberg, H. & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London*, Series B(358), 303–14.

Tager-Flusberg, H. & Sullivan, K. (1995). Attributing mental states to story characters: a comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics* **16**, 241–56.

Tomblin, J. B. (2011). Co-morbidity of autism and SLI: kinds, kin and complexity. *International Journal of Language and Communication Disorders* **46**(2), 127–37.

Verhoeven, J., Romme, N., Prodi, E., Leemanso, A., Zink, I., Vandewalle, E. … Sunaert, S. (2012). Is there a common neuroanatomical substrate of language deficit between autism spectrum disorder and specific language impairment? *Cerebral Cortex* **22**(10), 2263–71.

Waterhouse, L. & Fein, D. (1982). Language skills in developmentally disabled children. *Brain and Language* **15**(2), 307–33.

Wechsler, D. (1991). *Wechsler Intelligence Scales for Children*, 3rd ed. San Antonio, TX: Pearson Assessments.

Wechsler, D. (1999). *Wechsler Abbreviated Scales of Intelligence*. San Antonio, TX: Pearson Assessments.

Wechsler, D. (2004). *Wechsler Intelligence Scales for Children*, 4th ed. San Antonio, TX: Pearson Assessments.

Whitehouse, A. J. O., Barry, J. G. & Bishop, D. V. M. (2008). Further defining the language impairment of autism: is there a specific language impairment subtype? *Journal of Communication Disorders* **41**(4), 319–36.

APPENDIX

Semantic–pragmatic relevance index. Two points were given for each utterance containing all of the underlined points. An utterance containing some, but not all, underlined information received one point. Maximum score was therefore 66.

1. There was a boy and a dog
2. at a pond.
3. Boy **or** boy and dog are going fishing **or** frog catching
4. with a net and a bucket.
5. Boy **or** boy and dog saw a frog.
6. Boy **or** boy and dog tried/want to catch it.
7. Boy **or** boy and dog tripped **or** fell.
8. The boy and the dog fell in the water.
9. The boy comes up with the bucket on his head.
10. The frog is (use of emotion word).
11. The boy tried to grab/reaches for **or** catch/chases after the frog.
12. The frog jumped away **or** somewhere. (The frog jumped = 1 point)
13. The frog is on a log.
14. The boy has a reaction. (They looked at each other = 1 point)
15. The boy told the dog to go somewhere.
16. The boy sneaked up on the frog from the other end.

(Note: if 15 and 16 are combined using the word 'surround', 'cornered', or 'close in from both sides' award 4 points)

17. The boy uses the net to catch the frog. (The boy catches the frog = 1 point)
18. He catches the dog instead **OR** But he caught the dog. (The dog is in the net = 1 point)
19. The frog is in the water.
20. The frog has a reaction.
21. The boy yelled at the frog/He said I'll get you / He got very mad at the frog.
22. The boy and the dog walked away **or** left **or** went home
23. The frog is sad.
24. The boy and the dog leave footprints.
25. The frog was lonely. (The frog was sad/bored/etc. = 1 point)
26. He followed the footprints.
27. They led him to the boy's house.
28. He followed them to the boy's bathroom. (He goes to the boy's bathroom = 1 point)
29. The boy was taking a bath.

30. The frog was happy to see boy **or** boy and dog.
31. The boy was happy to see the frog.
32. So the frog got into the tub.
33. The boy, the dog, and the frog were happy to be together **or** friends.