

The Information Score as a Measure of Oral Discourse Comprehension in the Early School Years

PLEASE CITE THIS PAPER AS:

Claessen, M., Tucker, M., Dawes, E. & Leitão, S. (2020) The Information Score as a Measure of Oral Discourse Comprehension in the Early School Years, *International Journal of Speech-Language Pathology*. 22(3), 338-346.

<https://www.tandfonline.com/doi/full/10.1080/17549507.2020.1745888>

Mary Claessen, Megan Tucker, Emily Dawes, Suze Leitão
School of Occupational Therapy, Social Work and Speech Pathology,
Curtin University

Corresponding author: Mary Claessen, M.Claessen@curtin.edu.au

ABSTRACT

Purpose: Oral discourse comprehension is a fundamental skill that is predictive of future reading comprehension. The ability to identify important semantic propositions, in a narrative text, and their attachment to a mental model is an essential component of discourse comprehension. This study aimed to pilot a measure of free-recall in a narrative retell across three age groups to explore development. A secondary aim compared measures of free-recall and cued-recall.

Method: 132 typically developing children aged 4 to 7 years provided a retell of the Squirrel Story narrative and completed the Narrative Comprehension Assessment (NCA). The Squirrel Story Information Score (SSIS) was used as a novel measure of free-recall calculated from the narrative retells. The NCA, which involves literal and inferential comprehension questions, was used as a measure of cued-recall. Scoring reliability was calculated to ensure test-retest validity.

Results: Correlation analysis found excellent reliability of the SSIS measure. The difference between 4 and 5 year olds, and 5 and 6 year olds, was significant, with a large effect size. The difference between the 5 and 6 year olds was not significant. A large positive correlation was found between the measures of SSIS and NCA.

Conclusion: The measure of SSIS proved sensitive to developmental progression, with a positive relationship found between the measures of free-recall and cued-recall. Variability between the measures highlights the need to consider both measures to ensure an accurate representation of discourse comprehension skills.

Key words: oral comprehension, narrative, information score

Oral discourse comprehension involves processing both the explicit and implicit information within spoken discourse, and integrating this with stored knowledge in order to create an accurate mental representation of what is heard (e.g., Bishop, 2014; Cain & Oakhill, 2007; Florit, Roch, & Levorato, 2011; Stein & Albro, 1996). Early developing lower-level language skills, such as syntax and vocabulary, play a fundamental role in the oral comprehension of younger children (Cain & Oakill, 2007; Hogan, Bridges, Justice & Cain, 2011). These language skills serve as the foundation for higher-level, later developing skills such as inference making (Cain & Oakill, 2007; Hogan et al., 2011). Florit et al. (2011)

reported that typically developing childrens' ability to comprehend both explicit and implicit information, within an orally presented narrative, develops between the ages of 4 and 6 years, with explicit information easier to recall than implicit. Drawing inferences from implicitly provided information has been closely linked to text comprehension (e.g., Filiatrault-Veilleux, Bouchard, Trudeau, & Demarais, 2015; van Kleeck, Woude, & Hammett, 2006). The ability to make inferences, through drawing upon world and contextual knowledge, is a critical skill for effective discourse comprehension (Bishop, 2014; van Kleeck, 2008).

Cain, Oakhill, Barnes, and Bryant (2001) investigated the role of inference making in skilled and unskilled comprehenders (aged 7 to 8 years) within an oral narrative task and found less skilled comprehenders had difficulty at an early stage in the comprehension process, often failing to recall the information required to make the inference. This suggests that, even prior to an inference being made, the essential propositions were not encoded, highlighting that the ability to encode relevant semantic propositions is an important component of inferencing.

The Importance of Building a Mental Model

An essential skill of comprehension at the discourse level (e.g., narrative) involves building or constructing a mental model that integrates all the relevant semantic information to support understanding (e.g., Bishop, 2014). Effective comprehension of a story told orally is not achieved by comprehending each utterance in isolation, rather by building a successive mental representation of situations, characters and events, and the relationships between them, effectively summarising the presented information (Bishop, 2014; Cain & Oakhill, 1996).

Kintsch and van Dijk (1978) proposed a model of text comprehension that aims to describe the processes used in discourse comprehension. The schemas formed and drawn upon during discourse comprehension, as described by Kintsch and van Dijk (1978), are in keeping with the model described by Bishop (2014), whereby presented information is summarised to its gist during the construction of the mental model. Narratives provide a familiar, coherent structure that children are typically exposed to in early school years. In western cultures, knowledge of story grammar (Stein & Glenn, 1979) provides the listener with a specific and predictable schema, which allows for propositions deemed relevant to be mentally organised in a sequential structure (Kintsch & van Dijk, 1978). This knowledge provides a supportive structure for causal inferencing, whereby prior story grammar knowledge provides the foundation to infer relationships between the semantic propositions, thus allowing for the construction of a coherent mental model (van Kleeck, 2008; Van den

Broek, 1990). In this way, familiarity with, and knowledge of, story grammar guides the formation of the macrostructure and facilitates the generation, memory, and reproduction of narratives (Kintsch & van Dijk, 1978). Macrostructure knowledge, such as story grammar, has been shown to impact the ability to create an accurate mental model for comprehension (Cain, 2003; Dawes, 2017). Failure to construct a coherent and integrated mental model not only impairs memory for the semantic propositions, but also overall story comprehension (Bishop & Adams, 1992).

Narrative Assessment

Narratives are less contextualised than conversation, and often resemble the de-contextualised language demands of the classroom (Paul & Smith, 1993). Thus narrative is an ideal naturalistic context for assessing higher-level language skills essential for success in the classroom. Narrative assessment is commonly used in clinical practice as a context to measure both expressive language skills such as syntax, semantics, morphology, and phonology, and oral comprehension, providing clinicians with information that can be compared to developmental norms and used to inform goal setting and therapy (Westerveld and Claessen 2014; Lawson, 2012; Fey, Catts, Proctor-Williams, Tomblin & Zhang, 2004).

Discourse comprehension skills, such as the comprehension of narrative texts, have been shown to predict future reading comprehension, necessary for educational success, and can be used to identify children with language disorders and those at risk of later reading comprehension difficulties (Hogan et al., 2011; Hogan, Cain, & Bridges, 2012; Fey et al., 2004). Effective reading comprehension results from the integration of oral comprehension skills with the decoding of written text (Hogan et al., 2011; Kendeou, van den Broek, White, & Lynch, 2009). The ability to read words does not alone predict reading comprehension; successful reading comprehension is underpinned by oral comprehension skills (Nation 2019) As such it is crucial that clinicians are able to identify children with poor oral comprehension who are at risk of later reading comprehension deficits.

To assess oral comprehension at the discourse level, researchers and clinicians typically use two approaches; a child's ability to answer comprehension questions after hearing a story (cued-recall) and/or an analysis of a narrative retell (free-recall) (e.g., Fey et al., 2004; Peterson, Gilliam, & Gilliam, 2008).

Comprehension questions (cued-recall). A child's ability to answer questions related to aurally presented information (e.g. paragraph or story) is often used to assess oral comprehension (Filitrault-Veilleux et al., 2015; Joffe, Cain, & Maric, 2007; Bishop & Adams, 1992). It is argued that the ability to answer questions "off-line" (after hearing a

story) is representative of a child's ability to create an accurate mental representation of the narrative (Burriss & Brown, 2014; Lynch & Van den Broek, 2007). Whilst cued recall has been shown to be a reliable measure of narrative comprehension, by asking questions a cue is provided, which serves to prompt recall of specific encoded information (e.g., Cain, et al., 2001; Paris & Paris, 2003). There are several existing assessments that use off-line comprehension questions as an index of oral comprehension. The Developmental Neuropsychological Assessment (NEPSY-II; Korkman, Kirk, & Kemp, 2007), a comprehensive psychological assessment, includes a norm-referenced narrative memory assessment that includes measures of cued-recall by way of comprehension questions as well as free-recall in a narrative retell task. The narrative memory task is designed to compare memory ability across different cued supports, but no Australian norms are available. Another existing measure is the standardised norm-referenced Test of Narrative Language (TNL; Gillam & Pearson, 2004) that has been widely used by researchers and clinicians and has been shown to be valid and reliable in identifying language disorder. The TNL is designed for children aged 5 to 12 years and currently no Australian norms are available (Gillam & Pearson, 2004). A recently developed measure of cued-recall, based on the framework developed by Paris and Paris (2003), is the Narrative Comprehension Assessment (NCA), which includes literal and inferential comprehension questions for the Squirrel Story Narrative (Carey, Leitão, & Allen, 2006; Dawes, 2017), designed for younger children aged from 4 to 7 years. Australian data are currently being collected for typically developing children, in addition to children with Developmental Language Disorder (DLD; Dawes, 2017; Dawes, Leitão, Claessen, & Lingoh, 2019).

Narrative retell (free-recall). Assessment of free-recall (narrative retelling) as an index of oral comprehension is also a useful tool for the paediatric population (Morrow, 1985; Wenner, 2004). The ability to generate a comprehensive retell in free-recall is underpinned by the ability to encode, comprehend, and express the salient points of aurally presented material (Kendeou et al., 2005). In a narrative retell, accurate and sufficient information needs to be recalled in a temporally ordered sequence. Kintsch and van Dijk (1978) note the importance of information being stored in the correct temporal sequence.

A significant skill in the production of a retell is the construction of a macrostructure or a semantic discourse plan. A critical component of discourse comprehension is the ability to identify main ideas, or semantic propositions (Johnston & Afflerbach, 1985). The Bus Story Test (Renfrew, 1997) is a commonly used assessment of narrative discourse (Westerveld and Claessen 2014; Bolderson, Dosanjh, Milligan, Pring, & Chiat, 2011;

Hutchinson & Clegg, 2011). The Bus Story includes the calculation of an information score for the retell, whereby points are awarded for the inclusion of predefined ideas that involve expressing semantic relationships (Bishop & Edmundson, 1987). To achieve maximum points, coherence within the narrative retell is also considered, with failure to provide a referent at each change of ‘actor’ resulting in a point being subtracted (Renfrew, 1997). The ability to correctly reference characters in narratives has been shown to develop over the early school years, and identified as a deficit in children experiencing language difficulties (Ripich & Griffith, 1988). Bishop and Edmundson (1987) found that a child’s ability, at age 4, to provide an accurate recount of a sequence of events in a simple story, when measured by the information score, was predictive of persistent language disorder. The information score measure correctly identified 83% of the sample, and was a better predictor than micro-level skills such as syntax, morphology or phonology (Bishop & Edmundson, 1987). These results are consistent with the theoretical view that macrostructural elements play a pivotal role in discourse comprehension.

What does an information score measure? The information score is a semantic measure that reflects the ability to recall and retell the semantic propositions within a narrative (Norbury & Bishop, 2003). Renfrew (1997) suggests that the information score in the Bus Story can identify difficulties with verbal comprehension and sequencing (Renfrew, 1997). Cosentino, Adornetti, and Ferreitti (2013) propose that an information score is a reflection of a child’s ability to comprehend the macro or superstructure needed to organise the semantic content of a narrative using top-down processing, rather than the micro-structure reflected in the child’s ability to produce grammatically correct utterances. This top-down hypothesis implies that a prior understanding of the structure or flow of discourse drives sentence comprehension (Bishop 2014). Knowledge of story grammar enables the listener to predict the macrostructure (executive function of action planning) and, therefore, allows for efficient and effective attachment of the semantic propositions of the narrative to a mental model for comprehension.

Analysing free-recall as an index of oral discourse comprehension

It appears that a child’s ability to identify relevant information and encode the semantic propositions, coupled with schema knowledge should support an accurate and sequential recount, reflected in a measure such as a narrative retell information score. In free-recall, semantic propositions could be omitted for several reasons; failure could arise due to poor memory, resulting in reduced number of propositions being included or, alternatively, there may have been failure to encode the proposition in the first place (Cain et al., 2001).

What is not fully known is what proportion of encoded semantic propositions a child will produce, and if this changes with age. It is therefore important to build our understanding of the development of the inclusion of semantic propositions.

The Current Study

The aim of the present study was to investigate free-recall as a measure of oral narrative comprehension, across three age bands of typically developing children. Specifically, the study aimed to: (a) pilot a novel measure of oral comprehension through analysis of free-recall, measured by an information score obtained from a narrative retell, (b) establish inter-rater reliability of the piloted measure, (c) establish validity through investigating the information score's sensitivity to developmental progression. A secondary aim was (d) to explore the convergent validity by comparing two measures of oral comprehension, cued-recall and free-recall, through investigating the correlation between an information score and the ability to answer comprehension questions.

With early intervention essential for optimal outcomes, the study aims to provide normative reference data that are able to capture developmental differences across the early school years. This may assist in differential diagnosis, strengthening the ability to recognise children who may be experiencing language difficulties in order to provide timely intervention.

Hypotheses

The hypotheses were as follows:

1. The novel measure of free-recall will be a reliable measure of oral narrative comprehension.
2. The free-recall measure will be sensitive to developmental progression, with a significant difference in scores across the three age groups.
3. There will be a positive correlation between the information score (a measure of free-recall) and the comprehension questions score (a measure of cued-recall).

Method

Participants

Ethics approval was obtained from the Curtin University Human Research Ethics Committee (PSYCH SP 2014-15-12) and the Western Australian Department of Education prior to participant recruitment. One hundred and thirty two participants between the age of 4

and 7 years were recruited from primary schools in the Perth metropolitan area for this study. See table 1 for participant details.

Table 1

Participant details

AGE	MALE	FEMALE	TOTAL
4;0-4;11	20	24	44
5;0-5;11	14	17	31
6;0-6;11	28	29	57

Inclusion criteria comprised English as a first language, and typical language and learning skills as determined by the participants' teachers. Informed consent was obtained from the parents of all participants and assent from the participants themselves.

G*Power sensitivity analysis revealed that a sample size of 132 had sufficient power to detect small to medium effect ($f = .27$) at alpha .05 (Cohen, 1988; Faul, Erdfelder, Lang & Buchner, 2007).

Measures

Squirrel Story Narrative. The Squirrel Story is a criterion-referenced narrative retelling assessment that was developed by speech pathologists to assess narrative abilities in young children aged 3 – 6 years (Carey et al., 2006). The story was visually presented on an iPad with the accompanying verbal script heard via the iPad using the male-voice, Australian accent setting.

Information score (free-recall). The Squirrel Story Information Score (SSIS), was developed and piloted for use with the Squirrel Story based on the scoring systems used in the Bus Story and NEPSY-II, and the theoretical models described in Bishop (2014) and Kintsch and van Dijk (1978), see Appendix A for SSIS protocol. The measure was designed to capture the number of semantic propositions the children included in their retells. Two points are awarded for semantic propositions that are considered critical to the story, with one point given when only part of the information is provided. Subsidiary items worth one mark are also included; these are considered extra, non-critical pieces of information which add to the detail of the retell. Sixty-four points in total are available for main items and 21 points for subsidiary items, providing an overall maximum score of 85. The scoring protocol also involves one point being subtracted if the referent is not established at each change of referent or is incorrect, as per the Bus Story (1997). The scoring method was piloted with the

assistance of two final year speech pathology students. Feedback and peer discussions were used to operationally refine and define the scoring protocol.

Narrative comprehension questions (cued-recall). The Narrative Comprehension Assessment (NCA) includes comprehension questions based on the Squirrel Story Narrative (Dawes, Leitão, Claessen, & Lingoh. (2019). The NCA is based on research by Paris and Paris (2003) whose comprehension task reflected significant positive correlations across three picture books which indicated generalizability of the approach. The NCA includes 5 literal and 14 inferential comprehension questions. Previous research has found that the NCA was able to differentiate between typically developing and children with developmental language disorder (DLD), with significant differences found between the groups ($p < .001$) (Dawes, Leitão, Claessen, & Lingoh. (2019). The Squirrel Story NCA questions and scoring guide can be freely downloaded via:

<https://www.languageandliteracyinyoungpeople.com/developmental-language-disorder>

Procedure.

The participants were required to: (a) listen to the Squirrel Story on an iPad while looking at the accompanying story pictures, (b) answer literal and inferential questions off-line (NCA), (c) listen to the story again, and (d) retell the story to the researcher. Non-leading non-verbal and verbal prompts were provided when necessary, as per the protocols. The assessment sessions were audio-recorded directly onto the iPad and later uploaded to a MacBook Air 6.1 computer. Audio files were saved into the iPad in M4a format and converted into MP3 for transcribing using Express Scribe Transcription software.

Accuracy of transcription was checked with 10% of the sample randomly selected, de-identified and transcribed by a speech pathology student blinded to the study aims. Agreement was calculated to be 95.5%, with no errors found with content words. Errors typically occurred with utterance revisions and repetitions. The transcriptions were then collated and scores calculated for the NCA and SSIS.

Twenty percent of the sample were de-identified and blinded to examine inter-rater and intra-rater reliability on the developed task. A speech pathology student, who was not involved in pilot testing, tested inter-rater reliability on this data. One researcher rescored 20% of the sample in order to establish test-retest validity. Intraclass Correlation Coefficients (ICCs) were calculated to examine intra and inter-rater reliability of total scores.

To ensure robustness of the measure, ICCs were then calculated line by line on the SSIS assessment. Available points on the SSIS were divided into line items that represented

each target utterance. Thirty-seven target utterances were listed, with 0 – 3 points awarded for each target (see Appendix B). This process allowed the reliability of each line item to be investigated.

Results

Descriptive Statistics

Descriptive statistics for the SSIS across the three groups are presented in Table 2.

Table 2

Descriptive Statistics for Squirrel Story Information Score for Each Group

Group	n	M	SD	95% confidence interval	
				Lower bound	Upper bound
4;0-4;11	44	21.73	10.90	18.41	25.04
5;0-5;11	31	29.97	12.59	25.35	34.58
6;0-6;11	57	34.02	12.19	30.78	37.25

1. Reliability of the Squirrel Story Information Score

Internal consistency of the SSIS was assessed with intra-rater reliability calculated on 20% of the sample ($n = 26$). ICC was calculated on total scores with excellent reliability found between the repeated measures $ICC = .999$, $F(25) = 1654.90$, $p < .001$. Assumption of normality was supported. Fisher's skewness coefficient for the repeated measure was calculated from descriptive statistics (skewness/standard error of skewness) which indicated normal distribution, 0.956 and 0.973 respectively.

Inter-rater reliability of the SSIS was also investigated with ICC, calculated on the total scores, with excellent reliability found, $ICC = .994$, $F(25) = 308.46$, $p < .001$. Calculation of Fisher's skewness coefficient supported the assumption of normality, -0.063 and 0.081 respectively. Further analysis of inter-rater reliability was conducted by calculating ICCs for each line item of the SSIS. The scores ranged from .723 to 1.000 .

An ICC was also calculated to assess the inter-rater reliability of the NCA. Assumption of normality was supported (Fisher's skewness coefficient -0.429 and -0.476 respectively). Calculation of ICC revealed excellent reliability, $ICC = .991$, $F(13) = 219.87$, $p < .001$.

2. Sensitivity to Developmental Progression

A one-way between groups analysis of variance (ANOVA) was used to investigate the difference in information scores across the three age groups. The assumption of normality

was violated for scores in the 4;0-4;11 and 5;0-5;11 groups, however the ANOVA is robust to such violations. The homogeneity assumption was not violated as demonstrated by a non-significant Levene's statistic ($p = .63$). The ANOVA was statistically significant, supporting the hypothesis that the information score would differ between the three age groups, $F(2, 129) = 13.452, p < .001$ with a large effect size, $\eta^2 = .21$.

Post hoc analyses with Tukey's HSD demonstrated that the 4;0 to 4;11 group ($M = 21.73, SD = 10.90$) scored significantly lower ($p < .01$) than both the 5;0 to 5;11 group ($M = 29.97, SD = 12.59$) and the 6;0 to 6;11 group ($M = 34.02, SD = 12.19$) ($p < .000$). Effect sizes were $d = .52$ and $.91$ respectively. There was no significant difference between the SS information scores of the 5;0 to 5;11 group and the 6;0 to 6;11 group ($p = .281$), $d = .27$ (see Table 3).

Table 3
Performance on SSIS

Group	Main Items (/64)	Subsidiary items (/21)	Total (/85)
4;0-4;11	18.9	2.8	21.7
5;0-5;11	25.5	4.4	30.0
6;0-6;11	29.4	4.6	34.0

3. Correlation Between the Squirrel Story Information Score and Narrative Comprehension Task

To assess the relationship between the SSIS and the NCA, a bivariate Pearson's correlation coefficient was calculated to assess the convergent validity of the two measures. The assumptions of independence, normality, linearity and homoscedasticity were not violated. The bivariate correlation between the total scores for the two measures was positive and large, $r(130) = .58, p < .001$ (Cohen, 1988).

Post-hoc analysis explored the relationship between the total SSIS and NCA scores within each group, with significant correlations found in each age group (see Table 3). The relationships between total, literal and inferential scores on the NCA with the total SSIS score were also explored across the total sample and within each group (see Table 4).

Table 4
Results of Pearson's Correlation Coefficients, Including Post-hoc Exploration Within Groups and Individual Components of NCA Scores

Group	n	Total SSIS/Total NCA	Total SSIS/Literal NCA	Total SSIS/Inferential NCA
Total sample	132	.575**	.508**	.531**
4;0-4;11	44	.461**	.527**	.350.*
5;0-5;11	31	.528**	.379*	.461**
6;0-6;11	57	.399**	.236	.390**

Discussion

This study aimed to develop a reliable measure of free-recall by way of an information score as an index of oral discourse comprehension. The research design investigated whether the developed measure was sensitive to developmental progression by comparing results across three age groups of typically developing young school children. The study further explored the convergent validity of the novel free-recall measure by comparing results to the NCA measure of cued-recall.

Reliability of the Squirrel Story Information Score

Pilot testing of the newly developed Squirrel Story Information Score was followed by extensive investigation of inter-rater and intra-rater reliability calculated on total scores and each individual line item of the assessment. The ICCs calculated on the total and individual line items revealed excellent reliability. These results support the intra-rater and inter-rater reliability of the piloted measure, with correlations equivalent to studies that investigated the reliability of the commonly used assessment the Bus Story. The English version of the Bus Story does not provide formal reliability statistics; however other versions have addressed the lack of reliability testing. The American version of the Bus Story reports correlations ranging from $r = .7$ to $r = .92$ and the Italian version ICCs ranged from .82 to .96 for the information score measure (Mozzanica et al., 2016; Renfrew, 1997; Renfrew, Cowley, & Glasgow, 1994). The current study calculated inter-rater and intra-reliability on 20% of the transcripts. This was a comparatively large sample compared to the American version, which only rescored 6% of the sample, and smaller than the Italian version that rescored 32% of the sample. Whilst the Italian version reported excellent inter and intra-rater reliability, the analysis did not include calculation of ICCs for each item in the assessment. Overall, the results of intra-rater and inter-rater reliability of the SSIS are comparative to the results available for other similar assessment tasks given the robust process adopted in the development of the Squirrel Story Information Score.

Sensitivity to Developmental Progression

The Squirrel Story (Carey et al 2006) was developed as a measure of narrative skills in children aged 3 – 6 years. We hypothesised that the newly developed SSIS measure would show significant improvement across the three age groups: 4, 5 and 6 years. Our results partially supported the hypothesis with significant group differences found between ages 4 and 5, however not between 5 and 6. This suggested that the SSIS is sensitive to developmental progression in the early school years, however performance has begun to level out by age 6. These findings are consistent with previous research that employed an information score measure on a narrative retell (Mozzanica et al., 2016; Westerveld & Vidler, 2015). Mozzanica et al.'s (2016) study included children aged 4 – 8;11 and Westerveld and Vidler's (2015) participants were aged 5;3 – 8;9. Both of these studies reported information scores to increase significantly with age. However, similar to the present study, Westerveld and Vidler (2015) found no significant differences in the information scores in the older children comparing children aged 6;4-7;7, and 7;7-8;9 years.

Correlation Between the Squirrel Story Information Score and the Narrative Comprehension Assessment

A secondary aim of this study explored the convergent validity of the SSIS by comparing the novel measure of free-recall with that of cued-recall in the NCA. We predicted a significant correlation between the two measures, and the results supported the hypothesis with a statistically significant 33% of the variance in the total SSIS explained by the participant's ability to answer the comprehension questions in the NCA.

Post-hoc analysis sought to further explore the relationship between the measures by investigating the effect of age, and the relationship between the SSIS and the literal and inferential components of the NCA. The correlation between the total scores was shown to steadily decrease with age, whilst statistical analysis found a significant relationship in each age group, there was only 12% shared variance by age 6.

The NCA includes both literal and inferential comprehension questions, which tap into different skills. When examined separately, the correlation was significant for literal comprehension for 4 and 5 year olds, explaining 28 and 14 percent of variance respectively. However this reduced to a non-significant 6% for 6 year olds. One explanation for this finding might be that by age 6 children have reached ceiling for the literal questions, however this was not the case. Lower-level skills that are associated with discourse comprehension, such as vocabulary and grammar, are suggested to attribute to literal comprehension, particularly in younger children (Cain & Oakhill, 2007). This could be explained by a peak

increase in skills in the first years of schooling and the emergence of higher-level skills, such as inferencing, which become increasingly important in the development of oral discourse comprehension (Cain & Oakhill, 2007; Hogan et al., 2011; Paul & Norbury, 2012).

The lower-level skills associated with literal comprehension are suggested to be a precursor to developing inferential comprehension (Cain & Oakhill, 2007; Hogan et al., 2011). Research by Dawes (2018) on a cohort of pre-primary school children (aged 5-6 years) showed macrostructure, theory of mind, and literal comprehension to be significant predictors of inferential comprehension with 14.6%, 12.6% and 10.5% accounting for variance in inferential comprehension scores respectively. As inferential comprehension is a later developing skill, and related to literal comprehension, the results of this study suggest literal comprehension is more closely correlated to the ability to produce a narrative retell in younger children. For children aged 5-6 inferential comprehension becomes comparatively more correlated to the information score measure. Correlation between total SSIS and the inferential component of the NCA was .35 at age 4, reaching a peak of .46 at age 5 and dropping to .39 by age 6. The significant correlations between the SSIS and inferential comprehension scores between 4 to 6 years in this study align with the significant improvements in inferential comprehension found between 3 and 6 years in Filiatrault-Veilleux et al.'s study (Filiatrault-Veilleux et al., 2015).

Whilst the research suggests good narrative macrostructure and the creation of a mental model are important for inferential comprehension, the results of this study show a smaller amount of shared variance than the researchers predicted. It was thought that the developed measure of an information score would indicate a child's semantic ability, reflected by the integration of orally presented information into a schema in order to provide a temporally correct retell.

So what other skills are being reflected in the information score? Memory is an obvious skill required in a narrative retell, as it would be possible to produce an accurate retell simply by relying on memory alone (Cain et al., 2001; Korkman et al., 2007). It is also possible that a child who has difficulty with free-recall may perform better with cued support of comprehension questions, which may be indicative of declarative memory difficulties (Korkman et al., 2007).

Phonological short-term memory (PSTM) was not identified as a significant predictor of performance in the NCA cued-recall measure of oral comprehension (Dawes, 2017). However, free-recall by nature would rely on memory to a higher degree than cued-recall (Korkman et al., 2007). Efficiency in a range of processes including storage of semantic

information, ordering the information, rehearsal, retrieval and the reconstruction of stored propositions are said to be reflected in the development of PSTM (Gathercole, 1999). Interestingly, follow-up of participants in the Bishop and Edmundson (1987) study report that those with persistent difficulty and those whose language impairment had resolved both performed significantly lower on measures of sentence repetition, non-word repetition and spoonerisms compared to controls, skills that are related to PSTM ability (Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998). The processes associated with PSTM are essential in producing a comprehensive narrative retell, the development of which may explain the variability of SSIS scores to a greater extent than the hypothesised impact of semantics and schematic macrostructure development in this population. Future investigations should explore this potential relationship by measuring PSTM as well as the NCA and SSIS.

Limitations

This study was designed to establish the SSIS as a valid and reliable measure. Statistical analysis revealed excellent reliability and sensitivity to developmental progression. The Bus Story assessment has shown to be a valid predictor of school performance, with normative data provided by way of mean and standard deviations at six-month intervals. With rapid language development occurring in early school years, the collection of additional normative data will allow for the analysis of differences between smaller age brackets. Thus, future research should be directed at establishing further normative data to increase the validity and sensitivity of the measure. In this study the highest score achieved on the SSIS was 62/85; as such it would be clinically useful to have an understanding of the progression of scoring in older children.

Clinical Implications

This study collected normative reference data on typically developing Western Australian school children. The developed measure was able to capture significant differences between groups, proving it to be sensitive to developmental progression. The use of local participants goes some way to increasing the external validity of the SSIS in this population. The implementation of the common Australian curriculum framework (Australian Curriculum Assessment and Reporting Authority [ACARA], 2012) suggests that similar results would be expected in other states and territories.

It was considered of theoretical and clinical importance to compare the measures of cued-recall and free-recall. The results of the correlation analysis suggest that different

cognitive processes underpin the two measures, and as such assessment of oral narrative comprehension should include both measures. Broadly, we know that the calculation of an information score has good predictive validity for general academic performance, in comparison to inferencing ability, which has been linked more specifically to reading comprehension. Elicitation of a retell also allows clinicians to examine expressive language skills such as syntax, morphology and phonology, and further provides data for the calculation of measures like number of different words (NDW) and mean length of utterance (MLU). Therefore the elicitation of a narrative retell and the calculation of an information score provides detailed information regarding a child's language ability at the micro and macro level, and provides clinicians with information that can be compared to developmental norms and can be used to set therapy goals (Lawson, 2012; Fey et al., 2004).

Future Research Directions

This study identified PSTM as a possible contributing factor to the variance of the SSIS. It would be valuable for future research to investigate the relationship between PSTM and the SSIS to further identify the skills that account for the variability in the SSIS scores.

References

- Australian Curriculum Assessment and Reporting Authority. (2012). *National report on schooling in Australia 2012*. Retrieved from:
<http://www.acara.edu.au/reporting/national-report-on-schooling-in-australia-2012>
- Bishop, D. V. M. (2014). *Uncommon understanding: Development and disorders of language comprehension in children* (Classic Edition). Hove, UK: Psychology Press.
- Bishop, D. V. M. & Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential meaning. *Journal of Speech and Hearing Research, 35*. 119-129.
- Bishop, D. V. M, & Edmundson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*, 156-173. doi:10.1044/jshd.5202.156

- Bolderson, S., Dosanjh, C., Milligan, C., Pring, T., & Chiat, S. (2011). Colourful semantics: A clinical investigation. *Child Language Teaching and Therapy*, 27(3), 344-353. doi:10.1177/0265659011412248
- Burris, S. E., & Brown, D. D. (2014). When all children comprehend: Increasing the external validity of narrative comprehension development research. *Frontiers in Psychology*, 5, 168. doi:10.3389/fpsyg.2014.00168
- Cain, K. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *British Journal of Developmental Psychology*, 21(3), 335-352. doi:10.1348/026151003322277739
- Cain, K., & Oakhill, J. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology*, 14(2), 187–201. doi:10.1111/j.2044-835x.1996.tb00701.x
- Cain, K., & Oakhill, J. V. (2007). *Children's comprehension problems in oral and written language: A cognitive perspective*. New York, NY: Guilford Press.
- Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, 29(6), 850–859. doi:10.3758/bf03196414
- Carey, J., Leitão, S. & Allen, L. (2006). *Squirrel Story Narrative Assessment*. Keighley: Black Sheep Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.), Hillsdale, NJ: Erlbaum.
- Cosentino, E., Adornetti, I., & Ferretti, F. (2013). Processing narrative coherence: Towards a top-down model of discourse. *Workshop on Computational Models of Narrative*, 61-75. doi:10.4230/OASICS.CMN.2013.61
- Dawes, E., Leitão, S., Claessen, M., & Lingoh, C. (2019). Oral literal and inferential narrative comprehension in young typically developing children and children with developmental language disorder. *International Journal of Speech-Language Pathology*, 21(3). doi:10.1080/17549507.2019.1604803
- Dawes, E. (2017). *The hidden language skill: oral inferential comprehension in children with developmental language disorder* (Doctoral thesis, Curtin University, Perth, Western Australia). Retrieved from: <https://espace.curtin.edu.au/>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191. doi:10.3758/bf03193146

- Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, J. B., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research, 47*(6), 1301-18. doi: 10.1044/1092-4388(2004/098)
- Filiatrault-Veilleux, P., Bouchard, C., Trudeau, N. & Desmarais, C. (2015). Inferential comprehension of 3-6 year olds within the context of story grammar: A scoping review. *International Journal of Language & Communication Disorders, 50*(6). 737-749. doi:10.1111/1460-6984.12175
- Florit, E., Roch, M. & Levorato, M. C. (2011). Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills. *Discourse Processes, 48*(2). 119-138, doi:10.1080/0163853X.2010.494244
- Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*(11), 410-419. Retrieved from <https://search-proquest-com.dbgw.lis.curtin.edu.au/docview/85516406?accountid=10382>
- Gillam, R.B., & Pearson, N.A. (2004). *Test of Narrative Language*. Austin, TX: Pro-ed.
- Hogan, T.P., Bridges, M.S., Justice, L.M, & Cain, K. (2011). Increasing higher level language skills to improve reading comprehension. *Special Education and Communication Disorders Faculty Publications. 79*. Retrieved from: <http://digitalcommons.unl.edu/specedfacpub/79>
- Hogan, T. P., Cain, K., & Bridges, M.S. (2012). Young children's oral language abilities and later reading comprehension. In T. Shanahan & C. J. Lonigan (Eds), *Early childhood literacy: The National Early Literacy Panel and beyond* (pp. 217-232) Baltimore, MD: Brookes.
- Hutchinson, J., & Clegg, J. (2011). Education practitioner-led intervention to facilitate language learning in young children: An effectiveness study. *Child Language Teaching and Therapy, 27*(2), 151-164. doi:10.1177/0265659010397232
- Joffe, V. L., Cain, K., & Maric, N. (2007). Comprehension problems in children with specific language impairment: Does mental imagery training help? *International Journal of Language & Communication Disorders, 42*(6), 648-664. doi:10.1080/13682820601084402
- Johnston, P & Afflerbach, P. (1985). The process of constructing main ideas from text. *Cognition and Instruction, 2*, 207-232, doi:10.1080/07370008.1985.9648917
- Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of decoding

- and oral language skills. *Journal of Educational Psychology*, *101*, 765–778.
doi:10.1037/a0015956
- Kendeou, P., Lynch, J. S., van, d. B., Espin, C. A., White, M. J., & Kremer, K. E. (2005). Developing successful readers: Building early comprehension skills through television viewing and listening. *Early Childhood Education Journal*, *33*(2), 91-98.
doi:10.1007/s10643-005-0030-6
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. doi:10.1037/0033-295x.85.5.363
- Korkman, M., Kirk, U., & Kemp, S. (2007). *A developmental neuropsychological assessment* (2nd ed.). San Antonio, TX: Psychological Corporation.
- Lawson, K. (2012). The real power of parental reading aloud: Exploring the affective and attentional dimensions. *Australian Journal of Education*, *56*(3), 257-272.
doi:10.1177/000494411205600305
- Lynch, J. S. & Van den Broek, P. (2007). Understanding the glue of narrative structure: Children’s on- and off-line inferences about characters. *Cognitive Development*, *22*(3), 323-340. doi:10.1016/j.cogdev.2007.02.002
- Morrow, L. M. (1985). Retelling stories: A strategy for improving young children's comprehension, concept of story structure, and oral language complexity. *Elementary School Journal*, *85*(5), 647-661. doi.org/10.1086/461427
- Mozzanica, F., Salvadorini, R., Sai, E., Pozzoli, R., Maruzzi, P., Scarponi, L., . . . Schindler, A. (2016). Reliability, validity and normative data of the Italian version of the bus story test. *International Journal of Pediatric Otorhinolaryngology*, *89*, 17-24.
doi:10.1016/j.ijporl.2016.07.029
- Nation, K. (2019). Children’s reading difficulties, language, and reflections on the simple view of reading. *Australian Journal of Learning Difficulties*, *24*(1), 47-73.
- Norbury, C. F., & Bishop, D. V. (2003). Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders*, *38*(3), 287-313. Retrieved from <https://search-proquest-com.dbgw.lis.curtin.edu.au/docview/61980468?accountid=10382>
- Paris A. H., & Paris S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly*, *38* (1), 36–76. doi:10.1598/RRQ.38.1.3
- Paul, R., & Norbury, C.F. (2012). *Language disorders from infancy through adolescence* (4th ed.). St Louis, Missouri: Elsevier.

- Paul, R., & Smith, R. L. (1993). Narrative skills in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech and Hearing Research, 36*(3), 592-598. doi:10.1044/jshr.3603.592
- Peterson, D. B., Gillam, S. L., & Gillam, R. G. (2008). Emerging procedures in narrative assessment. *Top Language Disorders 28*(2), pp. 115-130. doi:10.1097/01.TLD.0000318933.46925.86
- Renfrew, C. E. (1997). *The Bus Story (4th ed.)*. United Kingdom: Speechmark Publishing Ltd.
- Renfrew, C., Cowley, J., & Glasgow, C. (1994). *The Renfrew bus story language screening by narrative recall (American edition)*. The Centerville School Delaware.
- Ripich, D. N., & Griffith, P. L. (1988). Narrative abilities of children with learning disabilities and nondisabled children: Story structure, cohesion, and propositions. *Journal of Learning Disabilities, 21*(3), 165-173. doi:10.1177/002221948802100309
- Stein, N. L., & Albro, E. R. (1996). The emergence of narrative understanding: Evidence for rapid learning in personally relevant contexts. *Issues in Education, 2*(1), 83-98. Retrieved from <https://search-proquest-com.dbgw.lis.curtin.edu.au/docview/85595241?accountid=10382>
- 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 Processesing, Vol. 2*, (pp 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 and Hearing Research, 41*(2), 407. doi:10.1044/jslhr.4102.407
- Van den Broek, P. (1990). Causal inferences and the comprehension of narrative texts. *Psychology of Learning and Motivation, 175-196*. doi:10.1016/s0079-7421(08)60255-8
- van Kleeck, A. (2008). Providing preschool foundations for later reading comprehension: The importance of and ideas for targeting inferencing in storybook-sharing interventions. *Psychology in the Schools, 45*(7), 627-643. doi:10.1002/pits.20314
- van Kleeck, A., Woude, J. V. & Hammett, L. (2006). Fostering literal and inferential language skills in head start preschoolers with language impairment using scripted book-sharing discussions. *American Journal of Speech-Language Pathology, 15*. 85-95. doi:10.1044/1058-0360(2006/009)

- Wenner, J. A. (2004). Preschoolers' comprehension of goal structure in narratives. *Memory (Hove, England)*, 12(2), 193-202. doi:10.1080/09658210244000478
com.dbgw.lis.curtin.edu.au/docview/66702936?accountid=10382
- Westerveld, M. F., & Claessen, M. (2014). Clinician survey of language sampling practices in Australia. *International Journal of Speech-Language Pathology*, 16(3), 242-249.
- Westerveld, M. F., & Vidler, K. (2016). Spoken language samples of Australian children in conversation, narration and exposition. *International journal of speech-language pathology*, 18(3), 288-298.