Research Papers

What Guides Inference Generation? A Study of Young Hindi Learners Studying in Challenging Contexts in India

*Lina Mukhopadhyay, **Vasim Salim Tamboli, ***Kankan Das, ****Anusha Balasubramanian and *****Ianthi Maria Tsimpli

Abstract

In a vast multilingual country like India, primary education is offered in various languages as Mediums of Instruction (MoI) to support the state languages and foster social justice in education. An important milestone of primary education is to help learners develop comprehension skills in MoI to process information and lay a strong foundation for education. Comprehension involves 'inference generation skill' that helps learners formulate multiple possible answers. Teaching and assessment in India have a product-based content testing approach and teachers are not trained to deal with individual differences in responses in a constructive manner. In this paper, learners are assessed in oral and print mode to understand what gives rise to individual differences in comprehension through Hindi as MoI. A group of 30 bi or multilingual learners, 7 to 12 years old, attending Class IV in state run primary schools in Bihar (India) participated in the study. A quantitative analysis of learner performance shows that inference generation is affected by

*Professor, English as Foreign Lanuages University, Hyderabad (e-mail: linamukhopdhayay@efluniversity.ac.in).

**MultiLiLa Research Associate, The English and Foreign Languages University, Hydrebad.

***Assistant Professor, Department of English Studies, Christ (Deemed to be University), Bengaluru.

****Postdoctoral Research Associate at the University of Cambridge, Cambridge

*****Professor and Chair of English and Applied Linguistics at the University of Cambridge and Fellow of Fitzwilliam College, Cambridge.

modality (oral or print), gender and the complexity of inferences. A qualitative analysis of individual variations shows that of the total number of inappropriate responses to comprehension questions, many refer to experiential or world-knowledge inferences but fail to link them to the specific story-based information. This indicates difficulties with inference generation and the ability to select only the relevant parts of the response. The findings have implications for pedagogical methods of promoting inference generation skills using world knowledge in combination with text-based information to offer meaningful feedback.

Keywords: *Narrative comprehension, Medium of Instruction (MoI), inference generation, propositional inference, pragmatic inference, world knowledge.*

सार

भारत जैसे विशाल बहुभाषी देश में स्थानीय भाषाओं को महत्व देने हेतु प्राथमिक शिक्षा स्थानीय राज्य की भाषाओं में प्रदान की जाती है, ताकि शिक्षा में सामाजिक न्याय को बढ़ावा दिया जा सके। प्राथमिक शिक्षा का एक महत्वपूर्ण आयाम सूचनाओं को संसाधित करने और शिक्षा के लिए मजबूत नींव रखने हेतु शिक्षार्थियों की शिक्षण के माध्यम में 'समझ कौशलों' को विकसित करने में मदद करना है। इस समझ में 'अनुमान जनन कौशल' शामिल है जो शिक्षार्थियों को कई संभावित उत्तर प्रतिपादित करने में मदद करता है। भारत में शिक्षण और मूल्यांकन उत्पाद-आधारित सामग्री परीक्षण उपागम पर निर्भर है और शिक्षकों को रचनात्मक तरीके से प्रतिक्रियाओं में व्यक्तिगत भिन्नताओं से निपटने के लिए प्रशिक्षित नहीं किया जाता है। इस शोध पत्र में, शिक्षार्थियों का मौखिक और प्रिंट माध्यम से परीक्षण किया गया जिससे यह ज्ञात हो कि मातृभाषा हिंदी द्वारा समझ विकसित करने के कारकों का ज्ञान हो सके। बिहार राज्य में संचालित प्राथमिक विद्यालय के कक्षा IV में पढ़ने वाले 7 से 12 वर्ष की आयु के 30 बहुभाषी शिक्षार्थियों के एक समूह ने इस अध्ययन में भाग लिया। शिक्षार्थी के निष्पादन का मात्रात्मक विश्लेषण दर्शाता है कि साधन (मौखिक या प्रिंट), लिंग और अनुमानों की जटिलता से अनुमान जनन प्रभावित होता है। व्यक्तिगत विभिन्नताओं के गुणात्मक विश्लेषण से पता चलता है कि समझ से संबंधित प्रश्नों के प्रति अनुचित प्रतिक्रियाओं की कुल संख्याओं में से कई अनुभवात्मक या विश्व-ज्ञान के संदर्भों का उल्लेख करते हैं लेकिन उन्हें विशिष्ट कहानी-आधारित जानकारी से जोड़ने में विफल रहते हैं। यह अनुमान जनन और उत्तरों के केवल प्रासंगिक भागों का चयन करने की योग्यता के साथ जुड़ी कठिनाइयों को बताता है। अध्ययन के परिणाम अर्थपूर्ण प्रतिपुष्टि प्रस्तुत करने के लिए पाठ-आधारित जानकारी के साथ विश्व ज्ञान का उपयोग करके अनुमान जनन कौशलों को बढ़ावा देने के हेतु शैक्षणिक विधियों का उपयोग करने पर बल देता है।

Introduction

In a vast multilingual country like India, primary education is offered in different languages as Mediums of Instruction (MoI) to support the state languages and foster social justice in education (Mohanty 2018; Enever 2018). However, a country that has a vast linguistic diversity with 462 spoken languages, MoI is available only in 31 languages, and has been reduced from 67 languages in the 1970s, as reported in a study by Meganathan (2011).

A large number of Indian schools, run by central or state governments, at the primary level offer at least two or three languages. But learning needs to happen mostly through the medium of instruction. Many learners from low SES families fail to reach grade appropriate comprehension skills in MoI because there is less or no overlap between home and school language, which creates a linguistic barrier and leads to poor learning outcomes. This has been widely documented in the recent years by ASER survey reports (Pratham, 2014, 2017) and research in Indian classrooms (Meganathan, 2011). Another critical learning problem came out in recent research in Indian classrooms is that independent study skills are underdeveloped because the classes are teacher-centric with a lot of read aloud and repetition exercises and assessment being based on rote learning of prescribed answers to questions that test content knowledge (Alcott and Rose, 2017; Banerji, 2017; Lightfoot et al., submitted). Furthermore, since children and teachers are multilingual, language mixing in classrooms is an often pedagogical strategy, though much of the multilingual practices are not planned to support learning the target language; they largely stay at the level of literal translation (Clarke, 2001, 2003; Wei, 2017; Brinkmann, 2018; Mathew et al., submitted). So both home environment and classroom practices are not conducive to the development of knowledge of MoI in children from low SES families (Jhingran, 2009; Mohanty, 2018).

Challenges in Developing Comprehension Skills in Primary Schools in India

Primary schools running by state governments in India frequently face the challenge to develop comprehension skills in the medium of instruction (MoI). Language textbooks contain a higher number of narratives and a few expository texts that serve as inputs based on which decoding and comprehension skills in the MoI are

expected to develop. Comprehension is an important school skill as it helps learners process information and understand concepts across subjects and lay a strong foundation of education. However, to develop comprehension skills it becomes incumbent on one's vocabulary, phonological awareness and morpho-syntactic knowledge in the Medium of Instruction (MoI) as well as word decoding skills. A majority of the urban poor and rural Indian children who come from low SES families live in multilingual environments, but their parents are not able to support them with literacy practices at home. Quite naturally their academic proficiency in MoI is low and development of higher order study skills such as decoding and comprehension, either oral or print, is negatively affected (Jhingran, 2005).

In addition to low proficiency in MoI, another factor that hinders development is instructional and assessment practices. Indian teachers in low cost government schools are found to heavily depend on textbooks and focus more on memorisation of factual information, read aloud sessions and translation of content instead of developing independent comprehension and critical reading skills (Clarke, 2003; Banerji, 2017; Mukhopadhyay et al., 2020). Indian school education policy documents bring out the lack of transacting knowledge through comprehension in class. A compelling document identifying that 'the load of non-comprehension' in school going Indian children 'is a pernicious burden' causing poor quality of education was brought out in the 1993 government of India report, Learning Without Burden (also known as Yashpal Committee Report). Similar arguments echo in the *Position Paper of the National Focus Group on Indian Languages* and *Position Paper on the National Focus Group on English* (NCERT, 2006), and the *National Curriculum Framework* (*2005*). Poor word decoding and reading comprehension outcomes have also been reported by ASER studies conducted with 6,00,000 children across India. The ASER reading test has shown that more than half of all children in Class V could not read a Class II level text fluently (Pratham, 2017). This results in poor learning outcomes and school dropouts, affecting girls more than boys (*Unesco's Education Report, 2015*). In practice this problem continues to impact students in low cost government schools. This paper makes an attempt to study critical issues in the development of comprehension ability of learners in schools that have Hindi as the medium of instruction

in the state of Bihar. The study would look at comprehension difficulties of young Indian learners who are from low SES families and lack literacy exposure and parental support at home.

Assessing Comprehension

Successful comprehension, either in the oral or print mode, involves a higher-level skill—inference generation—which allows learners to develop a deeper understanding of the text. It helps in expanding and paraphrasing messages in more than one way, learning to understand new content, and make connections between old and new information. As teaching and assessment in India are largely approached as product based learning (Clarke, 2003; Chand, 2011; Smith et al., 2005), teachers are not always trained to assess learners' comprehension skills in a constructive manner. So they fail to draw estimates of individual differences in comprehension levels. Thus, when learner responses do not match the teacher key, they are often considered erroneous and therefore discarded (Brinkmann, 2018; Milian Winch, 2019). It is essential for teachers to appreciate that being able to infer is proof of growth in the learner's cognitive-linguistic skills because they can now 'access' and 'generate' ideas to fill-in information that is implicit or unsaid in a text (Kendeou, 2015; Cain et al., 2001). It also enhances a child's social cognition, as inferences can be guided by their experiential knowledge helping her assign mental states to herself and others (Astington and Jenkin, 1995). So, inference generation skills are likely to give rise to individual differences in comprehension skills and most importantly teachers need to recognise and handle such differences effectively.

Assessing individual differences in comprehension skills and improving pedagogical practices to develop such skills has been a neglected area of research in the Indian context. This paper attempts to explore this gap by assessing comprehension skills and analysing variations in responses to understand how to give feedback to develop inference generation skills. So, in this paper individual differences in comprehension will be examined in primary level low SES Indian children for the first time. Through an in-depth analysis of learner attempts to derive inferences within a range of variation from the constructed key of the narrative tool, the paper in a novel approach aims to also raise teacher awareness of how variation in individual responses can provide material for constructive, meaningful feedback.

The MultiLiLa Project

Recent research in India led by the University of Cambridge in collaboration with Indian universities has revealed that children from low SES and challenging contexts have better learning outcomes when the MoI overlaps with at least one of their home languages (Tsimpli et al., 2020). This research is carried out within the Multilingualism and Multiliteracy (MultiLiLa) project (Tsimpli et al, 2019) which aims to increase awareness and knowledge about the link between learning levels of literacy and numeracy, cognitive abilities and the languages of instruction that multilingual children in India are exposed to. The project is a four-year longitudinal study, beginning in 2016, and operates across three research sites— Delhi, Hyderabad (Telangana) and Patna (Bihar). In addition to administering a series of language, numeracy and cognitive tasks to the children, the project also involves observations of lessons in the schools these children attend. Broadly, these observations seek to explore within 30 minutes of lessons— (i) what languages are being used by the teacher and the learners, (ii) at what stages during the lesson and (iii) accompanying what types of activities.

Evidence from the MultiLiLa project from language comprehension, oral and written, of a large cohort of 498 learners in Class IV enrolled in Hindi MoI schools in Patna and in Delhi has revealed interesting discrepancies between the two modalities. The children were tested on reading comprehension using the ASER literacy tool, which was supplemented with two novel comprehension questions about a story corresponding to the highest level of reading skill. The results from the reading assessment show that learners show good letter and word decoding skills (85 per cent accuracy) but drop in sentence and text reading (72 per cent) while score lowest in story comprehension (67 per cent) (Balasubramanian et al., 2020). In the same study, the same cohort of children was tested in oral comprehension skills using picture-retelling tasks taken from the MAIN manual (Gagarina et al., 2012). In contrast to reading comprehension scores, oral narratives showed better comprehension skills.

The present study focuses on a small set of 30 children from the larger cohort of children studying in Patna schools with the aim to compare narrative oral comprehension with reading

comprehension responses more closely. The study focuses on inference generation, a higher order comprehension skill, and how children use a combination of text-based situation model and world knowledge to draw inferences (Perfetti and Stafura, 2015).

Inference Generation

During reading, the reader constructs meaning at various levels, which might be accurate or near accurate with respect to the text's content (Chikalanga, 1992; Cain et al., 2001). Text understanding is reflected through a variety of inferences according to the learner abilities to make connections between ideas presented in a text (text-based inferences) or to retrieve and connect background knowledge (from long term memory) during the meaning making process (world knowledge based inferences). Thus, inference generation involves three cognitive processes— (a) access background or world knowledge, (b) integrate it with text information to (c) generate links that are not explicitly stated in the text (Kendeou, 2015; Cain et al., 2001; Perfetti and Stafura, 2015).

Inference generation can be further subdivided to help language teachers design questions for assessing this skill by taking into account different comprehension needs and levels of learners. To substantiate this need of designing inference generation questions at different levels of complexity, Chikalanga proposes inference generation taxonomy with two categories (p. 698):

- Propositional inferences (ProIs) also called text-based inferences are considered 'true-inferences' because these are generated from the content of the text. ProIs are 'convergent' in nature since only one acceptable answer is available based on the linguistic input provided in the text; thus within a contextual setting or story structure, disagreements on ProIs are less likely.
- Pragmatic inferences (PraIs) or world knowledge based inferences, on the other hand, are considered 'not always true' because they depend upon prior knowledge or world experience which are arbitrary; PraIs are 'invited inferences' as these inferences are beyond the linguistic input given in the text, and 'divergent' in nature since there is always a possibility of more than one answer or variation in the inferences because of prior knowledge, individual differences, and shared communicative context.

Since PraIs rely upon the understanding of both the text and the context of the text, they involve a higher level of cognitive

processing. This is also one of the reasons that PraIs are more difficult than inferences that are ProIs. Therefore, although PraIs are considered 'not always true', they can assess learners' higher level of meaning making abilities while comprehending a text. But such inferences can also give rise to individual differences in inference generation abilities.

Assessing Inference Generation Abilities

The human mind has information about people, actions, events, objects, nature and abstract concepts. This information is based on a person's family, social and academic experience and referred to as one's world knowledge structure. This knowledge contributes to one's inference generation abilities. While comprehending a text on a familiar topic, a reader derives world knowledge based inferences as they relies on various information sources like prior knowledge or real-world experience.

Similarly, the reader forms a situation model, i.e., constructs a micro-world of the text information with links in between ideas (e.g., characters, actions, events), which helps to select and track what is important in the text to comprehend. This creation of a micro-world resembles everyday experience from the real-world which nobody can take away or detach the reader from. As a result, if these structures are activated in the process of text comprehension and the reader incorporates world-knowledge related content, chances are higher for accurate text-based inference to take place (Perfetti and Stafura, 2015). Therefore, to aid learners' comprehension abilities, the key assumption is the need to create a substantial number of inference generation opportunities to help them make a rich representation of the text in mind. Also, to validate the inferences generated by learners by reflecting on their world knowledge structures, there is a need to create or use tasks that account for their world knowledge as well as take into account different responses of learners (Aukerman et al., 2017). Using a verbal protocol analysis to understand learners' engagement with their thought processes (Cote et al., 1999), asking them questions to have an estimate of their understanding of the text in parts (Long and Golding, 1993), and assessing representation of story grammar in the extended text (Gagarina et al., 2012; Taylor 2013) are a few ways to assess skills of both propositional and pragmatic inferences.

To summarise, in using inference generation learners are likely to show individual differences and it is crucial for teachers to

gain awareness about how to accommodate variations in learner responses to recognise different levels of these abilities and provide meaningful feedback (Cain et al, 2001; Aukerman et al., 2017). This paper presents, through an in-depth analysis of learner responses, how individual differences in inference generation can be identified and treated.

The Study

The study reported here is part of the four year MultiLiLa project presented above. Here we report on the comprehension of 30 Hindi speaking learners studying in four primary schools in Patna, the capital of Bihar, India. In Bihar state, Hindi is the local language, although people also speak Magahi, Maithili, Bhojpuri, and Urdu, which make many of the learners bi or multilingual (44 per cent).

Research Questions

The study addresses the following research questions:

1. Do language comprehension questions presented in the oral and the print modality help learners generate inferences equally well?
2. Do comprehension skills differ in boys and girls?
3. Do learner responses vary across different types of inference generation?
4. Are there individual differences in inference generation for narrative comprehension and if so, which are they?

Method

Participants

Thirty learners (M=16; F=14) from Class IV participated in this study. Their age ranges from 8 to 11 years (mean age = 9.44; SD = 0.89). Of the 30 learners, 27 reported as Hindi to be their L1 (90 per cent), while 1 reported Bhojpuri (3 per cent) and 2 reported Urdu (6 per cent) to be their L1. Note that though 90 per cent of the children reported that they speak hindi at home, the variety is different from the standard variety used as MoI in schools they are enrolled. The variety of hindi used for school instruction is considered '*shudh*' or standard and the learners are expected to develop academic proficiency in it. The absence of use of standard variety of hindi and use of other dialectical variations of it at home is likely to create a gap between home and school language and

this may put the children at a disadvantage. Of the 30 learners, 13 learners (44 per cent) used more than one language at home, whereas 17 learners (56 per cent) stated they knew only one language, namely hindi.

Table 1: Learner profile

City (State)	Medium of Instruction (MoI)	Number of Children (N)	Age (Range in years)	Age Mean (SD)	Child bilingualism	MoI (Hindi) overlap with home language	Parent occupation (with literacy practices)
Patna (Bihar)	Hindi	30 (F=14; M=16)	8 years 4 months to 11 years 8 months	9.44 (0.89)	44 %	100%	53.33%

In the study, a child questionnaire was used with subsections on child bilingualism and on socio-economic details and it was found that the learners were from lower socio-economic (SES) background families whose parents work mostly as daily labourers, vegetable vendors, auto rickshaw drivers and so on. These professions are without much involvement of literacy practices and serve as a low indicator of parental literacy.

Tasks Used

Narrative Comprehension Task

The Cat Story from the MAIN (Gagarina et al., 2012, pp. 111–117) was adapted to an informal oral variety of Hindi and an Indian male voice was used to record the audio input. The story had a setting (time and place) and three short episodes each comprising story structure elements such as—Goal, Attempt and Outcome. The three episodes were presented through a sequence of three picture panels and each panel had two pictures in it. So, the three episodes were depicted through six pictures (refer to Appendix 1 for the story details). The story had nine comprehension questions to draw upon inference generation abilities (refer to Table 5 in Appendix 1).

For assessing comprehension skills, scoring sheet was used to assess performance on the comprehension questions and the key provided in the MAIN manual (refer to Tables 2 and 3 below).

These comprehension questions were in sync with story structure elements (refer to Table 5 in Appendix 2). For example, three questions (Questions 1, 3 and 4) targeted three goals (Cat, Boy, and Cat). The other six questions elicited answers around internal state terms, i.e., terms expressing physiological (e.g., hungry, thirsty), perceptual (e.g., see, hear), emotional (e.g., happy, sad), linguistic (e.g., say, tell), mental (e.g., think, believe) and consciousness terms (e.g., alive, awake, asleep) (see Gagarina et al, 2012 for details).

For the purpose of examining two types of inference generation, the questions were classified into two categories:

- Propositional inference (ProI): Questions 1, 2a, 3, 4, and 5a
- Pragmatic inference (PraI): Questions 2b, 5b, 6a, and 6b

The comprehension questions to which answers can be logically derived from the audio and video input were classified as propositional inference. For example, for Question 3 (Why does the boy hold the fishing rod in the water?) the answer can be logically derived from the text (At the same time the boy began pulling his ball out of the water with his fishing rod). The comprehension questions which did not have explicit answers in the audio text are categorised under the PraI category. For example, Question 6b (Imagine that the boy sees the cat, how does the boy feel?) does not have its explicit answer in the linguistic input provided in the text.

The researcher read out the questions to each learner in hindi one at a time and participants answered in the same language. If any answers or responses were not convergent with the answer key, the researchers noted down such responses under the comment section in the scoring sheet. Based on the differential answers provided by the children in response to the questions, we will analyse these answers and consider if such answers can be improved with the help of feedback.

ASER Literacy Task

The ASER literacy tool in hindi (Pratham 2014, 2017) was used which measures: letter naming (10 items), single word reading (10 words), reading of sentences (4 sentences), reading of a story (9 sentences), and two novel propositional inference questions (2 questions). This paper reports, results from the two propositional inference questions each carrying one mark.

Results

Results on learner performance according to the four research questions (RQs) are reported.

Research Question 1 and 2: Impact of modality and gender on comprehension

Learner performance on narrative comprehension across two modalities—oral and print (ASER test) in boys and girls are presented in Figure 1.

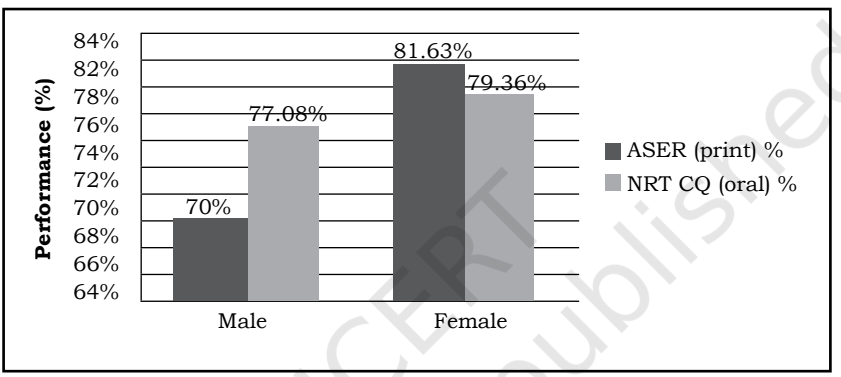

Figure 1: Performance on ASER test and oral narrative comprehension

Figure 1 shows a gender effect on narrative and reading comprehension scores. Specifically, girls performed better at 81 per cent than boys at 74 per cent across oral and written modality although this difference was not found to be significant on a one way within group ANOVA test (Gender X Test type) on the ASER comprehension (df =1.28, F =0.109, p<0.74) or the narrative comprehension (df =1.28, F = 0.802, p<0.37) tasks. Performance on narrative comprehension (78.14 per cent) was higher than performance on reading comprehension (75.42 per cent); this difference was found to be significant on a t-test (df=29, t= 23.05*, p<0.05). But performance in the two tasks was not correlated (r = 0.05, n.s.) meaning that a higher level of oral comprehension did not predict better performance in print comprehension, contrary to previous findings that oral skills correlate with print skills corroborating the simple view of reading (Cain et al. 2001).

Research Question 3: Impact of levels of difficulty of inference on comprehension

Learner performance on narrative comprehension under two types— (i) propositional inference and (ii) pragmatic inference are compared with (iii) propositional inferences from the ASER test and presented in Figure 2.

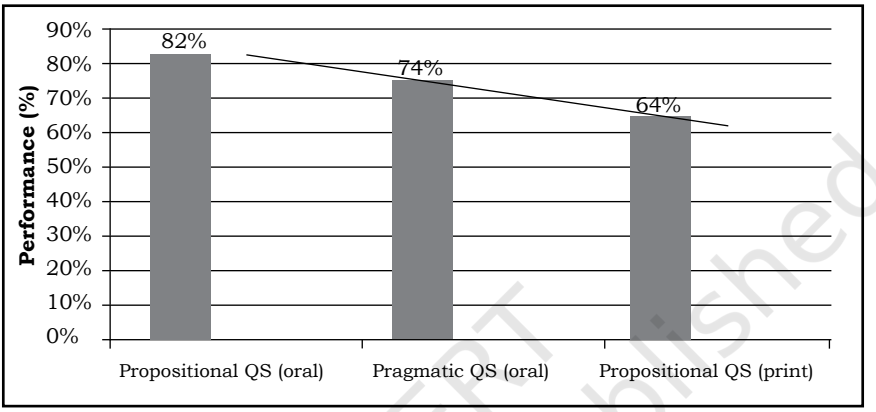

Figure 2: Performance on propositional (oral vs print) and pragmatic inference

Propositional inferences in the narrative task was highest (82 per cent) followed by pragmatic inferences in the same task (74 per cent) while lowest performance was in propositional inferences in the reading task (64 per cent). The three measures significantly differed in a one-way GLM repeated measures test (1X1X1: inference type) (df=2, F=4.16*, p<0.05). Oral propositional inferences correlated significantly positively with oral pragmatic inference (r=0.55* , p<0.02). But surprisingly propositional inferences in the oral and reading tasks were not correlated (r=0.14, n.s). This implies that within the same modality, here oral, there is a positive correlation between a lower and higher level of inference generation performance.

Research Question 4: Individual differences in learner responses in inference generation

An in-depth qualitative analysis of instances of the range of learner responses—from more to less accurate—was made to understand what gives rise to individual differences in inference generation.

The range of responses were analysed to look at the proportion of responses that do not converge with the answer key and how this can form material for teachers to provide constructive feedback. Note that this analysis is with respect to only the oral narrative comprehension task. Since, the print comprehension task had only propositional inference, such an analysis was not undertaken. The analysis of learner errors across the questions for propositional inference (Figure 3) and pragmatic inference (Figure 4) are presented.

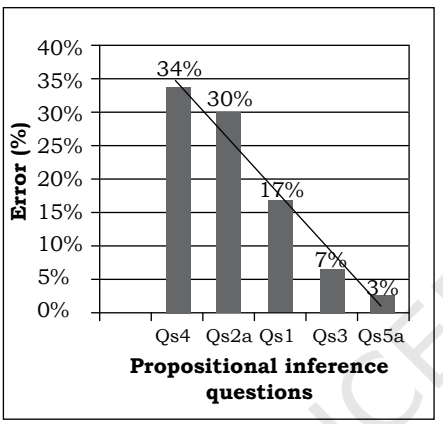

Figure 3: Propositional inference

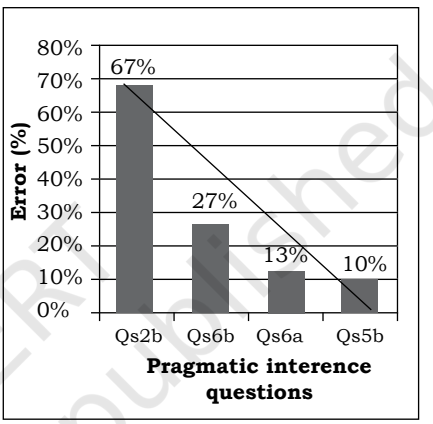

Figure 4: Pragmatic inference

The errors are presented as percentage frequency count in a decreasing order across the two inference categories. Note that learners are less accurate with more demanding inferences as indicated in Figure 4 in comparison to Figure 3 above.

Let us now look at some examples of actual learners responses from propositional and pragmatic questions in a qualitative manner to understand what gave rise to individual variations in narrative comprehension.

Qualitative Analysis of Propositional Inference Questions

Table 2 contains the frequency and range of answers for each propositional question divided into four categories—(i) correct (ii) partial situation knowledge, (iii) world knowledge and (iv) irrelevant or incorrect answers.

Table 2: Learner differences in propositional inference questions

Propositional Questions	Correct	Partial situation knowledge	World knowledge	Irrelevant
Q1: Why does the cat jump or leap forward? Ans: wants/to get/ catch/chase the butterfly/to play with the butterfly	25	On seeing the butterfly	It liked the butterfly	2
		On seeing the butterfly		
Q2a: How does the cat feel? Ans: Angry/ bad/ disappointed/ hurt	21	Will catch it again	It was hungry	1
		To catch the butterfly	It felt like eating the butterfly	
		I will catch the butterfly again	It felt like eating	
			It is tasty	
			To eat the butterfly	
Q3: Why does the boy hold the fishing rod in the water? Ans: Wants/ to get his ball back	28	0	0	2
Q4: Why does the cat grab the fish? Ans: Decided/wants to eat/have/steal the fish/takes the chance/ opportunity when the boy is not looking	20	It was hungry	Because the cat likes the fish	0
		It became greedy after seeing the fish	It likes the fish so much	
		It could not get the butterfly so started eating fish	It felt good to see the fish	
		It was hungry	Because cat was feeling good to see the fish	
		It was feeling sad	The felt good to see the fish	

Q5a: how does the boy feel? Ans: good/ fine/ happy/ please	29	0	0	1

Of the five questions, answers to two questions—2a and 4—have more instances of divergent answers. For instance, for Question 2a, 30 per cent of the learners were either not able to integrate reaction of the cat when it fails to catch the butterfly; they either expressed a partial situation model that the cat would try to catch the butterfly again or they used world knowledge of another episode where a cat can eat a fish and linked it to the butterfly and this is a faulty inference. For Question 4, the intension or goal of the cat is to be expressed. But 34 per cent of the learners were not able to build this inference that the cat in his mind decides to steal and eat the fish. They either gave the mental state of the cat being hungry or greedy, which is a partial situation model integration or they used world knowledge that cats like fish.

Qualitative Analysis of Pragmatic Inference Questions

Table 3 contains individual differences in answering the pragmatic questions presented in four categories as mentioned for propositional inference based responses in Table 2.

Table 3: Learner differences in pragmatic inference questions

Pragmatic Questions	Correct	Partial situation knowledge	World knowledge	Irrelevant
2b: Why do you think the cat is feeling angry/bad? Ans: Because cat could not catch the butterfly/fell into the bush/it hurts to fall into a pricky bush	10	The cat had thought of eating the butterfly	Because the butterfly was yellow in colour	6
		So that it can eat	Because there was a beautiful butterfly	
		How to eat it	Because the butterfly was very beautiful	
		Because cat was still hungry		

		The cat wants to eat		
		The cat wanted to catch the butterfly to eat		
		Because it will be very delicious to taste and eat		
		Became greedy by seeing the fish		
		Because there was a fish		
		Because the butterfly was sitting on the bush		
		So that the cat eat		
5b: Why do you think the boy is feeling good/happy? Ans: Because he has got the ball back	27	So that he can play again with that ball		1
		He used his fishing rod aptly.		
6a: How does the boy feel? Ans: Bad/angry/ mad	26 = angry: 20; bad: 5; shows displeasure:1		sad: 4	0
6b: Why do you think the boy feels bad/ angry/mad, etc.? Ans: Because the cat ate/ is eating/ took/ has taken his fish	22	Because he struggled hard	It is not good manners to eat others' things	
		Because he struggled to get the fish and wanted to eat		
		Because he struggled to get, but the cat enjoyed		

		The boy was catching the fish for a long time but could not eat it		
		Because he worked hard to catch fish and that cat came		
		Because he struggled hard to catch the fish		

To answer the pragmatic questions leaners needed to access and link world knowledge to the text or the episodes to complete the situation model. These questions were, therefore, more difficult to answer than the propositional inference ones. A question like 2b shows variation in inference generation at 47 per cent—there are responses listed in partial situational model where learners have related the causality of the cat's reaction to eating the fish, which is faulty. The fish is irrelevant in this episode. The cat is feeling bad because the butterfly escaped and the cat fell on the bush. Mentioning the fish or eating the butterfly basically shows that the child was unable to focus on the previous episode. The child already knew how the story ended so the child could not inhibit herself from making reference to a later episode of the story and focus only on the episode that the question was about. Some of the responses listed under world knowledge show that these learners have been able to only access how the butterfly looks like and fail to generate episode specific inference. Again in a question like 6b, learner variation in inference happens because they focus on the boy's failed attempts and not link the boy's reaction to the act of the cat (stealing and eating the fish). These are instances of building partial text model.

Feedback

Learner responses on the propositional and pragmatic inferences that show partial and world knowledge inferences can be taken up as partially correct answers. They can also serve as material for constructive feedback. Teachers can guide learners to generate more accurate inferences by helping learners 'notice' what is

missing or what needs to be selected. Thus, the proportions of learner responses on which feedback may be provided are listed in Table 4.

Table 4: Reassessment of Performance

	Percentage of correct responses	Percentage of responses that require feedback	Percentage of correct responses
Propositional Inference			
1	83%	11%	6%
2a	70%	27%	3%
3	94%	-	6%
4	67%	33%	-
5a	97%	-	3%
Pragmatic Inference			
2b	34%	47%	20%
5b	90%	7%	3%
6a	87%	13%	-
6b	73%	24%	3%

In Table 4 it can be seen that the learners' use of alternative forms of propositional inferences has led to a decrease in total percentage of incorrect responses for three questions and they are as follows— Qs4 (34% to 0%), Qs2a (from 30% to 3%), and Qs1 (from 17% to 6%). The error rate for pragmatic inference has decreased for the questions 2b (67% to 20%), 6a (13% to 0%) 6b (27% to 3%) and 5b (10% to 3%).

Discussion

A reassessment of learners' inference generation abilities was brought about on the basis of our analysis of the learners' ability to generate alternate responses as interim solutions before forming accurate inferences. Let us look at why some of the inferences are incomplete.

One reason for providing deviant answers is the inability to form a complete and appropriate episode wise situation model to answer questions like 1, 2a, 4 (propositional inference) and 2b and 6b (pragmatic inference). The learners have been able to create a partial representation of the model and therefore their answers are not fully accurate. If given feedback to notice the situation details

and complete their responses bridging world knowledge with the specific events in the narrative or providing additional information to complete the justification of the response, learners are likely to improve. We also observe that some of the deviant answers refer to later episodes in the story instead of the one the question is about (e.g., when the child responds to question 2b with reference to the cat noticing the fish which is a later part of the story). Given that the children have re-told the narrative, they are aware of the full story and seem to focus on the more recent developments rather than focus on the event asked about. This is also part of feedback that could be offered to the children that would enhance their ability to focus on details of the story as it unfolded and inhibit the more salient reference to the final episode where the story concludes.

A second reason for giving responses that do not match with the key are because the learners use world knowledge that is more generic in nature (e.g., cat liking the butterfly, or finding the fish tasty, etc.). Therefore answers to questions 1, 2a and 4 (pragmatic inference) and 2b and 6b (pragmatic inference) are based on learners' world knowledge. In such cases, learners need feedback to pay attention to combining the background or world knowledge they access with the situation specific or episode specific features to formulate answers that are coherent.

Figure 5 presents the answers that need feedback due to partial responses or world knowledge in contrast to the correct answers. The answers that need feedback are now separated from the incorrect responses. So we find that the percentage of errors has come down because 17 per cent of the responses can be improved based on the teacher feedback.

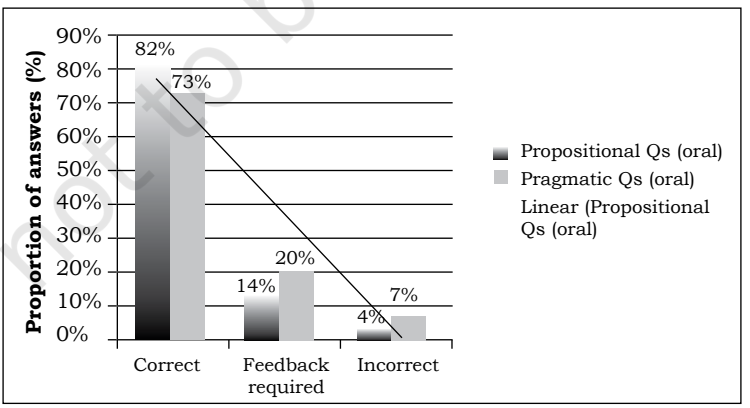

Figure 5: Differences of inference generation abilities

By comparing Figure 5 with Figures 3 and 4, it can now be concluded that a consideration of learners' partial situation model and/or world knowledge through which they access text comprehension clearly shows that their inference generation abilities are at work but at three levels:

correct < feedback required < incorrect

A study by Shepard-Carey (2019) shows a few methods in which emergent bilingual learner responses can be developed by way of scaffolding to arrive at appropriate text based and world knowledge based inferences. She uses pre-reading activities in small groups to help the learners predict the content to be read, focus on lexis to express inferences and activate background knowledge required to draw inferences for expository texts. Similar scaffolding methods can be adopted to give structured feedback once teachers are able to identify trends in responses. Without training to handle inference-based responses, teachers may look at divergent answers as instances of inference failure. The reason for them to do this would be that the variation in responses does not match with the key. Thus, variations in learner responses as we have presented through Tables 2 – 4, can be materials on which teachers can build constructive feedback and create learning opportunities.

In sum, the qualitative analysis of individual variations in this paper shows that of the total number of inappropriate responses to comprehension questions, many refer to experiential or world-knowledge inferences but fail to link them to the specific story-based information. This indicates difficulties with inference generation and the ability to select only the relevant parts of the response is what gives rise to individual variations in inference generation in young learners.

Conclusion: Pedagogical Implications of the Assessing Inference Generation Abilities

In this paper variations in responses of the learners have been analysed to demonstrate how comprehension development and feedback can be built into instructional practices of teachers. The kind of analysis of individual differences that this paper has reported is a novel approach to assess and develop comprehension skills and has not been attempted before. A narrative based comprehension task from the MAIN manual was used to find that the learners were able to tap into the mental states of the characters and were able to bring out inferences within the text (e.g., propositional inference),

new inferences that build up new events in the story (e.g., pragmatic inference) (Chikalanga 1992), interpret intentionality and make inferences about aspects of stories that support the creation of a situation model of comprehension (Perfetti and Stafura 2015).

Taken together, performance on both comprehension tasks seems to be good at an average of 77 per cent. This high performance on comprehension can be attributed to the high degree of overlap between the learners' home languages with the MoI (Hindi) at 90 per cent (refer to Table 1) (Tsimpli et al. 2020). It was also observed that 70 per cent of the learners who were successful in sentence and paragraph decoding ability also showed good comprehension skills in the print and oral.

The fact that oral propositional inference is the most successful type of inferential comprehension gives evidence that the process of comprehension is guided by the levels of difficulty of inferences (Cain et al. 2005) and the modality of comprehension. A reason for higher success in oral narrative comprehension could be because oral inference skills are found to develop earlier than print comprehension owing to one's variety of life experiences (Kendeau 2015); also the text had visual support, whereas in the print (ASER) test-t comprehension was more abstract as there was no visual support and it involved decoding skills. Thus, print modality added more challenge than oral modality.

A reason why oral propositional inference did not correlate highly with print propositional inference could be because absence of sentence level decoding skills in 30 per cent of the learners, which was an impediment to conscious attention towards reading comprehension. So, though this ability is found to be present in the learners in the oral modality and they are able to transfer it to the print modality, albeit not at a very high rate, given the fact that the print story was at a difficulty level appropriate for grade two learners.

Based on the findings in the qualitative analysis of individual differences in answering causal questions we can conclude that in comprehending a text better, a learner may be encouraged to create more fine-tuned situation models along with accessing their world knowledge. Teachers need to help learners focus on specific details of episodes and inhibit reference to final outcomes of the story. Teachers also need to gain awareness to treat learner differences in a constructive manner. These would create conditions for a gradual development of text-based as well as world knowledge-

based inferential skills in primary level of education in the Indian context (Swinney and Osterhout, 1990; Aukerman et al. 2017; Shepard-Carey 2019).

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

Alcott, B. and P. Rose. 2017. Learning in India's primary schools: How do disparities widen across the grades? *International Journal of Educational Development*. 56. pp. 42–51.

Astington, J.W. and J.M. Jenkins. 1995. Theory of mind development and social understanding. *Cognition and Emotion*. 9. pp. 151–165.

Aukerman M, Schuldt L. and L. Aiello. et al. 2017. What meaning-making means among us: The intercomprehending of emergent bilinguals in small-group text discussions. *Harvard Educational Review.* 87(4). pp. 482–511.

Balasubramanian, A., M. Vogelzang., I. Tsimpli., M. Panda., S. Alladi., A. Reddy., L. Mukhopadhyay., J. Treffers-Daller. and Marinis, T. 2020. Reading Skills and Reading Comprehension in a Cohort of Multilingual Children from Disadvantaged Contexts in India. The First South Asian Forum on the Acquisition and Processing of Languages presentation: Germany.

Banerji, R. 2017. Language and Learning: The Challenges of Primary Education in India. In Coleman, H. (Ed). *Multilingualisms and Development: Selected Proceedings of the 11th Language and Develpoment Conference*. New Delhi. New Delhi: British Council.

Brinkmann, S. 2018. Teachers' Beliefs and Educational Reform in India: from 'learner-centred' to 'learning-centred' education. *Comparative Education*. 55(1). pp. 9–29. https://doi.org/10.1080/03050068.2018.1541661

Bruner, J.S. 1986. Actual Minds, Possible Worlds. Cambridge MA: Harvard University Press.

Chand, V. 2011. Elite Positionings Towards Hindi: Language Policies, Political Stances and Language Competence in India. *Journal of Sociolinguistics*. 15(1). pp. 6–35.

Chikalanga, I. 1992. A Suggested Taxonomy of Inferences for the Reading Teacher. *Reading in a Foreign Language*. 8. p. 697.

Clarke, P. 2001. *Teaching & Learning: The Culture of Pedagogy*. Sage Publications. New Delhi.

———. 2003. Culture and Classroom Reform: The Case of the District Primary Education Project, India. *Comparative Education*. 39(1). pp. 27–44.

Coté, N. and S. R. Goldman. 1999. Building Representations of Informational Text: Evidence from Children's Think-aloud Protocols. In H. van Oostendorp and S.R. Goldman (Eds.). *The Construction of Mental Representations During Reading.* pp. 169–193. Lawrence Erlbaum Associates Publishers.

Enver, J. 2018. Politics and Policy in Global Primary English. Oxford University Press. Oxford.

Gagarina, N., D. Klop., S. Kunnari., K. Tantele., T. Välimaa., I. Balciuniene and J. Walters. 2012. Multilingual Assessment Instrument for Narratives (MAIN). ZAS *papers in linguistics*. 56. pp. 1–140.

Graesser, A.C. and J. Roger. 2009. A Theory of Inference Generation During Text Comprehension. *Discourse process*. 16. pp. 145–160.

Jhingran, D. 2005. Language Disadvantage: *The Learning Challenge in Primary Education*. APH Publishing Corporation. New Delhi.

———. 2009. Hundreds of Home Languages in the Country and Many in Most Classrooms: Coping with Diversity in Primary Education in India. Social Justice through Multilingual Education. pp. 263–282.

Kendeou, P. 2015. A General Inference Skill. In: O'Brien EJ, Cook AE and Lorch RF (eds) Inferences During Reading. Cambridge, MA: Cambridge University Press. pp. 160–181.

Government of India. 1993. Learning without burden. Report of the National Advisory Committee. Ministry of Human Resource Development: Department of Education. New Delhi.

Lightfoot, A., A. Balasubramanian., I. Tsimpli., L. Mukhopadhyay and J. Treffers-Daller. (submitted). Measuring the Multilingual Reality: Lessons from Classrooms in Delhi and Hyderabad. *International Journal of Bilingual Education and Bilingualism.*

Long, D.L. and J.M. Golding. 1993. Superordinate Goal Inferences: Are they Automatically Generated during Comprehension? Discourse Processes. 16(1–2). pp. 55–73.

Matthew, R., A. Lightfoot., L. Mukhopadhyay and I. Tsimpli. (submitted). Multilingual Practices in Indian Classrooms: Pedagogical and Policy Implications for English Medium Contexts. In L. Adinolfi & U. Bhattacharya (eds.) Global Lessons from Multilingual Education in South Asia: At the Intersection of Policy and Practice. Routledge series in Language and Content Integrated Teaching & Plurilingual Education.

Meganathan, R. 2011. Language Policy in Education and the Role of English in India: From Library Language to Language of Empowerment. In Coleman, H (Ed). Dreams and Realities: Developing Countries and the English Language. London: British Council.

Miliand Winch, C. 2019. Teaching through Textbooks: Teachers as Practitioners of a Discipline? Theory and Research in Education. 17(2). pp. 181–201.

Mohanty, A. 2018. The Multilingual Reality: Living with Languages. Multilingual Matters.

MUKHOPADHYAY, L. 2020. Translanguaging in Primary Level ESL Classroom in India: An Exploratory Study. *International Journal of English Language Teaching*. 7(2).

NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING. 2006. Position Paper of the National Focus Group on Teaching of English. New Delhi.

NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING. 2005. NCF. New Delhi.

PERFETTI, C.A. AND J.Z. STAFURA. 2015. Comprehending Implicit Meanings in Text without Making Inferences.

PRATHAM. 2014. Annual Status of Education Report (Rural) 2013. New Delhi: Pratham. http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202014/fullaser2014 ainreport_1.pdf

———. 2017. Annual Status of Education Report (Rural) 2016. New Delhi: Pratham.

SHEPARD-CAREY, L. 2019. The Inference-Making of Elementary Emergent Multilinguals: Access and Opportunities for Learning. *Journal of Early Childhood Literacy*. 0(0). pp. 1–39.

SMITH, F., F. HARDMAN. AND J. TOOLEY. 2005. Classroom Interaction in Private Schools Serving Low-income Families in Hyderabad, India. *International Education Journal*. 6(5). pp. 607–618.

SWINNEY, D.A. AND L. OSTERHOUT. 1990. Inference Generation during Auditory Language Comprehension. In Psychology of Learning and Motivation. Vol. 25. pp. 17–33. Academic Press.

TAYLOR, L. 2013. Testing Reading through Summary: Investigating Summary Completions Tasks for Assessing Reading Comprehension Ability. Studies in Language Testing, 39. Cambridge: UCLES/Cambridge University Press.

TSIMPLI, I., VOGELZANG, M., BALASUBRAMANIAN, A., MARINIS, T., ALLADI, S., REDDY, A. AND PANDA, M. 2020. Linguistic Diversity, Multilingualism and Cognitive Skills: A Study of Disadvantaged Children in India. Languages. 5(1).

TSIMPLI, I., L. MUKHOPADHYAY., J. TREFFERS-DALLER., S. ALLADI., T. MARINIS., M. PANDA., A. BALASUBRAMANIAN. AND SINHA, P. 2019. Multilingualism and Multiliteracy in Primary Education in India: A Discussion of Some Methodological Challenges of an Interdisciplinary Research Project. Research in Comparative and International Education. 14(1). pp. 54–76. doi:10.1177/1745499919828908

WESTBY, C.E. 2005. Assessing and Remediating Text Comprehension Problems. In Catts, H.W. and Kamhi, A.G. (eds) (2nd). Language and Reading Disabilities. Boston, MA; Pearson Education, Inc. pp. 157–232.

UNESCO. 2015. Education for All Global Monitoring Report: Achievements and Challenges.

Appendix 1

The Cat Story (MAIN Manual, pp. 111–117)

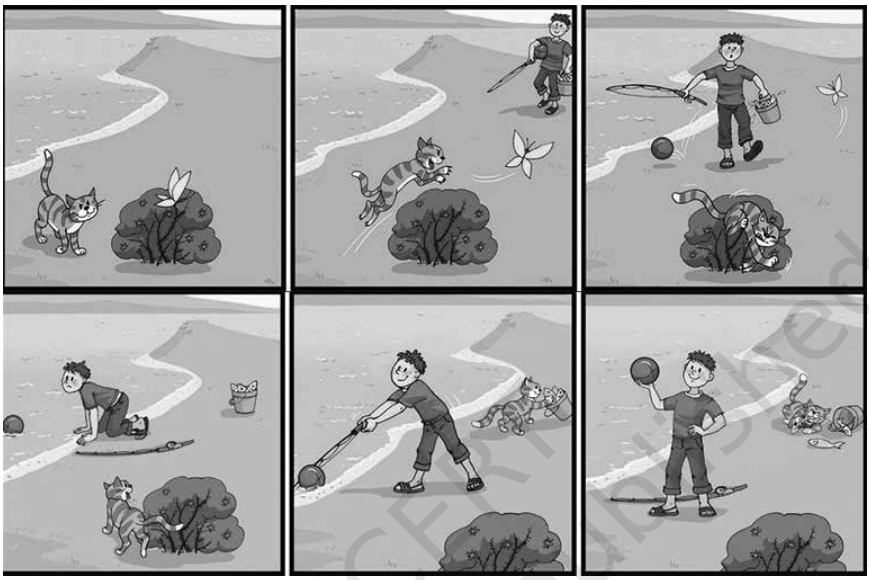

Table 5: Types of Inference Generation through Narrative Comprehension Questions

Episodes	NRT—Story Structure Elements	Comprehension Questions	Type of Inference
Episode 1	Goal— Cat (e.g., cat wanted to catch/get the butterfly)	Question 1 (Why does the cat jump/ leap forward?)	Propositional (Logical Explanatory)
	IST as Reaction (e.g., cat was angry/ bad)	Question 2a (How does the cat feel?)	Propositional → Pragmatic (Evaluative)
	Attempt (e.g., cat jumped forward/up)	Question 2b (Why do you think the cat is feeling angry/bad?)	Pragmatic (Evaluative)
Episode 2	Goal— Boy (e.g., boy wanted to get his ball back)	Question 3 (Why does the boy hold the fishing rod in the water?)	Propositional (Logical Explanatory)

	IST as reaction—Boy (e.g., boy was glad/ happy)	Question 5a (How does the boy feel?)	Propositional (Evaluative)
	Outcome— Boy (e.g., boy got his ball back)	Question 5b (causal) (Why do you think the boy is feeling good/happy?)	Pragmatic (Logical Explanatory)
Episode 3	Goal— Cat (e.g., cat wanted to eat the fish)	Question 4 (Why does the cat grab the fish?)	Propositional (Elaborative Explanatory - Causative)
Additional*	Extrapolative	Question 6 (perspective taking + causal) Imagine that the boy sees the cat 6a How does the boy feel? 6b Why do think the boy feels bad/angry/mad etc.?	Pragmatic (Evaluative) Pragmatic (Evaluative)