

The Construct Validity of a Reading Test Based on Narrative Texts

Godefroid B. Katalayi
University of the Western Cape, South Africa

Sivakumar Sivasubramaniam
University of the Western Cape, South Africa

Abstract—The study examines the construct validity of a reading test based on narrative texts. It aims to evaluate the extent to which the test tasks target the narrative macrostructure and microstructure elements in both the text passage and the test questions as well as the extent to which the order of test questions is sensitive to the narrative story line. The study uses 40 test items randomly selected from the DR Congo English state examination (ESE). The results indicate that the construct validity of the ESE based on narrative texts is threatened by the inclusion of a limited number of episodes in the narratives, the absence of some elements of episodic structure, the use of truncated narratives that fail to include coherence elements, the failure to examine all the critical aspects of the narrative construct, the inclusion of some test items that require examinees' abilities that are irrelevant to narrative story grammar, and the ordering of information requested by the test items that does not relate to the way information is ordered in the narratives.

Index Terms—construct validity, reading test, narrative texts, story grammar, narrative macrostructure, narrative microstructure

I. PREAMBLE

Text comprehension is an interaction between a reader and a text where the reader has to relate ideas from the text to his/her background knowledge in order to construct text meaning. A part of this interaction process requires that the reader understands text structure; that is, how the main ideas and details of the text are organized (Amer, 1992). The narrative text is one of the most familiar and most studied text structures. Narratives are stories that generally tell what happened. In language learning environment, these stories can be first-hand narratives about daily events, they can be retellings of previous experiences, or retold favourite tales, or again oral fantasy narratives (Paris, & Paris, 2003).

This paper investigates the construct validity of a reading test based on narrative texts. It draws on the theoretical underpinning from the story grammar model (Stein & Glenn, 1979) that proposes rules for describing the narrative features and that suggests that all narratives have a setting and an episodic structure meant to capture the general organizational pattern of a narrative (Trabasso & Nickels, 1992). The story grammar model premises that reader's knowledge of the narrative structural organization facilitates his/her comprehension of the narrative as this organization describes the story parts, the arrangement of these parts and how they are related (Mandler, 1980). Story grammar theorists argue that, since story grammar is an approach that attempts to construct consistent narrative features necessary for identifying the narrative parts and specifying how these parts are related (Mandler, 1980; Dymock, 2007), a story grammar facilitates narrative comprehension by describing the higher order structure that is used to encode, represent, and retrieve information from narratives (Ohtsuka & Brewer, 1992; Graesser, Singer, & Trabasso, 1994).

Although there are many different story grammars relating to different stories (see Mandler & Johnson, 1977; Rumelhart, 1975; Stein & Glenn, 1979; for example), it is generally agreed that a narrative begins with a setting and then progresses through the episodes where each episode consists of a goal statement for the main character, an attempt by the main character to reach the goal, an outcome of the attempt, and internal states that initiate the goal and express reactions (Gagarina, *et al.*, 2012).

The assessment of narrative comprehension

The assessment of narrative comprehension provides an opportunity to examine a wide range of students' abilities; including story structure, discourse features, and vocabulary and syntax (Gagarina, *et al.*, 2012). Research on the assessment of narrative comprehension recognizes the characteristic feature of narratives that consists of information at two levels: microstructure and macrostructure levels (Gagarina, *et al.*, 2012; Heilmann, Miller, & Nockerts, 2010). Microstructure analysis focuses on the linguistic structures used in the construction of coherent and cohesive discourse, while macrostructure analysis focuses on higher-order hierarchical organization that includes the story grammar components (Heilmann, Miller, & Nockerts, 2010). Studies that have explored the macrostructure and microstructure abilities underlying narrative comprehension (Gagarina, *et al.*, 2012; Liles, Duffy, Merritt, & Purcell, 1995) have hypothesized that, since narratives have a specific structural organization that is reflected in the narrative story grammar, students should be mindful of narrative schema in order to successfully attempt test tasks. The narrative schema, which

is the cognitive structure a reader creates when comprehending a story, is "an idealized internal representation of the parts of a typical story and the relationships among those parts" (Mandler & Johnson, 1977, p. 111). Therefore, in order to process a narrative and complete test tasks, students should be able to understand the main events of stories; including the setting, characters, and episodic structure. Furthermore, they should be able to understand characters' internal responses; including their mental processes, states, and experiences (Gagarina, *et al.*, 2012).

Studies that have investigated the narrative macrostructure elements in reading tests have examined students' ability to identify and/or understand the setting, the episodes and the internal state terms (Gagarina, *et al.*, 2012; Heilmann, Miller, & Nockerts, 2010). The setting provides information about the story time and place. The episodes provide the narrative story line, and each episode is marked by an initiating event (the problem or event that sets the story in motion), an attempt (goal-directed action characters take in response to the initiating event), and a consequence (the result of the characters' action). The internal state terms provide an awareness of physical and emotional cause-effect relationships necessary for recognizing characters' goals, the reasons for these goals and reactions following attempts to reach the goals (Hedberg & Westby, 1993; Gagarina, *et al.*, 2012). Briefly, in assessing narrative macrostructure elements, the students should be able to elicit the theory of mind/inferencing by their abilities to infer meaning about the story as a whole (Gagarina, *et al.*, 2012).

On the other hand, studies that have investigated the narrative microstructure have examined students' ability to identify and/or understand the linguistic structures that provide story actions, sequence and details; including adjectives, adverbs, and prepositions; as well as the linguistic markers of coherence; including transitions and connectives (Gagarina *et al.*, 2012; Heilmann, Miller, & Nockerts, 2010). Furthermore, these studies have investigated students' abilities to understand the narrative lexis and morpho-syntax (Gagarina, *et al.*, 2012). Such an assessment scope suggests that, in adopting the narrative macrostructure-microstructure framework, central to narrative comprehension is the ability to understand the narrative structure as well as the narrative language (Trabasso & Nickels, 1992).

However, studies that have adopted Kintsch and van Dijk (1978)'s model of text comprehension (Sheehan & Ginther, 2001; Shin, 2002; Kintsch, 1994; Gorin, 2005) have highlighted the sequencing of information in both the text and test questions as crucial to narrative assessment. Sheehan and Ginther (2001) reported that the location of relevant information within a narrative was associated with the difficulty of MC questions on the TOEFL- 2000 reading test. Shin (2002) found that students poorly performed on most test items that could not follow the story line of the narrative as such questions did not call for operations examinees were likely to use to construct narrative macrostructure. On their part, Kintsch (1994) and Gorin (2005) reported that changing the order in which the information in the narrative is presented impacted on the construction of coherent narrative representation, and this was found to increase the difficulty of test questions pertaining to the newly ordered information.

Although research insights suggest that the assessment of narrative comprehension should focus on both the microstructure and macrostructure abilities (Gagarina, *et al.*, 2012;), and that it should be sensitive to the ordering of information in both the text and the test tasks (Sheehan & Ginther, 2001; Kintsch, 1994; Gorin, 2005) so as to represent the underlying narrative discourse competence, there is an abject paucity of literature that incorporates these three aspects within a single validation framework. This study therefore aims to fill this gap. It is grounded in the assumption that narrative assessment validation should call for a widely drawn-up, integrative framework, which would include the macrostructure and microstructure elements in both the text and the test tasks and which is sensitive to the narrative story line.

The construct validity of reading tests based on narrative texts

This paper is underpinned by Messick's (1989, p. 41) modern conceptualization of validity as a unitary concept that consists of "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationale support the adequacy and appropriateness of inferences and actions based on test scores". It considers construct validity as an overarching concept that relates to the extent to which the inferences and decisions that are made on the basis of test scores are meaningful and appropriate (Kane, 2002, 2006). Therefore, in order to be meaningful and appropriate, scores on a reading test based on narrative texts must provide an indication of students' abilities to comprehend the narrative and reconstruct its story line. It is hoped that by using a wide-scoped, integrative framework that examines the macrostructure and microstructure elements of the narratives in both the text and the test questions, and that is sensitive to the narrative story line, we can better evaluate the extent to which students' scores provide information about their narrative comprehension abilities.

II. CONTEXT, AIM AND SIGNIFICANCE

The paper examines the construct validity of the DR Congo English state examination (ESE), a multiple-choice reading test administered to final year secondary school students as part of their certification. The ESE uses the conventional multiple-choice format where the student is required to choose one answer from five answers suggested.

Since the focus is on the reading texts based on narratives, the aim of this study is to evaluate the construct validity of the ESE within the epistemological stance that acknowledges the narrative story grammar as essential to narrative comprehension. More specifically, the study aims to evaluate the extent to which the ESE tasks target the narrative macrostructure and microstructure elements in both the text passage and the test questions as well as the extent to which the order of test questions is sensitive to the narrative story line.

By investigating the construct validity of the ESE based on narrative texts within this broader scope, the insights generated by this study can be gainfully deployed in the educational practice of reading assessment, and they can be used by reading assessment researchers to inform their validation studies.

The study therefore intends to answer these three research questions:

- To what extent are the ESE narratives structured around the critical elements of the narrative story grammar?
- To what extent do the ESE questions target the critical elements of the narrative story grammar?
- To what extent does the ordering of ESE test questions tap into the narrative story line?

III. METHOD

This study examines the content of test papers as an informed basis in order to evaluate the extent to which the ESE tasks target the narrative macrostructure and microstructure elements in both the text passage and the test questions as well as the extent to which the order of test questions is sensitive to the narrative story line.

Instrument

We randomly selected four test papers from the 2011 and 2012 ESE editions. Each test paper contained a narrative text whose length varied between 461-477 words; and it was followed by ten questions based on the comprehension of the narrative; making a total of 40 test items.

Framework for analysis

Gagarina, *et al.* (2012)'s framework of narrative analysis, grounded on Westby (1991)'s theoretical approach to narrative production and comprehension, serves as basis for evaluating the extent to which the narratives in the four test papers include the macrostructure and microstructure elements; and the extent to which the 40 test items target information relating to narrative macrostructure and microstructure elements.

In order to investigate the macrostructure elements in the four narratives, each narrative is analysed in terms of its story structure (time and place setting, episodic structure) and internal state terms used. Each episode is evaluated in terms of the initiating event (or story problem), the characters' goals, their attempts to reach the goal, the outcome of their attempts, and the reaction from the episode outcome. Furthermore, in order to investigate the internal states in the five narratives, the paper draws on research into the theory of mind that explains the mental and internal states attributed to the self and others (Lorusso, Galli, Libera, & Gagliardi, 2007). Therefore, the paper adopts Gagarina *et al.* (2012)'s taxonomy of internal state terms that include (a) emotive terms, (b) perceptual terms, (c) physiological terms, and (d) metacognitive terms.

In order to investigate the microstructure elements included in the five narratives, the paper builds on insights from Gagarina, *et al.* (2012)' framework of narrative production and comprehension, Kintsch's model of text comprehension (Kintsch, & van Dijk, 1978) and it examines (a) the density of the five narratives; including (i) the number of words included in each narrative, (ii) the average number of words in a sentence, (iii) the average number of sentences in paragraphs, (iv) the number of paragraphs and (v) the number of sentences included in the narrative); (b) the vocabulary load; (c) the syntactic complexity (the number of simple/compound/complex sentences), (d) the use of adverbs, adjectives and prepositions that help visualise the narrative story scene, and (e) the use of discourse coherence markers; including linguistic markers and connectives.

In order to investigate the extent to which the 40 test items target the macrostructure elements of the narrative story line, the four test papers are examined in terms of the number and percentage of test items that target information relating to (a) story structure; including the (i) setting (time and space), (ii) characters (their identification, emotions, personality and relationships), and (iii) episodes (Initiating event /story problem, goal, attempt/s, outcome and reaction), and (b) the internal state terms that assess examinees' ability to interpret physical and emotional cause-effect relationships that can reflect their ability to infer meaning about the story (Gagarina, *et al.*, 2012; Hedberg & Westby, 1993).

In order to investigate the extent to which the extent to which the 40 test items target the microstructure elements of the narrative, the four test papers are examined in terms of the number and percentage of test items that target information relating to (a) the story scene descriptive elements; including (i) adjectives, (ii) adverbs and (iii) prepositions, (b) the understanding of the lexis, and (c) the use of coherence markers; including (i) transitions and (ii) connectives.

Finally, in order to investigate the extent to which the order of test items relates to the order of information in the narratives, the paper builds on Shin (2002)'s and Gorin (2005)'s methodological suggestions that consist of mapping the order of information in the text with the order of information in the test questions.

IV. ANALYSIS

Structure of the narratives

Test paper #1 [T1]

The narrative featured in this test paper starts with the opening sentence that sets up the story place (*At the rural market of Burungangu*) and time (*In 1997*), and introduces the major character (*Yasekuru*). Then the story progresses around two episodes. Episode 1 introduces the narrative problem (*Yet, she [Asha] stopped gaining weight and*

contracted one infection to another) that breaks the family stability and results in Asha's death (*At the age of three, Asha died from AIDS*). The initiating event in Episode 2 is about Kirongozi's sickness (*Kirongozi who was a taxi driver from Worobe to Elaka also began to get sick*), his death as the episode outcome (*One day he collapsed and was taken to Wango Medical Centre. [...] Yasekuru's husband of eight years died of Aids-related complications*) and the reaction of his wife Yasekuru (*"I have got HIV. I will not get married. I will not have any more children"*).

This narrative uses some words and expressions that indicate characters' states of mind. Emotive terms are used to signal a polarity between positive and negative emotions characterized in the stability of the family as a result of the birth of Asha (*Yasekuru and Kirongozi were overjoyed*) and the family disintegration as a result of Asha and Kirongozi death (*Her [Yasekuru] cheerful smile hides a tragic story*). However, the narrative does not include any perceptual terms, physiological terms and metacognitive terms.

With regards to density, this narrative is a relatively short text with 461 words, five short paragraphs with an average of three sentences per paragraph. The narrative is structured around a total of fifteen sentences, of which eleven (11) are simple sentences, three (3) are compound sentences and only (1) one complex sentence. Many words and expressions used may be understood through reader's knowledge of vocabulary or context clues (For example: *AIDS, HIV prevention, cheerful smile, tragic story, ... were overjoyed at the birth of their daughter, to be in good health, she stopped gaining weight*, etc.). The narrative uses some time adverbials to indicate time sequence (*In 1997, Yasekuru...; At the age of three, Asha died...; A few years later, Kirongozi...; Yasekuru now lives...*) and adjectives to provide a description of characters' physical appearances, emotions and personality (*She appears happy and healthy*). However, only one transitional construction is used (*A few years later, [...]*) to link paragraph #2 and paragraph #3, while the use of connectives is very limited (*Yet, she stopped gaining weight and contracted one infection to another*).

Test paper #2 [T2]

The narrative featured in this test paper starts with the opening sentence that sets up the story time (*That Wednesday morning*) and place (*along MABANGA ROAD over the National Commercial School's fence*). Then the story progresses around two episodes. Episode 1 introduces the major characters (students, police) and minor characters (passengers, Apolosa) and presents the narrative problem that consists of trouble between the students and the police (*facing the police, flocks of furious young men were feverishly throwing to their opponents any kind of harmful objects*), the outcome of this trouble on the traffic (*Some [taxi drivers] stopped abruptly. Passengers were puzzled and wondered whether they could continue their trip*) as well as the reaction by the police to restore order (*men in dark blue uniform were launching tears' gaz from their guns in order to disperse those crazy students*). In Episode 2, the narrator uses a flashback technique where he introduces the narrative main problem that relates to the students' leader who is beaten by a group of students from a neighbouring school (*the [...] students' leader was deadly beaten in one ladies' residential flat when he went to pay a visit to his girl-friend two days ago*); the goal of his mates to avenge him (*his mates might have promised to invade that school*); the different attempts made by the students to reach this goal (*[they] were feverishly throwing at their opponents any kind of harmful objects: pieces of stones, bricks, iron sticks*) and the reaction that ensued (*The Chief District of police then decided on the protection of the school from the revenge of the [...] students*).

Narrative 2 uses many words and expressions that describe characters' states of mind. Emotive terms are used to provide an indication of a troubled situation (*Passengers were puzzled / flocks of furious / those crazy students / the anger of their close neighbours*), and the consequences that ensued (*the lorry driver was mortally wounded / Apolosa was seriously injured / students' leader was deadly beaten*). Perceptual terms are used to visualise the intensity of the incident (*irritating black smoke / flocks of furious*), and the objects and people involved in the incident (*any kind of harmful objects / men in dark blue uniform*). Metacognitive terms are used to provide an indication of characters' state of mind (*Passengers wondered / students hardly stand the police presence / The Chief District of police then decided the protection of the school*).

Narrative 2 is a relatively short text that comprises 463 words used in three paragraphs with an average of five sentences per paragraph. The narrative is structured around a total of sixteen sentences, of which twelve (12) are simple sentences, two (2) are compound sentences and two (2) are complex sentences. Most words and phrases can be classified as part of students' active vocabulary (For example: *drivers, taxi buses, trip, police, crazy students, sexual abuses, residential flats, phone cells, golden jewelry, girl-friend*, etc.). Furthermore, the narrative includes adverbs, adjectives and prepositions that are associated to the description of story scene and event. More specifically, the descriptive adjectives are used to describe characters (*men in dark blue uniforms / flocks of furious young men / crazy students*); and the adverbials of manner are used to indicate the intensity of the story event and outcome (*A lorry [...] was badly damaged. Its driver was mortally wounded. Apolosa, [...], was seriously injured*). However, no linguistic markers of coherence are used; instead, short sentences are juxtaposed as in the following four choppy sentences (*A lorry bumped a tree and was badly damaged. Its driver was mortally wounded. APOLOSA, a well-known policeman, was seriously injured. One student died*). Furthermore, since there are few compound and complex sentences, the use of connectives is very limited.

Test paper #3 [T3]

The narrative featured in this test paper starts with the opening sentence that sets up the story time (*On April 18th at 9.15 hrs.*) and place (*on the corner of Mbala Avenue and Kitoko road*). Then the story progresses around a single

episode with the anonymous *I* narrator as a major character and three minor characters (*Mujinga, Mbela, Mafuta*). The initiating event is a car crash (*Mr. Mbela had crashed into the side door of the white Mazda belonging to Ms. Mujinga*). This crash results in an argument between the two women (*the two drivers had got out of their vehicles and were arguing noisily*), each woman attempting to accuse the other of careless driving (*Mr. Mbela had rudely accused Ms Mujinga of being careless and not signalling. Ms. Mujinga said she had signalled but she was upset*). However, since no damage is observed, the narrator calms down the two women and each of them continues on her way (*Since nobody was injured they continued on their way*).

The narrative uses some internal state terms that portray characters state of mind. The use of emotive terms (*Mr. Mbela had rudely accused Ms Mujinga / [Mr. Mujinga] she was upset*) provide the state of mind of the two women during the argument. The use of some perceptual terms (*I heard a loud crash / I saw two cars / I did not see the crash itself / her view may have been obscured by the row of trees*) provides the picture of the scene and how the crash occurred as well as the intensity of the incident. The use of metacognitive terms indicates how the characters perceived the crash (*I suspected [...] that / she [Ms Mujinga] admitted this was true / She [Mrs Mafuta] seemed to sympathize with Ms Mujinga*).

Narrative 3 is a relatively short text of 467 words used in five (5) short paragraphs with an average of three (3) sentences per paragraph. The narrative is structured around a total of fifteen (15) sentences, of which ten (10) are simple sentences, three (3) are compound sentences and two (2) are complex sentences. The text vocabulary is appropriate to the town context (*loud crash, not signalling, turn left at the traffic lights, driving to work, gave her own account of the accident, walking along the pavement, nobody was injured*) and it can be encoded through context from reader's background knowledge. The narrative includes very limited number of adverbs, adjectives and prepositions that can help visualize the story scene and event (*I was standing on the corner of Mbala Avenue and Kitoko road; I heard a loud crash; the two drivers were arguing noisily*). However, the narrative does not include any linguistic markers of coherence as the sentences are only juxtaposed, and the use of connectives is extremely limited.

Test paper #4 [T4]

The narrative featured in this test paper starts with the opening sentence that sets up the story time (*Once upon a time*) and story place (*a farm*). Then the story progresses around two episodes. Episode 1 introduces the two major characters (*the fox and the geese*) and the minor character (*the farmer*) and sets the story problem (*He [the fox] hadn't eaten for a long time, and he was just dying for a good meal*), the goal the fox wants to achieve (*So, he [the fox] went out of his wood*), and the attempt the fox makes to reach the goal (*He [the fox] walked towards a farm where he knew there were some geese, chicken and ducks*). Then the story presents the outcome when the fox meets the geese in the field (*In a field near the farm the fox came upon a flock of beautiful fat white geese*) and two reactions from the two antagonists, the fox, and the geese (*"I am going to eat you!" said the fox. The geese were terrified and they didn't know what to do*). The initiating event in episode 2 becomes clear when one goose tries to deceive the fox (*One of the geese [...] said "well, Mr Fox, you've got the better of us this time, you are going to eat us, we have no way of escape*) with a goal to ask the fox for a favour (*Can we [geese] please ask one last favour?*). In order to escape, the geese attempt to make much noise while praying so as to draw the attention of the farmer (*they [geese] say in very loud voices, and they opened their mouths and cackled and cackled as loudly as they could*), and this results in the farmer coming out with a gun (*The farmer [...] took his gun, [...] rushed down to the field*) and the fox running into the forest (*the fox saw the farmer and [...] ran away*).

The narrative makes use some emotive terms that indicate the state of mind of the geese when they were caught up by the fox (The geese were *terrified*) and the reaction of the fox when geese were saying their prayers (what a *terrible* noise!); perceptual terms (a flock of *beautiful fat white* geese / when the fox *heard* this / there he *saw* the fox), physiological terms (*The fox was very hungry / the fox ran into the forest as hungry as he had left*) and metacognitive terms (*he [the fox] knew there were some geese / they [the geese] didn't know what to do / he [the fox] was sure the geese couldn't get away / the fox sat and thought / The farmer knew something was wrong*).

Narrative 4 is a relatively short text that includes 477 words organized in three (3) paragraphs with an average of five (5) sentences per paragraph. The narrative is structured around a total of fifteen (15) sentences, of which nine (9) are simple sentences, four (4) are compound sentences and two (2) are complex sentences. The narrative contains many words and expressions that might appear familiar to them (*fox; geese; ducks; came upon; you've got the better of us this time; crackled/cracklings*). Furthermore, the narrative includes some prepositions that indicate the spacial organization of the story event (*The farmer [...] took his gun, rushed out of the farmhouse, rushed down to the field and there he saw the fox*); and adjectives that describe characters' physical appearances (*a flock of beautiful fat white geese*). However, no linguistic markers are used to create coherence between paragraphs, and the use of connectives is very limited as most sentences are simple sentences.

Information targeted by test items

Table 1 presents the information relating to the type of text information targeted by the 40 test items.

TABLE 1

Text information targeted by test items		T1	T2	T3	T4	Total	
A. Narrative macrostructure elements		N=10	N=10	N=10	N=10	N	%
A1. Story structure							
Setting	Time	0	0	1	0	1	2.5
	Space	0	0	0	1	1	2.5
	Sub-Total 1	2					5.0
Characters	Characters' identification	1	0	1	0	2	5.0
	Characters' emotions	1	1	0	0	2	5.0
	Characters' relationships	0	0	0	1	1	2.5
	Sub-Total 2	5					12.5
Story episodes	Initiating event (Story problem)	0	1	0	0	1	2.5
	Goal	0	0	1	0	1	2.5
	Attempt	0	1	0	0	1	2.5
	Outcome	0	1	0	0	1	2.5
	Reaction	0	1	0	1	2	5.0
	Sub-Total 3	6					15.0
Total N items story structure		2	5	3	3	13	32.5
A.2. Internal state terms		3	2	2	2	9	22.5
Total N items macrostructure elements		5	7	5	5	22	55.0
B. Narrative microstructure elements							
Story scene descriptive elements	Adjectives	0	0	1	0	1	2.5
	Adverbs	1	0	0	0	1	2.5
	Prepositions	0	0	0	1	1	2.5
Subtotal: Story scene descriptors		0	0	0	0	3	7.5
Lexis		1	0	1	0	2	5.0
Coherence markers	Linguistic markers	0	0	0	0	0	0.0
	Connectives	0	0	0	0	0	0.0
Subtotal: Coherence markers		0	0	0	0	0	0.0
Total N items microstructure elements		2	0	2	1	5	12.5
Total 1: N items related to story grammar		7	7	7	6	27	67.5
C. Questions not related to story grammar							
Identifying the text title		1	1	0	2	4	10.0
Identifying where a specific idea is located		1	0	0	1	2	5.0
Understanding grammar	Tag questions	1	0	1	0	2	5.0
	Pronouns and reference	0	2	1	0	3	7.5
	Verb use	0	0	1	1	2	5.0
Subtotal N items grammar		7					17.5
Total 2: N items not related to story grammar		3	3	3	4	13	32.5
TOTAL		10	10	10	10	40	100

From the information presented in this table, we wish to make the following observations:

(a) Of the 40 items included in the study sample, only 27 items (67.5%) tapped into the narrative story grammar while 13 items (32.5%) targeted text information that was irrelevant to the narrative story grammar. However, of the 27 items that tapped into the narrative story grammar, there were more items that targeted macrostructure elements (22; 55%) than those that targeted microstructure elements (5; 12.5%).

(b) Of the 22 items that targeted macrostructure elements, there were more items that targeted story structure (13; 32.5%) than those that targeted the internal state terms (9; 22.5%). Furthermore, the items that targeted story structure related to the story setting (2; 5%), characters' identification, emotions and relationships (5; 12.5%), and the episodic structure, including i) the story problem (1; 2.5%), ii) goal (1; 2.5%), iii) attempts (1; 2.5%), and iv) outcome (1; 2.5%).

(c) Of the five test items that targeted microstructure elements, three items (3; 7.5%) related to the story scene descriptive elements including i) adjectives (1; 2.5%), ii) adverbs (1; 2.5%), and iii) prepositions (1; 2.5%); two test items (2; 5%) related to the use of lexis; while no items (0; 0%) targeted coherence elements.

(d) Of the thirteen test items that did not tap into the narrative story grammar, there were four items targeting the identification of the text title (4; 10%), three items relating to the location of a specific idea in the text (2; 5%); and seven items targeting the general aspects of grammar (7; 17.5) including i) the use of tag questions, ii) pronouns and references, and iii) verbs.

Information order in the narratives and the test questions

Table 2 maps the information order in the four narratives and the 40 test questions.

TABLE 2:
INFORMATION ORDER IN THE NARRATIVES AND THE TEST QUESTIONS

	T1	T2	T3	T4
Question 1	P4	All text	S1	P3
Question 2	P2 &3	P2	All text	All text
Question 3	P3, S3	P3, S2	P1, S4	P3, S5
Question 4	P4, S1	P1, S4	P2	P3, S3
Question 5	All text	P3, S3	P3	P2
Question 6	P2, S3	All text	P1	All text
Question 7	All text	All text	P1	P1, S3
Question 8	P2	All text	P4, S1	P2
Question 9	P1	P2	All text	P3
Question 10	All text	P2	All text	All text

T1=Test paper #1; T2=Test paper #2; T3=Test paper #3; T4=Test paper #4;
S1= Sentence #1; P1=Paragraph #1; P2=Paragraph #2, P3=Paragraph #3; P4=Paragraph #4

In light of the information presented in this, the following comments can be made:

In test paper 1, question #1 targeted information located in paragraph #4; question #2 targeted information located in paragraphs #2 and #3; question #3 targeted information located in paragraph #3; question #4 targeted information located in paragraph #4; question #5 targeted information in the whole text; question #6 targeted information located in paragraph #2; question #7 targeted information in the whole text; question #8 targeted information located in paragraph #2; question 9 targeted information located in the first paragraph; and finally question 10 targeted information scattered throughout the text.

In test paper 2, question #1 targeted information in the whole text; question #2 targeted information located in paragraph #2; question #3 targeted information located in paragraph #3; question #4 targeted information located in paragraph #1; question #5 targeted information located in paragraph #3; questions #6, #7 and #8 targeted information located in the whole text; and question 9 and question 10 targeted information located in the second paragraph.

In test paper 3, question #1 targeted information located in the opening sentence; question #2 targeted information in the whole text; question #3 targeted information located in paragraph #1; question #4 targeted information located in paragraph #2; question #5 targeted information in paragraph #3; question #6 targeted information located in paragraph #1; question #7 targeted information in paragraph #1; question #8 targeted information located in paragraph #4; and question 9 and question 10 targeted information spread through the whole text.

In test paper 4, question #1 targeted information located in paragraph #3; question #2 targeted information in the whole text; question #3 targeted information located in paragraph #3; question #4 targeted information located in paragraph #3; question #5 targeted information in paragraph #2; question #6 targeted information in the whole text; question #7 targeted information in paragraph #1; question #8 targeted information located in paragraph #2; question 9 targeted information in paragraph 3; and finally question 10 targeted information spread through the whole text.

V. DISCUSSION

This paper, as we have mentioned earlier, aimed to evaluate the extent to which the ESE tasks target the narrative macrostructure and microstructure elements in both the text passage and the test questions as well as the extent to which the order of test questions is sensitive to the narrative story grammar. The analysis of the content of the four test papers reveals that the structure of the four narratives is generally consistent with the basic features that characterize most narratives; suggesting that students' knowledge of the narrative structural organization can enable them to comprehend the narrative by identifying the different parts, the arrangement of these parts and how they are related (Mandler, 1980). However, although all the four narratives have a clear setting (space and time), with characters and an episodic structure; the analysis of the structure of the test papers reveals that the four narratives have a limited number of episodes with either one or two episodes. Besides, except for the narrative in test paper #4, the other narratives do not have all the five elements of an episode (initiating event, goal, attempts, outcome and reaction). Yet, the absence of some elements of episodic structure is likely to contribute to narrative complexity and pose comprehension problems (Gillam, Fargo, Petersen, & Clark, 2012) as examinees hardly construct a mental representation of the story characters, actions, and events (Graesser, Singer, & Trabasso, 1994).

Furthermore, the results indicate that, although limited in number, all the four narratives included emotion terms. However, except the narrative in test paper #4, the other narratives do not include one or some of other internal state terms; namely perceptual terms, metacognitive terms and physiological terms. This finding suggests that examinees may have difficulty to comprehend the narratives as they hardly understand characters' internal responses, including their mental processes, states, and experiences that are necessary for constructing narrative macrostructure (Gagarina, *et al.*, 2012) as examinees hardly perceive characters' states of mind that are necessary for interpreting characters' intentions and making inferences about aspects of stories (Westby, 2005).

From the structural complexity stance, although all the four narratives are relatively short and they include limited and accessible vocabulary so as to allow for normal processing, the use of story scene descriptive elements (adjectives, adverbs, and prepositions) necessary for visualising the story scene is extremely limited. Therefore, some examinees can find it difficult to comprehend these narratives as they hardly construct some representations of the story characters,

events, and actions (Zwaan, Langston, & Graesser, 1995). Furthermore, all the four narratives include many short simple sentences which are simply juxtaposed; with limited use of connectives and almost no use of transitions. Yet, a good narrative needs to be written in such a way that coherence relations are made explicit by linguistic markers in order to connect different text segments, such as cause-consequence, problem-solution, and chronological order (Graesser *et al.*, 2004). Therefore, the non-use of transitions and other linguistic signaling devices can be detrimental to comprehension as examinees hardly construct coherent representations and have no guiding cues necessary for understanding how ideas in one clause, sentence, and paragraph relate to the ideas in adjacent clauses, sentences, and paragraphs (Graesser *et al.*, 2004).

As to the issue of information targeted by test questions, the results reveal that, although the majority of items fairly target some aspects of the story grammar (story structure, internal state terms); there are some story grammar aspects that are less targeted (story scene descriptors) or simply ignored (linguistic markers of coherence). This prompts the issue of construct underrepresentation as the ESE fails to include some other critical aspects of the narrative construct; suggesting that examinees' scores on the test hardly reflect their comprehension of the narrative and their abilities to reconstruct the narrative story line (Katalayi, 2014). Furthermore, the results indicate that a relatively great number of test items target information irrelevant to narrative story grammar; therefore, this poses the issue of construct irrelevant variance on the ground that examinees' scores appear to be inflated with some variance as some items require abilities that are irrelevant to narrative comprehension (Katalayi, 2014).

Finally, regarding the information order in the text and test items, the results suggest a total mismatch of information order in the text and in the test questions. Yet, research (Sheehan & Ginther, 2001; Kintsch, 1994; Gorin, 2005) suggests that the sequencing of information in both the text and test questions is crucial to the validity of narrative assessment as examinees generally perform poorly on most test items that do not follow the story line of the text (Gorin, 2005), and the mismatch of the order of information in the text and test items can impact on the construction of coherent text representation and increase the difficulty of test questions (Kintsch, 1994; Gorin, 2005).

VI. CONCLUSION

Generally, the four narratives are well structured; they give an indication of some awareness of narrative structural organization on the part of ESE constructors. However, the quality of the narratives is hampered by the inclusion of a limited number of episodes, the absence of some elements of episodic structure, and the use of truncated narratives that fail to include coherence elements. Therefore, in order to enhance its construct validity, the ESE constructors need to ensure that the narratives include an acceptable number of episodes that use most episodic structure elements. Besides, the narratives need to include as many internal state terms as possible as well as many story scene descriptive elements (adjectives, adverbs, and prepositions) necessary for understanding characters' internal responses (Gagarina, *et al.*, 2012) and visualise the story scene and event (Kintsch, 1994).

Regarding the information targeted by the ESE items, we are inclined to conclude that the two issues that weaken the validity of the five tests are the failure to examine all the critical aspects of the narrative construct, and the inclusion of some test items that require examinees' abilities that are irrelevant to narrative story grammar. Therefore, we suggest that the ESE based on narrative texts be designed in such a way to include tasks that tap into the narrative story grammar so as to generate scores that reveal the actual comprehension of narratives.

Finally, regarding the order of information in the text and test questions, we are inclined to conclude that the order of information requested by the test items does not relate to the way information is ordered in the narratives. Therefore, the construct validity of the ESE based on narratives appears to be threatened by the mismatch of information order in both the text and the test items. We therefore suggest that the ESE based on narrative texts be adjusted so as to relate the order of the test questions to the order of the text information so as to reflect the story line used by the writer to construct the story.

In sum, we need to emphasize that validity is about the meanings and interpretations assigned to test scores (Messick, 1989; Kane, 2002). Therefore, by adopting an integrative validation framework that includes the macrostructure and microstructure elements in both the text and the test tasks and that highlights the way information requested in the test questions relates to the ordering of information in the text, students' scores on reading tests based on narratives can provide a sound indication of their relationship to three relevant narrative constructs; including the quality of narratives, the quality of test items in terms of the kind of text information they target, and the ordering of the test items in relation to the narrative story line.

REFERENCES

- [1] Amer, A. (1992). The effect of story grammar instruction on ESL students' comprehension of narrative, *Reading in a Foreign Language*, 8(2), 711-720.
- [2] Dymock, S. (2007). Comprehension strategy instruction: Teaching narrative text structure awareness. *The Reading Teacher*, 61(2), 161-167.
- [3] Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics* 56 (2-35).

- [4] Gillam, S. L., Fargo, J., Petersen, D. B., & Clark, M. (2012). Assessment of structure dependent narrative features in modeled contexts: African American and European American children. *English Linguistics Research*, 1(1), 17-34.
- [5] Gorin, J.S. (2005). Manipulating processing difficulty of reading comprehension questions: The feasibility of verbal item generation. *Journal of Educational Measurement*, 42(4), 351-373.
- [6] Graesser, A.C., McNamara, D.S., Louwerse, M.M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36, 193-202.
- [7] Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395.
- [8] Hedberg, N. L., & Westby, C. E. (1993). Analyzing storytelling skills. Theory to practice. Tucson, AZ: Communication Skill Builders.
- [9] Heilmann, J., Miller, J. F., & Nockerts, A. (2010). Sensitivity of narrative organization measures using narrative retells produced by young school-age children. *Language Testing*, 27, 603-626.
- [10] Kane, M. T. (2002). Current concerns in validity theory. *Journal of Educational Measurement*, 38(4), 319-342.
- [11] Kane, M.T. (2006). Validation. In R.L. Linn (Ed.) *Educational measurement* (4th ed., pp. 17-64), Westport, CT: Praeger.
- [12] Katalayi, G. B. (2012). The DR Congo English state examination: Some fundamental validity issues (Context validity evidence), Unpublished Master Thesis, Department of Language Education, Faculty of Education, University of Western Cape.
- [13] Katalayi, G. B. (2014). Fundamental validity issues of an English as a foreign language test: A process-oriented approach to examining the reading construct as examined by the DR Congo English state examination. Doctoral Dissertation, Department of Language Education, Faculty of Education, University of Western Cape.
- [14] Kintsch, W.M., & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- [15] Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49, 294-303.
- [16] Liles, B. Z., Duffy, R. J., Merritt, D. D., & Purcell, S. L. (1995). Measurement of narrative discourse ability in children with language disorders. *Journal of Speech and Hearing Research*, 38, 415-425.
- [17] Mandler, J.M., & Johnson, N.S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9, 111-151.
- [18] Mandler, J. M. (1980). A code in the mode: The use of a story schema in retrieval. *Discourse Processes*, 1, 14-35.
- [19] Messick, S. (1989). Validity. In R. L. Linn (Ed.). *Educational measurement* (3rd ed., pp 13-103). New York: American Council on Education, Macmillan.
- [20] Ohtsuka, K., & Brewer, W. F. (1992). Discourse organization in the comprehension of temporal order in narrative texts. *Discourse Processes*, 15, 317-336.
- [21] Paris, A. H., & Paris, C. J. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly* 38(1), 36-76.
- [22] Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A.M. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211-236). New York: Academic.
- [23] Sheehan, K.M., & Ginther, A. (2001). What do passage-based MC verbal reasoning items really measure? An analysis of the cognitive skills underlying performance on the current TOEFL reading section. Paper presented at the 2000 Annual Meeting of the National Council of Measurement in Education.
- [24] Shin, S. (2002). Effects of subskills and text types on Korean EFL reading scores. *Second Language Studies*, 20(2): 107-130.
- [25] Stein, N.L., & Glenn, C.G. (1979). An analysis of story comprehension in elementary school children. In R.O. Freedle (Ed.), *New directions in discourse processing: Advances in discourse processing* (Vol. 2, pp. 53-120). Norwood, NJ: Ablex.
- [26] Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes*, 15, 249-275.
- [27] Westby, C. E. (1991). Assessing and remediating text comprehension problems. In A. Kamhi & H. Catts (Eds.), *Reading disabilities: A developmental language perspective*. Boston, MA: Allyn & Bacon.

Godefroid B. Katalayi lectures Academic literacy and Method of English modules in the Faculty of Education, University of the Western Cape, South Africa. He also lectures Language testing and TEFL modules in the Department of English at “Institut Supérieur Pédagogique de Kananga”, Democratic Republic of the Congo. His research interest includes language assessment, the validity of reading tests, the assessment of narratives, and the teaching of grammar and reading to EFL learners.

Sivakumar Sivasubramaniam is currently Associate Professor and Head of the Language Education Department in the Faculty of Education at the University of the Western Cape, Republic of South Africa (RSA). He also serves the Editorial Board of the Journal of English as an International Language (EILJ) as Chief Editor and the Editorial Board of Asian EFL Journal (AEJ) as Associate Editor. He has been a foreign language/ second language educator for over thirty years now and has taught English in India, Ethiopia, Thailand, Bahrain, Armenia, and U.A.E prior to relocating to the Western Cape. He holds an MA in English Literature from the University of Madras, India, an MA in (Linguistics) TESOL from the University of Surrey, U.K and a PhD in Applied Linguistics from the University of Nottingham, U.K. His research interests include response-centred reading/ writing pedagogies, literature-based language pedagogies, constructivism in EIL, second language advocacy, narratives in language education and text-based approaches to academic and social literacy practices.