

2. Study of Response Validity of the IELTS Writing Subtest

*Peter Mickan, The University of Adelaide
Stephen Slater and Carol Gibson
Centre for Applied Linguistics
University of South Australia*

Publishing details

**International English
Language Testing System (IELTS)
Research Reports 2000
Volume 3**

Editor: Robyn Tulloh

IELTS Australia Pty Limited
ABN 84 008 664 766
Incorporated in the Australian Capital Territory
Web: www.ielts.org

© 2000 IELTS Australia.

This publication is copyright. Apart from any fair dealing for the purposes of private study, research or criticism or review, as permitted under the Copyright Act, no part may be reproduced by any process without written permission. Enquiries should be made to the publisher.

National Library of Australia
Cataloguing-in-Publication Data
2000 ed
IELTS Research Reports 2000 Volume 3
ISBN 0 86403 036

2. Study of Response Validity of the IELTS Writing Subtest

*Peter Mickan, The University of Adelaide
Stephen Slater and Carol Gibson
Centre for Applied Linguistics
University of South Australia*

Abstract

This paper reports on an investigation into the response validity of the IELTS Writing Subtest. The general purpose of the study was to identify factors in the assessment context which may influence candidates' achievement on the writing test.

The first part of the study examines the readability of test prompts and identifies discourse and pragmatic features which may influence candidates' comprehension of the prompts.

The second part of the study describes and analyses the test taking behaviours of intending IELTS candidates on the IELTS Writing Subtest. The analysis in this section is based on researchers' observations of subjects, on verbal protocols recorded as subjects wrote their responses, and on post-test interview data recorded with the same subjects.

The analysis for readability suggests that candidates' understanding of the test prompts is influenced by factors such as the purpose described for the tasks and the lexico-grammar of the tasks. The analysis of the test-taking behaviours of subjects points to socio-cultural influences on candidates' demonstration of their writing ability.

The study suggests the usefulness of using qualitative procedures such as verbal protocols for the investigation of response validity.

1.0 The Study

1.1 Research Context

This study was planned as an exploration of issues raised by a number of assessors about the quality of responses on the IELTS Writing Subtest. The assessors informally suggested that a substantial number of candidates do not produce responses which reflect their academic writing ability, and that they may be underachieving for a variety of reasons. In a preliminary analysis of the causes for this underachievement the assessors suggested three possible explanations:

1. candidates' misinterpretation of the writing task prompts
 2. candidates' lack of knowledge of the task topics
 3. candidates' inadequate control of the discourse-semantic and lexico-grammatical resources for the composition of appropriate responses.
-

The subjects for this study were non-native speakers of English. The selection of subjects took account of their cultural backgrounds in order to include subjects with different first language backgrounds. These were:

German	Indonesian	Korean	Swedish	Thai
2	2	1	1	3

Subjects were estimated by their teachers to be at an overall ability level in English of about 6 (IELTS). This approximate level was chosen so that subjects' command of spoken English would be adequate for participation in the study. However, subjects were given the option of reporting either in their first language or in English. One subject chose to give her verbal report in her L1 and for the interview to be conducted in L1. Her verbal response was transcribed and translated by an accredited interpreter.

The nine candidates were selected from those who intended to take the Writing Subtest assessment (Task 1 and Task 2) during scheduled IELTS test administrations from December 1997 to February 1998.

2.2 Data Collection

For this study we used test items from the IELTS non-live (no longer in use) Writing Subtest versions post April 1995. The versions were selected for the study by University of Cambridge Local Examinations Syndicate (UCLES). The actual questions are confidential, and for the purposes of this study we have not reproduced the actual versions, but paraphrased them. However the analysis of the prompts incorporates selectively the original wording of the selected test versions. During the analysis informal reference was made to live versions in order to make some general comparisons across different kinds of tasks. The test item prompts were analysed for readability from both pragmatic and lexicogrammatical perspectives.

For the collection of data on the test-taking behaviours of candidates we audio-recorded nine intending candidates' verbal protocols while composing responses to the three selected non-live versions of IELTS Writing Subtest Tasks 1 and 2. At the same time the subjects were observed by the researchers, who noted subjects' verbal comments and writing behaviours in relation to the composition of their responses. We also conducted interviews with candidates after they had completed their responses. The recorded verbal reports were transcribed verbatim. The concurrent verbal reports were analysed in order to identify common patterns in the composing actions of the subjects. The post-test interview data was used as a reference to check out the observations made by the researchers of candidates' test-taking actions.

3.0 Readability of Test Prompts

Readability of text prompts is crucial for task fulfilment. The nature and extent of candidates' understanding of test prompts influences their capacity to meet assessment criteria. Mickan and Clennell (1996) refer to the comprehensibility of examination questions as 'an essential condition for the elicitation of appropriate responses from students'.

Horning (1993) defines readability as: '... the ease of understanding a particular text for particular readers, based on how the text is written'. A text has meaning potential (Halliday

1985), and a reader assigns meaning to a text at a particular time in the process of reading. In general the brevity of test prompts limits the contextual information through which a reader is able to assign meaning to it. The formulation of test prompts is therefore of critical importance for candidates' interpretation of their meaning potential.

Various measurements of readable writing such as length of sentences, word counts and ratio of content to function words (Horning 1993) have been used as measures of readability. We considered the use of such measures for the study of a selection of non-live and live versions of the writing subtest prompts, however our analysis using this approach showed that the task prompts varied in length and in the number of embedded clauses, but the analysis did not appear to illuminate aspects of readability. Horning (1993) notes that the measurement of such features of texts 'fails to tap discourse-based and reader-based features of text'. Therefore for the analysis of readability of test prompts we took account of lexicogrammatical, discourse, contextual and sociocultural factors for their potential influence on candidates' capacity for understanding test prompts.

3.1 Pragmatic Features of Test Prompts

An analysis of the pragmatic features of test prompts focuses on the purposes for composing a text or texts. Texts serve social purposes and the explicitness of the social purposes of test prompts contribute to candidates' understanding of what they are to do in response to the prompt. Explicit social purpose serves as a guide to a writer in the selection of an appropriate genre and in the choice of the lexicogrammatical and discourse semantic elements necessary for the realisation of a socially acceptable text form.

From a pragmatic viewpoint the prime surface purpose of test prompts is to provide input for the assessment of writing. The test prompts used in this study however contain elements which make this purpose problematic.

The task instructions state for example the purpose as either:

- a. writing to a lecturer for Task 1, or
- b. to an educated, non-specialist audience for Task 2.

The title of the subtest however, is academic writing, so it is reasonable to assume that an overall purpose is to produce academic writing.

For Task 1 a common, stated purpose is to describe in written form information that is presented in a statistical, graphic or visual form. This purpose satisfies the need to elicit written text but raises the issue of task authenticity in the sense that in many actual situations in which academic writers abstract statistical information already shown in a graph or table it is in the context of comment, interpretation or extended discussion, as for example in the commentary accompanying newpoll surveys reported in the press, or in students' oral class presentations, where statistical information might be displayed on an overhead transparency to illustrate oral analysis.

For Task 2 a commonly stated purpose is to elicit from the candidate a written argument on a topic of general interest. The unstated but assumed intention is to assess a writer's language ability. The pragmatic meaning then operates at several levels, including linguistic choice,

topic and task authenticity. The candidate needs to decode the relative importance of competing purposes and audiences in the composition of a response.

Such pragmatic elements potentially reduce for candidates the clarity of the task. Task 1 of the subtest material used in this research project state 'Write a report for a university lecturer...'. This suggests that candidates' schemata embody a clear construct of university lecturer and can furnish a genre identity for 'report' as used in this instance. The university lecturer and educated, non-specialist audience are intended fictions used to simulate an academic context. In actual academic contexts, academic writing is generally based on content material that is studied prior to doing an assignment.

Given the significant role of such social-cultural factors as audience and purpose on text design, candidates' conceptualisations of audience and purpose and therefore their selection of text features for the composition of their responses may be assumed to be influenced by such instructions. The explicitness of the social purpose of a text enables a writer to select appropriate relationship, content and textual meanings exhibited in conventional features of a socially acceptable text form.

3.2 Lexico-grammar and the Readability of Test Prompts

Mickan and Clennell (1996) identified the grammatical structuring of questions, lexical choice and the contextualisation of assessment tasks as affecting students' comprehension of examination questions. The following analysis of the lexical and syntactic elements of task prompts considers how they might affect subjects' comprehension of the prompts. The analysis is of Task 2 prompts. The topic questions are from the Academic writing subtests. The actual wording for the questions remains confidential, so the task topics have been paraphrased instead.

Academic Writing Task 2 Version A (paraphrase)

The topic for this question is free education: it states that there are many people who believe education should be free, but there are countries where fees are paid by students in secondary schools. The task is to look at one or the other sides of the argument about free education, and to write this up as a report.

The first sentence of the prompt puts 'many people' as the actors (logical subject) who internally carry out the action through the sensor verb 'believe'. However, the psychological subject and theme is free education which is subordinated in a clause ('..that education should be free') carrying modality through 'should'.

The second sentence is in propositional contrast and exhibits ellipsis ('..in many countries'). Again the theme 'school fees' is not at the front end of the prompt.

The third prompt sentence uses an imperative mood ('Write..') to address the reader and offers a binary choice 'for or against'. Again the theme ('paying school fees') is subordinated elliptically at the rear of the sentence.

Lexically the item writer's choice of 'many' as in 'many people believe..' has enabled the opening proposition of the task prompt to read more strongly as a popular belief (contrast

with 'Some people believe..'). This perhaps makes it less likely that the candidate will be confident enough to argue for school fees.

The actors ('secondary school students') in the second prompt sentence apparently 'pay fees' which might confuse some candidates who take that proposition literally and wonder why it isn't the parents of those same students who pay the fees. The selection of 'many' as in 'in many countries' adds weight to the generalisability of the proposition. Again, depending on a candidate's cultural experience and knowledge, the strength of the proposition might be problematic.

The third prompt sentence asks a candidate to 'Write a report...' suggesting again a genre (a report) with which the candidate can identify and inhabit with appropriate text. The authenticity of the task seems somewhat problematic given that the appeal for a report is in the context of 'academic writing' on a topic of which the candidate has had no prior warning nor prior opportunity to consider his or her perspective.

The option in the task prompt to 'write a report for or against...' suggests a text structure that is open to polarised perspective (for or against). This is a variation which subjects in this study were not expecting. The three subjects who did this task presented arguments both for and against free education in their written responses. It appears that the subjects chose a framework for their texts with which they were familiar. They interpreted the task as: Discuss the advantages and disadvantages of free education.

Academic Writing Task 2 Version B (paraphrase)

The topic for this task is overpopulation. The question states that expert opinion is that overpopulation contributes to damage to the environment and to levels of poverty. The task is to consider what to do about overpopulation and also to identify possible problems in taking proposed actions.

Sentence 1 of the prompt uses actor as subject: 'Experts...' and the senser verb 'believe...' with the complex subordinate clause containing as its final words the themes 'large population and a high birthrate'. The sentence is heavily nominalised. (Experts/ poverty/ destruction/ countries/ pressures/ population/ birthrate).

The second input sentence pivots from an interrogative mood with 'How do you think...?' followed by an elliptical clause carrying the themes 'overpopulation and high birthrates' right at the end. The modal 'could' introduces the sense of levels of likelihood and hypothesising.

The third sentence of the task prompt uses an imperative mood ('Indicate.... '), an elliptical clause and a modal 'might...'. The use of 'such measures' operates cohesively and anaphorically with the initial question form ('How do you think..?').

The initial part of the first sentence ('Experts now believe...') positions the reader into a deferential acceptance of the proposition since the candidates know that they are not expected to be an expert as they are writing for a non-specialist audience. The noun 'destruction' is emotive and carries a sense of wilfulness on the part of the absent destroyers (actors). The agents of the actions are not identified (those whose actions lead to large populations and high birthrates).

The second sentence asks the reader in personal terms ('How do you think...?') for suggested solutions to prejudged problems (overpopulation and high birthrates). The notion of 'overpopulation' again carries a built-in value judgment rather than being seen as problematic in general terms. The notion of 'countries' is both nonspecific (any or all countries) and rendered animate, since the following verb renders countries capable of the action of 'reducing...' (countries could reduce..) and the modal invites the reader/writer to hypothesise.

The third prompt sentence uses an imperative mood and implies that there are likely to be difficulties ('Indicate any difficulties...'). The lexis chosen is impersonal and suggestive of social engineering rather than human consequences ('.implementing such measures'). In countries which have not bureaucratised such changes the metaphor of 'implementing' may be problematic conceptually in this context even if recognised for core meaning. The choice of the verb 'might arise..' is suggestive of difficulties happening of their own accord rather than through the agency of government or human decision-making. This adds to the impersonality and abstractness of perspective in the prompt.

Academic Writing Task 2 Version C (paraphrase)

The topic lists the yearly incomes of seven professions and asks the question whether the variation in income is an indicator of merit and social significance.

The first prompt sentence opens in a style characteristic of spoken language ('Here is....'). The second part lists seven professions and salaries ('possible earnings') in pounds. It is arranged in order of salary size with the smallest salary at the top and so on down.

The third part of the prompt is a Yes/No interrogative ('Do differences.....reflect....?') tempting the reader to interpret the item writer as seeking a definitive answer rather than an argument.

The lexical choices in the first prompt sentence have unusual features. The specification of 'England' as the location of the jobs and earnings in the list creates cultural specificity. The use of the adjective 'different' as in 'different professions' is redundant since the list makes the differences self evident. The choice of 'possible' as in 'possible earnings' is ambiguous since it is not clear whether it is being used to mean actual, potential earnings or merely hypothetical earnings. The choice of the expression 'professions' is interesting since 'footballer' is not always considered a profession, a term which in some uses implies prescribed years of training involving a body of knowledge and examinations.

The second part of the prompt chooses predominantly middle class occupations which implies class-based notions of value to society and makes cultural assumptions of the importance of certain professions in society.

The third part of the prompt chooses the word 'salary' as a partial synonym for the earlier term 'earnings' but that choice underscores the middle classness of the listed occupations since 'salary' is not usually used for income of factory workers or unskilled labourers. The use of the expression 'really' as in 'Do differences in....really reflect...?' seems to convey an appeal to an authoritative reality or to require a decisive answer to a 'Do...?' question when in fact the issue is one of judgement. Some candidates may believe that they are being asked to approach that authoritative reality in their response, a cross culturally intimidating thought. The choice of the expression '..importance in society' is a generic sociological perspective and as such is highly abstract suggesting an equally abstract response that deals in sociological universals. An alternative and less abstract formulation would have been '..importance in *your* society',

thus also enabling candidates to feel permission to connect primarily with their own cultural context, despite the use of 'England' as the context for the list in the prompt.

3.3 Discussion

The pragmatic and lexicogrammatical analyses of selected Writing tasks suggest features of test prompts which may contribute to their readability. The test prompts reflect socio-cultural contexts and embody socio-cultural assumptions. Candidates' comprehension of the prompts involves unravelling the contextual meanings of a task in order to interpret the intentions of the writing task and the nature of the task topic.

Although this study did not include an analysis of subjects' comprehension of test prompts, subjects made reference to their understanding of task prompts in the recorded data. When candidates received the test prompts, their first actions involved reading and framing an understanding of the tasks. Three candidates reported some difficulty with the interpretation of prompts. HA's protocol for Task 2 commenced as follows:

"I was thinking um, of word I can't understand" [HA. T2. VP. 1]

Subjects spent time before writing their responses reading and rereading the prompts, focusing on those elements of the prompts which were not clear to them. RI was asked in her post-test interview whether she understood the test question.

She commented as follows:

R. "I seem to understand but I am not sure"

and

R. "I know the meaning of the words, but I don't understand the overall meaning difference ... reflects ... Does this mean the difference of the salary ... does it influence .. what is it .. I don't understand" [RI. I: 14]

The readability of task prompts creates a framework around which candidates build their responses.

The confident interpretation of a task prompt and topic creates conditions for candidates to respond with assurance, as planning and formulating text are enmeshed in the meaning of a prompt and in candidates' perceptions of task purpose. An understanding of the prompt is necessary for candidates' selection and composition of salient meanings in their response texts.

In this section we have considered some linguistic and pragmatic features of test prompts which may contribute to their comprehension and therefore influence the capacity of candidates to display their academic writing ability by producing appropriate responses to the prompts. The next section explores further this close connection between the task prompt and the composition of responses.

4.0 Subjects' Composing Actions

The data reported and examined in this section were collected with the aim of identifying factors which may impact on subjects' capacity to display adequately their ability to write academic English. The analysis is based on researchers' observations, on the verbal protocol data recorded as candidates wrote their responses, as well as on the recorded post-test interview data. The purpose for the collection and analysis of this data was to obtain on-line information on the actions and decisions of the writers as they composed their responses, and to interpret how these behaviours might influence their test-taking behaviours.

The perspective selected for the analysis was the identification of significant actions the writers took for the composition of responses to the test prompts. This perspective is used as an indicator of what the test is testing for the candidates involved (Alderson, Clapham and Wall 1995, Henning 1987).

The subjects' verbal reports and the researchers' observations indicate the complexity of the processes writers engage in for the composition of texts. Consistent with other studies of second language writing processes (Grabe and Kaplan 1996, Hudelson 1989, Krapels 1990), the subjects' main actions in this study included planning prior to writing, formulating text, and editing.

4.1 Interpretation of Task and Planning the Response

Planning has been identified as an activity which successful writers do before they begin writing (Hudelson 1989, Scardamalia and Bereiter 1987). In this study we observed three subjects who wrote notes as part of planning their responses. Although other subjects may have engaged in planning, they did not verbalise their actions so these processes are not recorded in their protocols.

The variation in time spent between receiving the test task and the commencement of writing shows that some subjects' apparently spent much more time interpreting the prompts and planning their responses than others. For all subjects less time was spent before writing on Task 1 (minimum 15 seconds - maximum 71 seconds) than on Task 2 (minimum 35 seconds - maximum 10 minutes). In the pre-writing time subjects were reading and rereading the prompt in order to comprehend what was being asked for in the assessment question. RI for example said as she read the question:

"I don't really understand what the question is" [RI. T2. VB. 13]

The ability to understand the task was linked to planning the response. As subjects read the assessment tasks they selected elements in the tasks for use in the development of appropriate responses. The actions were associated with selecting ideas and building task representations in preparation for writing responses. HA for Task 1 reported the following as he read the task and before he began to write:

"I'm thinking how does the graph show us er whi, er which way to explain hm the percentage of people unable to find work". [HA. T1. VP. 1]

The process of analysis of the question was contiguous with the selection of salient elements in the task and planning the response.

Similarly for Task 2 HA read the task prompt while planning the response. He commented in the post task interview:

"when I read the question then I was thinking of good example.
[Researcher: mhmh]
and er then er what is good for introduction." [HA. T2. I. 1]

Here the writer's reading of the task is connected with initial formulation of the response. Some subjects adopted words or phrases from the task prompt for building their responses.

WE for example reported at the beginning of her protocol for Task 1:

"Er .. I'm just looking at graph to to build some to build er words" [WE. T1. VP.1]

The wording of the prompts seemed to provide a stimulus or a starting point for the commencement of subjects' own texts. This was particularly the case for Task 1 where the subjects were required to reproduce meanings depicted in task prompts, so that they were to some extent dependent on the technical, lexical and syntactic elements in the prompts for building their responses. Subjects were able to take language items from the prompts which gave technical authority and a sense of accuracy for the composition of their texts.

The close relationship between the organisation of the information contained in the prompt and the process of planning and structuring the response was illustrated in DI's report for Task 1. As he read the prompt he analysed it and utilised information in the prompt for structuring his response. The following extract illustrates the juxtapositioning of task interpretation and planning.

"O.K. Just check it. This chart shows a graph we have a graph with three with the percentage of people who can't find work .. maybe I change it a bit. This chart shows a graph with three different countries ... no, that's OK. It's OK. OK. Now, what I, first what I'm doing, I .. yeah .. then I will go, I will take each, each country and I will describe what's happened between these years. So I maybe I just start with the first, with the United Kingdom and and will describe what happened between 19, so I will first, made the first, United Kingdom, Canada, and finally Japan. Then I will maybe give some trends, and in the end I make a conclusion of all." [DI: T1. VP. 2].

This section of the account shows the alignment of DI's understanding of the prompt with the formulation of options for inclusion in the response. He is involved in the interpretation of the prompt while he skilfully rehearses both selection of information to include in the response and also the structuring of the information for redeployment in his text. The candidate highlighted key concepts in the prompt. He also simultaneously brainstormed and structured systematically ideas for inclusion in his response.

4.2 Organising the Response Text

Subjects' accounts demonstrate the close association of reading and understanding a prompt with planning the response. Subjects' reported planning activities at macro- and micro-levels.

Subjects engaged in macro-level planning during pre-writing and while they composed their responses. WE in her Task 1 protocol spent 64 seconds in silence while reading the prompt and then said:

"I'm going to write the introduction" [WE. T1. VP. 1]

After writing the first sentence, she paused and said:

"Um right now I'm going to write the thesis statement of the introduction" [WE. T1. VP. 1]

Similarly HA after writing the opening to his Task 1 text ('The graph shows us') stated

"I'm trying to write introduction" [HM VP T1: 1]

A little later he said:

"I'm from now on I'm trying to write er the body" [HM VP T1: 1]

Towards the end of the protocol for Task 1 he reported:

"Er now I'm trying to er write at end of the body and er to write conclusion" [HM VB T1: 3].

This account records HM's awareness of text organisation not just at the planning stage, but throughout the composition of the response. The successful construction of a text requires that subjects attend to text organisation at a macro-level. What is interesting is subjects' use of metalanguage to describe their actions for the organisation of their texts. The use of metalanguage to describe the structure of the response is recorded in most subjects' protocols. The subjects described the design or structure of their responses in conventional and general concepts, in the above examples as introduction, body and conclusion. Other subjects used the term 'thesis' to describe a section of the text, as in WE's account above. Subjects appeared to find the metalinguistic concepts useful for the construction of their responses.

Of the nine subjects in this study, three wrote written plans before beginning writing their responses. These were produced for Task 2. No subjects wrote plans for Task 1. Subjects' planning processes varied. RU for example did not plan or write notes for Task 1. He read the question for 20 seconds and then began to compose his text. For Task 2 (Overpopulation topic) he read the prompt for two minutes, then brainstormed for six minutes, writing a page of notes as a plan for his response.

The written plans mirrored subjects' understanding of the task prompt as well as documenting ideas for inclusion in their responses. This is apparent in the plan of one of the candidates (SU) for Task 2 on the topic of free education. The plan is a typed reproduction of the original.

Agree disagree

- support education - develop humanity - free until secondary - force planing edu -
more income - can't support all of them

SU sketched this plan at the top of the examination paper before commencing her text. The plan, although brief, depicts her understanding of what she planned to do for the composition of her response. She interpreted the instruction in the prompt's rubric 'Present a written argument or case' by listing points under two headings "Agree" and "disagree". The subject's plan illustrated her reframing of the task and her initial structuring of a response at a macro-level. As has already been discussed above (3.2) the task did not require argumentation for and against free education. Candidates were instructed to write a 'report for or against' the payment of school fees.

A third subject, DI, spent ten minutes planning his response for Task 2. The plan exhibits systematic planning. The plan provided a structure for the identification of ideas for inclusion in the text. The content of this and the two other plans is not the abstract design of a response. The plans consist of ideas or content concepts related to the topic of the task. It seems that the foregrounding of ideas in this way makes them available for uptake during the process of composition. It may also assist writers in the creation of coherence in the sequencing and in the prioritising of ideas.

Once subjects had begun writing, their attention focused on the formulation of text at the micro-level. They generated ideas at the point of inscription. In the following extracts from RU's protocol the process of text generation is exhibited:

".. now I will give er the example or reason for what I have got there." [RU. T2. VP. 5]

"Er, I will give a reason, I will explain what happens when someone stops immigration," [RU. T2. VP. 5]

"I have two reasons for, er two examples for a stop the overpopulation in the world .." [RU. T2. VP. 6]

"I will compare the developing countries" [RU. T2. VP. 6]

RU formulated what he intended to write while he composed the text. This was an ongoing process of selection of ideas during the formulation of text.

The formulation of text included the selection of vocabulary. Subjects commonly reported looking for words as they produced their responses.

"Er thinking about the right words" [WE. VP. T2: 8]

Sometimes subjects translated terms from their first language:

"Trying to translate a word from Swedish to English" [SO. T1. VP. 2]

The selection of lexical items was linked to the attempt to express accurately intended meanings, as in the next extract:

"Trying to find another, mm better word for biggest" [SO. VP T1: 2]

Formulation of text required choosing appropriate grammatical elements at the discourse level. For example

"I have two reasons for, er two examples for a stop the overpopulation in the world, and I think about the correct er sentence structure that I can put the examples into the sentence". [RU. T2. VP. 6]

As subjects composed their responses they also edited what they had already written. Editing included changing spelling, making grammatical changes and rephrasing segments of text. In the following instance the subject used her knowledge about language to explain the editing action she was taking as she composed her response:

"I need to use the preposition to link the when I write and analyse on the top and I need to join the word to er to er to look smooth to follow the sentence for the task" [S. VP. T1: 8]

SU's concern here is with cohesion in the construction of her text.

For Task 2 the subjects engaged in knowledge transforming activities. These were evident in subjects' written plans, and also in verbal references in their protocols. Three candidates made specific reference to brainstorming ideas for their responses. For example DI said:

"So birth control. I make a little brainstorm, so I can organise my thoughts. Birth control. Education. Information, especially women, because men don't think too much" [DI. T2. VP. 2]

Task 2 engaged subjects in active decision-making in the generation of information. It required writers to plan and structure knowledge according to recognisable text structure. Successful task completion relied to a greater extent on world knowledge than Task 1. This allowed for candidates to draw on their own experience and knowledge for the formulation of their responses.

The distinction Scardamalia and Bereiter (1987) make between knowledge telling and knowledge transforming in the process of composition of texts suggests that the tasks measure different aspects of writing ability. The Task Ones in this study required subjects to process information contained in the prompt in a text structure ('write a report') which is not clearly defined. Task 2 required writers to generate ideas on a topic, and to plan and structure the representation of knowledge in a conventional structure ('present a written argument or case').

5.0 Discussion of Findings

In this study we have examined the readability of IELTS Writing Subtest prompts and used observations and verbal reports to obtain insights into the text-taking behaviours of candidates taking the IELTS Writing Subtest. At a general level the data shows the complexity of test-taking for individual subjects. Subjects approached the composition of responses in different ways, some planning their texts in detail while others began writing their texts a few seconds after receiving the task prompts. Assessing writing ability constitutes a complex literacy event and this study points to factors in the assessment context which may contribute to candidates' effectiveness in displaying their writing ability. Both the readability of prompts and subjects' skills in the formulation of appropriate responses impact on the response validity of the test items.

In the analysis of prompts for readability we described potential difficulties for candidates embedded in the lexico-grammar of the prompts as well as the potential influence of pragmatic factors such as audience and socio-cultural experience on the interpretation of prompts.

The analysis highlighted the interaction of subjects' socio-cultural knowledge and their capacity to display their writing ability. Subjects' planning and formulation of responses was influenced by the socio-cultural knowledge - that is world knowledge and textual knowledge and experience which they were able to access for the composition of their responses. Task 2 in particular, based on the data in this study, tested the capacity of subjects to argue a case, and as well assessed the socio-cultural information they could draw upon to do so. The plan of one European subject illustrated how he was able to draw upon his experience of cultural and social issues for the composition of his response.

In a previous study of the composing actions of L2 writers Mickan (1996) found that subjects utilised socio-cultural information and experience for structuring their responses according to generic conventions. In a comparable way in this study salient elements in the prompts appeared to signal to candidates acculturated responses which they used for the creation of their written responses. The subjects needed to access knowledge of text conventions for the composition of texts appropriate to the social purposes of the communicative event.

The composition of contextually appropriate texts requires the use of obligatory textual elements (Halliday and Hasan 1985), which conform to cultural conventions. The obligatory elements are realised through structural elements in the organisation of texts, as well as in the lexicogrammatical elements of the texts. In addition to the obligatory elements the writers have optional elements to choose from to create their texts.

As the obligatory elements culturally distinguish types of texts or genres, that is they convey cultural meanings, we might assume that it is these elements which are foregrounded in assessors' evaluation of the degree of success in candidates' construction of texts. Hence the importance of subjects' identification of the kind of text to produce in response to a prompt, a process which includes analysis of the social purpose of the response and the audience of the response. This process was evident in the planning and formulating actions of the writers. They read the prompts and scanned them for clues for the design of their texts. For Task 1 subjects had to analyse and reproduce the information in the prompts, whereas for Task 2 they had to utilise to a greater extent experiential knowledge for the formulation of their responses. With information from their analyses subjects made selections from discourse-semantic and lexico-grammatical resources to construct their responses. In the process they edited their texts, although time limitations restricted editing activity.

6.0 Implications of the Study

The purpose of this study was to explore the test-taking behaviours of IELTS candidates in order to identify factors in the assessment context which may influence candidates' achievement on the IELTS Writing Subtest. The study focused on an analysis of the readability of the task prompts and on the reported test-taking behaviours of subjects.

The analysis of the test prompts highlighted pragmatic and lexicogrammatical factors which influence readers' comprehension of the prompts. In particular the socio-cultural information embedded in the task prompts has an impact on their readability. As the prompts are brief texts, the contextual information which they exhibit and which is fundamental to candidates' understanding, is necessarily limited. Candidates therefore need to rely on their interpretation of limited cues for the appropriation of the intended meaning of the task prompts.

The analysis of the verbal protocols and of the post-test interviews showed the uncertainty felt by subjects when their understanding and analysis of the prompts was not secure. The detailed analysis of the readability of selected prompts in this study suggests that the formulation of task prompts influences the ability of candidates to produce appropriate responses.

As this study did not propose specifically the collection of data on subjects' interpretation of task prompts, this is an area for further research. What is needed is the scrutiny of candidates' processes of interpretation through concurrent verbal protocols as they read the prompts. This procedure should provide evidence of both the textual and the pragmatic factors which

influence candidates' interpretation of task prompts, and therefore their capacity to compose appropriate responses. The further investigation of how candidates process test prompts would be of practical use for future item writing.

The verbal protocols of subjects in this study provided evidence of how the interpretation of prompts impacted on the planning and formulation of responses, as subjects used information in the prompts for the creation of their responses. The writers drew upon internalised social information about the purposes and structures of texts and through the selection of lexicogrammatical items built their responses. Analysis of subjects' response texts was not a part of this study, but the study suggests the value of further investigation of this aspect of the assessment context. A suggested investigation is the functional analysis of the written responses of candidates in order to identify how candidates' interpretation of prompts guides the textualisation of meanings for the composition of appropriate responses. Such a study would describe and compare the textual or generic features of texts written in response to different task prompts. The investigation would allow for the targeted investigation of the influence of differences in socio-cultural experience and knowledge on candidates' capacity to exhibit their writing ability.

The protocols for this study showed that some subjects did little advanced planning of their responses, and instead spent most time formulating responses at the local level, particularly in the selection of lexicogrammar. The extent to which this influenced their compositions was not examined. Further investigation of how candidates use plans for the construction of their texts and an analysis of the affect of both macro-level and micro-level planning on text realisation would provide insights into effective test-taking behaviours.

At the macro-level of text formulation subjects appeared to have been guided by general formulae for structuring their texts. Subjects' made reference in their accounts to the structure of their texts using terms such as 'introduction', 'body' or 'thesis' and 'conclusion'. The terms were used to describe the construction of responses for both Tasks 1 and 2, even though the social purposes of the two tasks and therefore their linguistic realisations differed. For Task 1, subjects were reliant on accurate analysis of the prompts' content, as the task required the reproduction of information in the prompt. Task 2 placed greater demands on subjects as they needed to access quickly ideas, examples and arguments, which they then had to actualise in a particular text type or generic form. The data suggest that for the preparation of candidates the differences of texts called for by Task 1 and Task 2 could be described more accurately in genre specific terms rather than in general formulae. For example explicit instruction in the generic features of a 'written argument or case' (Task2) would provide textual resources for those candidates who are not experienced in the building of this kind of text. This suggestion has implications for the preparation of candidates.

To sum up, this exploratory study has provided insights into response validity of the IELTS Writing Subtest: it has identified factors which influence the readability of task prompts, and it has described composing behaviours which influence the composition of responses to tasks. The use of verbal protocols for the identification of subjects' writing behaviours as they composed their responses was found to be valuable, however the ability of subjects to verbalise their actions as they wrote differed considerably. Future studies could address this through training. An alternative is for subjects to report in their first language, as was done successfully in this study with one subject. The real value in investigating candidates' test-taking behaviours through such procedures are the insights obtained thereby into what factors influence individual test-takers' behaviours and how they impact on candidates' achievements under test conditions. The documentation of candidates' processing of test prompts in

conjunction with the composing of responses has the potential to inform testing processes and lead to the enhancement of the assessment capabilities of tests of writing.

Bibliography

- Alderson, J. Clapham, C. & Wall, D. (1995) *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Bachman, L. (1990) *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Brandt, D. (1992) *The Cognitive as the Social. Written Communication*. 9,3, 315-55.
- Chin, E. (1994) Redefining "Context" in Research on Writing. *Written Communication* 11,4: 445-82.
- Grabe, W. and Kaplan, R. (1996) *Theory and practice of writing*. London: Longman.
- Graham, J. G. (1987) English language proficiency and the prediction of academic success. *TESOL Quarterly* 23,1: 505-521
- Green, A. (1998) *Verbal protocol analysis in language testing research*. Cambridge: Cambridge University Press
- Halliday, M.A.K. (1985) *An introduction to functional grammar*. Edward Arnold, London
- Halliday, M.A.K. and Hasan, R. (1985) *Language, Context and text*. Victoria, Australia: Deakin University Press
- Henning, G. (1987) *A guide to language testing*. Cambridge MA, Newbury House
- Horning, A.S. (1993) *The psycholinguistics of readable writing: a multidisciplinary exploration*. Norwood NJ: Ablex Publishing Company
- Hudelson, S. (1989) *Write on: children writing in ESL*. Englewood Cliffs, NJ: Prentice Hall
- Huot, B. (1996) Toward a new theory of writing assessment. In *College Composition and Communication* 47,4: 549-566
- Johns, Ann. (1990) L1 composition theories: implications for developing theories of L2 composition. In Kroll (ed). 24-36
- Krapels, A. (1990) An overview of second language writing process research. In Kroll (ed): 37-56
- Kroll, B. (ed). (1990) *Second language writing*. Cambridge: Cambridge University Press
- McNamara, T. (1996) *Measuring Second Language Performance*. Essex: Addison Wesley Longman.
- Mickan, P. (1996) Composing, negotiating meaning and second language learning. Paper presented at the 21st annual conference of the Applied Linguistics Association of Australia, Sydney: University of NSW, 3-6 October
- Mickan, P. and Clennell, C. (1996) *Special provisions research and linguistic limitation in English language proficiency. Research report for Senior Secondary Assessment Board of South Australia*. Adelaide: CALUSA
- Nystrand, M. (1986) *The structure of written communication*. London: Academic Press

Scardamalia, M. and Bereiter, C. (1987) Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (ed). *Advances in applied psycholinguistics, Volume 2*. Cambridge: Cambridge University Press. 142-175

Smagorinsky P. (ed.) (1994) *Speaking About Writing* . London. Sage Publications.