CRELLA and the socio-cognitive approach to test validation

Dr Sathena Chan
CRELLA, University of Bedfordshire, UK

HKEAA, 31-10 2013

CRELLA
University of Bedfordshire

CRELLA (Centre for Research in English Language Learning and Assessment)

www.beds.ac.uk/crella

CRELLA Research Staff

Prof Cyril Weir: Director of CRELLA

Prof Stephen Bax: Professor in Applied Linguistics

Prof Tony Green: Professor in Language Assessment

Dr Vladimir Zegarac: Reader in Language and Communication

Dr John Field: Senior Lecturer in Cognition in Language Learning and Assessment

Dr Lynda Taylor: Senior Lecturer in Language Assessment

Dr Fumiyo Nakatsuhara: Senior Lecturer in Language Assessment

Dr Sathena Chan: Post-doctoral Research Fellow

Dr Chihiro Inoue: Post-doctoral Research Fellow

Prof Liz Hamp-Lyons: Visiting Professor

Prof Roger Hawkey: Visiting Professor

Rebecca van der Westhuizen: Research Administrator

What we do

- Research projects and consultancies for international examination boards, ministries and universities
  * The British Council (IELTS, ILA, Aptis)
  * Cambridge English Language Assessment (FCE, CAE, CPE)
  * English Language Testing (Password)
  * English Profile
  * The Language Training and Testing Center (GEPT)
  * Trinity College London (ISE, GESE)
  * Eiken Foundation of Japan (EIKEN, TEAP)

What we do (cont)

- Training in language assessment

Recent clients include:

Cambridge ESOL Examinations

LTTC Taiwan GEPT examinations

Association of Language Testers in Europe

The British Council

Society for Testing English Proficiency, Japan

University of Central Lancashire Examinations Board

The Ministry of Education, Singapore

Russian Federation TEMPUS partnership

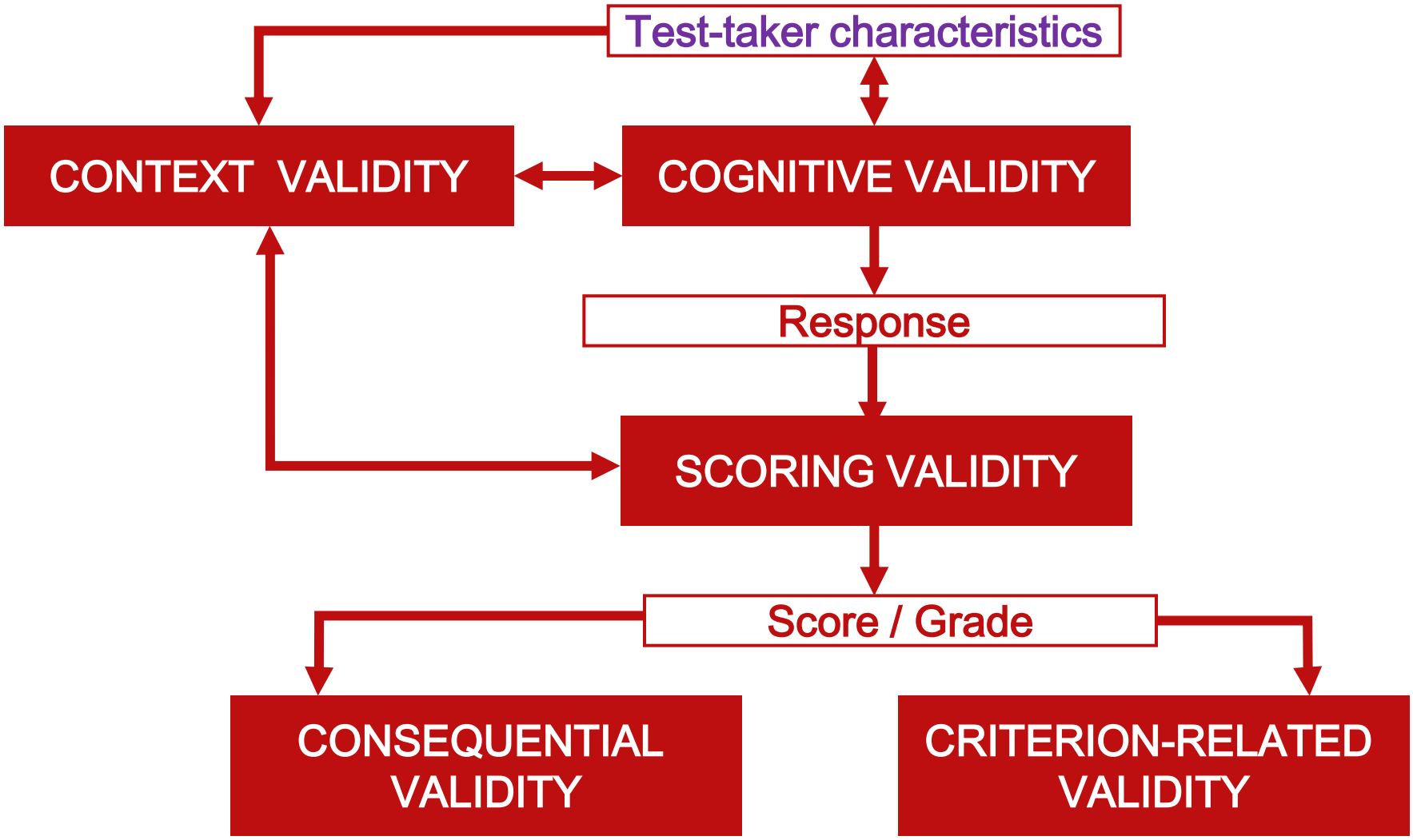
What we do (cont)

- PhD supervision (over 30 students) and MA courses
- Language testing library at Putteridge Bury
- Postgraduate forum (weekly)
- CRELLA research seminars
- Publications

The socio-cognitive framework for language test development and validation

CRELLA
University of Bedfordshire

Weir's (2005) socio-cognitive framework for test validation



Critical questions to be addressed

1. [Test taker characteristics] How are the physical/physiological, psychological and experiential characteristics of candidates catered for by the test?
2. [Context validity] Are the characteristics of the test tasks and their administration appropriate?
3. [Cognitive validity] Are the cognitive processes required to complete the test tasks appropriate?
4. [Scoring validity] How far can we depend on the scores, which result from the test?
5. [Consequential validity] What effects do the test and test scores have on various stakeholders?
6. [Criterion-related validity] What external evidence is there that the test is measuring the construct of interest?

Example validation studies drawing upon the socio-cognitive framework

CRELLA
University of Bedfordshire

1. Research studies on context validity

- Different reading-into-writing test formats (Chan, 2013)
  - Essay task with 2 passages
  - Essay task with 2 passages and 2 non-verbal (e.g. charts, diagrams)
- Different speaking test formats (Nakatsuhara, 2013)
  - Information-gap task
  - Ranking task
  - Free discussion task

Research studies on context validity

Contextual parameters for reading-into-writing tests (Chan, 2013; Khalifa and Weir, 2009; Shaw and Weir, 2007)	

SETTING: Task
•Time
•Clarity of purpose
•Topic domain
•Genre
•Text length
•Overall cognitive demands
•Language functions to perform
•Writer-reader relationship
•Clarity of knowledge criteria
 	
SETTING: ADMINISTRATION
•Writing mode
•Security
	DEMANDS: Input
	•Format (verbal and/or non verbal)
	•Genre
	•Text length
	•Discourse mode
	•Idea density
	•Concreteness of ideas
	•Explicitness of textual organisation
	•Cultural specificity
	•Linguistic complexity
	  •Lexical complexity
	  •Syntactic complexity
	  •Degree of coherence

Reading-into-writing test

UoB Reading-into-Writing Test

Your lecturer has asked you to write a short essay - drawing upon all the information (verbal and non-verbal) contained in the two articles below.

Key points for your essay

1. Give your essay a title which summarises the content of the two articles.
2. Describe the problem and its causes as identified in the articles.
3. Summarize the different solutions that are suggested in the articles.
4. State in a conclusion what of these you think is the most effective solution and give your reasons.

Marking criteria

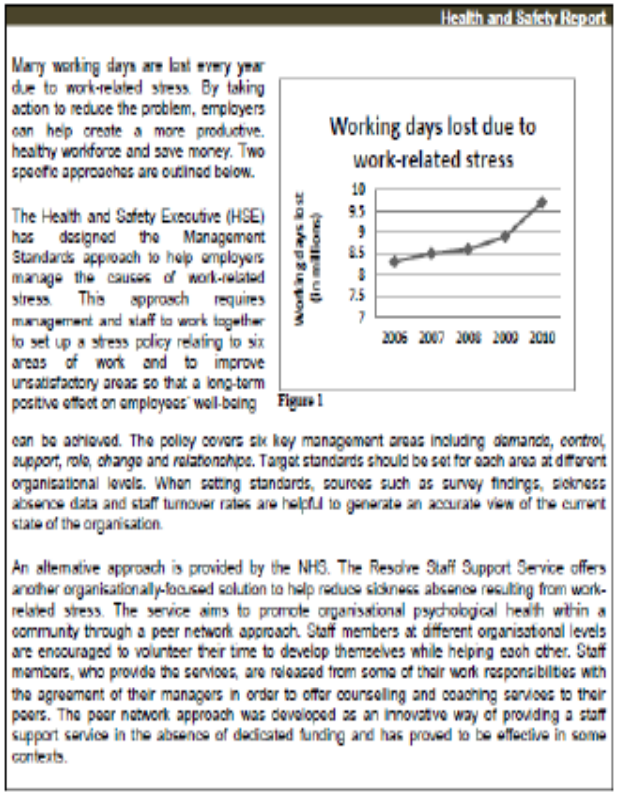
You must use your own words. You will fail if you copy chunks from the articles (more than 5 continuous words in a sentence).

Your essay will be scored on

1. Content
   Are all the 4 key points listed above covered in your essay?
2. Organization
   Is the overall structure of your essay clear and logical?
3. Language
   Have you used a wide range of appropriate grammar and vocabulary?

Instructions on timing

1. You will have 15 minutes to read the articles and then plan your essay.
2. You will then have 35 minutes to write your essay. You should write between 180 and 200 words.
3. You will have another 10 minutes to review and revise your essay. During this time, check that you have fulfilled all the key points stated in task instructions 1-4 above.



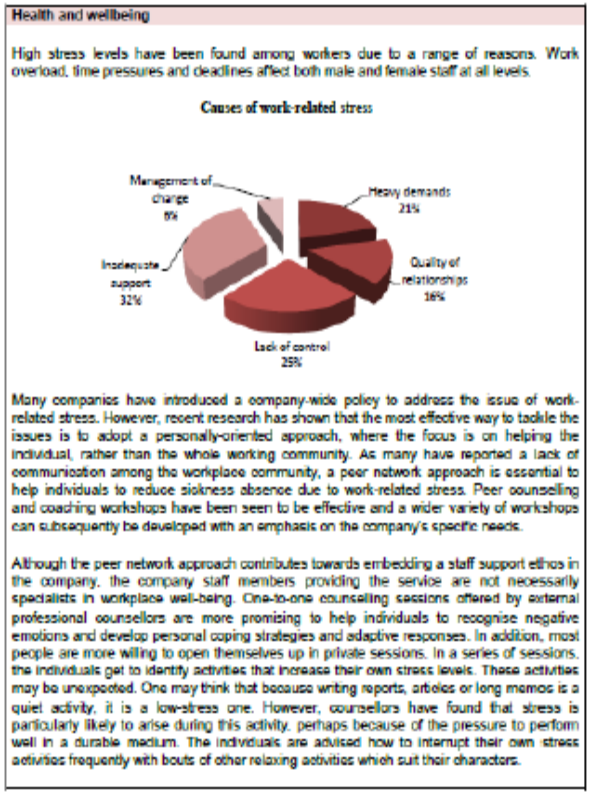
Health and Safety Report

Many working days are lost every year due to work-related stress. By taking action to reduce the problem, employers can help create a more productive, healthy workforce and save money. Two specific approaches are outlined below.

The Health and Safety Executive (HSE) has designed the Management Standards approach to help employers manage the causes of work-related stress. This approach requires management and staff to work together to set up a stress policy relating to six areas of work and to improve unsatisfactory areas so that a long-term positive effect on employees' well-being can be achieved. The policy covers six key management areas including *demands, control, support, role, change* and *relationships*. Target standards should be set for each area at different organisational levels. When setting standards, sources such as survey findings, sickness absence data and staff turnover rates are helpful to generate an accurate view of the current state of the organisation.

Figure 1

An alternative approach is provided by the NHS. The Resolve Staff Support Service offers another organisationally-focused solution to help reduce sickness absence resulting from work-related stress. The service aims to promote organisational psychological health within a community through a peer network approach. Staff members at different organisational levels are encouraged to volunteer their time to develop themselves while helping each other. Staff members, who provide the services, are released from some of their work responsibilities with the agreement of their managers in order to offer counselling and coaching services to their peers. The peer network approach was developed as an innovative way of providing a staff support service in the absence of dedicated funding and has proved to be effective in some contexts.



Health and wellbeing

High stress levels have been found among workers due to a range of reasons. Work overload, time pressures and deadlines affect both male and female staff at all levels.

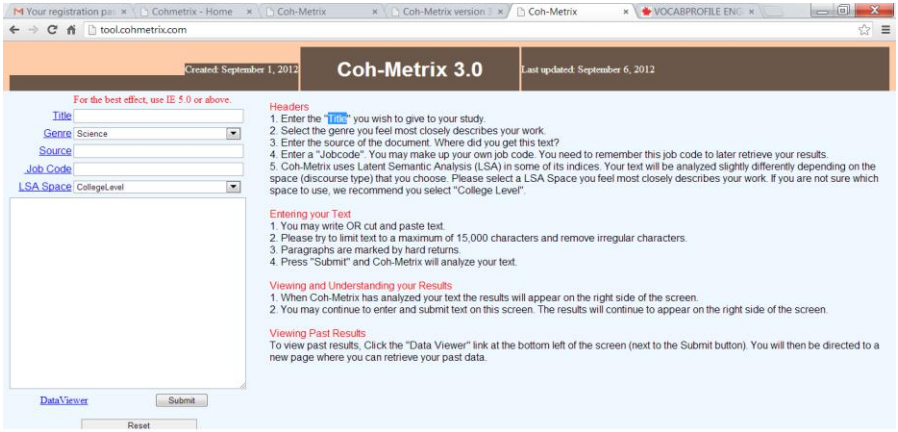
Causes of work-related stress

Many companies have introduced a company-wide policy to address the issue of work-related stress. However, recent research has shown that the most effective way to tackle the issues is to adopt a personally-oriented approach, where the focus is on helping the individual, rather than the whole working community. As many have reported a lack of communication among the workplace community, a peer network approach is essential to help individuals to reduce sickness absence due to work-related stress. Peer counselling and coaching workshops have been seen to be effective and a wider variety of workshops can subsequently be developed with an emphasis on the company's specific needs.

Although the peer network approach contributes towards embedding a staff support ethos in the company, the company staff members providing the service are not necessarily specialists in workplace well-being. One-to-one counselling sessions offered by external professional counsellors are more promising to help individuals to recognise negative emotions and develop personal coping strategies and adaptive responses. In addition, most people are more willing to open themselves up in private sessions. In a series of sessions, the individuals get to identify activities that increase their own stress levels. These activities may be unexpected. One may think that because writing reports, articles or long memos is a quiet activity, it is a low-stress one. However, counsellors have found that stress is particularly likely to arise during this activity, perhaps because of the pressure to perform well in a durable medium. The individuals are advised how to interrupt their own stress activities frequently with bouts of other relaxing activities which suit their characters.

Automated textual analysis tools to research linguistic features

1. CohMetrix analysis
2. Vocab profile
3. Text inspector
4. WordSmith Tools



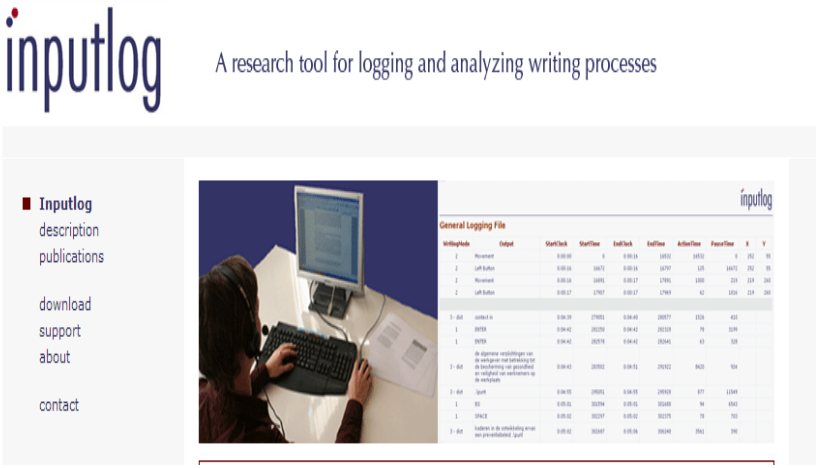
2. Research studies on cognitive validity

- Onscreen summary and essay test tasks with screen record software (Chan, 2010) (Funded by Pearson)
- Essay test task with multiple reading input (Chan and Wu, 2013) (Funded by LTTC)
- Computerised vs paper-based essay test task (Chan et al, forthcoming) (Funded by the British Council, Cambridge English Language Assessment)
- Onscreen reading tests with eye tracking technology (*Bax and Weir; 2012)* (Funded by the British Council, Cambridge English Language Assessment)

Cognitive validity parameters in reading-into-tests (Chan, 2013; Khalifa & Weir, 2009; Shaw and Weir, 2007)

Cognitive Phases	Cognitive processes
Conceptualisation	• Task representation • Macro-planning
Meaning and discourse construction	• Careful global reading • Selecting relevant ideas • Connecting and generating
Translating and micro-planning	• Translating ideas into linguistic forms • Micro-planning
Organising	• Organising ideas in relation to input texts • Organising ideas in relation to own texts
Low-level monitoring and revising	• Low-level editing while writing • Low-level editing after writing
High-level monitoring and revising	• High-level editing while writing • High-level editing after writing

keystroke logging

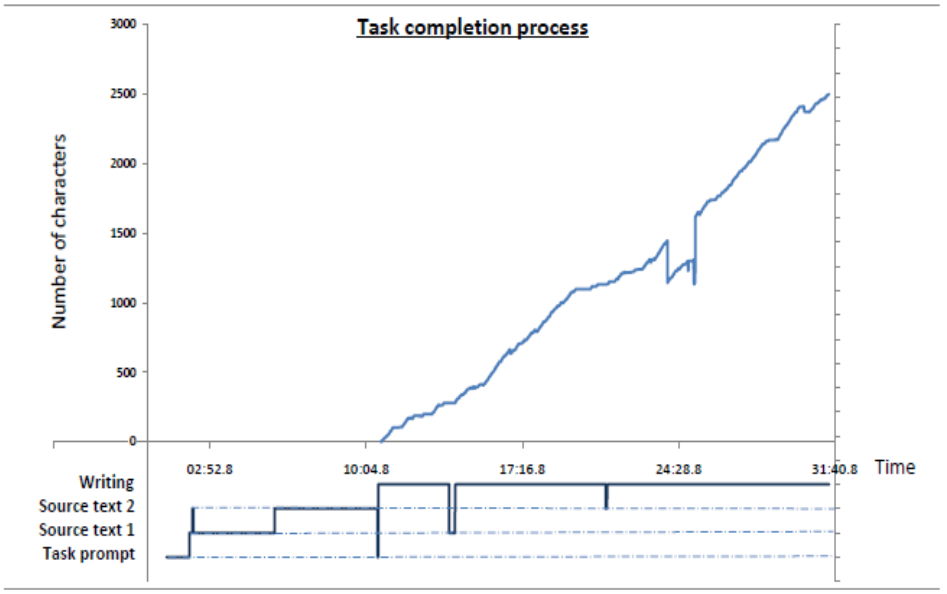


Keystroke log

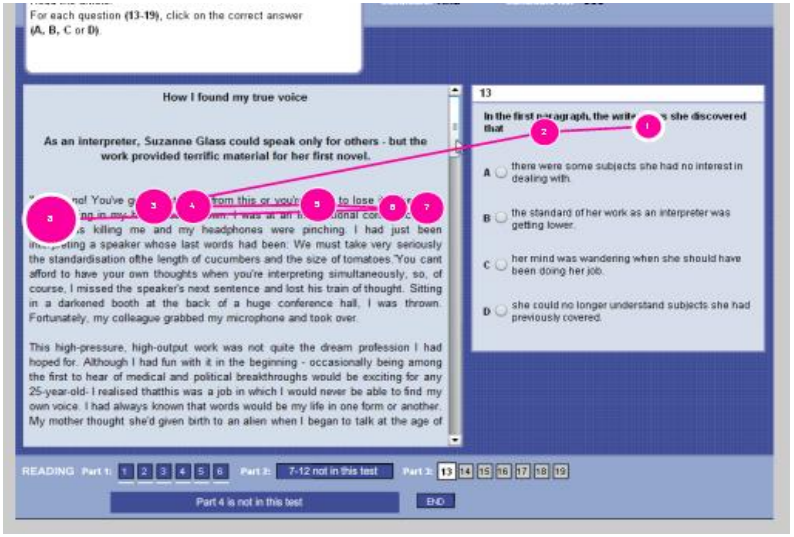
- [Movement][LEFTClick][Movement][LEFTClick][Movement][LEFTClick][Movement][RSHIFT]This·eass[BACK3]ssa·[BACK]y·is·about·the·arguement·of·whether·[RSHIFT]Eng[BACK]dangered·[RSHIFT]Lanu[BACK]guages·should·be·sav[BACK]f[BACK]ved·from·extinction·or·not.·[RSHIFT]From·both·articles·read,·[RSHIFT]Language·is·a·form·of·communicate[BACK]ion·e[BACK]either·[BACK 7]between·[BACK]·[BACK 8]within·the·human·race.

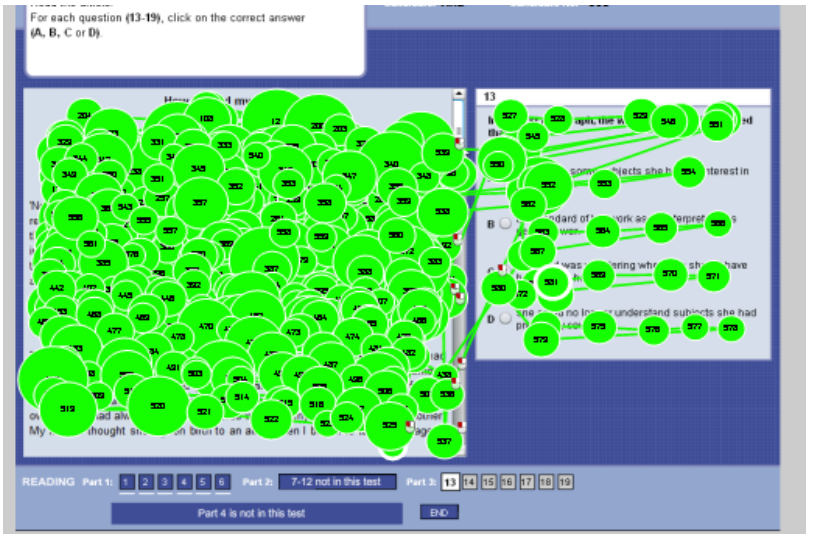
Summary data

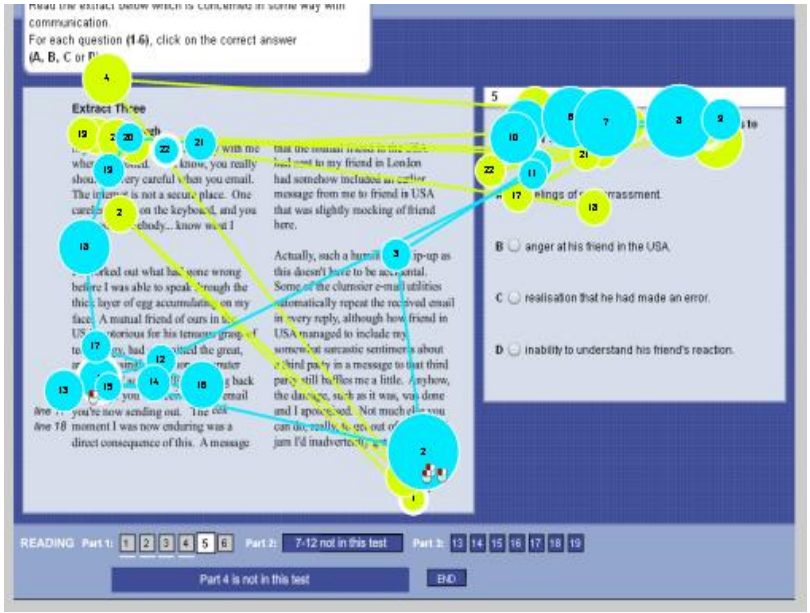
Total Words	486
Average Word Length (WL)	4.99382716
Total Sentences	22
Average Characters/Sentence (C/S)	155.6818182
Total Paragraphs	8
Average Words/Paragraph (W/P)	60.75
Average Sentences/Paragraph (S/P)	2.75



Eye-tracking







For Chinese language tasks/tests?

- Context validity: defining the difficulty level of Chinese tests and task features
- Cognitive validity: investigating the processes of using Chinese language
- EYE and PEN?

For more information:

- Geranpayeh, A. & Taylor, L. (eds.) (2013). Examining Listening, Cambridge: CUP.
- Khalifa, H. & Weir, C.J. (2009) Examining Reading, Cambridge: CUP.
- Shaw, S.D. & Weir, C. J. (2007) Examining Writing, Cambridge: CUP.
- Taylor, L. (ed.) (2011) Examining Speaking, Cambridge: CUP.
- Weir, C. J. (2005) Language Testing and Validation: an Evidence-Based Approach, Basingstoke: Palgrave Macmillan.

Thank you!

Dr Sathena Chan
CRELLA, University of Bedfordshire, UK
sathena.chan@beds.ac.uk

CRELLA
University of Bedfordshire