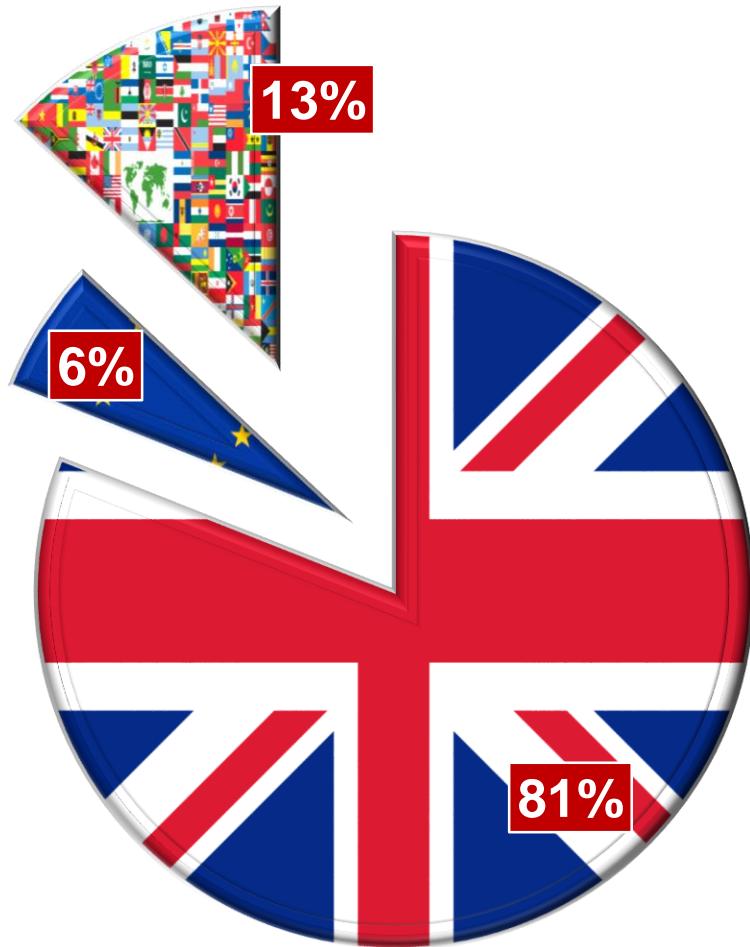# Assessing the Assessments: Tests of Academic Language Proficiency

Anthony Green

University of Bedfordshire

Students in UK HEIs
2016/17 HESA statistics

13%

6%

81%

2,317,880 students

442,375 international

of which 44% postgraduate

Business & administrative studies
121,675

Engineering and technology
52,545

Social studies 42,170

University of Bedfordshire

Introducing the Big Four

TOEFL iBT ®, IELTS™, PTE Academic, C1 Advanced

Availability and recognition
What claims do they make?

Test	Availability		Recognition	
	Countries	*Centres*	*Countries*	*Inst.*
Cambridge C1: www.cambridgeenglish.org	130	2,800	'Worldwide'	Over 8,000
IELTS: www.ielts.org	140	1,000	Over 145	Over 9,000
PTE-A: pearsonpte.com	Over 50	200	56	Over 1,700
TOEFL iBT: www.ets.org/toefl	165	Over 4,500	Over 130	Over 9,000

Four major tests of Academic language
What claims do they make?

The TOEFL iBT® test measures the 4 English language skills you will need to succeed at your institution.

> A Guide to the TOEFL® Test for Institutions, ETS

The IELTS Academic test is for people applying for higher education in an English speaking environment. It reflects some of the features of academic language and assesses whether you are ready to begin studying or training.   www.ielts.org

PTE Academic is [a] test of English for study abroad. Based on the real-life English skills test takers need to communicate effectively in their destination.   pearsonpte.com

C1 Advanced shows you have the language skills that universities are looking for.   www.cambridgeenglish.org

Predictions, predictions

How well do tests of academic language proficiency predict academic performance?
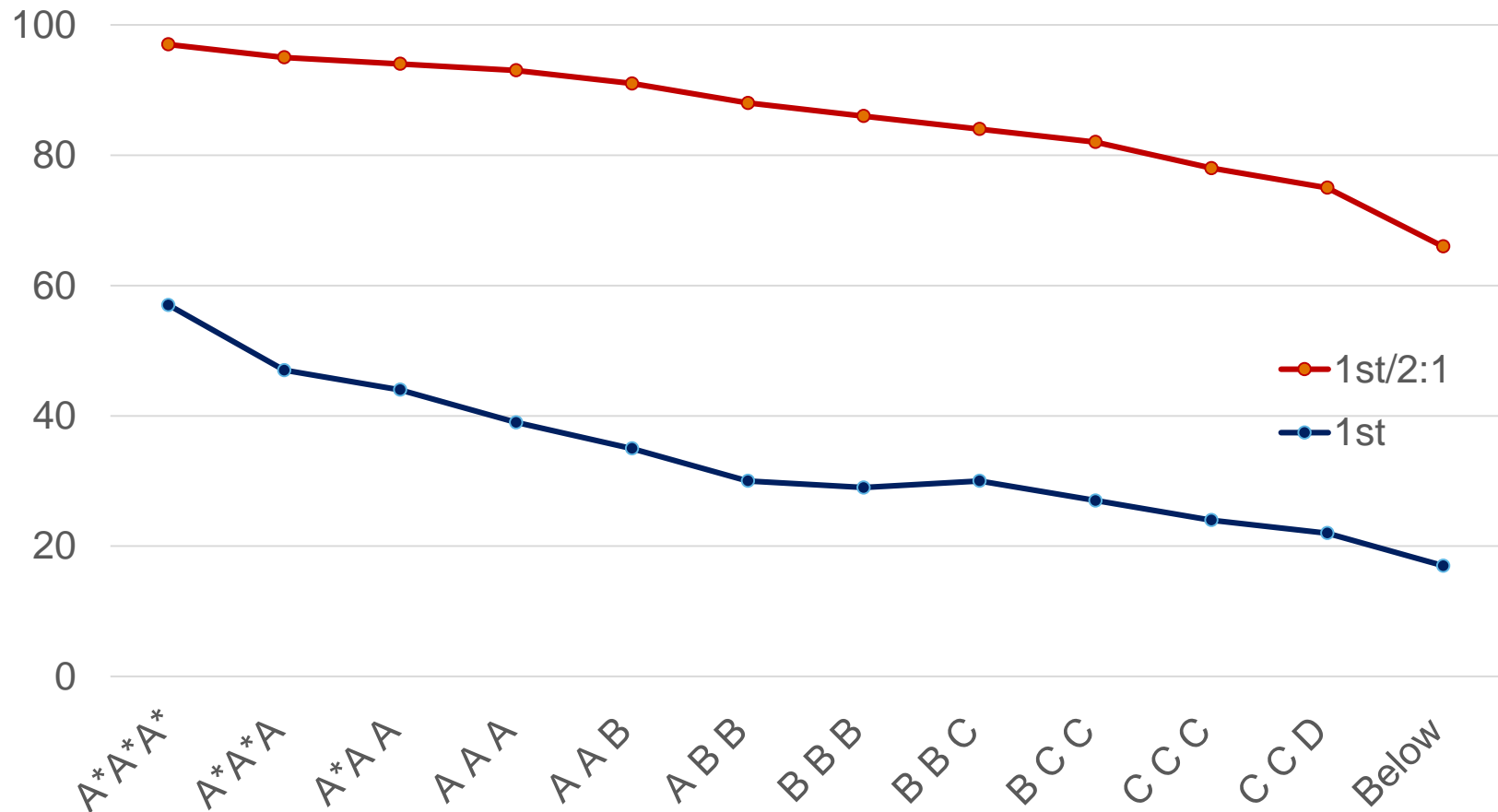
Context

Indicators of school academic achievement and intelligence

A levels and degree outcomes

How do A level grades impact chances of success?
www.officeforstudents.org.uk

Correlations with Grade Point Averages

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychological bulletin, 138(2), 353.

	N	k	r	95% CI	ρ
HS GPA	34,724	46	0.40	[0.35, 0.45]	0.41
SAT	22,289	29	0.29	[0.25, 0.33]	0.33
ACT	31,971	21	0.40	[0.33, 0.46]	0.40
A level points	933	4	0.25	[0.12, 0.38]	0.31
Intelligence	7,820	35	0.20	[0.16, 0.24]	0.21

Academic language proficiency

Correlations with Grade Point Averages

Predictive power of IELTS & TOEFL is limited – from no correlation to 0.4 for overall scores – but studies generally small-scale, local

Exceptions include

TOEFL

Cho and Bridgeman 2012   2,594 students: 10 US universities

Harsch, Ushioda & Ladroue 2017   483 students: UK

IELTS

Oliver, Vanderford & Grote 2012   353 students: Western Australia

Thorpe, Snell, Davey-Evans & Talman 2017   4,342 students: UK

Academic language proficiency

Correlations with Grade Point Averages

Cho, Y., & Bridgeman, B. (2012). Relationship of TOEFL iBT® scores to academic performance: Some evidence from American universities. *Language Testing*, *29*(3), 421-442.

Harsch, C., Ushioda, E., & Ladroue, C. (2017). Investigating the Predictive Validity of TOEFL iBT® Test Scores and Their Use in Informing Policy in a United Kingdom University Setting. *ETS Research Report Series*, *2017*(1), 1-80.

Oliver, R., Vanderford, S., & Grote, E. (2012). Evidence of English language proficiency and academic achievement of non-English-speaking background students. *Higher Education Research & Development*, *31*(4), 541-555.

Thorpe, A., Snell, M., Davey-Evans, S., & Talman, R. (2017). Improving the Academic Performance of Non-native English-Speaking Students: the Contribution of Pre-sessional English Language Programmes. *Higher Education Quarterly*, *71*(1), 5-32.
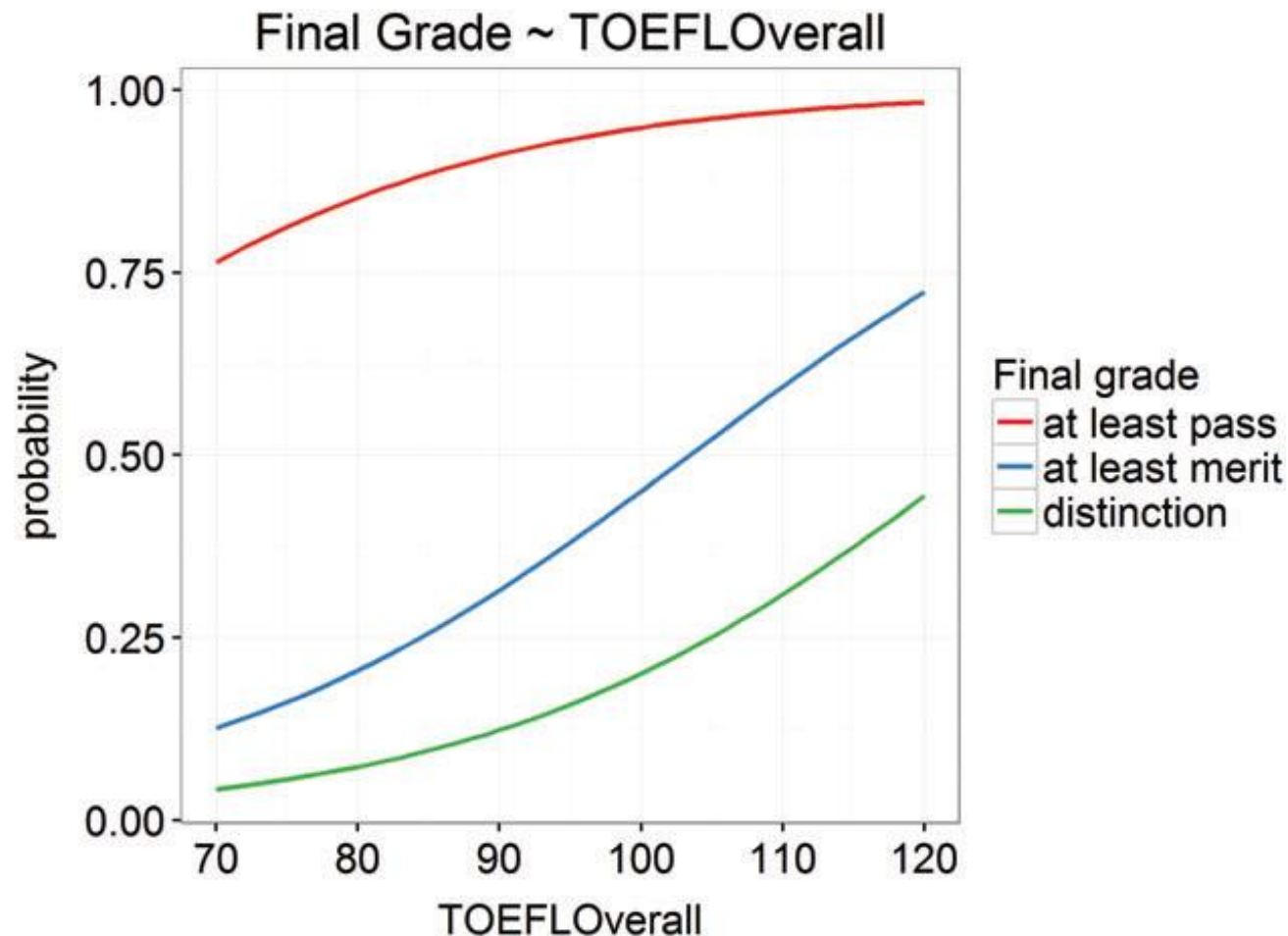
Academic language proficiency
Correlations with Grade Point Averages

University of
Bedfordshire

- Relationships between language tests and academic success vary across institutions, subject and country of student origin

- UG and PG NNES typically underperform compared to English-speaking students (with exceptions: Europe, Americas)

- Reading (& sometimes Writing) usually more closely related to academic performance than Listening or Speaking

- Students entering on the basis of low scores followed by pre-sessional English underperform compared to direct entrants

Academic language proficiency

Probability of degree outcomes: Harsch, Ushioda & Ladroue 2017

Final Grade ~ TOEFLOverall

Final grade
- at least pass
- at least merit
- distinction

Academic language proficiency
Shifting focus from GPA to language performance

- Paul, A. (2007). IELTS as a predictor of academic language performance, part 2. *IELTS Research Reports, 7.*

IELTS predicted language use in university work, but language deteriorates as cognitive task demands increase in university work

- Rea-Dickins, P., Kiely, R. & Yu, G. (2007) Student identity, learning and progression: The affective and academic impact of IELTS on 'successful candidates. *IELTS Research Reports, 7,* 59-136.

University tasks are more demanding than IELTS tasks: reading and writing identified as major sources of difficulty

Interposing test scores

- role of tests in admissions and their relationship to academic language use poorly understood by users

- both admissions staff and learners may understand a minimum test score as implying no problems with language post admissions

- language teachers can find it difficult to balance test practice, skills development and academic study skills

Ginther, A., & Elder, C. (2014). A comparative investigation into understandings and uses of TOEFL iBT®, IELTS, and PTE-A for Graduate Admissions in the United States and Australia. *ETS Research Report Series*, *2014*(2), 1-39.

Badger, R., and Yan X. "To what extent is communicative language teaching a feature of IELTS classes in China?." IELTS Research Reports Volume 13, 2012 (2012): 1.

Linking language examinations to the CEFR

Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual. (Council of Europe 2009)

Aims to 'make the results of the examination in question more transparent to both the users of examination results and to test takers themselves' (p.26)

Standard setting

Standard setting = "translation of policy to numerical cut score" (Reckase 2009)

1. **content standard**: the language abilities needed to support academic study: **Specification**

2. **performance standard**: the score (or constellation of scores) on a given test that best corresponds to the level of ability required for success: **Benchmarking/ standard setting**

Justifying links to the CEFR

The European Association for Language Testing and Assessment (EALTA) (2006) 10 requirements for defensible linking:

What evidence is there of the quality of the process followed to link tests and examinations to the Common European Framework?

3. Have the test content and the test specifications been analysed in relation to the CEFR descriptors?

8. What standard setting procedures have been used to establish cut-off scores for the relevant CEFR level(s)?

References for bilateral comparisons

- Council of Europe (2009). Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). Strasbourg: CoE.

- PTE Academic Score Guide November 2012 Version 4

  - pearsonpte.com

- Linking TOEFL iBT™ Scores to IELTS® Scores – A Research Report

  - ets.org

- Lim, G. S., Geranpayeh, A., Khalifa, H., & Buckendahl, C. W. (2013). Standard setting to an international reference framework: Implications for theory and practice. *International Journal of Testing*, *13*(1), 32-49.

Content

Linking language examinations to the CEFR's descriptive scheme

Logical fallacies in linking
Language abilities are not unitary

A learner needs at least a C1 level of language to follow a course in law

$$\neq$$

A learner with a certificate showing a C1 level of language has the *right* language skills to follow a course in law

Can express him/herself fluently and spontaneously without much obvious searching for expressions.

$$\neq$$

Can understand a wide range of demanding, longer texts, and recognise implicit meaning.

Linking language examinations to the CEFR's descriptive scheme

Contexts for language use

- *Domains*

- *Communication themes*

- *Situations*

- *Conditions and constraints*

Communicative language abilities

- *Activities*

- *Competences*

- *Strategies*

Contexts for language use: Domains
spheres of action or areas of concern in which social life is organised

- the **personal** domain, in which the person concerned lives as a private individual, centred on home life with family and friends, and engages in individual practices such as reading for pleasure, keeping a personal diary, pursuing a special interest or hobby, etc.;

- the **public** domain, in which the person concerned acts as a member of the general public, or of some organisation, and is engaged in transactions of various kinds for a variety of purposes;

- the **occupational** domain, in which the person concerned is engaged in his or her job or profession;

- the **educational** domain, in which the person concerned is engaged in organised learning, especially (but not necessarily) within an educational institution.

Contexts for language use: Communication themes

topics which are the focus of attention in particular communicative acts

1. personal identification
2. house and home, environment
3. daily life
4. free time, entertainment

    4.1 leisure

    4.2 hobbies and interests

    4.3 radio and TV

    4.4 cinema, theatre, concert, etc.

    4.5 exhibitions, museums, etc.

    4.6 intellectual and artistic pursuits

    4.7 sports

    4.8 press

5. travel
6. relations with other people
7. health and body care
8. education
9. shopping
10. food and drink
11. services
12. places
13. language
14. weather

Contexts for language use: Situations

external situations which arise may be described in terms of…

1. the locations in which, and the times at which, they occur;

2. the institutions – the structure and procedures of which control much of what can normally occur;

3. the persons involved, especially in their relevant social roles in relation to the user/learner;

4. the objects (animate and inanimate) in the environment;

5. the events that take place;

6. the operations performed by the persons involved;

7. the texts encountered within the situation

Contexts for language use: Situations
Table 5. External contexts of use

Domain	Locations	Institutions	Persons	Objects	Events	Operations	Texts
Personal	Home: : house, rooms, garden own of family of friends of strangers Own space in hostel, hotel The countryside, seaside, etc.	The Family Social networks	(Grand)Parents, Offspring, Siblings, Aunts, Uncles, Cousins, In-laws, Spouses Intimates, Friends, Acquaintances	Furnishing & furniture Clothing Household equipment Toys, tools, personal hygiene Objets d'art, Books, Pets, Wild /domestic animals, Trees, Plants, Lawn, Ponds, Household goods, Handbags, leisure/sports equipment	Family occasions Encounters Incidents, accidents, Natural phenomena Parties, visits Walking, cycling, motoring Holidays, excursions Sports events	Living routines (dressing, undressing cooking, eating, washing, etc.) DIY, gardening Reading, Radio & TV Entertaining Hobbies Games & sports	Teletext Guarantees Recipes Instructional material Novels, magazines, Newspapers Junk mail Brochures Personal letters Broadcast and recorded spoken texts
Public	Public spaces: street, square, park, etc. Public transport Shops (super)markets Hospitals, surgeries, clinic Sports stadia, fields, halls Theatre, cinema, entertainment Restaurant, pub, hotel Places of worship	Public authorities Political bodies The law Public Health Services clubs Societies Political parties Denominations	Members of the Public Officials Shop personnel Police, army, security Drivers, conductors, Passengers Players, fans, spectators Actors, audiences Waiters, barpersons Receptionists Priests, Congregation	Money, purse, wallet Forms, Goods Weapons Rucksacks Cases, Grips Balls Programmes Meals, Drinks, Snacks Passports, Licences	Incidents Accidents, Illnesses Public meetings Law-suits, Court trials Rag-days, Fines, Arrests Matches, contests Performances Weddings, Funerals	Buying and obtaining public services Using medical services Journeys by road/rails/ship/air Public entertainment and leisure activities Religious services	Public announcements and notices Labels & packaging Leaflets, Graffiti Tickets, Timetables Notices, Regulations Programmes Contracts Menus Sacred texts Sermons, Hymns
Occupational	Offices Factories Workshops Ports, railways Farms Airports Stores, shops, etc. Service industries Hotels	Firms Civil Service Multinational Corporations Nationalised industries Trade Unions	Employers/ees Managers Colleagues Subordinates Workmates Clients Customers Receptionists, Secretaries Cleaners, etc.	Business machinery Industrial machinery Industrial & craft tools	Meetings Interviews Receptions Conferences Trade fairs Consultations Seasonal sales Industrial accidents Industrial disputes	Business admin. Industrial management Production operations Office procedures Trucking Sales operations Selling, marketing Computer operation Works office Maintenance	Business letter Report Memorandum Life & safety notices Instructional manuals Regulations Advertising material Labelling & packaging Job description Sign posting Visiting cards, etc.
Educational	Schools: Hall Classrooms, Playground, Sports fields, corridors Colleges Universities Lecture Theatres Seminar rooms Student Union Halls of Residence Laboratories Canteen	School College University Learned societies Professional Institutions Adult education bodies	Class teachers Teaching staff Caretakers Assistant staff Parents Classmates Professors, lecturers (Fellow) Students Library & laboratory staff Refectory staff, cleaners Porters, Secretaries, etc.	Writing material School uniforms Games equipment & clothing Food Audio-visual equipment Black-board & chalk Computers Briefcases & School bags	Return to school/entry Breaking up Visits and Exchanges Parents' days / evenings Sports days, Matches Disciplinary problems	Assembly Lessons Games Playtime Clubs & societies Lectures, Essay writing Laboratory work Library work Seminars & tutorials homework Debates & discussions	Authentic texts (as above) Textbooks, Readers Reference books Blackboard text OP text Computer screen text Videotext Exercise materials Journal articles Abstracts Dictionaries

Situations

Table 5. External contexts of use

Domain	Locations	Institutions	Persons
Educational	Schools: Hall Classrooms, Playground, Sports fields, corridors Colleges Universities Lecture Theatres Seminar rooms Student Union Halls of Residence Laboratories Canteen	School College University Learned societies Professional Institutions Adult education bodies	Class teachers Teaching staff Caretakers Assistant staff Parents Classmates Professors, lecturers (Fellow) Students Library & laboratory staff Refectory staff, cleaners Porters, Secretaries, etc.

Situations

Table 5. External contexts of use

Objects	Events	Operations	Texts
Writing material School uniforms Games equipment & clothing Food Audio-visual equipment Black-board & chalk Computers Briefcases & School bags	Return to school/entry Breaking up Visits and Exchanges Parents' days / evenings Sports days, Matches Disciplinary problems	Assembly Lessons Games Playtime Clubs & societies Lectures, Essay writing Laboratory work Library work Seminars & tutorials homework Debates & discussions	Authentic texts Textbooks, Readers Reference books Blackboard text OP text Computer screen text Videotext Exercise materials Journal articles Abstracts Dictionaries

Conditions and constraints

imposed by external conditions under which communication occurs

Physical conditions:

a) for speech:

　clarity of pronunciation;

　ambient noise (trains, aircraft, 'static', etc.);

　interference (crowded street, markets, pubs, parties, discos, etc.);

　distortions (poor telephone lines, radio reception, public address systems);

　weather conditions (wind, extreme cold, etc.).

b) for writing:

　poor reproduction of print;

　difficult handwriting;

　poor lighting, etc.

Social conditions:

　number and familiarity of interlocutors;

　relative status of participants;

　presence/absence of audience;

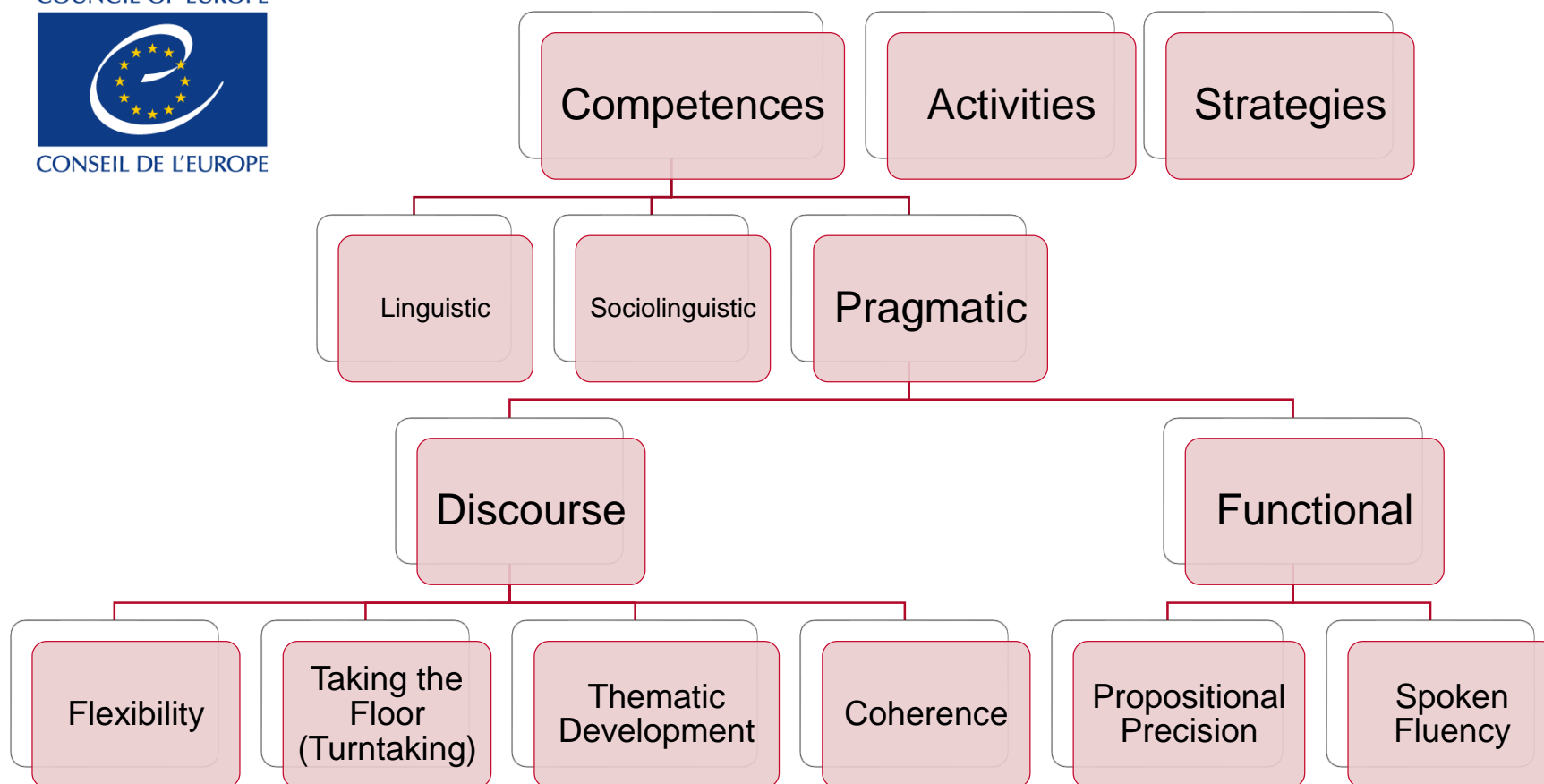　social relationships between participants.

Time pressures:

　different pressures for speaker/listener and writer/reader;

　preparation time for speeches, reports, etc.;

　limitations on time allowed for turns and interactions (e.g. by rules, expense, competing events and commitments, etc.);

Other pressures:

　financial; anxiety-producing situations, etc.

Communicative language abilities:
Illustrative scales
Three perspectives on communicative language use

University of Bedfordshire

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

- Competences
  - Linguistic
  - Sociolinguistic
  - Pragmatic
    - Discourse
      - Flexibility
      - Taking the Floor (Turntaking)
      - Thematic Development
      - Coherence
    - Functional
      - Propositional Precision
      - Spoken Fluency
- Activities
- Strategies

Refection on when, where, why a language is used (CEFR p.44)

University of Bedfordshire

What skills will they need to have developed? How can they still be themselves without being misinterpreted?

What knowledge of the world or of another culture will they need to call on?

Under what conditions will they have to act?

What sort of things will they be listening to or reading?

What tasks will they have to accomplish?

What themes will they need to handle?

Will they have to speak, or simply listen and read with understanding?

What objects will they need to refer to?

What sort of people will they have to deal with?

Specification

Recommended procedures for capturing test content in the CoE manual

- Specification forms

  o 67 pages of forms (before completion)

  o 24 Forms: A1 to A24

  o 3 Grids for analysis

  o 3 Grids for presentation of outcomes

- Graphical profile

  o Form A23
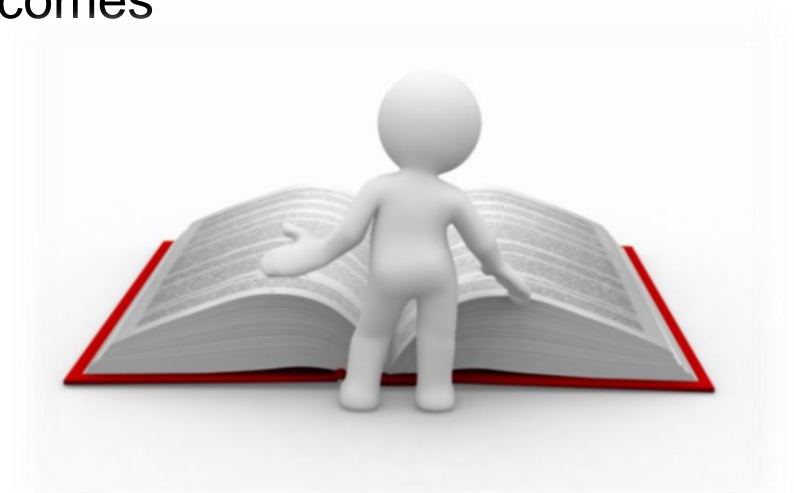
Presentation grids

Not widely used in presenting test content

Outcomes not readily comparable

Standards for identifying 'links' are vague

Unclear from the forms what the test taker is asked to do

The CEFR Grid for Speaking Tasks

1 Report on analysis of
2 Target language

1 GENERAL INFORMATION (whole speaking test)

3 N°. of tasks in speaking component
4 Integration of skills
5 Total duration of speaking component
6 Target performance level
7 Channel
8 Test purpose

2 TASK INPUT/PROMPT for task n°./name

9 Language of instructions/rubric
10 Channel
11 Language level of instructions/rubric
12 Task duration (minutes)
13 N°. assessors present
14 Recorded?
15 Control/guidance by task
16 Control/guidance by interlocutor
17 Specification of content
18 Interaction type
19 Discourse mode (genre)
20 Audience (real)
21 Audience (imagined, as in role play)
22 Type of prompt
23 Topic
24 Planning time
25 Setting (imagined)

3 RESPONSE (the expected spoken response elicited by the prompt(s)/input)

26 Length of response
27 Text type
28 Rhetorical function(s)
29 Register
30 Domain
31 Grammatical level
32 Lexical level
33 Discourse features
34 Situational authenticity
35 Interactional authenticity
36 Cognitive processing
37 Content knowledge
38 Task purpose

4 RATING OF TASK

39 Known criteria
40 Rating method
41 Assessment criteria
42 N°. of raters
43 Use of moderator

5 FEEDBACK TO

44 Quantitative fee
45 Qualitative fee

COUNCIL   CONSEIL
OF EUROPE   DE L'EUROPE

Graphic profile
Presents the outcome of the specification process
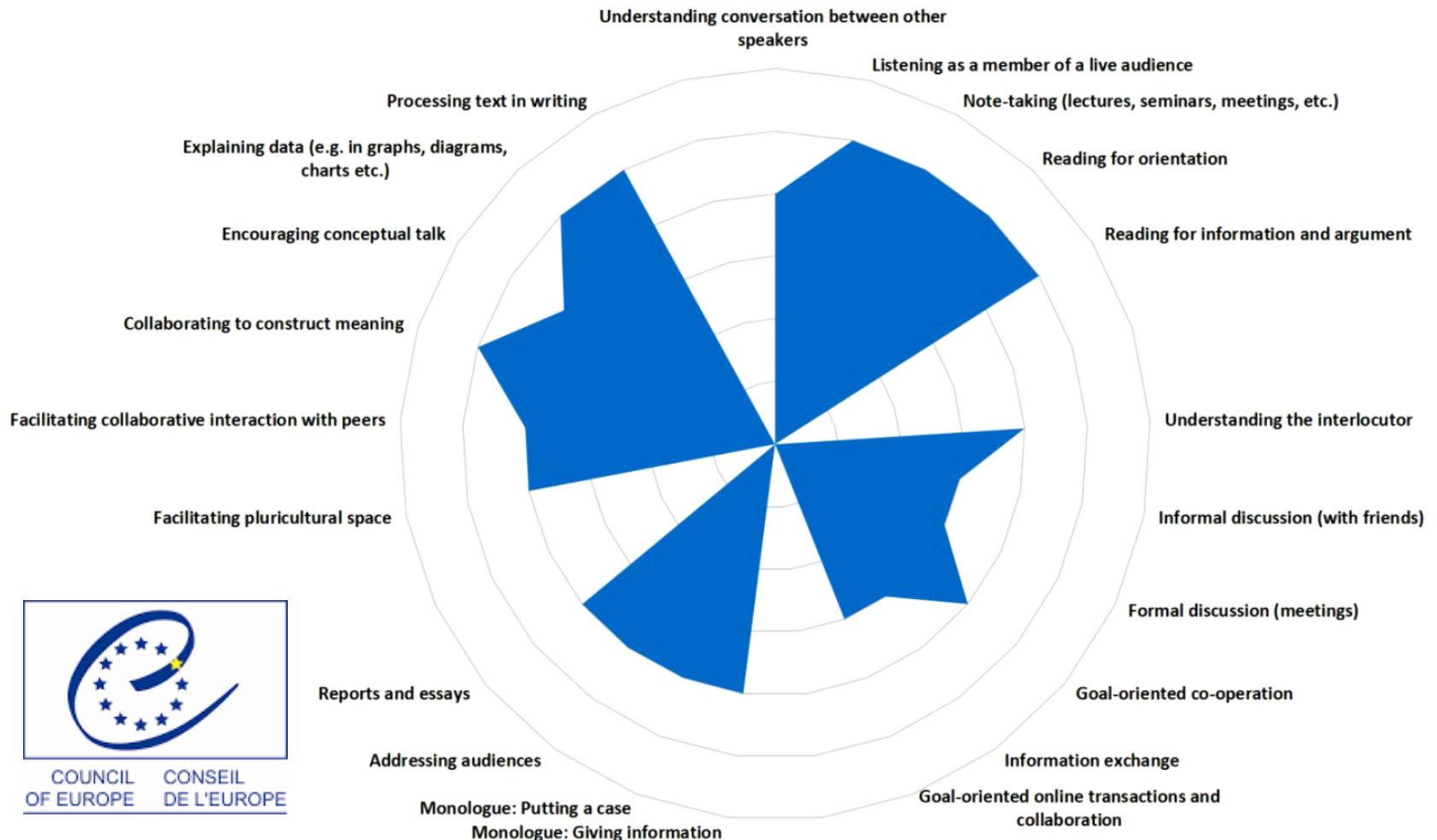
Content: How does the content coverage of the four tests relate to the CEFR?

Test providers do not use the Council of Europe (2009) profiling tools to communicate about the content of their tests

Comparing content

Taylor, L. & Chan, S. (2015).
IELTS Equivalence Research Project (GMC 133)
www.gmc-uk.org
Appendix 1

IELTS		Advanced		PTE-A		TOEFL iBT	
Task 1	Task 2	Part 1	Part 2	1.7	1.8	Task 1	Task 2
Graph/ table	Prompt	Short text (< 150 words)	Choose 1 of 4 tasks. Read text (< 80 word) (set reading text)	written text < 300 words	Prompt	Read a short passage, listen to related material (230-300 words each)	Prompt
Describe data	Discursive essay	Article/ report/ proposal/ letter	Article/ competition entry/ essay/ information sheet/ letter/ proposal/ report/ review	Summarize in one sentence	Discursive essay	Describe how the information relates	Discursive essay
20'	40'	90'		10'	20'	20'	30'
150	250	180-220	220-260	75	200-300	150-225	300+

Levels

Linking language examinations to the CEFR's common reference levels

Levels: test providers' claims

IELTS	CEFR level	TOEFL iBT	PTE-A	C1 Advanced	Cambridge English Scale
	C2	-	E80 A85 D95	Grade A	200–210
	C1	R24 L22 S25 W24 T**95**	E67 A76 D84	Grade B/C	180–199
	B2	R18 L17 S20 W17 T**72**	E51 A59 D75	Level B2	160–179

IELTS scale: 9, 8, 7.5, 7, 6.5, 6, 5.5, 5, 4.5, 4

B1

Outcomes of standard setting exercises and bilateral comparisons

Pearson

		PTE-A	IELTS	TOEFL iBT	Advanced
PTE (avg.)	B2	**59**	⚖ 6.5	87	
IELTS	B2	42	**5.5**	54	
ETS	B2	49	⚖ 6.0	**72**	

ETS

		PTE-A	IELTS	TOEFL iBT	Advanced
IELTS	B2		**5.5**	46	
ETS	B2		⚖ 6.0	**72**	

Cambridge

		PTE-A	IELTS	TOEFL iBT	Advanced
IELTS	C1		7.0/ high 6.5		**C**

Content: How does the content coverage of the four tests relate to the CEFR?

Test providers do use the Council of Europe (2009) tools to communicate about the level of their tests

But they

- use different approaches

- report in different ways

- disagree on the outcomes

Some conclusions

In the Higher Education context,

- The CEFR has not increased transparency of test content for score users, or facilitated comparisons between tests – this is needed

- Test providers have related tests to the CEFR, but outcomes reveal limited agreement

- Claimed CEFR links make a very poor basis for setting cut scores for entry – each test should be judged on its own merits:

  - Content: How well does this test reflect the language demands faced by students? How well is it likely to predict success in meeting those demands?

  - Level: What score on each test represents an acceptably high probability of success in meeting language demands?

The way ahead

Think bigger: join up the data

Look deeper: student needs in a time of change

Educate: help the university to understand the potential and limits of language tests