



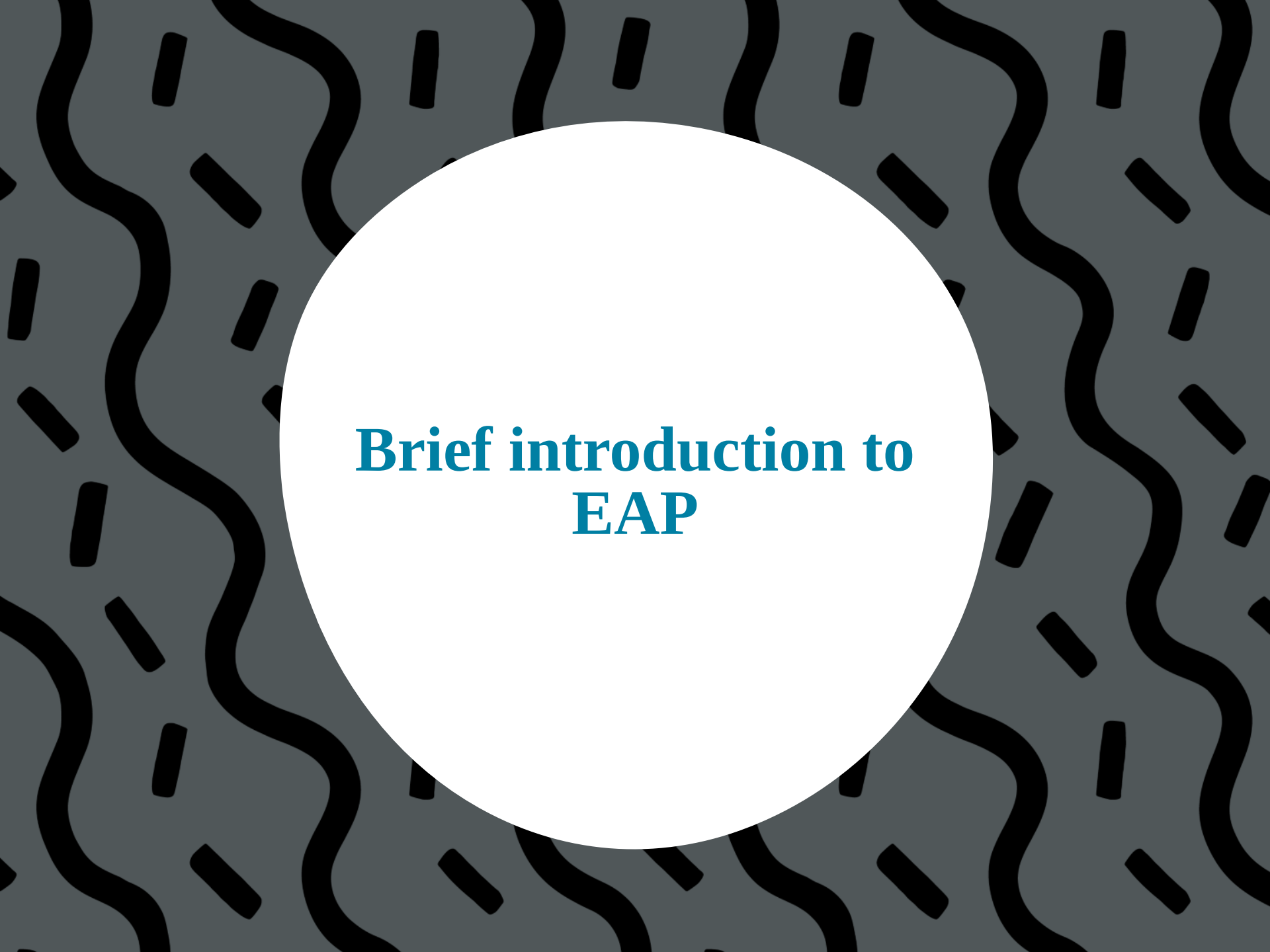
The CEFR in higher education: developing descriptors of academic English

Prof. John de Jong & Dr Veronica Benigno

Language Testing Forum 2016 - University of Reading

Outline

- Brief introduction to EAP
- The suitability of the CEFR in the academic context
- Developing descriptors of Academic English



Brief introduction to EAP

English for Academic Purposes

- EAP “refers to language research and instruction that focuses on the specific communicative needs and practices of particular groups in academic contexts” (*Hyland and Hamp-Lyons, 2002*)
- EAP is used in a variety of settings (in English-speaking countries or outside) and for different purposes (international or home students, at academic or secondary school level)
- Needs of mainstream students ignored by a “deficit-approach” to EAP (*Wingate and Tribble, 2012*)
 - Writing instruction for international students
 - Remedial writing courses, typically in learning support or study skills units

EAP and EGP

- EAP is different from EGP because of its more focused scope
 - Generally seen as a subcomponent of ESP (*Johns, 2013*)
 - Based on a need analysis (*Flowerdew & Peacock, 2001*)
 - Major focus is on study skills and specific tasks which are relevant to the academic context, not the language (*Alexander et al., 2008*)
 - Bias towards written skills, perhaps because written genres are seen as “more central to professional success” (*Feak, 2013*)
 - High stakes and limited time (*Alexander, 2012*)

Testing EAP and validity claims

- Weir (2005) calls for a test to elicit authentic processing (cognitive validity) and to accurately reflect the targeted language (content validity)
- In reality, the construct represented in tests is much more narrow and restricted than real language use:
 - Length of task
 - Background knowledge
 - Chances to reedit responses
- This may lead us to conclude that tasks should be relevant but they can never be authentic!

Using EAP descriptors in teaching and testing

- In teaching, EAP descriptors could help:
 - apply a systematic and standardized approach to curriculum development
 - bridge the gap between teaching and assessment
- In assessment, EAP descriptors could help:
 - produce more accurate test specifications to describe the content of the test
 - judge performance and progress in relation to scaled learning objectives and therefore give meaning to the scores

Back to the basics: the EAP construct

- Fulcher (1999) expressed concerns about the definition of the construct of EAP, particularly in relation to the degree of agreement between raters on what constitutes academic English
- The CEFR is used as the starting point to align tests to what is considered to be a valid standard of scaled proficiency – but how suitable are the CEFR descriptors for developing a test of academic English?

**The suitability of the
CEFR in the academic
context**

Limitations of the CEFR

- Uneven distribution of descriptors across the four skills (a large part (65%) of information is about Speaking)
- Lack of descriptors at the lower and higher levels (A1, C1, C2) and no information below A1
- Too wide levels (A1-C2) and opaque meaning of bands: what does it mean to be at a level?
- Very little information about professional/academic use of language

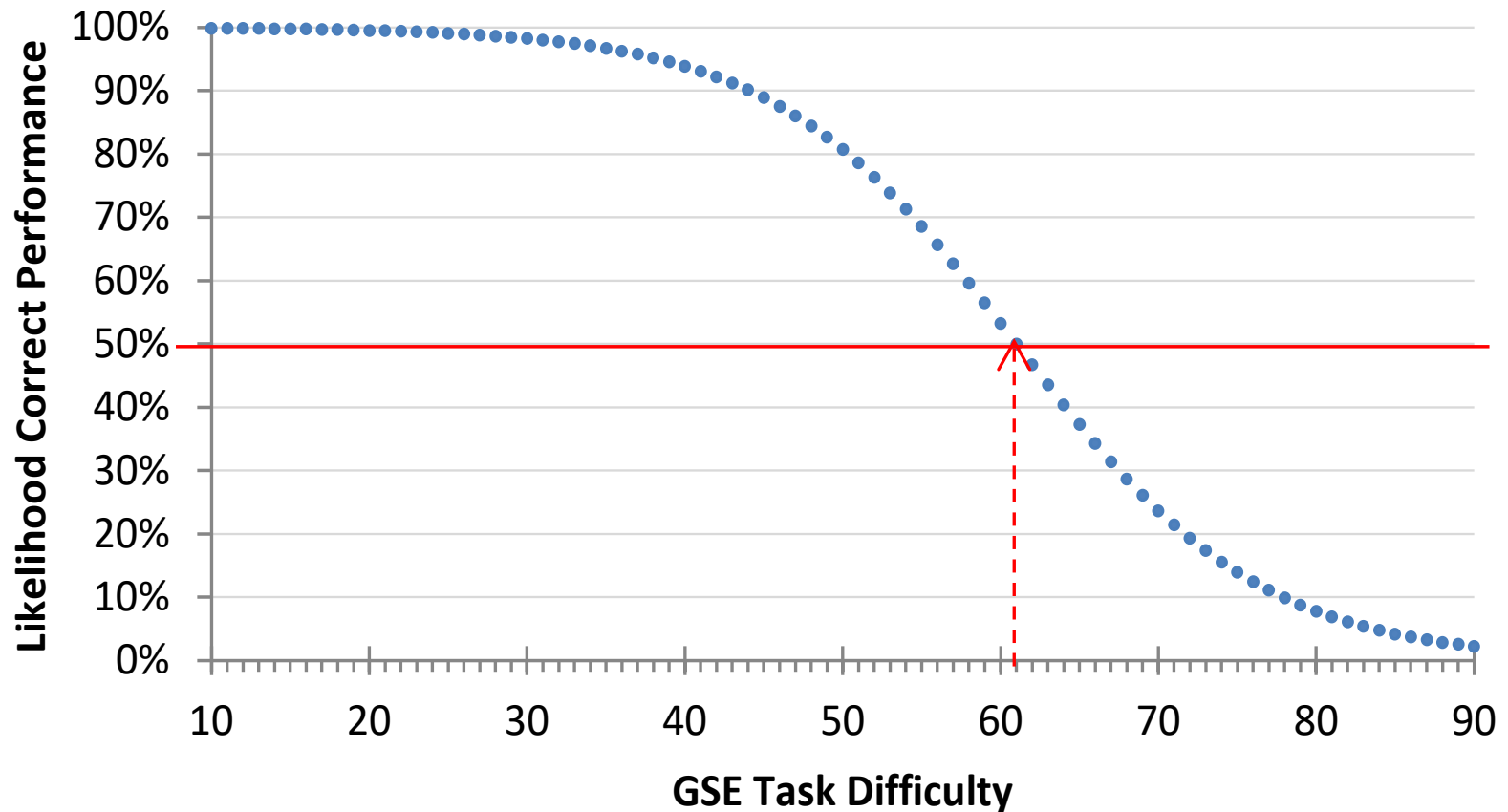
Probabilistic interpretation of “being at a level”

- Each CEFR descriptor was originally calibrated using the Rasch model
- In the Rasch model, when a person's ability is equal to the difficulty of the item, there is by definition a 0.5 probability of a correct response
- Therefore **being at a level means having a 50% probability** of being able to perform the entire set of descriptors defining that level

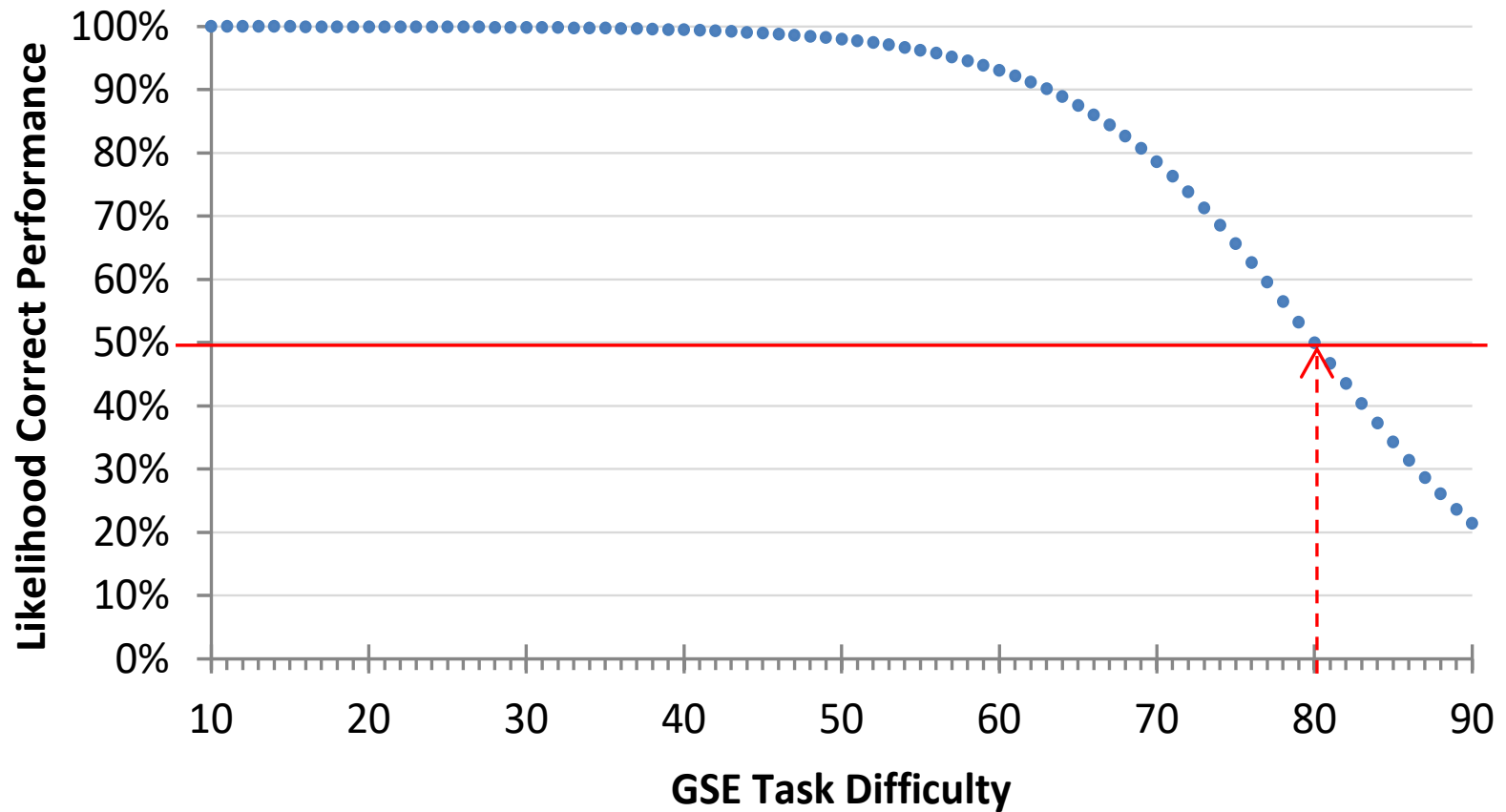
Table: Probability of success at CEF levels for learners at different levels – given on the original theta values identified by Brian North

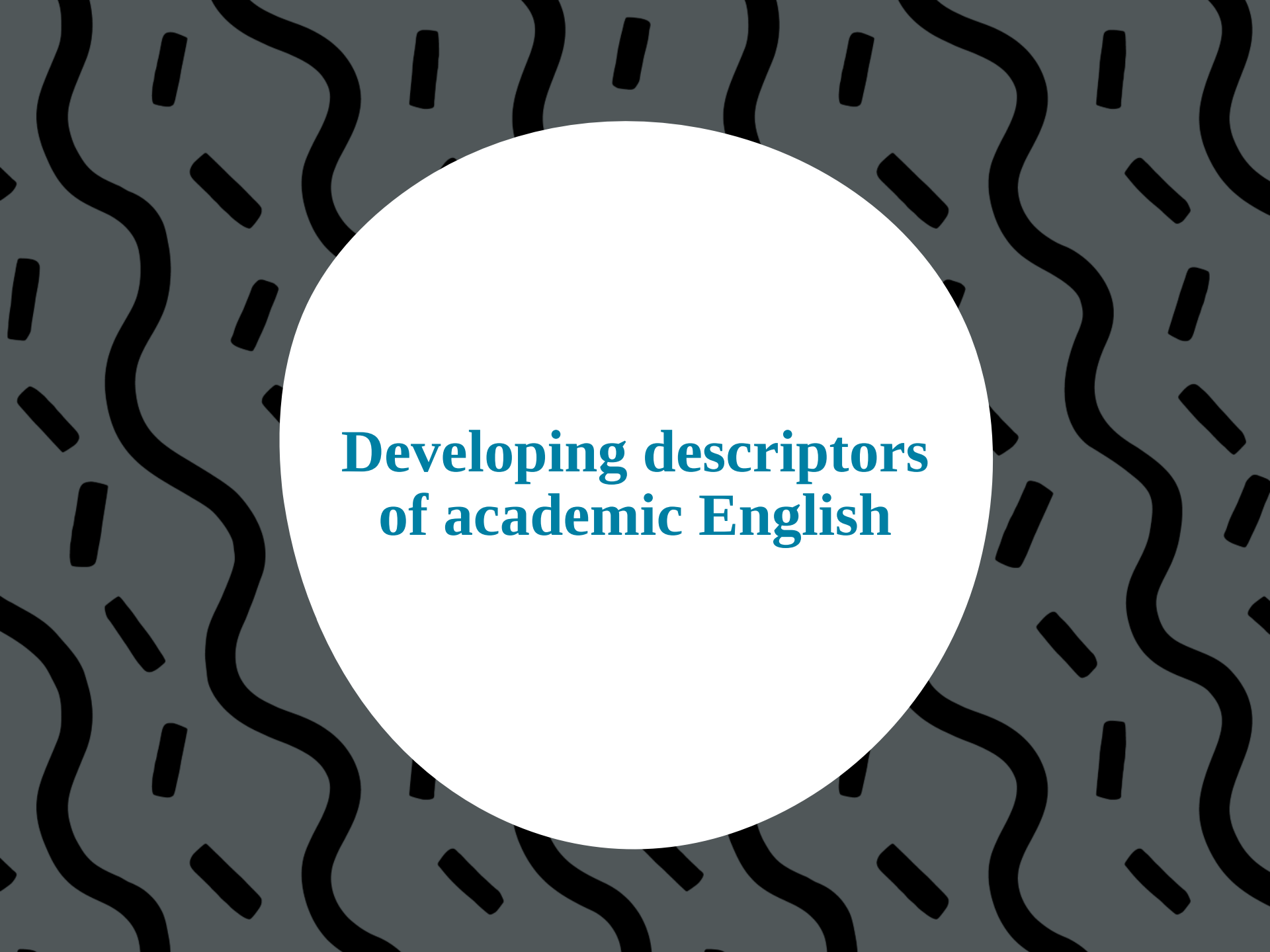
		Learners at level				
		A1	A2	B1	B2	C1
Tasks at level	C1	0.00	0.00	0.03	0.17	0.50
	B2	0.00	0.02	0.12	0.50	0.83
	B1	0.03	0.12	0.50	0.88	0.97
	A2	0.18	0.50	0.88	0.98	1.00
	A1	0.50	0.82	0.97	1.00	1.00

A learner at 61 on GSE



A learner at 80 on GSE





Developing descriptors of academic English

Rationale

- The CEFR contextualises language activities in four domains: the public domain, the personal domain, the educational domain and the occupational domain” (CoE, 2001, p. 14).
- In creating specialist learning objectives, we are addressing the different language needs that arise in different domains.
- The GSE Learning Objectives for Academic English address the needs of learners in the educational domain, with a focus on academic study at the tertiary/post-secondary level.
- All learners need to acquire a core of proficiency in English, and for that reason, the GSE Learning Objectives for Academic English include the learning objectives developed for adult learners of general English.

The CEFR and academic language needs

- Can understand recordings in standard dialect likely to be encountered in social, professional or **academic** life and identify speaker viewpoints and attitudes as well as the information content.
- Can follow the essentials of lectures, talks and reports and other forms of **academic**/professional presentation which are propositionally and linguistically complex.

The majority (8) are placed at level B2 and a minority (5) at level C1, none at Level C2.

In the ALTE appendix descriptors mentioning ‘academic’ are at levels 4 and 5, claimed to be at C1 and C2 respectively.

Sourcing Learning Objectives

The GSE Learning Objectives were mainly sourced from:

- British and American English course materials (e.g. North Star, Academic Connections, Language Leader)
- Syllabuses from various Ministries of Education

In the published set they are coded - to show their origin

- (C) Common European Framework descriptor, verbatim, © Council of Europe
- (Ca) Common European Framework descriptor, adapted or edited, © Council of Europe
- (N2000) North (2000) descriptor, verbatim
- (N2000a) North (2000) descriptor, adapted or edited
- (N2007a) North (2007) expanded set of C1 and C2 descriptors, adapted or edited
- (P) New Pearson descriptor

Creating Learning Objectives

Descriptors created in conformity with specific guidelines:

- Performance: the language performance itself (e.g. *Can answer the telephone*): **what** someone can do
- Criteria: the intrinsic quality of the performance (e.g. *using a limited range of basic vocabulary*): **how well** someone can do something
- Conditions: any extrinsic constraints or conditions defining the performance (e.g. *with support, if spoken slowly and clearly, etc.*)

Rating Learning Objectives

- Descriptors were distributed over 4 batches in an overlapping fashion
- Batches were evenly attributed to groups of raters
- Within batches descriptors were grouped by skill, but a) the order of the skills and b) the order of descriptors within each skill was randomised so that each batch was presented differently

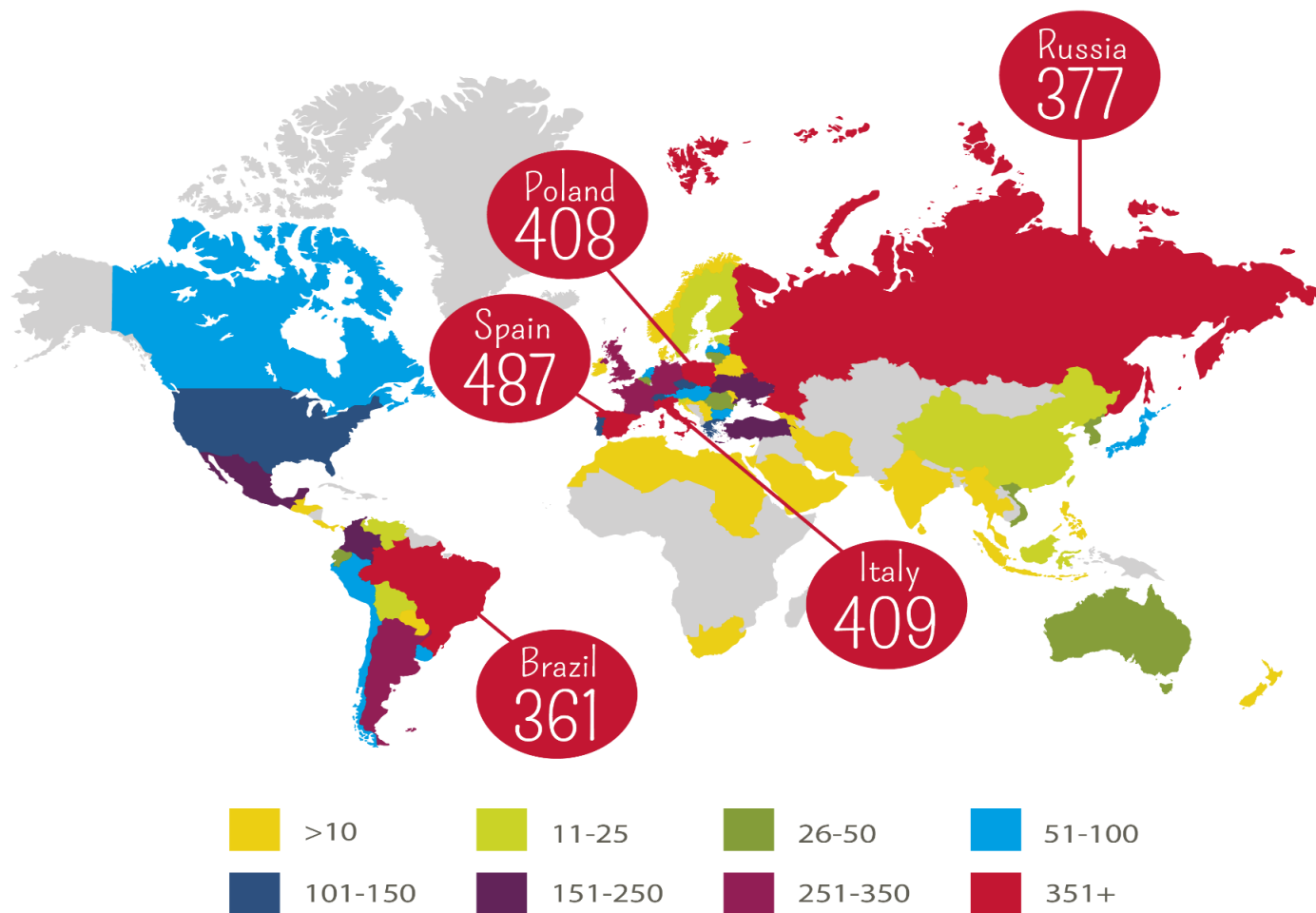
Rating Learning Objectives

- Expert raters (around 80 -100 per batch): at least 2 years teaching experience with target learner group, attended training workshops, rated on both the CEFR and the GSE
- Online raters (N=558) with at least 2 years teaching experience, attended online training, rated on the CEFR only

Procedure

- Rating of 100-120 descriptors for expert raters (2-3 hours); rating of 30-40 descriptors for online raters (1 hour)
- At least 10 anchors (mainly from North, 2000) were required.

Global research: over 6,000 teachers across 50 countries

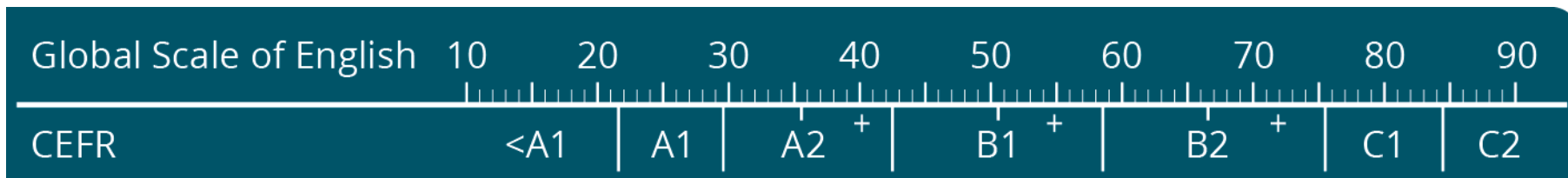


Sample training task (for experts)

SAMPLE DESCRIPTOR

Can leave phone messages containing detailed information.

- First, decide which CEFR level you think a student would need to be at in order to have a 50% chance of successfully completing the task
- Then, decide where within that level you think the learning objective sits. Use the GSE value to indicate start, middle or end
- Answer, e.g.: B1(+) [=GSE 51-58] → 54



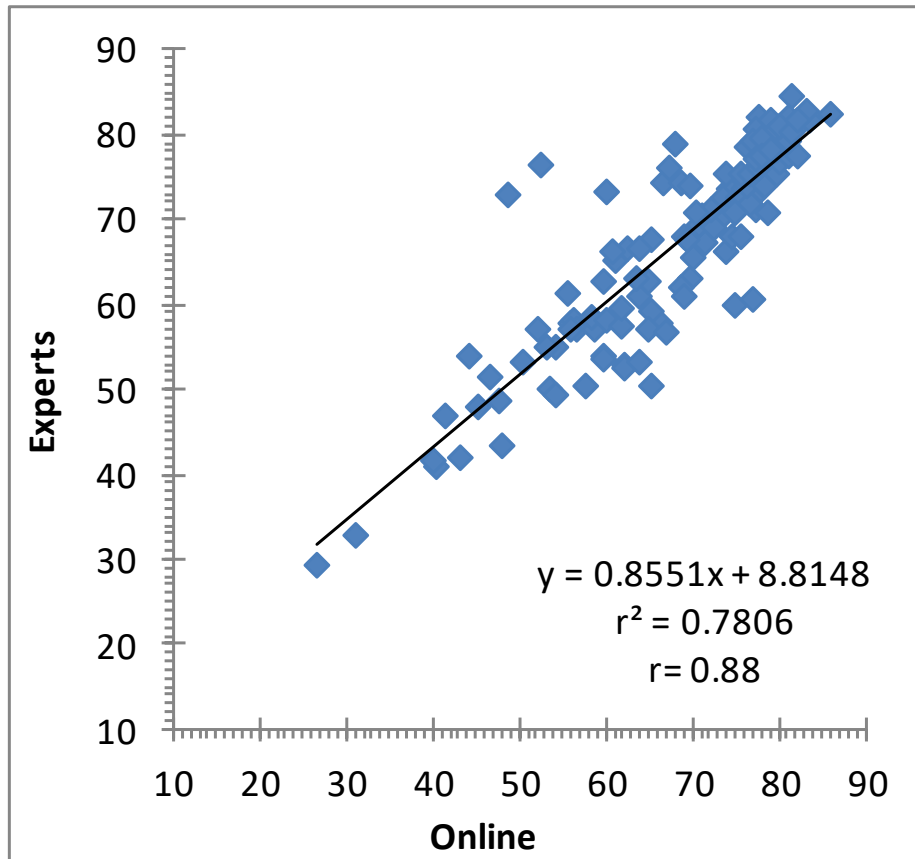
Calibrating Learning Objectives

Statistical analysis and data cleaning to remove

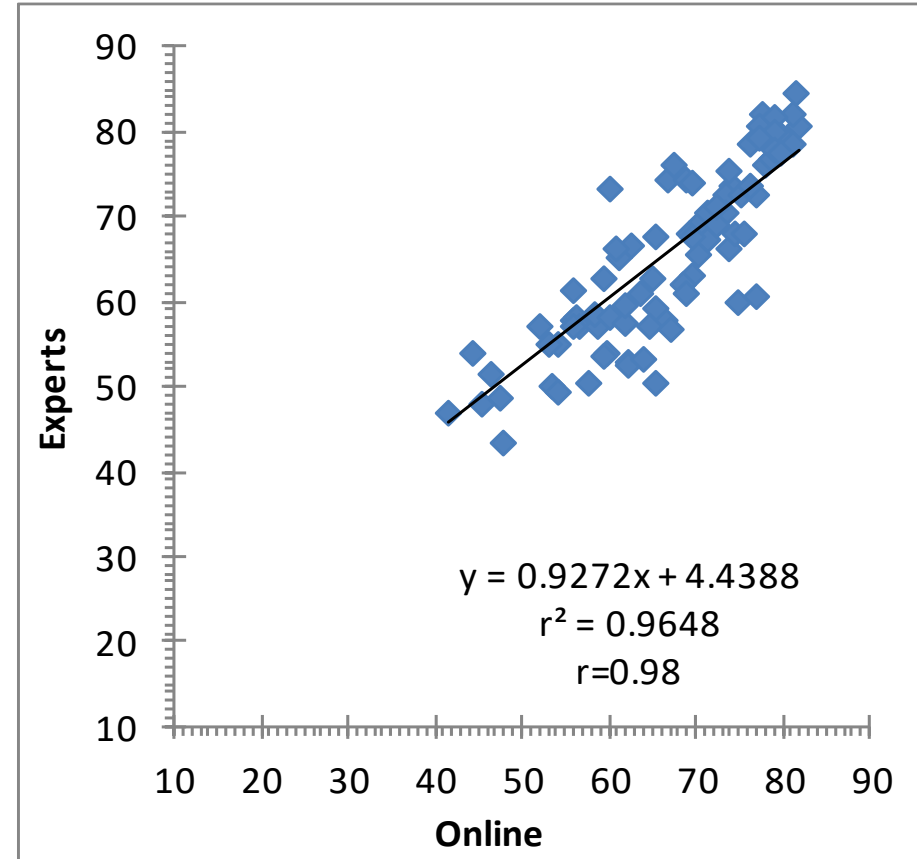
- “Misfitting” raters, e.g. raters who rated less than 75% descriptors or with a low SD indicating they did not use the full range of scale or with an average score which was too distant from the group average
- “Poor” Learning Objectives, e.g. descriptors with an overall degree of rater agreement $< .70$
- **The final GSE values showed a 0.99 correlation with values based on the classical analyses by North (2000)**

Agreement between Expert and Online raters

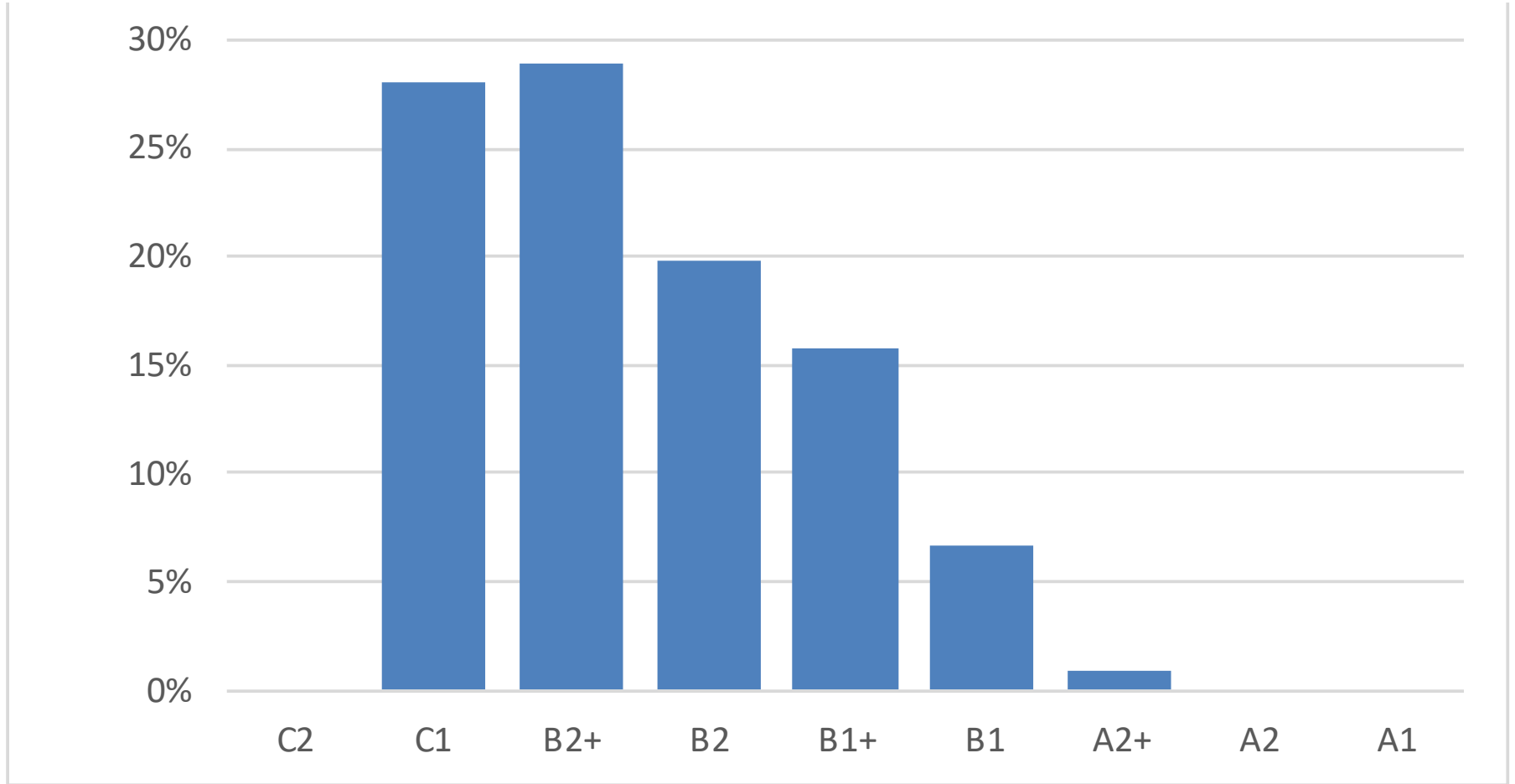
Prior to data cleaning



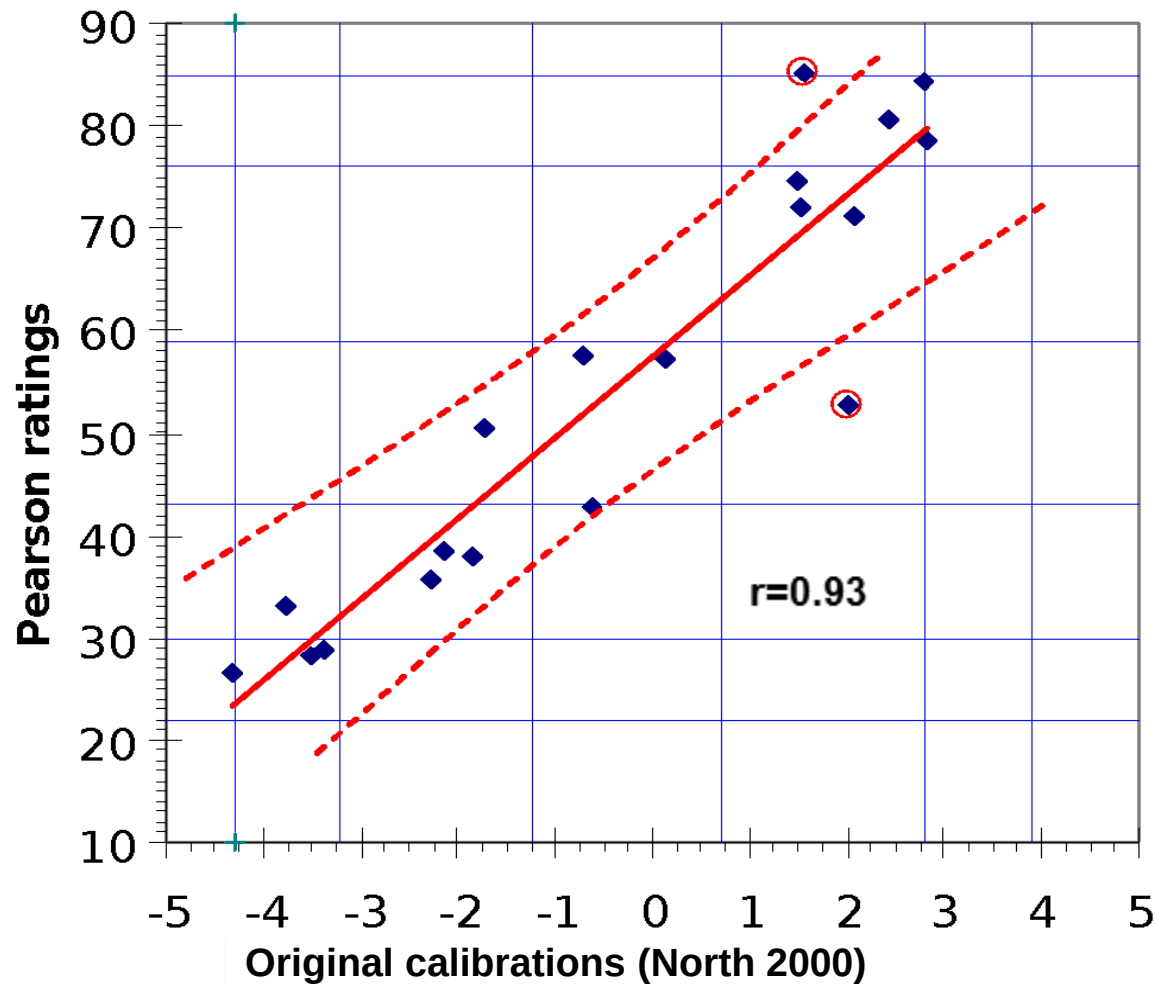
After data cleaning



Distribution of Academic Learning Objectives



Compare 19 anchor items



Example Academic Descriptors

S	Descriptor	GSE	CEFR	Agree	z-dif
L	Can recognise markers that introduce supporting examples.	57	B1+	0.87	0.23
L	Can recognise markers that signal the main parts of a lecture.	57	B1+	0.85	-0.16
R	Can evaluate information in an academic text using specific criteria.	75	B2+	0.91	0.80
R	Can identify examples from an academic text to support an argument.	69	B2+	0.85	-1.46
S	Can effectively discuss the meaning and implications of research data.	80	C1	0.85	-0.26
S	Can effectively request information from a professor outside of class.	58	B1+	0.77	0.45
W	Can use appropriate tone and register when writing academic texts.	78	C1	0.89	0.38
W	Can synthesise information from two or more academic texts.	78	C1	0.76	-0.58

Learning Objectives for Academic English

- In collaboration with teachers, ELT authors and language experts from around the world
 - ✓ Created on the CEFR model (North, 2000)
 - ✓ Describe performance abilities in academic English
 - ✓ Can be used to create the level-appropriate content syllabus for examinations and to identify assessment criteria



There's so much more to learn

Find out more about us at
[English.com/gse](https://www.english.com/gse)

For any queries about the GSE
Teacher Toolkit, please contact
Veronica Benigno at
veronica.benigno@pearson.com

ALWAYS LEARNING