

How prejudices affect rating (case study Slovene)

8 November 2019



Fairness

Slovene

Research

Petra Likar Stanovnik

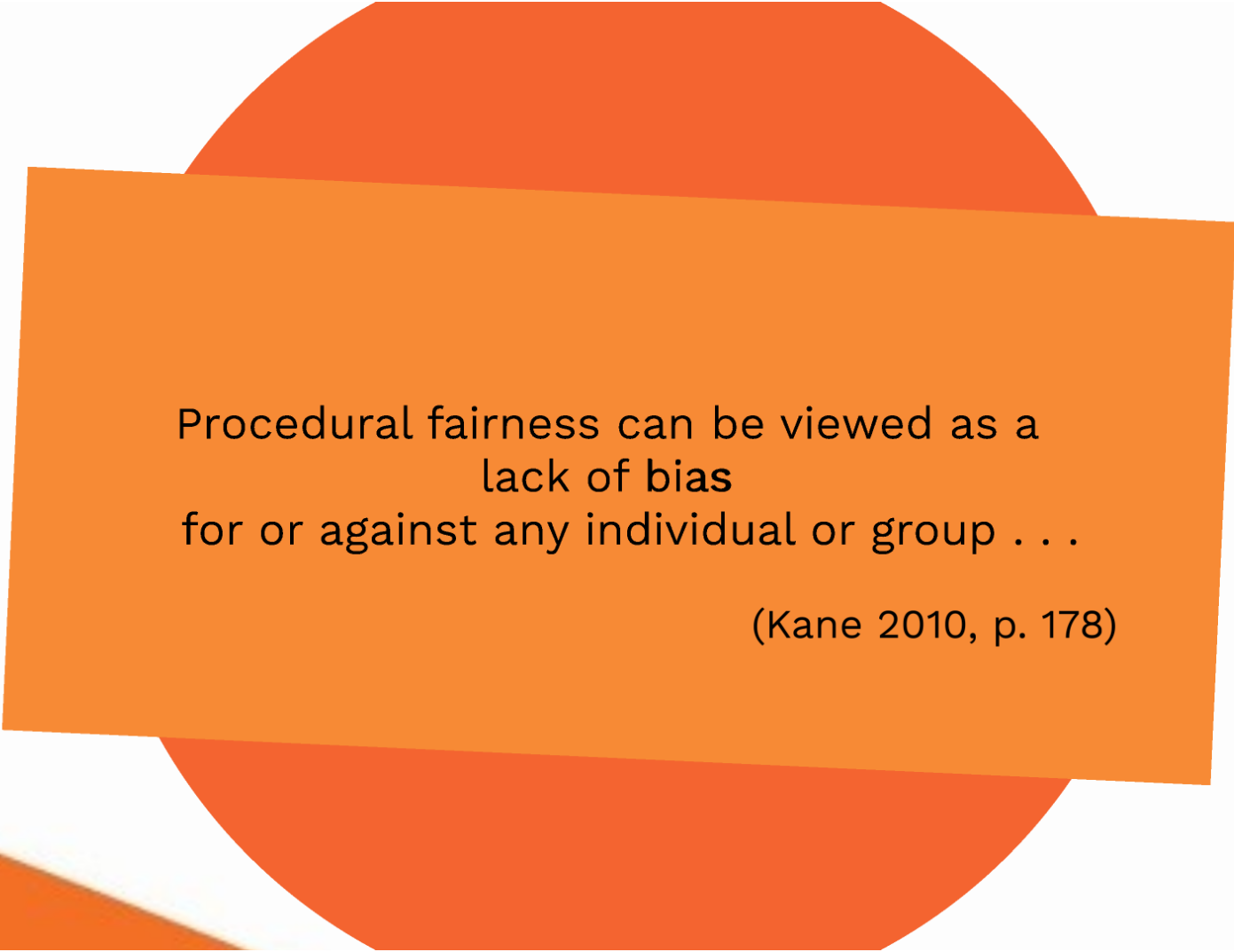
“ Ethnic prejudice is an antipathy based upon a faulty and inflexible generalization. It may be felt or expressed. It may be directed toward a group as a whole, or toward an individual because he is a member of that group ”

(Allport 1966, p. 9)

prejudices



Fairness



Procedural fairness can be viewed as a
lack of bias
for or against any individual or group . . .

(Kane 2010, p. 178)

bias → reliability → test quality

sources



bias

prejudice

How prejudices affect rating (case study Slovene)

8 November 2019



Fairness

Slovene

Research

Petra Likar Stanovnik



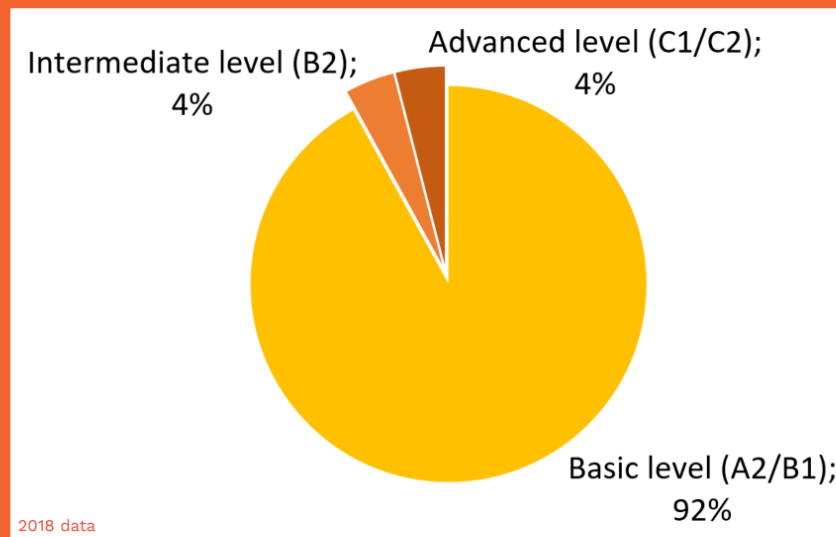
Slovene

cannot be learned

Slovene

less valuable

Slovene language exams



Basic
Level
Exam

Candidates

Results

Basic Level Exam
A2/B1 CEFR

60% 85%



Reading



Listening



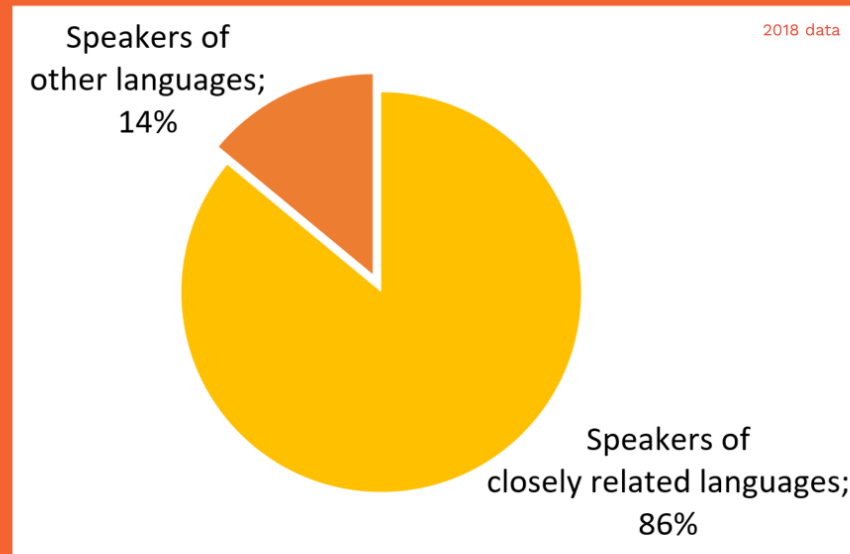
Speaking



Writing

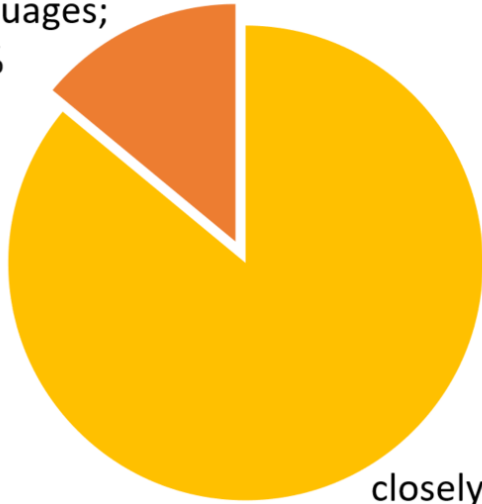


Candidates



Candidates

Speakers of
other languages;
14%

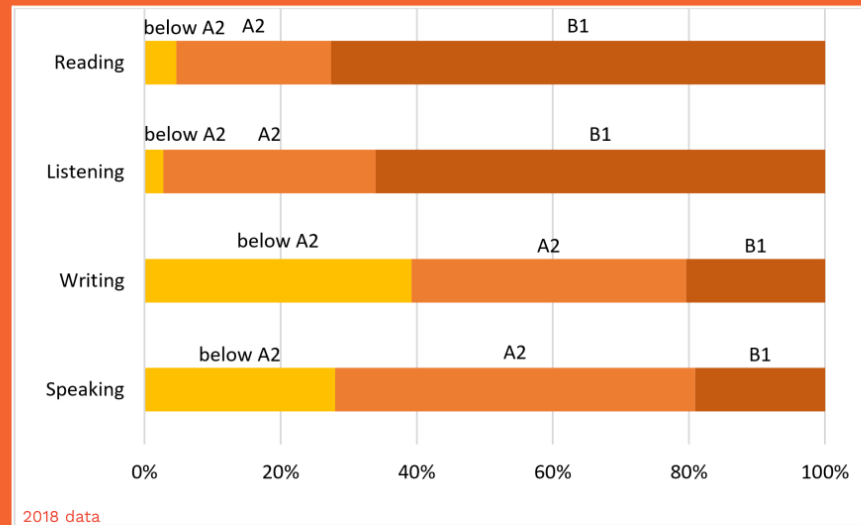


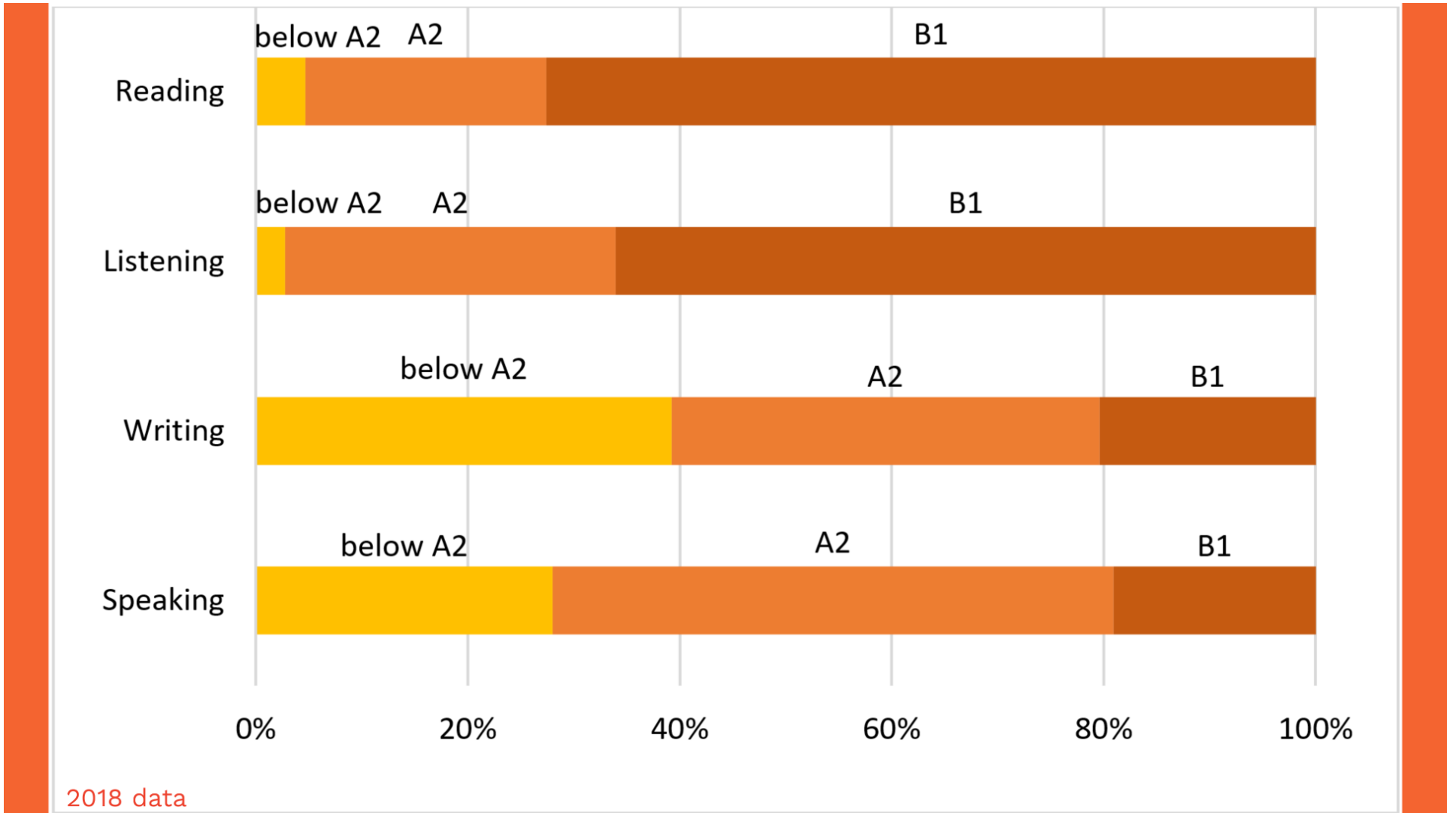
Speakers of
closely related languages;
86%

2018 data



Test results





How prejudices affect rating (case study Slovene)

8 November 2019



Fairness

Slovene

Research

Petra Likar Stanovnik

Research



Rating
analysis

Attitudes
on rating
criteria



Test

standardized
statistical analysis



classical test analysis
IRT analysis



Test takers

low motivation



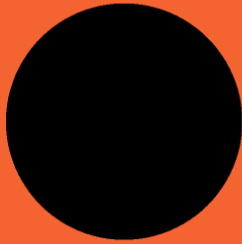
lack of time
similarity of languages
low education

...

(Ferbežar 2012)



Raters



2) Attitudes on rating criteria



1) Rating analysis

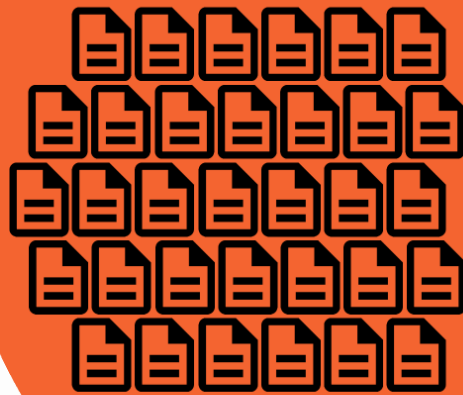


Research: 2018; Ferbežar, Likar Stanovnik

Rating analysis

March 2018

681
writing performances



41
raters



Rating

3rd rates



Writing



Test takers



Raters



Test takers

Two guided tasks:

- a practical text in a hypothetical situation
 - a reply to a short letter
- 40-50 words each



Raters

Rating:

- analytical scale
- categories:
text content, vocabulary,
accuracy, coherence and style
- 10 points per text
- 20 points for writing subtest:
12+ points = A2
17+ points = B1





External
raters (34)

Centre's
raters (7)

1st rating

2nd rating



final score
82%

1st rating

2nd rating

3rd rating
18%



Results

1st rating



2nd rating



3rd rating

18%

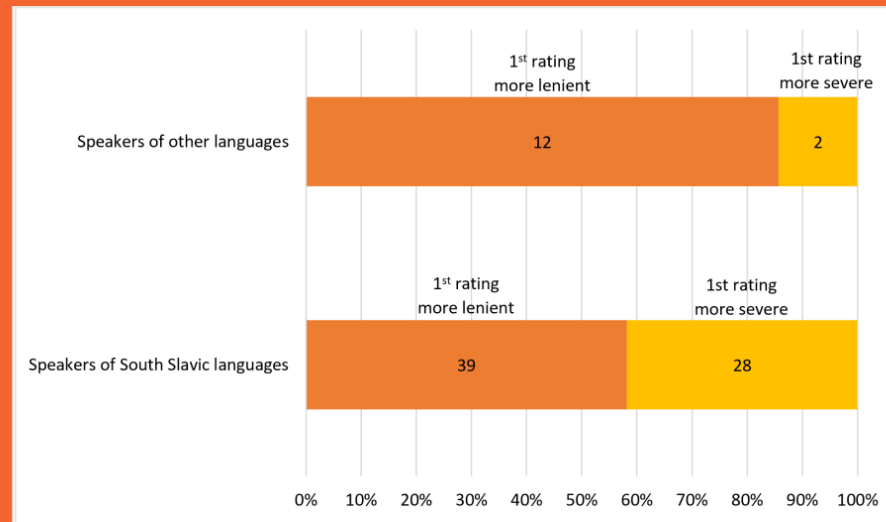


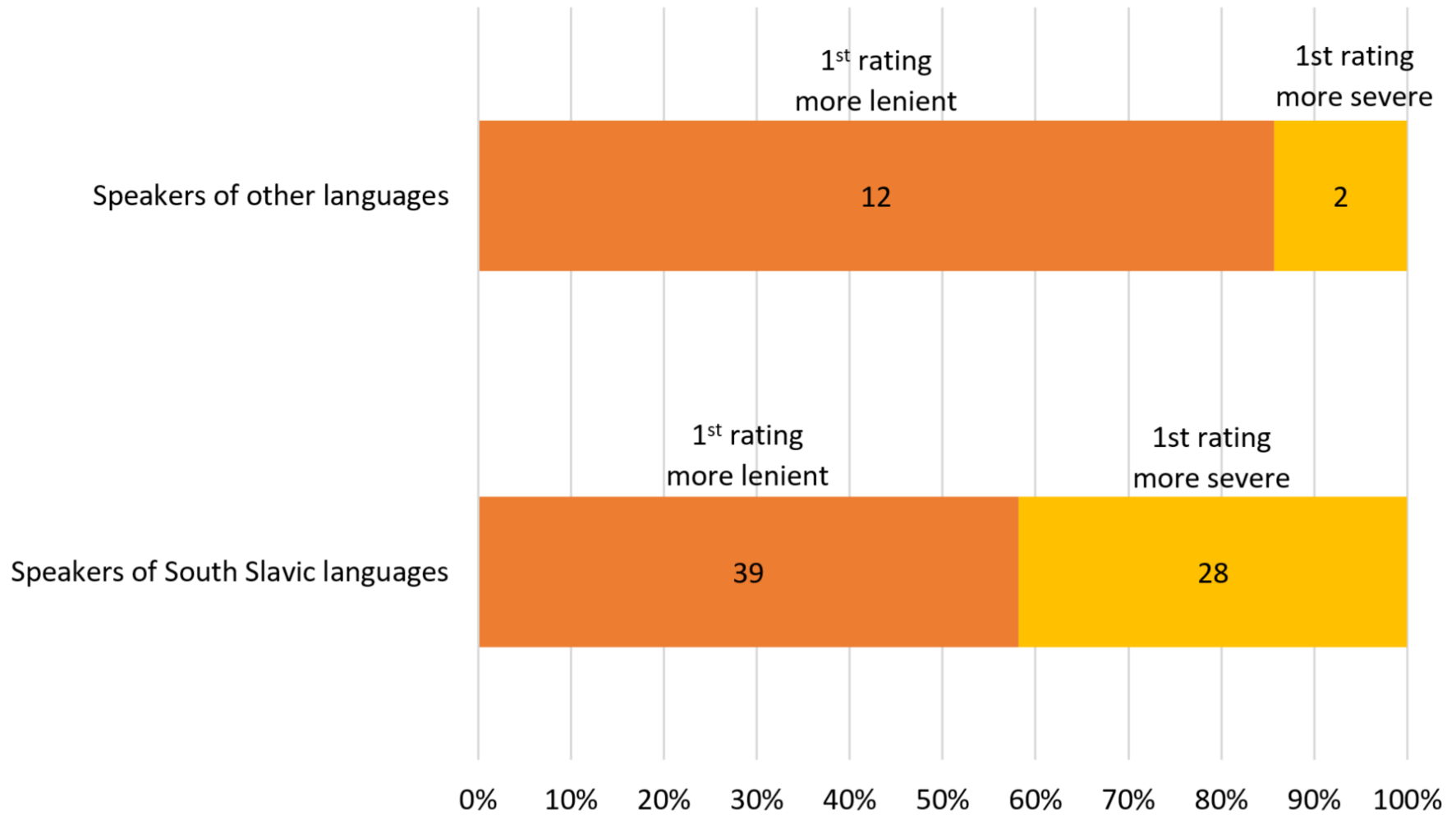
Difference:

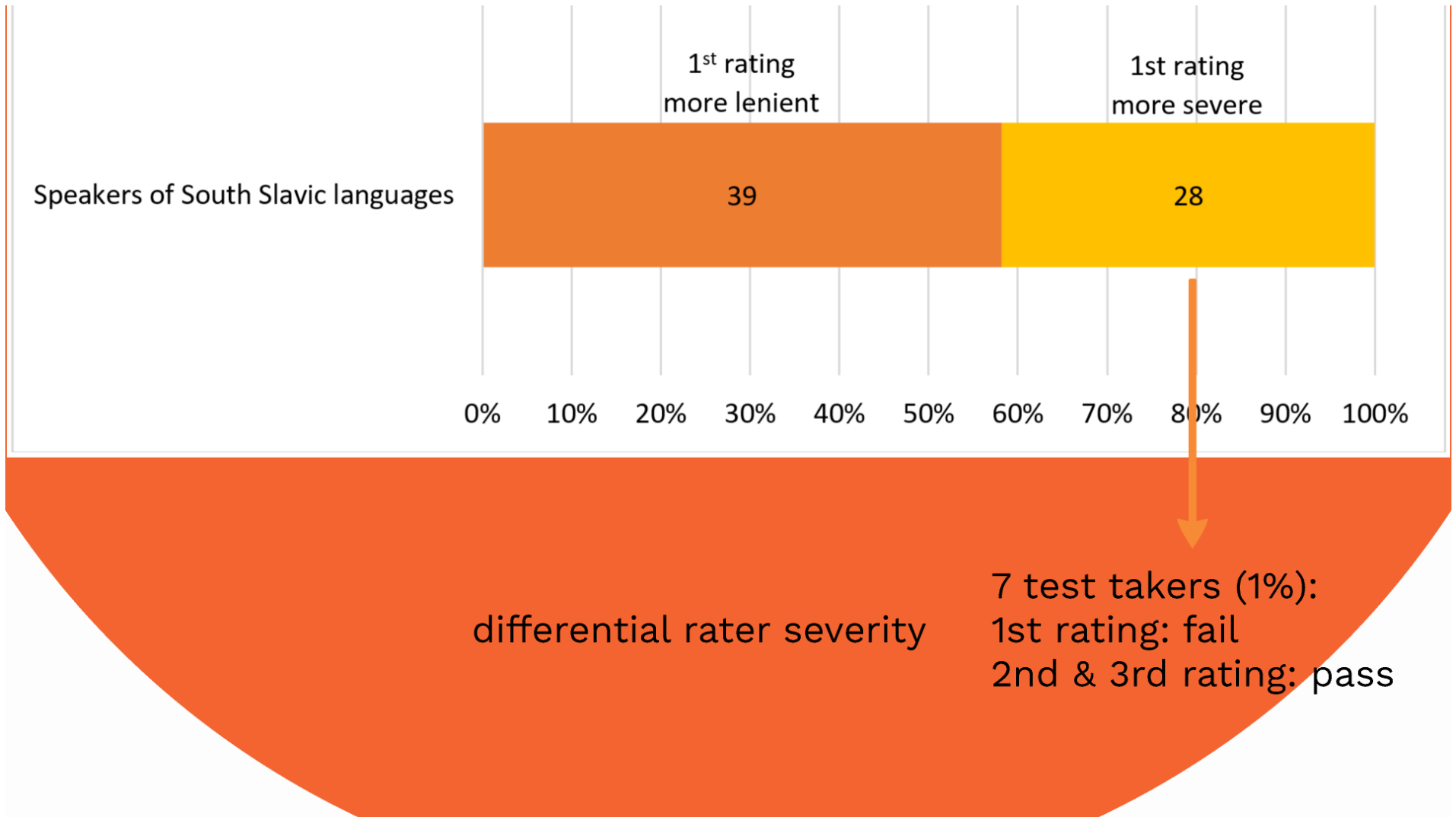
- number of points awarded (3+)
- pass/fail (below A2/A2)
- level (A2/B1)

Difference in scoring

81 ratings (12%)







Questionnaire

April-June 2018

87
raters

External
raters (68)

Centre's
raters (19)



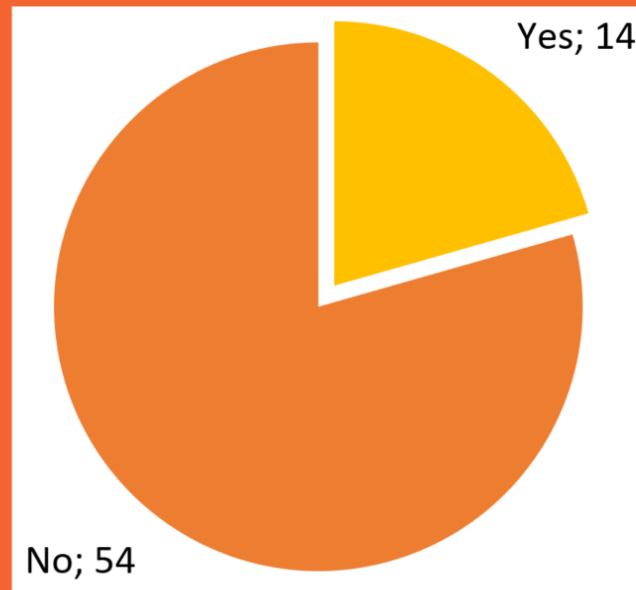
Q1

Q2

Q3

Should speakers of Slovene as a L1
be the reference for rating criteria?

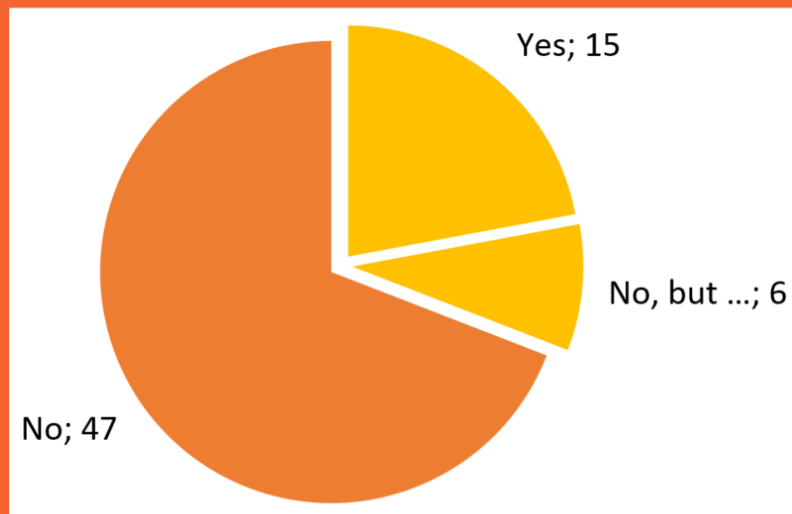
68 external raters



- native speakers know the language best and are most suited for setting the norm
- non-native speakers will never be capable of speaking Slovene at the level of a native speaker
- importance of social and symbolic role of the Slovene language

Should non-Slavic speakers be rated according to different criteria?

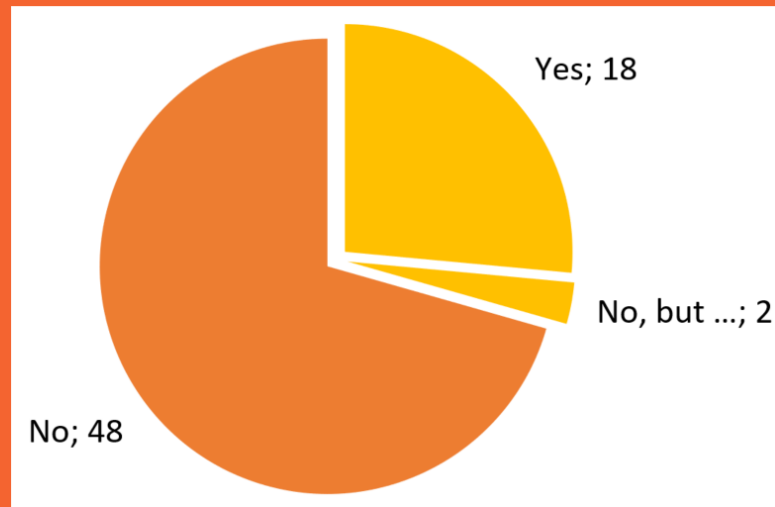
68 external raters



- Slavic speakers have the advantage of not having to put much effort into learning Slovene and can be better understood on account of linguistic proximity
- South Slavic language speakers are being privileged
- favorable inclination towards non-Slavic speakers, because it is harder for them to learn Slovene

Should the test taker's "effort" or "path traveled" be taken into account when rating?

68 external raters



- far greater efforts are required of non-Slavic speakers
- the test takers who speak closely related languages do not prepare for the exam and thus show their dismissive attitude towards the Slovenian language and culture
- South Slavic language speakers have no desire to learn Slovenian even after living in Slovenia for 10 or more years



Fairness

Measures:

- 16-hour introductory examiner and rater training
- regular training for raters
- a book of regulations on administering and rating exams
- standardization prior to rating speaking and writing performance
- rater monitoring
- regular checking of intra- and inter-rater agreement

Measures:

- 16-hour introductory examiner and rater training
- regular training for raters
- a book of regulations on administering and rating exams
- standardization prior to rating speaking and writing performance
- rater monitoring
- regular checking of intra- and inter-rater agreement

Research:

- rater bias
- resistant to change
(Eckes 2012, McNamara 1996, Baker 2012, Schaefer 2008)

- unreliable scoring of written performance
(Brown 2015, Wind, Peterson 2018)



Fairness

Measures:

- 16-hour introductory examiner and rater training
- regular training for raters
- a book of regulations on administering and rating exams
- standardization prior to rating speaking and writing performance
- rater monitoring
- regular checking of intra- and inter-rater agreement
- **raising raters' awareness of prejudices**
- **development of intercultural competence**