

SIG on technology

Needs survey (2018)

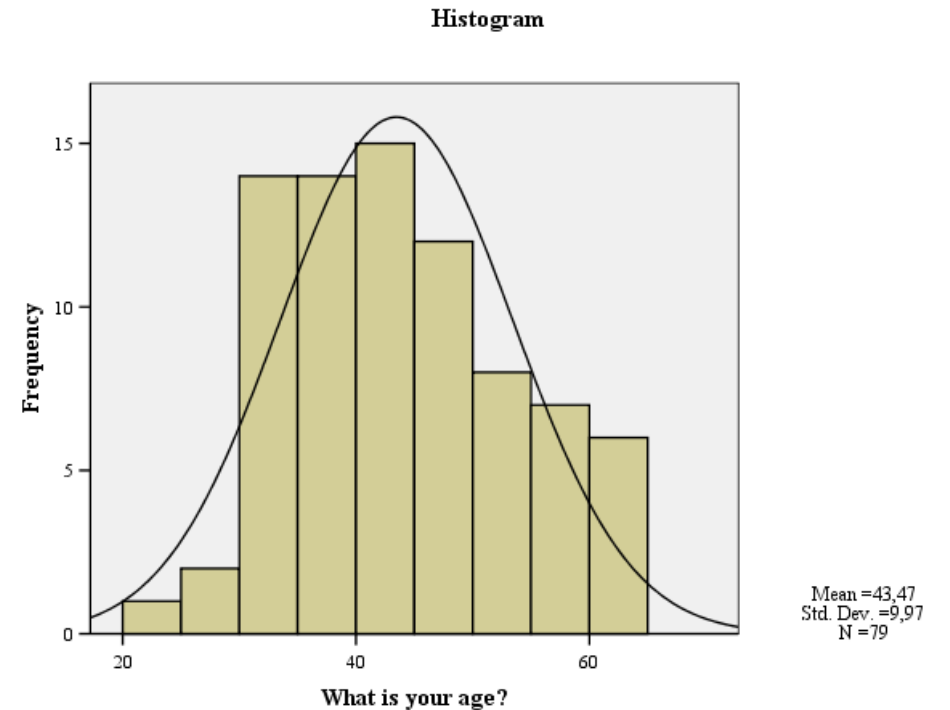
Salamanca meeting

Chair : Vincent FOLNY (CIEP)
Co-chair: Dominique Casanova (CCI)

Demography of respondents

- 79 respondents have filled in the survey.
- Final sample :42 (fully completed)

Statistics		
What is your age?		
N	Valid	79
	Missing	0
Mean		43,47
Median		43,00
Std. Deviation		9,970



Thanks to all of you for your answers!

Gender distribution

(respondents and missing responses)

What is your gender?						
		Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	Female	54	68,4	68,4	68,4	
	Male	25	31,6	31,6	100,0	
	Total	79	100,0	100,0		

			Missing
What is your gender?	Female	Count	30
			81,1%
	Male	Count	7
			18,9%
Total		Count	37
			100,0%

New public

Distribution
of Job titles

Which of the following most closely matches your job title?

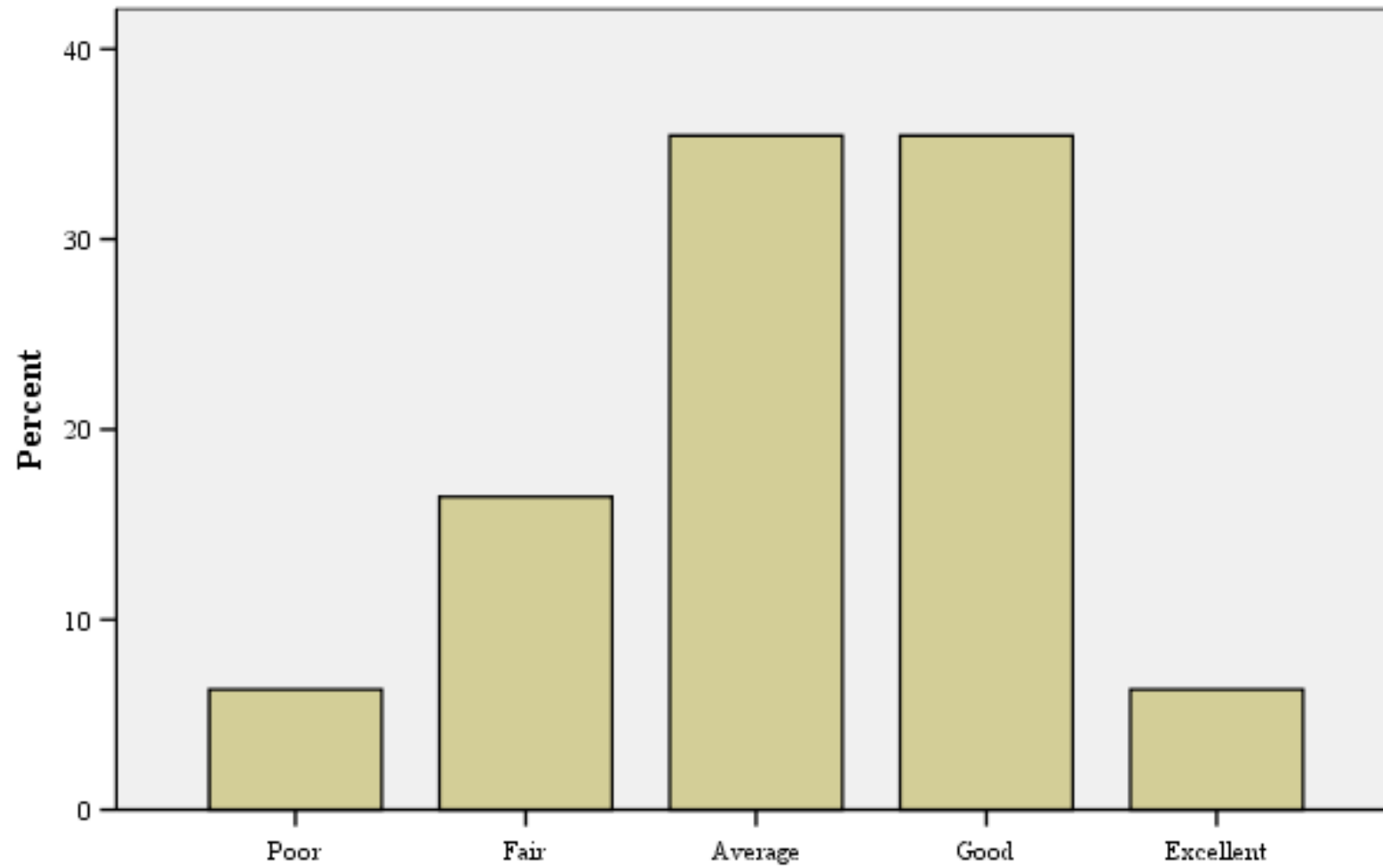
		Frequency	Percent	Valid Percent
Valid	In charge of item writers or rater teams	26	32,9	32,9
	Rater	20	25,3	25,3
	Manager / executive	14	17,7	17,7
	Item writer	9	11,4	11,4
	Other	7	8,9	8,9
	psychometrician	3	3,8	3,8
	Total	79	100,0	100,0

Job title x gender

Which of the following most closely matches your job title? * What is your gender? Crosstabulation

			What is your gender?		
			Female	Male	Total
Which of the following most closely matches your job title?	In charge of item writers or rater teams	Count	19	7	26
			35,2%	28,0%	32,9%
	Rater	Count	17	3	20
			31,5%	12,0%	25,3%
	Item writer	Count	6	3	9
			11,1%	12,0%	11,4%
	Manager / executive	Count	9	5	14
			16,7%	20,0%	17,7%
	Other	Count	2	5	7
			3,7%	20,0%	8,9%
	psychometrician	Count	1	2	3
			1,9%	8,0%	3,8%
Total		Count	54	25	79
			100,0%	100,0%	100,0%

Your technical and technological knowledge linked with Language testing is :



Technical knowledge of people not responding

- Their technical knowledge seems to be quite good:
- 2 hypothesis (for partial answering) :
 - The questionnaire was too long for them;
 - They were very optimistic about their technical knowledge.

Your technical and technological knowledge linked with Language testing is : * Game based assessment Crosstabulation

		Game based	
		Missing	
Your technical and technological knowledge linked with Language testing is :	Poor	Count	3
		% within Game based assessment	8,1%
	Fair	Count	7
		% within Game based assessment	18,9%
	Average	Count	10
		% within Game based assessment	27,0%
	Good	Count	13
		% within Game based assessment	35,1%
	Excellent	Count	4
		% within Game based assessment	10,8%
	Total	Count	37
		% within Game based assessment	100,0%

Technical knowledge x gender

- Same level for female / male : Average / good
- But, males seem to be more confident than females (0% “poor” for male and only 3,7 % of “Excellent” for female)

Your technical and technological knowledge linked with Language testing is : * What is your gender? Crosstabulation

			What is your gender?		
			Female	Male	Total
Your technical and technological knowledge linked with Language testing is :	Poor	Count	5	0	5
		% within What is your gender?	9,3%	,0%	6,3%
	Fair	Count	9	4	13
		% within What is your gender?	16,7%	16,0%	16,5%
	Average	Count	18	10	28
		% within What is your gender?	33,3%	40,0%	35,4%
	Good	Count	20	8	28
		% within What is your gender?	37,0%	32,0%	35,4%
	Excellent	Count	2	3	5
		% within What is your gender?	3,7%	12,0%	6,3%
Total	Count	54	25	79	
	% within What is your gender?	100,0%	100,0%	100,0%	

Check good balance gender of webinars presenters will be a good practice and will impact positively on confidence

Reliability of the survey

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,968	,968	47

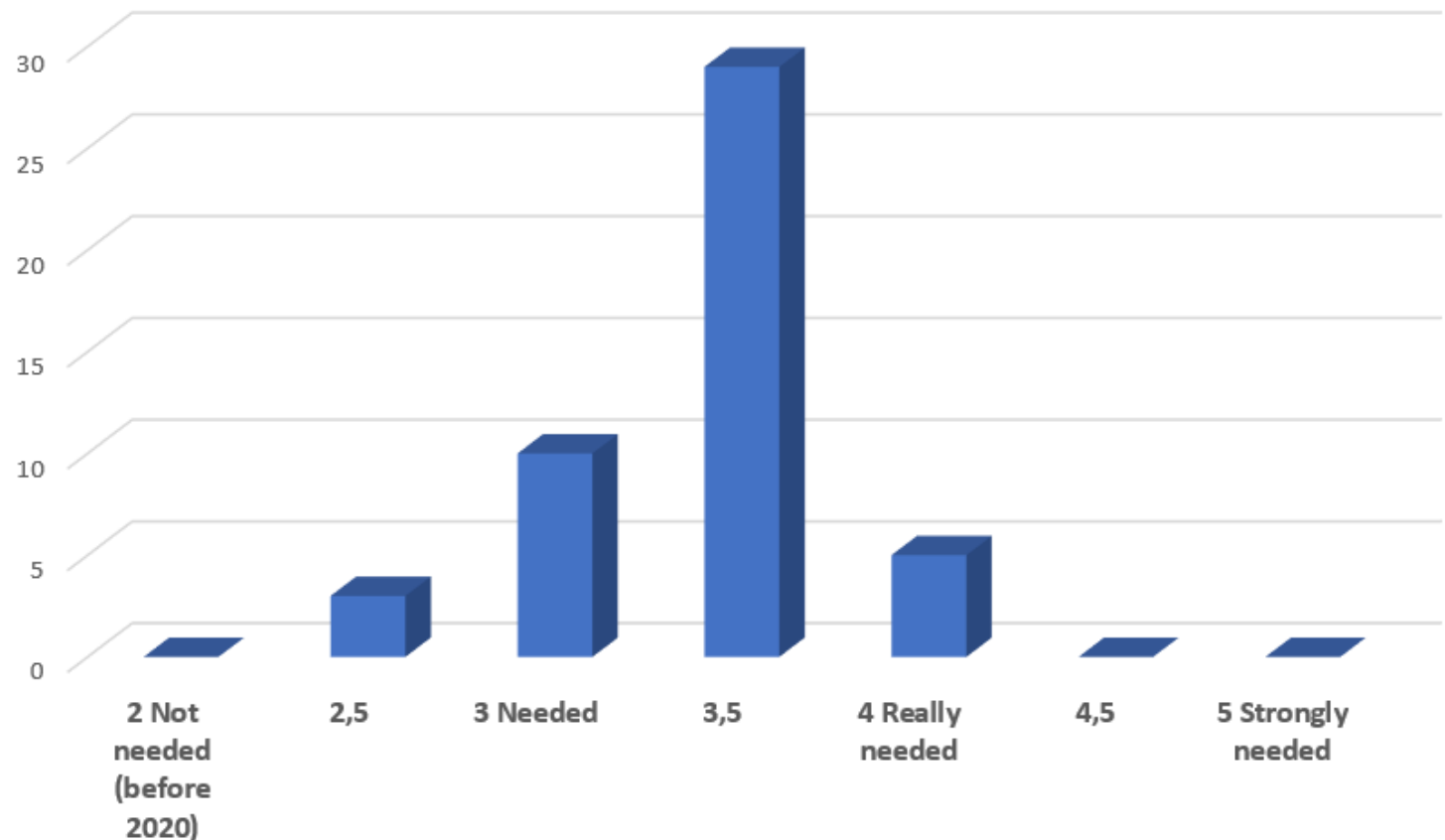
Level of need

- As Technology is an on going process, we can see that this is no “urgent” need if not the perception of “near future” need.
- So, we are on time and we don’t have to rush things with too much specialized topic.

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	3,150	2,143	3,714	1,571	1,733	,114	47
Item Variances	1,374	,843	1,905	1,062	2,260	,076	47

Distribution mean of items



15 most needed items

Rank	Item	Mean Statistic	Std. Error	Std. Deviation Statistic
1	Digital literacy for raters / examiners	3,7	0,2	1,1
2	Broad view on technology and digital tools for testing	3,6	0,2	1,0
3	Paper vs computer vs tablet	3,6	0,2	1,2
4	Generation of a data base and use of statistical indices for quality improvement	3,6	0,2	1,3
5	GDPR (General Data Protection Regulation) : issues and opportunities	3,5	0,2	1,1
6	Digitized items and tasks formats	3,5	0,2	1,1
6	Adaptive testing	3,5	0,2	1,0
6	Use of video-conferencing technology in the assessment of spoken language	3,5	0,2	1,3
9	Creation and analyses of a corpus (tools to use)	3,4	0,2	1,2
10	Multistage testing	3,4	0,2	1,0
10	Corpus analysis for validation	3,4	0,2	1,1
12	Listening / reading : textual analyses indices and statistical analyses (on text)	3,4	0,2	1,0
13	Rating audio recording vs video recording	3,3	0,2	1,2
13	Speaking / writing: textual analyses indices and statistical analyses (on text)	3,3	0,2	1,1
15	Writing tasks with internet access	3,3	0,2	1,3
15	Automated discourse analyses techniques (writing assessment)	3,3	0,2	1,4
15	Cognitive computing (Artificial Intelligence) helping rater to make decision on CEFR level	3,3	0,2	1,3

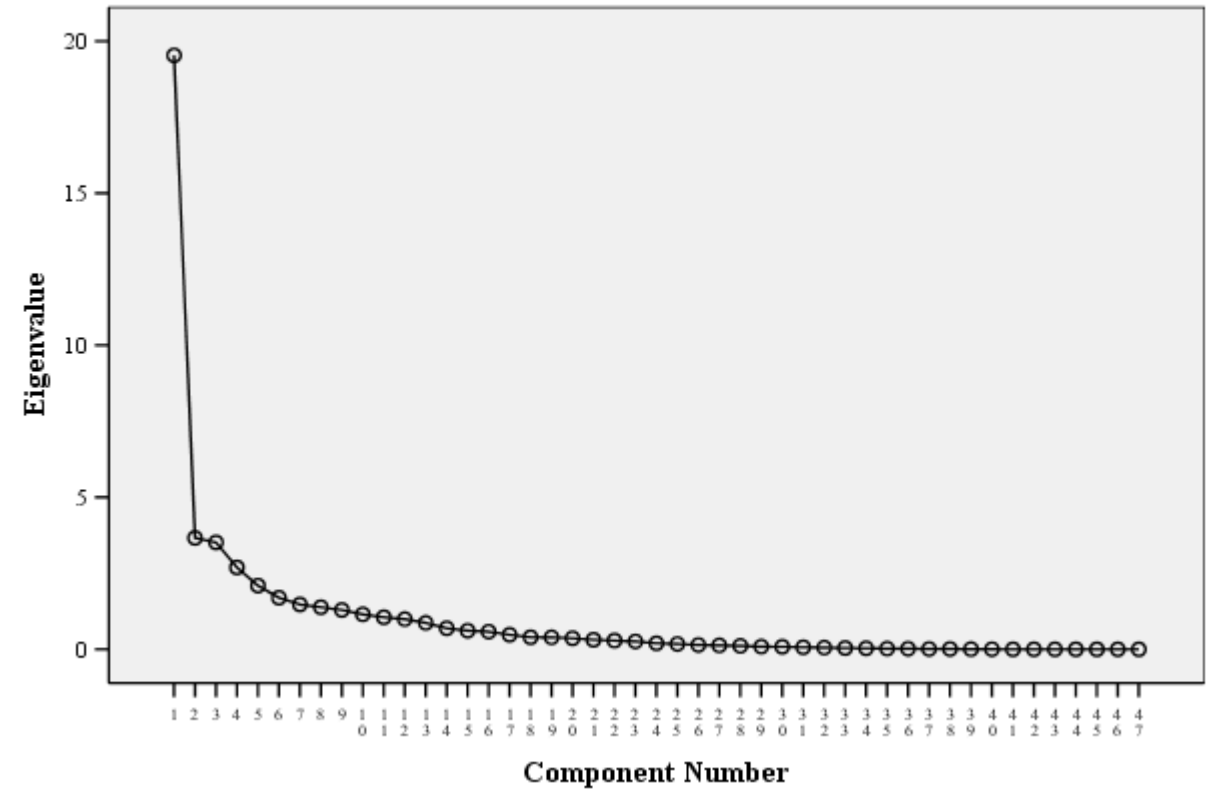
15 less needed items

Rank	Item	Mean Statistic
32	Artificial intelligence-based assessment systems	3,1
34	Language features extraction from written/oral productions	3,0
35	Automatic Item generation	3,0
36	Data mining in education Research	2,9
36	Software for multifaceted Rasch model	2,9
36	NPL for testing and assessment	2,9
39	Use of psychometric libraries with free programming languages (R, Python)	2,8
40	Implementation, strategy to improve its efficiency	2,8
41	Contribution of neuroscience tools to enhance construct definition (reading, listening)	2,8
42	Technical norms and assessment (dublin core, course-ware management interface, learning c	2,7
43	Eye-tracking (validation of reading and listening)	2,6
44	Game based assessment	2,5
45	Use of artificial neural networks for classification / assessment	2,5
46	Optical mark reading	2,5
47	Automatic sign language recognition	2,1

Dimensionality

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	19,533	41,560	41,560
2	3,664	7,795	49,355
3	3,514	7,477	56,832
4	2,691	5,726	62,558
5	2,095	4,458	67,016
6	1,696	3,609	70,625
7	1,474	3,137	73,763
8	1,383	2,942	76,704
9	1,293	2,752	79,456
10	1,151	2,448	81,904

Scree Plot



Mainly unidimensional but with significant secondary dimension

Dimension 1 : “automation”

Analysis based on
saturation for a
same factor

ITEM	1
Use of artificial neural networks for classification / assessment	0,86
Implementation, strategy to improve its efficiency	0,86
Automatic Item generation	0,84
Automated discourse analyses techniques (writing assessment)	0,80
Cognitive computing (Artificial Intelligence) helping rater to make decision on CEFR level	0,80
Speaking (validity argument)	0,79
Artificial intelligence-based assessment systems	0,76
Chatbot for oral interaction	0,76
Automated essay evaluation	0,75
Language features extraction from written/oral productions	0,74
Use of video-conferencing technology in the assessment of spoken language	0,72
Text classification for reading and listening	0,72
Digital literacy for teachers	0,71
Creation and analyses of a corpus (tools to use)	0,71

Dimension 2 : software /psychometrics

ITEM	2
Advanced software skills for classical test analyses / IRT	0,63
Basic software skills (classical and IRT)	0,63
Software for multifaceted Rasch model	0,54
Adaptive testing	0,44
Progress assessment / repeated measure	0,42
Use of psychometric libraries with free programming languages (R, Python)	0,30

Dimension 3 :

psychometrics /
corpus indices /
textual indices
(text)

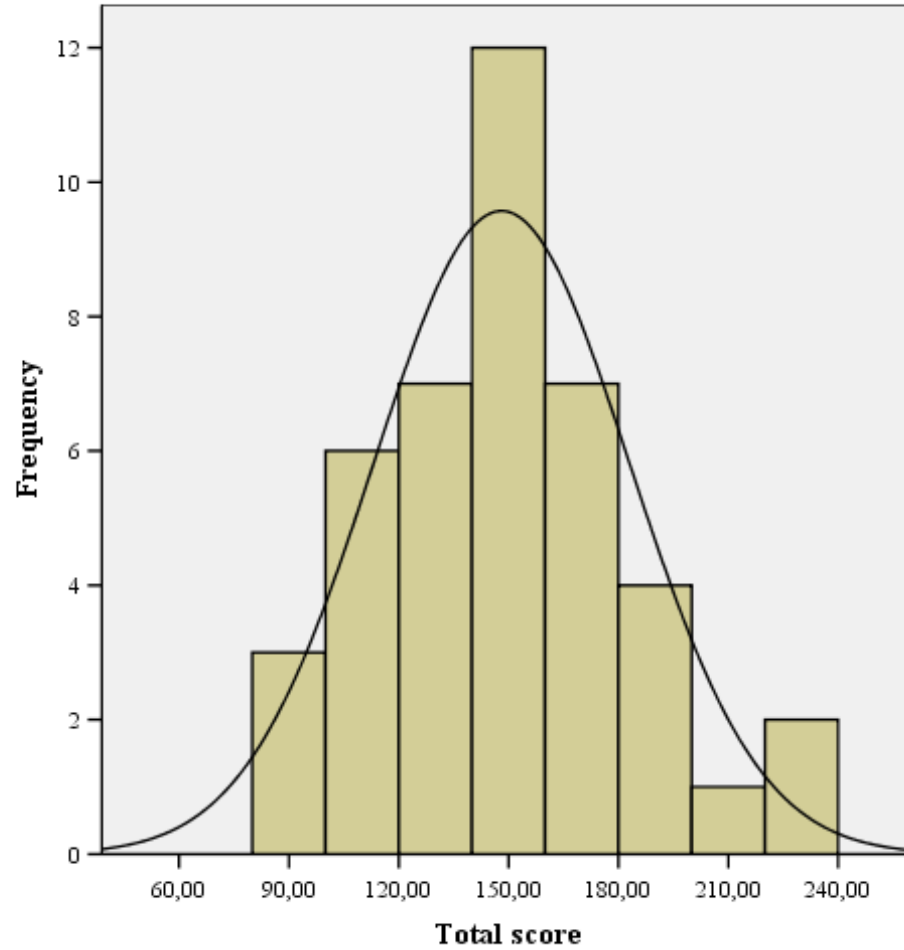
ITEM	3
Use of psychometric libraries with free programming languages (R, Python)	0,56
Speaking / writing: textual analyses indices and statistical analyses (on text)	0,54
Generation of a data base and use of statistical indices for quality improvement	0,54
Software for multifaceted Rasch model	0,50
Listening / reading : textual analyses indices and statistical analyses (on text)	0,50
Advanced software skills for classical test analyses / IRT	0,46
Corpus analysis for validation	0,42

Dimension 4 :

Digital tools and
delivery modes

ITEM	4
Paper vs computer vs tablet	0,58
Technical norms and assessment (dublin core, course-ware management interface, learning object metadata, instructional management system, IMS QTI, sharable content object reference model, MINE SCORM)	0,48
Smartphone and testing	0,41
Optical mark reading	0,39
Eye-tracking (validation of reading and listening)	0,29

Distribution
of total score
of needs

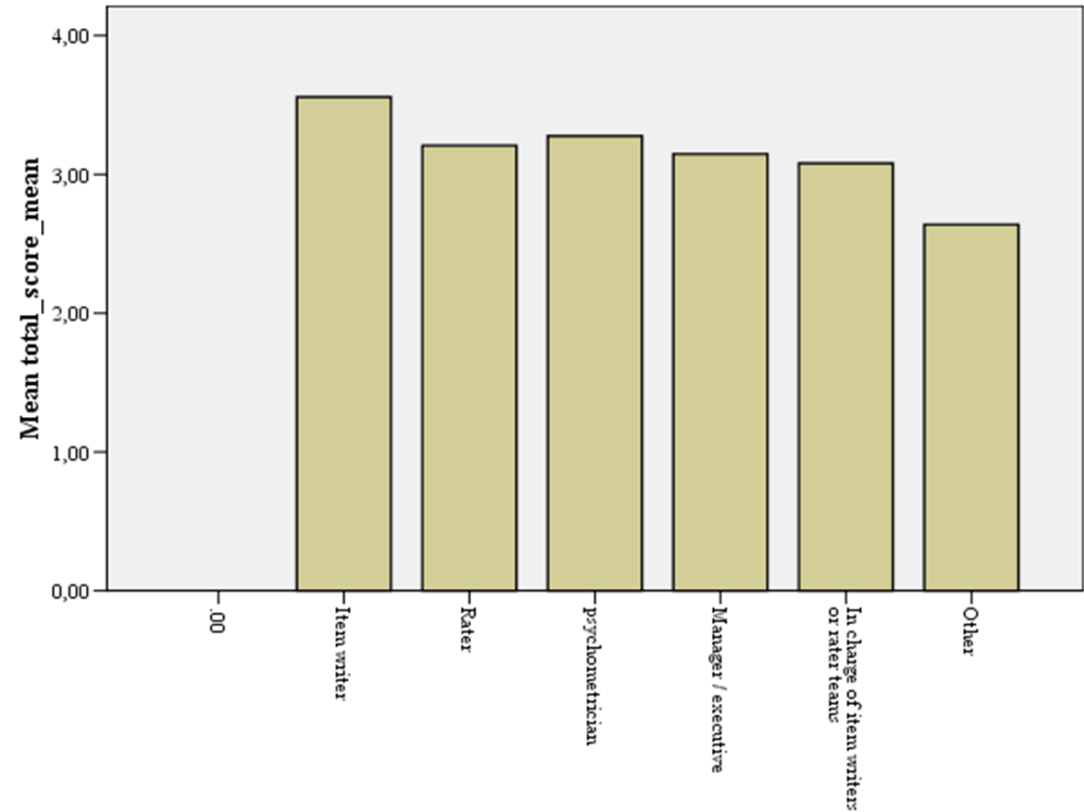


**Perception of
needs is not
homogenous.**

Mean =148,0714
Std. Dev. =35,0062
N =42

Level of needs x job title

- Item writers are the one expressing more needs;
- But data is difficult to interpret (due to the size of the data);
- The needs of male and female are strickly the same.

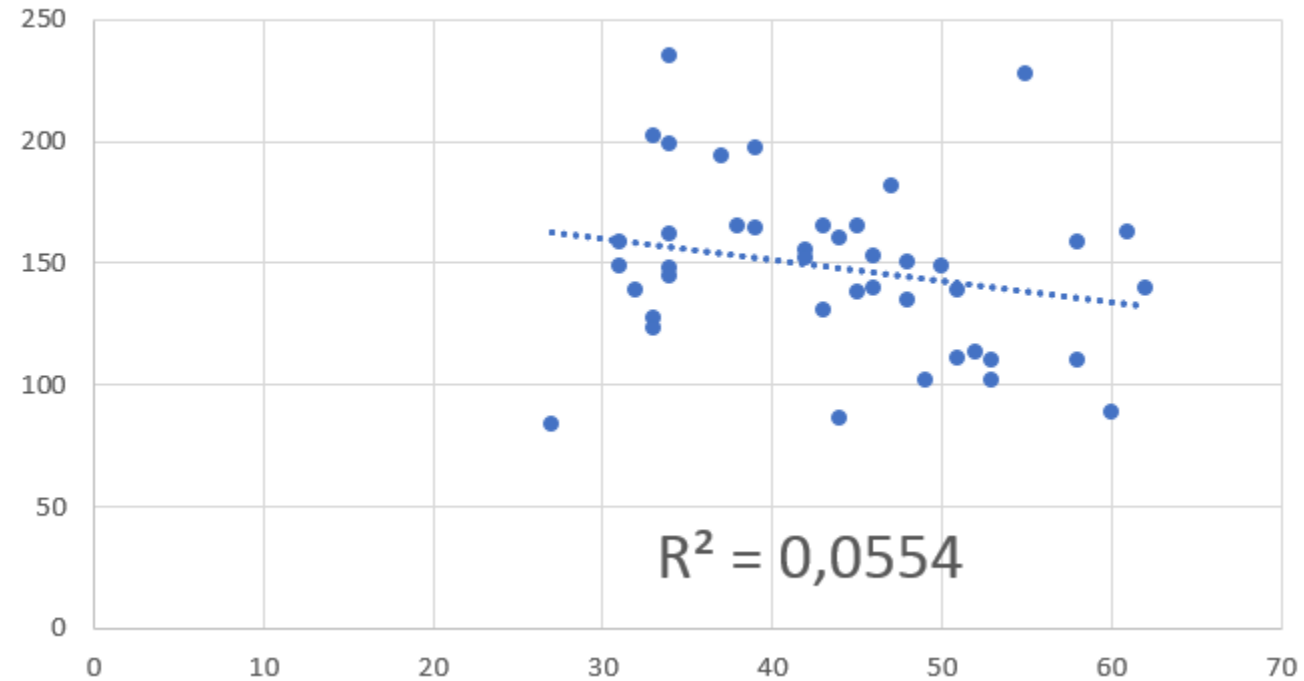


Report

Total score			
What is your gender?	Mean	N	Std. Deviation
Female	148,0833	24	28,28568
Male	148,0556	18	43,27813
Total	148,0714	42	35,00620

No correlation between the age and the level of need

Correlation between age and level of need



Findings

- Level of needs in high!
- We got answers from new ALTE colleagues (rater, item writers...)!
- Dissemination of the knowledge should be important as people interested are “in charge of other people” / manager.
- Rating is an area of big interest.
- Automation / use of software / corpus and text / delivery modes are the dimensions to be treated.

Findings

- We should start with webinars dealing with broad topics and to go further with webinars focused on more specific topic in the middle term (mode delivery ; data protection, item/test format; use of video and corpus analysis).
- We should avoid webinars on technical tools, specific points and non “easy” topics like artificial intelligence and very specific technical topic or highly skilled procedures.

Proposal / suggestion

- Starting with broad topics for webinars in 2019/2020 (aim : to get a good review of the state of art from the presenters).
- The trajectory will be to go for more specific topics in the future.
- Made a proposal of a winter or summer school to enter in more specific topics / technics as the use of software or analysis of corpus or digitalized tasks (we should make an effort not to come with a proposal too expensive; during the last day of the meeting we could have a webinar conference (half a day)).
- For webinar, we should ask new researchers, practitioners to work with us to renew our ALTE contacts and to get new ideas, different points of view.
- In a term of two years, we will access and revised what have been done to follow with the idea to understand what Alte members need.