

# Guaranteeing Fairness and Equality of Treatment using Parallel Simulation Exercises in Large Scale Assessment Centres: 2 empirical studies



**Stéphane Vanderveken**, M.A. Psychology  
Adviser at the European Personnel  
Selection Office

**Guy Van Biesen**, M.A. Psychology  
Adviser at the European Personnel  
Selection Office

34th Annual ACSG Conference  
10-14 March 2014  
South Africa

# EPSO selects staff for:

5000   
European  
Parliament

3549   
CONSILIUM  
Council

34 345   
European  
Commission

...and many  
other agencies of  
the EU

2015   
Court of Justice

895   
Court of Auditors

  
Ombudsman

  
Economic &  
Social  
Committee

  
Committee  
of the Regions

  
Data Protection  
Supervisor

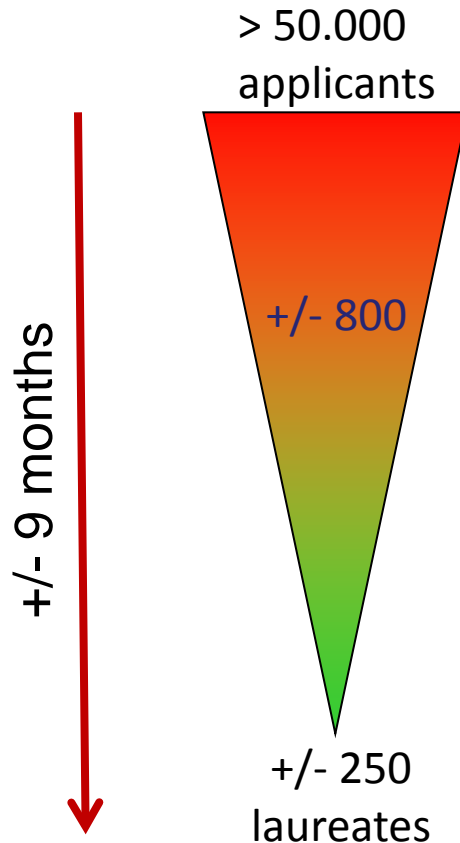
# Some figures

Since the creation of EPSO in 2003:

- more than 600 000 candidates have participated in EPSO competitions;
- around 700 open competitions organised;
- >20,000 successful candidates.



# Competition structure – an example



## 1. Computer-based (Admission) Tests (“CBT”)

- Verbal, numerical, abstract reasoning
- In the language of the candidate (L1)
- Situational Judgement Test

## 2. Assessment Centre (“AC”)

- In Brussels
- Fixed quota of invited candidates
- Tests in English, French or German (L2)
- General and field-specific competencies
- Usually tests over one day

## 3. Reserve list

- Fixed quota

# Testing for competency

## Assessment Center

	Competency based Interview	Oral presentation	Group exercise	Case study
Analysis and Problem Solving		X		X
Communicating		X		X
Delivering Quality and Results		X		X
Learning and Development	X		X	
Prioritising and Organising			X	X
Resilience	X	X		
Working with Others	X		X	
Leadership	X		X	
Technical skills				X

Duration: 1 day

6 candidates per session

4 assessors per session

Full capacity: 6 sessions / day



# Accountability

Article 90  
To EPSO

Complaint to  
EU Ombudsman

Article 91  
To the Court of Justice

## And exposure...

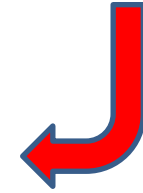
- many publications, websites and social networks aiming at sharing information regarding EPSO testing tools;
- all competitions open to all, always

# Guaranteeing

## Fairness and Equality of Treatment



Everybody  
should get the  
same  
opportunities



Everybody  
should face  
the same  
tests



Over an **extended testing period**, it requires the use of tests **different enough to assure fairness** but **similar enough to assure equality**

High exposure

+

Long testing period (Months)

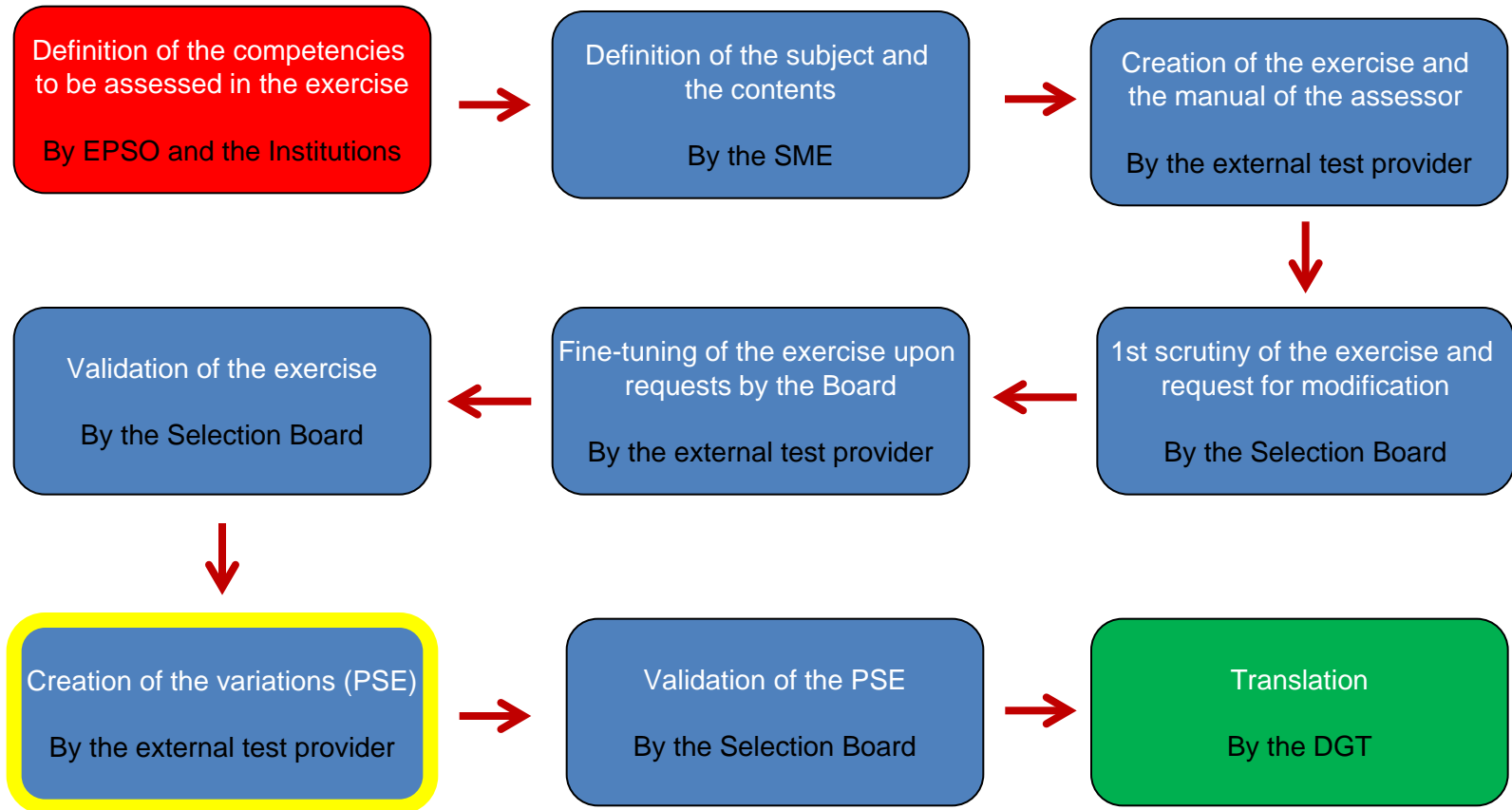
=

**Need to use Parallel Simulation exercises (PSE)**





# Simulation exercise development



# Literature on PSE

Brummel, Rupp, Spain, *Constructing Parallel Simulation Exercises for Assessment Centers and Other Forms of Behavioral Assessment*, *Personnel Psychology*, 62, 137-170, 2009

- There is a gap in scientific research on this subject;
- Difficult to get equivalent levels of complexity;
- Need for a rigorous methodology;
- Stimuli and behaviors to be elicited should be carefully listed;
- Minor variations are preferable

# Equality of treatment



# Study nr 1: Lawyers

- Year of competition: 2010
- Grade: AD5 (= junior)
- Domain: Law
- Number of candidates: 174
- Exercise analyzed: written case study
- Scoring system: 2 blind markers per copy, + third marking in case of Discrepancy
- Competencies assessed: 4 general + 1 job related
- Number of variations used: 4 major variations and their minor variations
- Testing period: 9 weeks



# Study nr 1: Lawyers

		N	Moyenne	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Limite inférieure	Limite supérieure
<i>sum_cs_scores</i>	<i>a</i>	38	21,21	5,39	,87	19,44	22,98
	<i>b</i>	41	19,73	6,43	1,00	17,70	21,76
	<i>c</i>	40	21,22	6,60	1,04	19,12	23,33
	<i>d</i>	32	21,00	5,54	,98	19,00	23,00
	<i>de</i>	9	19,89	9,33	3,11	12,71	27,06
	<i>fr</i>	14	21,71	7,34	1,96	17,47	25,95
	<i>Total</i>	174	20,80	6,30	,48	19,86	21,74

## ANOVA

		Somme des carrés	df	Mean Square	F	Significance
<i>sum_cs_scores</i>	<i>Between Groups</i>	80,87	5	16,17	,40	,85
	<i>Within Groups</i>	6779,09	168	40,35		
	<i>Total</i>	6859,96	173			

Nor is there any significant difference on any singular competency score

# Study nr 2: Assistants

- Year of competition: 2012
- Grade: AST3 (= assistant with experience)
- Domain: Project management
- Number of candidates: 123
- Exercise analyzed: group exercise
- Scoring system: 2 assessors per candidate
- Competencies assessed: 3 general
- Number of variations used: 4
- Testing period: 3 weeks



# Study nr 2: Assistants

## Descriptives

		N	Moyenne	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Limite inférieure	Limite supérieure
GR_sum	1	38	18,63	4,58	,74	17,13	20,14
	3	40	18,80	4,20	,66	17,46	20,14
	5	25	19,56	4,78	,96	17,59	21,53
	6	20	19,00	5,19	1,16	16,57	21,43
	Total	123	18,93	4,56	,41	18,12	19,75

## ANOVA

		Somme des carrés	df	Mean Square	F	Significance
GR_sum	Between Groups	14,08	3	4,69	,22	,88
	Within Groups	2525,40	119	21,22		
	Total	2539,48	122			

Nor is there any significant difference on any singular competency score

# Fairness





# Study nr 1: Lawyers

		N	Moyenne	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Limite inférieure	Limite supérieure
CS_sum	1	58	21,17	5,70	,75	19,67	22,67
	2	66	20,68	6,22	,77	19,15	22,21
	3	50	20,52	7,11	1,00	18,50	22,54
	Total	174	20,80	6,30	,48	19,86	21,74

Period 1 = 21/09 – 07/10/2010  
 Period 2 = 11/10 – 28/10/2010  
 Period 3 = 08/11 – 09/12/2010

## ANOVA

		Somme des carrés	df	Mean Square	F	Significance
CS_sum	Between Groups	12,89	2	6,44	,16	,85
	Within Groups	6847,07	171	40,04		
	Total	6859,96	173			

Nor is there any significant difference on any singular competency score

# Study nr 2: Assistants

## Descriptives

		N	Moyenne	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Limite inférieure	Limite supérieure
GR_sum	1	40	18,98	3,91	,62	17,72	20,23
	2	36	19,78	4,95	,82	18,10	21,45
	3	47	18,26	4,75	,69	16,86	19,65
	Total	123	18,93	4,56	,41	18,12	19,75

Period 1 = 18-19/06/2012  
 Period 2 = 25-26/06/2012  
 Period 3 = 2-9/07/2012

		Somme des carrés	df	Mean Square	F	Significance
GR_sum	Between Groups	47,35	2	23,67	1,14	,32
	Within Groups	2492,13	120	20,77		
	Total	2539,48	122			

Nor is there any significant difference on any singular competency score

# Conclusions

- Over an extended period of testing, **FAIRNESS** and **EQUALITY OF TREATMENT** may involve potentially antithetic constraints, and require the adoption of PSE;
- **FAIRNESS**: in both studies, it appears that the system is fair as the level of performance remains the same over time. PSE avoids that later participants benefit from knowledge potentially acquired from previous testing sessions;
- **EQUALITY**: in both studies, the use of different versions of the same exercise does not generate difficulty inequalities, as it does not induce a systematic difference of level of performance.

# Recommendation

- Whenever possible, it is advised to run pilot studies in order to test the versions before deployment;
- If this is not feasible, assessors must receive a statistical feedback on score differences during the testing period (EPSO practice);
- Every case is a case: studies on differences must be carried out for each selection process;
- Specific guidelines on PSE elaboration are still needed for exercises like group discussion, oral presentation and written case studies.

# Case F-127/1, De Mendoza versus European Commission

12/02/2014

The Civil Service Tribunal of the European Court of Justice states in point 52:

*"Par suite, il y a lieu de considérer que, au vu des circonstances susmentionnées, le fait que toutes les épreuves d'étude de cas n'ont pas eu lieu le même jour n'a pas, en l'espèce, entraîné un traitement différencié des candidats, susceptible d'en avantager certains par rapport à d'autres ni non plus un risque d'inégalité de traitement supérieur à celui inhérent à tout concours. "*

The use of PSE's in selection is lawful and does not constitute a breach in fairness over time as far as:

- What is tested is not knowledge but competencies;
- PSE's do not require possessing a particular previous knowledge;
- The material of PSE's is important (>20 pages);
- AC requirements of standardization are in place;
- Monitoring and feedback to the assessors are in place.

