

ΣΥΣΤΗΜΑ ΑΞΙΟΛΟΓΗΣΗΣ ΚΑΙ ΠΙΣΤΟΠΟΙΗΣΗΣ ΓΛΩΣΣΟΜΑΘΕΙΑΣ
ΕΡΓΟ ΤΟΥ ΑΡΙΣΤΟΤΕΛΕΙΟΥ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΘΕΣΣΑΛΟΝΙΚΗΣ

Αναλυτική έκθεση δεδομένων και αποτελεσμάτων

Υποέργο 3: Ετοιμασία Οργάνων Μέτρησης για την Γαλλική Γλώσσα
Πακέτο εργασίας 7: Έλεγχος ποιότητας των χρησιμοποιηθέντων οργάνων
μέτρησης

Υπεύθυνη Υποέργου: Καθ. Αριστέα-Νικολέττα Συμεωνίδου-Χριστίδου



ΥΠΟΥΡΓΕΙΟ ΕΘΝΙΚΗΣ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΕΑΕΚ
ΕΥΡΩΠΑΪΚΗ ΕΝΩΣΗ
ΣΥΓΧΡΗΜΑΤΟΔΟΤΗΣΗ
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
ΕΥΡΩΠΑΪΚΟ ΤΑΜΕΙΟ ΠΕΡΙΦΕΡΕΙΑΚΗΣ ΑΝΑΠΤΥΞΗΣ




Η ΠΑΙΔΕΙΑ ΣΤΗΝ ΚΟΡΥΦΗ
Επιχειρησιακό Πρόγραμμα
Εκπαίδευσης και Αρχικής
Επαγγελματικής Κατάρτισης



Γ' Κοινοτικό Πλαίσιο Στήριξης (ΚΠΣ)**Επιχειρησιακό Πρόγραμμα Εκπαίδευσης και Αρχικής Επαγγελματικής Κατάρτισης (ΕΠΕΑΕΚ)**

Άξονας Προτεραιότητας:	2	Προώθηση & Βελτίωση της εκπαίδευσης και της αρχικής επαγγελματικής κατάρτισης στο πλαίσιο της διά βίου μάθησης
Μέτρο:	2.1	Αναβάθμιση της ποιότητας της παρεχόμενης εκπαίδευσης
Ενέργεια:	2.1.2	Αξιολόγηση του εκπαιδευτικού έργου, των εκπαιδευτικών συντελεστών και λειτουργιών καθώς και των μαθητών
Κατηγορία Πράξεων:	2.1.2ζ	Σύστημα Αξιολόγησης και Πιστοποίησης Γλωσσομάθειας
Έργο:	111004	Σύστημα αξιολόγησης και πιστοποίησης γλωσσομάθειας στο ΑΠΘ
Ανάδοχος Έργου:		Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Φορέας Επίβλεψης:		Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Υποέργο:	03	Ετοιμασία Οργάνων Μέτρησης για την Γαλλική Γλώσσα
Υπεύθυνη Υποέργου:		Αριστέα-Νικολέττα Συμεωνίδου-Χριστίδου, (Καθηγήτρια Σημασιολογίας με έμφαση στη διαγλωσσική μελέτη Ελληνικής και Γαλλικής γλώσσας, Τμήμα Γαλλικής Γλώσσας και Φιλολογίας)
Πακέτο εργασίας:	7	Έλεγχος ποιότητας των χρησιμοποιηθέντων οργάνων μέτρησης
Παραδοτέα:	1	Αναλυτική έκθεση δεδομένων και αποτελεσμάτων

Πακέτο Εργασίας 7: Έλεγχος ποιότητας των χρησιμοποιηθέντων οργάνων μέτρησης

Τα χρησιμοποιηθέντα όργανα θα υποστούν και a posteriori έλεγχο προς αποφυγή σφαλμάτων και προς έλεγχο της ποιότητάς τους αντιπαραθετικά με τα όργανα στις υπόλοιπες γλώσσες. Συγκεκριμένα θα γίνουν τα εξής:

Περιγραφή:

- Έλεγχος αξιοπιστίας, με στατιστική ανάλυση κλασική ή συμβατική. Δηλαδή θα υπολογιστεί ο δείκτης ευκολίας του ερωτήματος, ο δείκτης διακριτικής ικανότητας, το ποσοστό επιλογής του κάθε «παραπλανητή» καθώς και η συμβολή της βαθμολογίας σε κάθε μία από τις 4 ενότητες στη διαμόρφωση της τελικής βαθμολογίας.
- Έλεγχος εγκυρότητας περιεχομένου (content validity) και εγκυρότητας δομής (construct validity).
- Δειγματοληπτικός έλεγχος συνάφειας αποτελεσμάτων εξέτασης και προσδοκιών τόσο των υποψηφίων όσο και των διδασκόντων που τους ετοίμασαν για την εξέταση.
- Τα αποτελέσματα των ελέγχων θα ερμηνευτούν, συγκριτικά με εκείνα των υπολοίπων γλωσσών.

Αναμενόμενα αποτελέσματα:

Δεδομένα που θα επιτρέπουν αφενός τη διαρκή βελτίωση της διαδικασίας κατασκευής των οργάνων μέτρησης και αφετέρου την επιτυχή αντιμετώπιση οποιουδήποτε ελέγχου από εθνικό ή διεθνή οργανισμό αξιολόγησης συστημάτων πιστοποίησης.

Παραδοτέο: Αναλυτική έκθεση δεδομένων και αποτελεσμάτων

Περιεχόμενα

1	Έλεγχος αξιοπιστίας	5
1.1	<i>Ανάλυση ερωτημάτων – Δοκιμασίες 1 και 3</i>	5
1.1.1	Παρουσίαση των μετρήσεων	5
1.1.2	Ερμηνεία των μετρήσεων	21
1.1.2.1	Δείκτες ευκολίας	21
1.1.2.2	Δείκτες διάκρισης	22
1.1.2.3	Ποιότητα των παραπλανητών	22
1.1.2.4	Σταθερότητα των αποτελεσμάτων και αντιστοιχία μεταξύ αποτελεσμάτων	22
1.1.2.5	Ομοιογενοποίηση	23
1.2	<i>Ανάλυση βαθμολογιών - Δοκιμασίες 2 και 4</i>	32
1.2.1	Παρουσίαση των μετρήσεων	32
1.2.2	Συνάφεια μεταξύ των βαθμολογητών	34
2	Έλεγχος εγκυρότητας	35
2.1	<i>Εγκυρότητα της δομής</i>	36
2.2	<i>Εγκυρότητα του περιεχομένου</i>	37
3	Έλεγχος συνάφειας αποτελεσμάτων εξέτασης και προσδοκιών	38
4	Ερμηνεία των αποτελεσμάτων των ελέγχων	39
	<i>Παράρτημα : Δείγματα των αναλύσεων των ερωτημάτων</i>	48

Στην παρούσα αναφορά αναλύεται η ποιότητα των οργάνων μέτρησης που χρησιμοποιήθηκαν κατά την αξιολόγηση που οδηγεί στην κατάκτηση του ΚΠΓ, μιας πιστοποίησης γνώσης ξένων γλωσσών. Η αναφορά που ακολουθεί με ακρίβεια τις οδηγίες που απαριθμούνται στο πακέτο 6 και βρίσκονται στις πρώτες σελίδες αυτού του τευχίδιου.

1 Έλεγχος αξιοπιστίας

Αξιοπιστία νοείται η ικανότητα ενός εργαλείου αξιολόγησης ή μέτρησης να πετυχαίνει με ακρίβεια τα χαρακτηριστικά που το ενδιαφέρουν. Υπάρχουν διάφορες μορφές αξιοπιστίας οι οποίες καταγράφονται στην κλασική θεωρία της μέτρησης. Οι κυριότερες μορφές αξιοπιστίας είναι

- η σταθερότητα των αποτελεσμάτων τα οποία δεν έχουν εξαχθεί ταυτόχρονα,
- η αντιστοιχία μεταξύ των αποτελεσμάτων τα οποία έχουν εξαχθεί με την παράλληλη χρήση του ίδιου οργάνου,
- η ομοιογένεια (ή εσωτερική συνεκτικότητα) των ερωτημάτων που συνθέτουν το όργανο μέτρησης.

Για πρακτικούς λόγους, οι αναλύσεις των ερωτημάτων που αποτελούν τις ερωτήσεις πολλαπλής επιλογής των δοκιμασιών 1 και 3 (απαντητικά φυλλάδια οπτικής αναγνώρισης) θα παρουσιαστούν ξεχωριστά από τις βαθμολογίες που δόθηκαν στις δοκιμασίες 2 και 4 (καταγραμμένες κρίσεις σχετικά με τα απαντητικά φυλλάδια οπτικής αναγνώρισης).

1.1 Ανάλυση ερωτημάτων – Δοκιμασίες 1 και 3

1.1.1 Παρουσίαση των μετρήσεων

Στους πίνακες που ακολουθούν, οι οποίοι συνάχθηκαν με βάση τα αρχεία ακατέργαστων δεδομένων που τέθηκαν στη διάθεση της ομάδας σύνταξης και εκπόνησης των θεμάτων, αναφέρονται για κάθε ερώτημα,

- ο δείκτης ευκολίας (*e.i.*), σε μορφή ποσοστού %
- ο δείκτης διάκρισης (*d.i.*), επίσης σε μορφή ποσοστού %
- ο δείκτης επιτυχίας των παραπλανητών, εφόσον υπάρχουν.

Η μέση τιμή των δεικτών ευκολίας υπολογίζεται

- ανά δραστηριότητα (δεξιά στήλη) και
- ανά δοκιμασία (κάτω από τον πίνακα, *average*).

Το ίδιο ισχύει και για τους δείκτες διάκρισης. Για κάθε δοκιμασία, η μέση μεταβλητή των δεικτών αναγράφεται στο κάτω μέρος του πίνακα (*st.dev.*)

Αποτελέσματα των υποψηφίων στις δοκιμασίες των γαλλικών της εξεταστικής περιόδου Μαΐου
2007 – Δοκιμασία 1

Επίπεδο Β1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	85	24	85	0	14						0	
2	97	3	1	97	1						0	
3	92	10	6	92	1						1	1
4	90	22	7	0	90						2	89,2
5	82	14	16	82	1						1	14,6
6	97	7	2	0	97	0	0				0	
7	83	29	6	0	0	83	9				0	
8	90	19	90	1	1	2	6				0	2
9	85	24	2	0	0	13	85				0	90,8
10	99	3	0	99	0	0	0				0	16,4
11	47	22	45	47	7						0	
12	59	45	59	33	7						0	
13	52	40	52	34	14						0	3
14	69	41	24	69	8						0	62,4
15	85	31	11	5	85						0	35,8
16	87	22	0	1	10	87	1				0	
17	56	55	15	2	56	6	21				0	
18	73	50	73	1	13	3	9				0	4
19	68	48	9	0	19	3	68				1	75,8
20	95	9	3	95	1	0	0				0	36,8
21	70	31	70	17	12						1	
22	86	29	4	10	86						0	
23	79	29	15	79	6						0	5
24	65	41	7	65	28						0	64,6
25	23	52	71	23	6						0	36,4
26	77	29	77	12	11						0	
27	50	34	2	48	50						0	
28	62	45	23	62	15						0	6
29	74	38	5	21	74						0	66,6
30	70	45	70	18	12						0	38,2
31	35	12	48	35	18						0	
32	69	33	5	69	26						0	
33	36	36	7	36	56						0	7
34	89	26	7	4	89						0	58,2
35	62	48	62	32	6						0	31,0
36	44	19	50	44	6						0	
37	59	57	30	59	11						0	
38	75	38	5	20	75						0	8
39	69	34	69	23	8						0	63,2
40	69	41	69	14	17						0	37,8
41	27	40										
42	28	57										
43	28	24										9
44	56	52										38,0
45	51	43										43,2
46	97	12										
47	87	24										
48	48	38										10
49	94	19										84,4
50	96	5										19,6
	69,32	30,98	average									
	21,01	14,70	st.dev.									

Επίπεδο Β2

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	92	14	1	92	2	1	3	1			0	
2	33	51	5	1	0	60	33	0			0	
3	24	41	35	2	3	24	33	3			0	
4	82	34	4	2	82	2	8	2			0	1
5	54	43	54	2	10	9	18	5			1	62,2
6	88	30	1	1	2	4	4	88			0	35,5
7	57	36	36	57	7						1	
8	85	21	85	6	9						0	
9	37	26	26	37	37						0	
10	66	35	66	15	19						0	2
11	69	37	9	69	21						0	61,2
12	53	18	20	53	27						0	28,8
13	69	48	22	69	9						0	
14	81	28	10	81	9						0	
15	73	34	73	5	21						0	
16	82	25	82	10	7						0	
17	82	26	82	9	9						0	
18	80	34	7	12	80						0	3
19	55	35	41	55	4						0	76,3
20	88	20	2	10	88						0	31,3
21	63	42	3	8	4	22	63				0	
22	44	48	34	5	4	44	14				0	
23	45	48	45	20	5	20	10				0	4
24	76	36	6	5	76	4	10				0	58,2
25	63	43	12	63	10	10	4				0	43,4
26	76	40	2	1	2	11	6	0	1	76	2	
27	75	45	75	1	0	1	13	8	2	0	0	
28	47	66	0	23	1	1	26	47	1	0	1	
29	47	66	0	24	47	3	3	3	20	0	1	
30	48	64	24	0	1	25	48	1	0	0	0	
31	51	60	0	51	2	8	1	11	26	0	0	5
32	48	62	0	0	22	1	1	26	48	0	1	55,1
33	49	62	0	1	25	49	3	2	0	20	0	58,1
34	60	60	1	3	25	3	2	6	60		1	
35	45	62	6	45	2	24	1	10	11		1	
36	92	21	1	2	0	1	92	3	1		0	
37	43	54	6	16	3	43	1	26	4		0	
38	57	59	10	4	57	1	1	12	15		0	6
39	37	64	7	24	2	23	2	37	5		0	57,7
40	70	44	70	6	9	4	1	5	3		0	52,0
41	74	40	15	74	11						0	
42	75	23	75	19	6						0	
43	90	20	6	5	90						0	7
44	78	36	78	13	9						0	75,8
45	62	34	62	25	13						0	30,6
46	69	16	28	69	2						0	
47	85	36	4	85	11						0	
48	28	37	37	28	35						0	
49	78	29	13	78	9						0	
50	60	29	60	13	28						0	
51	64	40	9	64	26						0	
52	80	25	80	12	8						0	
53	53	47	53	12	35						0	8
54	66	32	6	28	66						0	65,6
55	73	45	14	13	73						0	33,6
56	47	38	18	47	35						0	
57	41	35	46	13	41						0	
58	38	21	38	38	23						0	9
59	59	43	59	30	11						1	47,0
60	50	43	50	27	23						0	36,0
61	40	37									0	
62	93	11									0	
63	5	5									0	10
64	7	11									0	35,4
65	32	45									0	21,8
66	90	22									0	
67	63	45									0	
68	91	26									0	
69	45	28									0	
70	16	16									0	
71	54	51									0	
72	69	41									0	
73	23	31									0	11
74	49	32									0	51,8
75	18	27									0	31,9
	59,35	37,05	average									
	21,21	14,32	st.dev.									

Επίπεδο Γ1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	67	56	0	18	4	3	8	0	67		1	
2	52	62	0	5	10	5	52	6	21		0	
3	94	14	94	0	1	3	0	1	1		1	
4	77	52	1	0	3	9	9	77	1		1	
5	54	56	2	3	16	54	12	8	5		0	1
6	61	44	2	1	61	19	11	4	3		0	68,1
7	72	64	0	72	4	8	9	5	2		0	49,7
8	63	76	2	2	5	6	63	7	8	4	2	
9	87	26	87	2	4	1	1	1	5	0	0	
10	67	76	3	4	67	6	8	8	4	0	0	
11	88	18	1	4	2	1	2	2	2	88	0	
12	63	82	2	8	8	7	5	63	3	3	1	
13	66	78	6	2	6	8	5	6	66	0	1	2
14	73	62	0	73	5	6	8	4	3	2	0	71,4
15	64	72	0	4	4	64	10	8	8	2	1	61,3
16	81	42	15	81	4						0	
17	80	18	80	10	11						0	
18	31	14	31	28	40						1	
19	82	36	82	12	5						0	
20	63	30	5	32	63						0	
21	63	52	8	63	11	5	11				1	
22	79	42	4	5	7	79	5				0	
23	68	60	68	16	8	6	2				0	
24	60	46	15	12	60	8	5				0	
25	74	50	5	3	15	2	74				1	
26	65	60	3	29	2	65	1				0	
27	59	54	59	8	24	4	5				0	
28	54	76	14	54	13	12	6				1	
29	76	52	6	4	7	7	76				0	
30	52	62	17	6	52	13	11				0	
31	61	26	30	9	61						1	
32	82	34	9	82	9						1	
33	80	22	12	8	80						0	3
34	82	24	82	10	8						1	66,6
35	39	14	39	20	40						1	40,7
36	95	10	0	95	2	1	2				0	
37	89	28	5	2	89	1	3				0	
38	72	46	4	1	3	72	20				0	4
39	63	54	17	1	4	16	63				0	78,8
40	75	44	75	2	2	10	12				0	36,4
41	80	22	16	80	4						0	
42	51	16	51	26	23						0	
43	83	34	83	8	9						1	
44	75	40	14	11	75						1	
45	72	46	23	72	6						0	
46	72	20	3	72	3	15	6				1	
47	59	12	12	10	59	8	11				0	
48	62	38	9	4	11	13	62				1	
49	49	34	49	12	13	14	12				0	
50	48	44	27	2	15	48	8				1	
51	85	38	85	4	10						1	
52	67	30	13	67	18						1	
53	89	28	89	5	6						1	
54	69	48	10	69	20						1	
55	76	32	16	7	76						1	
56	76	52	4	5	5	9	76				1	
57	43	22	34	6	10	43	6				1	
58	42	34	9	42	32	12	4				1	5
59	33	24	33	26	13	22	6				1	63,6
60	40	52	21	19	40	12	6				1	33,3
61	11	18										
62	12	20										
63	5	14										6
64	20	26										23,6
65	70	32										22,0
66	20	40										
67	71	16										
68	47	70										
69	52	56										
70	2	2										
71	76	28										
72	75	46										
73	21	32										7
74	51	26										41,9
75	4	8										32,4
	61,08	39,12	average									
	22,06	19,17	st.dev.									

Αποτελέσματα των υποψηφίων στις δοκιμασίες των γαλλικών της εξεταστικής περιόδου Μαΐου
2007 – Δοκιμασία 3

Επίπεδο Α

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	100	-1	0	100	0	0	0				0	
2	100	1	0	0	0	100	0				0	
3	99	3	0	0	99	0	0				1	1
4	100	1	0	0	0	0	100				0	99,8
5	100	0	100	0	0	0	0				0	0,8
6	88	25	88	9	2						1	
7	95	11	2	3	95						1	
8	86	32	10	86	3						0	2
9	97	10	1	2	97						0	90,6
10	87	29	87	9	3						1	21,4
11	89	28	11	89	0						0	
12	65	58	26	65	9						0	
13	82	50	82	17	1						0	3
14	95	11	3	2	95						0	81,6
15	77	43	77	18	4						1	38,0
16	55	56	55	26	18						2	
17	66	53	26	66	8						1	
18	58	40	58	22	18						2	4
19	65	32	4	65	30						0	56,0
20	36	29	20	36	44						0	42,0
21	53	72										
22	62	58										
23	52	54										5
24	64	49										56,6
25	52	75										61,6
26	88	31	0	9	1	1	88				0	
27	87	38	2	3	87	3	5				0	
28	97	6	97	0	1	1	0				0	6
29	89	29	2	3	3	89	2				0	88,8
30	83	38	0	83	7	5	5				1	28,4
31	63	43	63	30	6						1	
32	71	51	9	71	18						1	
33	78	49	15	6	78						1	7
34	81	33	81	12	6						0	76,0
35	87	25	87	6	6						1	40,2
36	59	56	59	17	23						1	
37	68	40	16	68	16						0	
38	58	51	17	24	58						0	8
39	53	15	53	36	10						0	62,0
40	72	36	72	14	13						0	39,6
41	37	43	2	3	9	37	11	37			2	
42	45	39	1	5	12	23	45	14			0	
43	79	49	6	79	7	3	2	2			1	9
44	31	49	0	5	10	23	29	31			1	50,8
45	62	67	0	7	62	14	12	4			1	49,4
46	12	28										
47	34	44										
48	40	67										10
49	68	47										32,6
50	9	19										41,0
	69,48	36,24	average									
	22,84	19,58	st.dev.									

Επίπεδο Β1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	40	29	57	40	3						0	
2	76	27	5	17	76						1	
3	67	41	12	67	20						0	1
4	81	26	81	18	0						0	70,2
5	87	23	3	10	87						1	29,2
6	86	19	1	6	86	5	1				0	
7	63	65	16	13	7	2	63				0	
8	94	6	1	1	4	94	0				0	2
9	69	50	69	1	0	1	29				0	78,2
10	79	44	13	79	2	0	5				0	36,8
11	50	25	47	50	3						0	
12	89	31	5	6	89						0	
13	69	33	69	27	4						0	3
14	82	33	6	12	82						0	69,6
15	58	38	36	58	6						0	32,0
16	61	33	14	61	10	11	3				0	
17	58	52	18	9	10	58	5				0	
18	88	22	7	1	1	3	88				0	4
19	55	58	14	20	55	8	3				0	61,6
20	46	66	46	10	23	19	2				0	46,2
21	71	41	18	71	11						0	
22	84	20	84	2	13						1	
23	95	13	2	95	3						0	5
24	81	34	81	5	13						1	84,0
25	89	23	10	1	89						0	26,2
26	86	30	9	86	6						0	
27	65	43	65	25	10						0	
28	59	37	23	18	59						1	6
29	48	45	29	48	23						0	64,6
30	65	59	65	18	17						0	42,8
31	74	33	15	74	11						0	
32	67	49	67	17	16						0	
33	72	53	9	19	72						0	7
34	91	11	91	4	5						0	65,6
35	24	47	24	19	57						0	38,6
36	80	32	80	9	12						0	
37	53	29	39	53	8						0	
38	74	54	9	17	74						0	8
39	73	41	17	73	10						0	72,2
40	81	38	81	11	7						0	38,8
41	70	20										
42	96	7										
43	78	52										9
44	54	50										73,6
45	70	50										35,8
46	61	42										
47	44	41										
48	48	49										10
49	63	34										43,4
50	1	2										33,6
	68,30	36,00	average									
	18,42	14,95	st.dev.									

Επίπεδο Β2

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	88	23	0	1	6	88	4				0	
2	85	30	5	0	7	2	85				0	
3	91	18	91	2	2	1	4				0	1
4	96	11	2	96	0	0	1				0	88,8
5	84	28	2	1	84	9	5				0	22,0
6	47	10	19	47	34						0	
7	57	44	19	25	57						0	
8	80	36	80	8	12						0	2
9	61	51	33	61	6						0	64,6
10	78	36	6	16	78						0	35,4
11	51	41	51	35	14						0	
12	72	9	72	23	5						0	
13	53	49	53	24	23						0	3
14	59	47	32	59	8						1	64,8
15	89	19	4	7	89						0	33,0
16	61	42	4	4	14	17	61				1	
17	49	51	18	6	8	49	19				0	
18	63	38	63	14	5	11	7				0	4
19	73	33	9	73	5	11	2				0	63,0
20	69	49	6	4	69	10	11				0	42,6
21	73	40	9	73	15	1	0				0	
22	96	9	0	1	2	1	96				0	
23	67	52	14	17	67	1	1				0	5
24	74	36	74	9	15	1	0				0	81,0
25	95	10	0	1	1	95	3				0	29,4
26	76	45	2	3	9	9	76				0	
27	64	44	64	1	5	26	4				0	
28	88	31	2	88	2	5	2				0	6
29	73	50	9	3	73	3	12				0	71,6
30	57	49	22	5	11	57	5				0	43,8
31	81	22	81	13	6						0	
32	55	45	35	9	55						0	
33	82	21	17	82	1						0	7
34	83	29	83	13	4						0	77,2
35	85	28	3	85	12						0	29,0
36	71	44	71	8	21						0	
37	68	52	13	68	19						0	
38	25	31	13	61	25						0	8
39	68	36	14	68	18						0	63,2
40	84	32	13	4	84						0	39,0
41	12	20										
42	65	45										
43	62	43										9
44	8	19										30,2
45	4	10										27,4
46	29	14										
47	10	11										
48	12	6										10
49	16	26										13,6
50	1	1										11,6
	61,80	31,32	average									
	26,14	14,65	st.dev.									

Επίπεδο Γ1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	52	26	1	1	20	52	26				1	
2	40	36	3	1	49	7	40				0	
3	93	18	1	93	1	2	2				1	1
4	93	16	93	1	1	1	2				1	61,0
5	27	33	1	3	27	38	30				0	25,8
6	82	25	0	2	82	15	1				0	
7	97	5	0	1	0	0	97				1	
8	98	5	98	1	0	0	0				0	2
9	81	25	1	11	6	81	0				1	88,6
10	85	21	1	85	12	2	0				0	16,2
11	35	16	35	58	7						0	
12	85	8	85	10	6						0	
13	49	51	49	4	46						1	
14	38	12	13	38	49						1	
15	20	13	77	20	4						0	
16	91	22	0	0	91	7	0				1	
17	99	4	0	0	0	0	99				0	
18	64	32	12	64	3	20	0				1	
19	62	34	62	31	1	6	0				0	
20	66	37	24	5	4	66	1				0	
21	73	38	73	14	13						1	
22	49	32	12	37	49						1	
23	40	22	27	40	32						1	
24	58	32	25	16	58						1	
25	41	26	41	19	39						1	
26	85	37	6	2	85	3	4				1	
27	65	45	65	1	3	18	13				0	
28	53	45	11	8	4	23	53				0	3
29	88	21	1	88	2	6	3				0	60,5
30	49	58	17	1	7	49	26				0	29,3
31	47	39	32	20	47						1	
32	58	55	39	58	3						0	
33	38	29	38	41	21						0	
34	44	29	18	37	44						0	
35	80	30	80	11	9						1	
36	63	45	17	5	63	2	12				1	
37	77	28	3	5	2	77	12				1	
38	56	51	56	3	16	7	18				0	
39	83	39	7	83	4	2	3				0	
40	54	53	16	4	13	12	54				1	
41	96	13	1	96	1	1	0				1	
42	75	55	75	1	5	14	4				0	
43	89	26	5	1	2	2	89				0	
44	90	29	2	0	90	4	3				0	
45	79	49	16	1	2	79	2				0	
46	94	11	3	2	94						0	
47	82	1	12	82	6						0	
48	89	7	89	6	4						1	4
49	56	51	56	16	28						0	70,9
50	67	50	16	67	17						0	34,5
51	6	16										
52	15	32										
53	8	16										
54	17	32										
55	59	39										
56	14	22										
57	62	41										
58	28	17										5
59	23	36										26,3
60	31	36										28,7
	60,63	29,53	average									
	25,80	14,48	st.dev.									

Αποτελέσματα των υποψηφίων στις δοκιμασίες των γαλλικών της εξεταστικής περιόδου Μαΐου
2008 – Δοκιμασία 1

Επίπεδο Β1

	e.i.*100	d.i.*100	A	B	C	D	E	x		
1	33	52	12	14	33	5	36	0		
2	53	62	53	7	16	14	9	0		
3	40	40	12	5	20	25	40	0	1	
4	67	60	10	67	8	11	3	0	47,8	
5	46	71	14	9	21	46	10	0	57,0	
6	63	48	28	63	9			0		
7	43	55	43	36	20			0		
8	90	26	5	5	90			0	2	
9	82	12	5	82	13			0	69,4	
10	69	55	69	18	13			0	39,2	
11	70	72	2	10	13	5	70	0		
12	93	21	93	1	2	1	2	0		
13	78	55	0	78	9	4	8	0	3	
14	87	33	1	4	4	87	5	0	80,0	
15	72	69	4	8	72	3	13	0	50,0	
16	78	34								
17	55	55								
18	67	52							4	
19	77	33							62,0	
20	33	55							45,8	
	64,80	48,00	average							
	18,07	16,40	st.dev.							

Επίπεδο Β2

	e.i.*100	d.i.*100	A	B	C	D	E	x		
1	56	52	6	23	56	3	11	1		
2	84	30	84	7	2	1	6	1		
3	70	45	4	17	4	3	70	1	1	
4	50	64	4	50	37	3	7	0	70,0	
5	90	21	2	3	2	90	3	0	42,4	
6	24	41	24	57	19			0		
7	48	45	34	18	48			0		
8	36	30	39	36	25			0	2	
9	60	24	22	60	17			0	43,6	
10	50	32	50	18	31			0	34,4	
11	75	30	75	22	3			0		
12	56	28	56	27	17			0		
13	33	25	48	33	19			0	3	
14	39	33	41	20	39			0	54,6	
15	70	35	70	17	12			0	30,2	
16	49	41	49	33	18			1		
17	25	-2	31	44	25			0		
18	58	48	17	24	58			0	4	
19	73	40	73	24	3			0	55,2	
20	71	34	20	9	71			0	32,2	
21	48	55								
22	79	38								
23	65	51							5	
24	38	58							51,6	
25	28	41							48,6	
	55,00	37,56	average							
	18,30	13,54	st.dev.							

Επίπεδο Γ1

	e.i.*100	d.i.*100	A	B	C	D	E	x		
1	87	14	9	3	87			1		
2	56	38	9	36	56			0		
3	50	48	42	50	9			0		
4	67	32	27	67	5			1		
5	25	20	29	46	25			0		
6	50	44	5	50	45			0		
7	88	12	4	88	7			1		
8	95	4	0	5	95			0	1	
9	68	46	68	2	30			0	68,3	
10	97	6	1	97	2			0	26,4	
11	98	6	2	0	98	0	1	0		
12	96	8	96	1	2	1	0	1		
13	97	6	2	1	1	97	0	0	2	
14	98	6	0	98	1	1	1	0	97,4	
15	98	4	0	1	0	1	98	1	6,0	
16	80	52	80	3	13	4		1		
17	72	54	5	9	72	14		1		
18	97	10	1	1	1	97		0	3	
19	94	18	2	94	4	1		0	87,0	
20	92	24	92	2	6	1		0	31,6	
21	99	2								
22	87	32								
23	78	28							4	
24	84	34							80,6	
25	55	64							32,0	
	80,32	24,48	average							
	19,60	18,45	st.dev.							

Αποτελέσματα των υποψηφίων στις δοκιμασίες των γαλλικών της εξεταστικής περιόδου Μαΐου
2008 – Δοκιμασία 3

Επίπεδο Α

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	x	
1	93	21	3	2	93					2	
2	95	13	4	95	1					0	
3	90	20	3	5	90					1	1
4	83	32	83	1	16					1	90,0
5	89	20	89	4	6					1	21,2
6	39	55									
7	60	54									
8	86	31									2
9	55	59									63,4
10	77	49									49,6
11	57	72	57	9	5	18	10			1	
12	66	70	17	8	5	66	2			2	
13	77	58	9	77	2	2	9			1	3
14	76	54	10	2	5	5	76			1	71,8
15	83	44	6	2	83	5	3			1	59,6
16	48	58									
17	63	65									
18	78	51									4
19	75	65									65,2
20	62	77									63,2
	72,60	48,40	average								
	15,27	18,72	st.dev.								

Επίπεδο Β1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	x	
1	84	24	2	84	3	3	9			0	
2	51	56	5	11	28	4	51			1	
3	63	40	6	1	63	9	21			0	1
4	79	38	79	1	1	15	3			1	68,8
5	67	60	9	2	7	67	15			0	43,6
6	59	38	14	59	26					0	
7	58	50	25	16	58					0	
8	63	27	17	63	20					0	2
9	66	41	66	10	23					0	59,2
10	50	35	39	50	12					0	38,2
11	95	14	1	95	0	3	1			0	
12	92	20	92	1	1	1	4			0	
13	94	15	1	1	94	1	2			0	3
14	86	30	2	2	1	8	86			1	90,8
15	87	32	2	2	3	87	6			0	22,2
16	38	74									
17	33	56									
18	24	41									4
19	12	33									26,8
20	27	55									51,8
	61,40	38,95	average								
	24,39	15,40	st.dev.								

Επίπεδο Β2

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	x	
1	88	30	3	1	2	88	5			0	
2	85	33	85	4	4	2	4			0	
3	78	49	3	8	78	2	8			0	1
4	80	45	5	2	8	4	80			1	83,0
5	84	35	3	84	8	1	3			0	38,4
6	78	41	78	11	12					0	
7	86	31	11	3	86					0	
8	84	28	84	6	10					0	2
9	97	6	1	97	2					0	85,0
10	80	34	80	7	13					0	28,0
11	49	33	51	49	1					0	
12	71	50	71	11	18					0	
13	81	37	8	81	11					0	3
14	78	25	78	12	10					0	70,2
15	72	32	9	19	72					0	35,4
16	43	67									
17	12	27									
18	8	24									4
19	18	49									22,4
20	31	56									44,6
	65,15	36,60	average								
	27,00	13,05	st.dev.								

Επίπεδο Γ1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	x	
1	38	32	9	38	53					0	
2	44	51	19	37	44					0	
3	68	29	21	11	68					0	1
4	78	24	6	78	16					0	60,6
5	75	39	75	14	11					0	35,0
6	75	20	7	19	75					0	
7	87	12	5	87	8					0	
8	30	26	12	57	30					0	2
9	26	13	14	26	60					0	59,0
10	77	30	77	9	14					0	20,2
11	97	11	1	1	2	97	0			0	
12	96	12	1	1	0	0	96			0	
13	70	49	19	11	70	0	0			0	3
14	70	57	70	4	23	1	2			0	83,4
15	84	30	8	84	5	1	1			1	31,8
16	23	18	31	23	34	12				0	
17	78	32	3	4	15	78				0	
18	54	53	54	23	0	22				0	4
19	67	33	22	67	5	6				0	57,2
20	64	34	22	4	64	10				0	34,0
21	11	24									
22	60	25									
23	69	43									5
24	85	25									53,6
25	43	42									31,8
	62,76	30,56	average								
	22,88	12,90	st.dev.								

1.1.2 Ερμηνεία των μετρήσεων

1.1.2.1 Δείκτες ευκολίας

Η σύγκριση των μέσων όρων βαθμολογίας (επεξεργασμένων ή μη) που έχει λάβει το σύνολο των υποψηφίων σε κάθε δοκιμασία κάθε επιπέδου αποκαλύπτει μια μεταβλητή που είναι δύσκολο να αποφευχθεί στο μέτρο που η κατανόηση των αυθεντικών κειμένων περνά μέσα από δραστηριότητες αποκωδικοποίησης και αναγνώρισης οι οποίες, ίδιας φύσης αλλά διαφορετικής δυσκολίας και σπάνια προδίδει μια υπερβολική ευκολία (βλ. επίπεδο Γ1, δοκιμασία 3 στους παρακάτω πίνακες¹).

Ο υψηλός βαθμός ευκολίας των δοκιμασιών του επιπέδου Α δικαιολογείται από το γεγονός ότι τα θέματα απευθύνονται σε ένα κοινό αποτελούμενο από υποψηφίους επιπέδου Α1 ου Α2 και ότι, κατά το ήμισυ τουλάχιστον, οι δραστηριότητες πρέπει να αντιστοιχούν σε υποψηφίους επιπέδου Α1.

Μεταβλητές δεικτών ευκολίας στη δοκιμασία 1

	I.E. 2007	I.E. 2008
A	-	69,48
B1	69,32	68,30
B2	59,35	61,80
Γ1	61,08	60,63

Μεταβλητές δεικτών ευκολίας στη δοκιμασία 3

	I.E. 2007	I.E. 2008
A	-	72,60
B1	64,80	61,40
B2	55,00	65,15
Γ1	80,32	62,76

¹ Αυτή η μη φυσιολογική ευκολία μια συγκεκριμένης δραστηριότητας, εντοπίστηκε, και επανεξετάστηκε κατά τη σύνταξη των θεμάτων της εξεταστικής περιόδου Μαΐου 2008. Η βελτίωση αυτού του δείκτη ευκολίας ήταν εμφανής κατά την εξέταση των βαθμολογιών που δόθηκαν στη συγκεκριμένη εξεταστική περίοδο.

1.1.2.2 Δείκτες διάκρισης

Από την εξεταστική περίοδο του Μαΐου 2008, οι δείκτες διάκρισης έχουν σταθεροποιηθεί μεταξύ 30 και 40. Στη συνέχεια παρουσιάζονται οι προσπάθειες διερεύνησης των αιτιών αυτή της σχετικής μείωσης της ποιότητας των ερωτημάτων που συνθέτουν τη δοκιμασία 1 του επιπέδου Γ1.

Μεταβλητές δεικτών διάκρισης στη δοκιμασία 1

	I.D. 2007	I.D. 2008
A	-	36,24
B1	30,98	36,00
B2	37,05	31,32
Γ1	39,12	29,53

Μεταβλητές δεικτών διάκρισης στη δοκιμασία 3

	I.D. 2007	I.D. 2008
A	-	48,40
B1	48,00	38,95
B2	37,56	36,60
Γ1	24,48	30,56

1.1.2.3 Ποιότητα των παραπλανητών

Η ανάλυση του βαθμού καταλληλότητας των παραπλανητών, που υπολογίζεται με βάση την «επιτυχία» τους στα απαντητικά φύλλα, αποκαλύπτει μόνο κάποιες σποραδικές δυσλειτουργίες, οι οποίες δεν συγκεντρώνονται ποτέ σε μία και μόνο δραστηριότητα. Οι λόγοι για τους οποίους η επιλογή ορισμένων παραπλανητών μπόρεσε να μειώσει την αξιοπιστία ενός ερωτήματος έχουν εντοπιστεί.

1.1.2.4 Σταθερότητα των αποτελεσμάτων και αντιστοιχία μεταξύ αποτελεσμάτων

Από τις διαπιστώσεις αυτές προκύπτει ότι, αν παραβλέψουμε ορισμένες εσωτερικές δυσλειτουργίες που εντοπίστηκαν και διορθώθηκαν στις δοκιμασίες του Μαΐου, η ανάλυση των βαθμολογιών αποκαλύπτει μια ικανοποιητική σταθερότητα στα αποτελέσματα (μικρή διακύμανση από τη μια

χροιά στην άλλη) η οποία συνοδεύεται από μια αντιστοιχία μεταξύ των αποτελεσμάτων (μεταξύ δύο παράλληλων δοκιμασιών) καθώς το ζεύγος των δοκιμασιών, αντλούνται από μία τράπεζα η οποία τροφοδοτείται από την ίδια ομάδα συντακτών και οδηγούν σε αποτελέσματα με ικανοποιητικό βαθμό αντιστοιχίας.

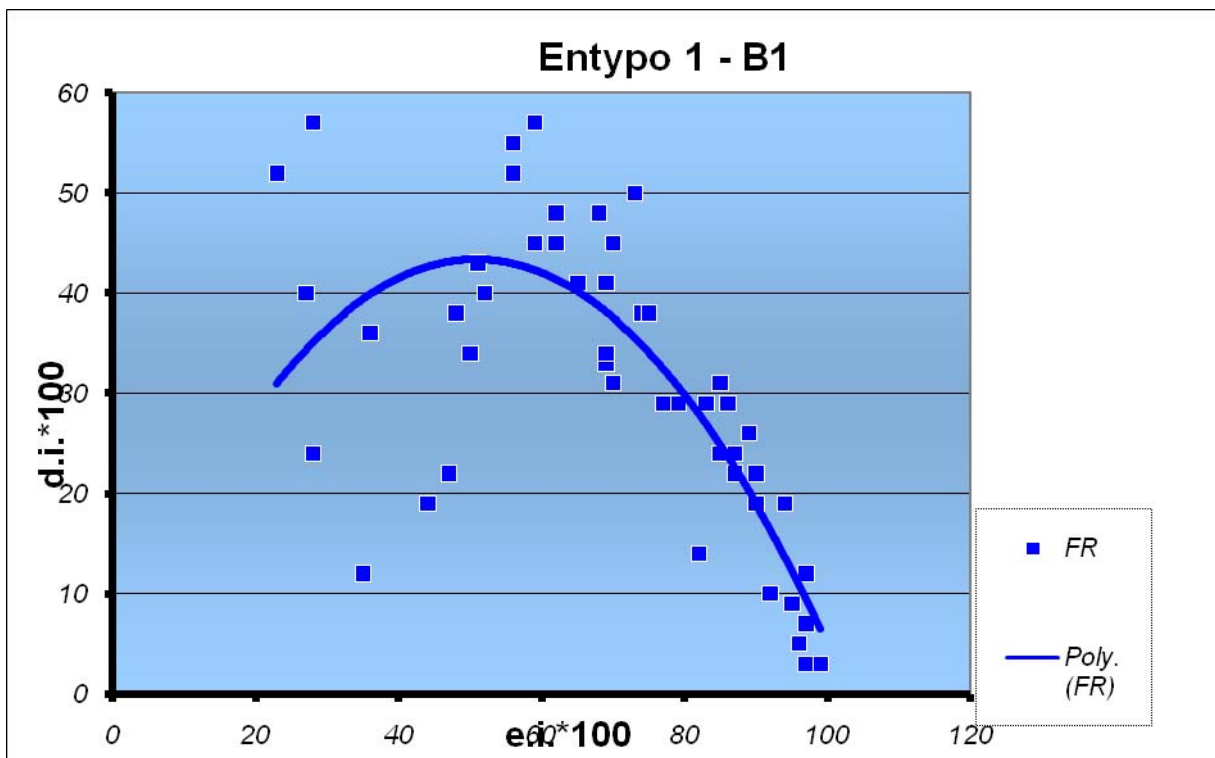
1.1.2.5 Ομοιογενοποίηση

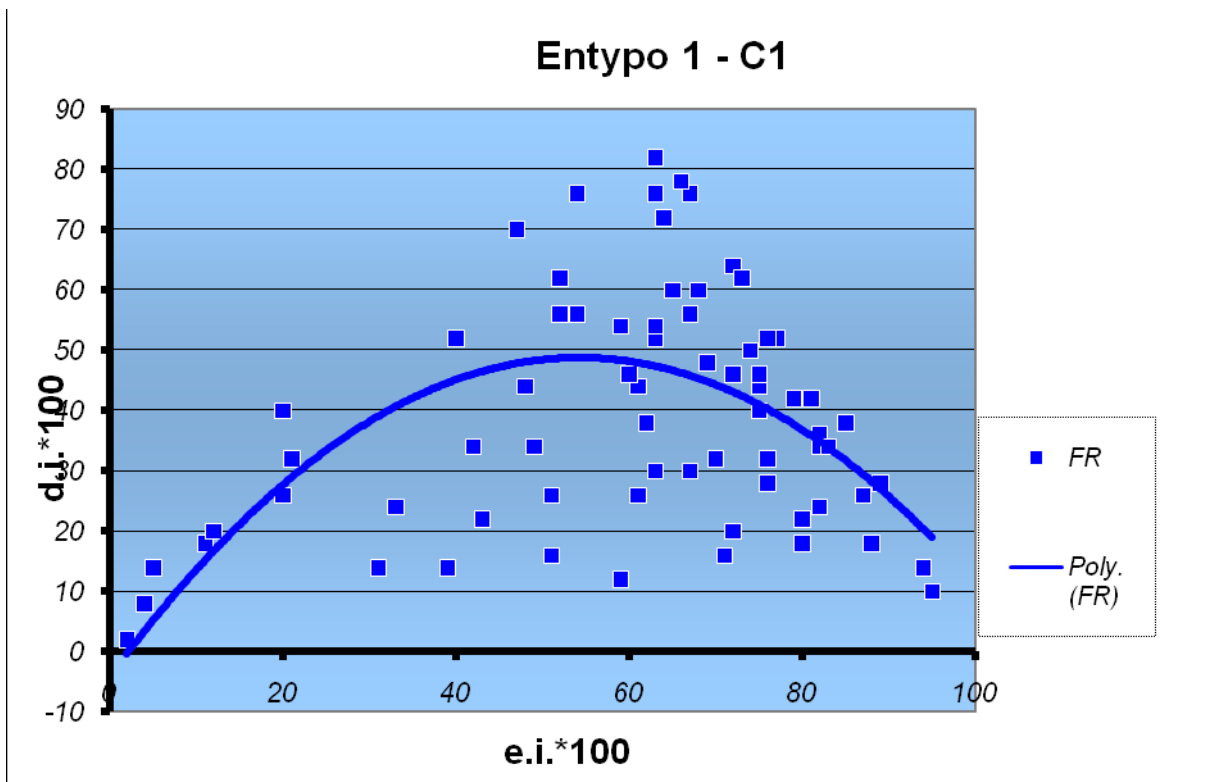
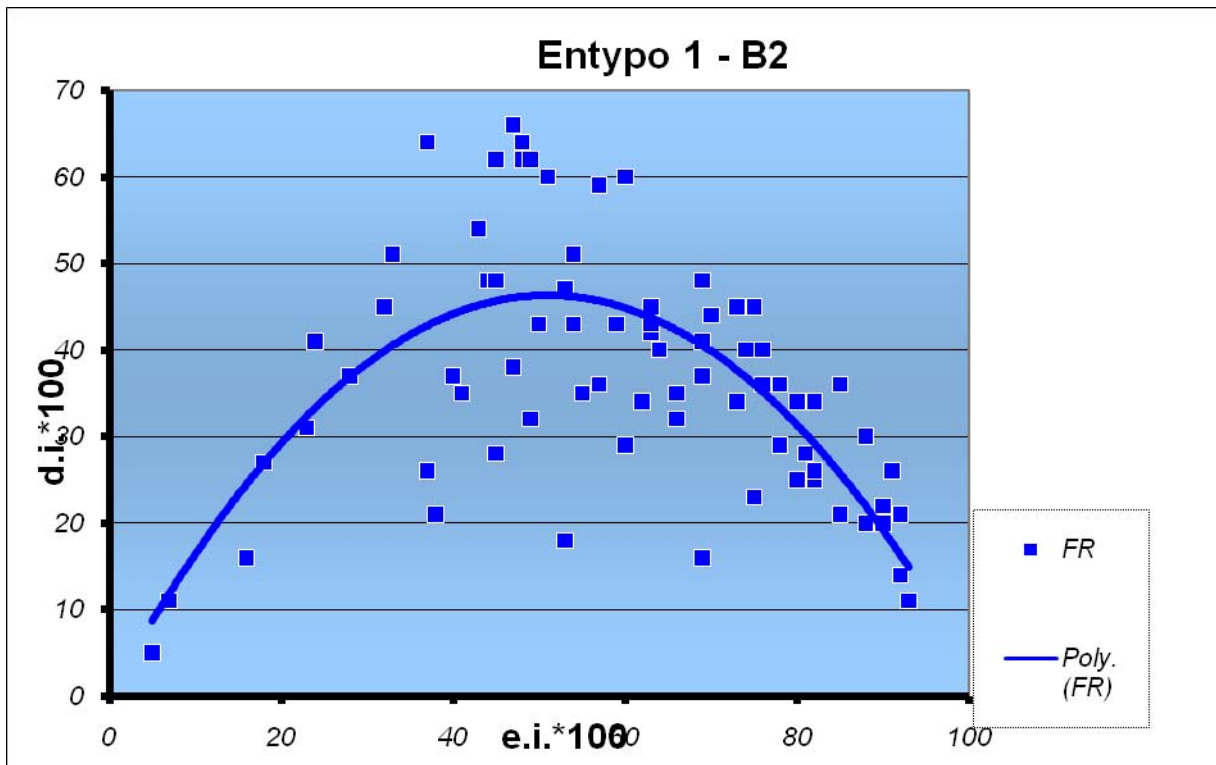
Οι φθίνουσες καμπύλες που ορίζονται από τη συσχέτιση μεταξύ των δεικτών ευκολίας και διάκρισης της δοκιμασίας 1 (εξεταστική περίοδος Μαΐου 2008) αποκαλύπτουν μία ικανοποιητική εσωτερική συνεκτικότητα εντός κάθε δοκιμασίας.

Όσον αφορά τη δοκιμασία 3, λαμβάνοντας υπόψη το μικρό αριθμό των ερωτημάτων που περιλαμβάνει, την ιδιαίτερη φύση των δεξιοτήτων που αξιολογούνται και των αστάθμητων τεχνικών παραγόντων, οι φθίνουσες καμπύλες που ορίζονται από τη συσχέτιση των δεικτών ευκολίας και διάκρισης αποκαλύπτουν επίσης μια ικανοποιητική ομοιογένεια

Συμπεριφορά των ερωτημάτων κατά τη δοκιμασία 1

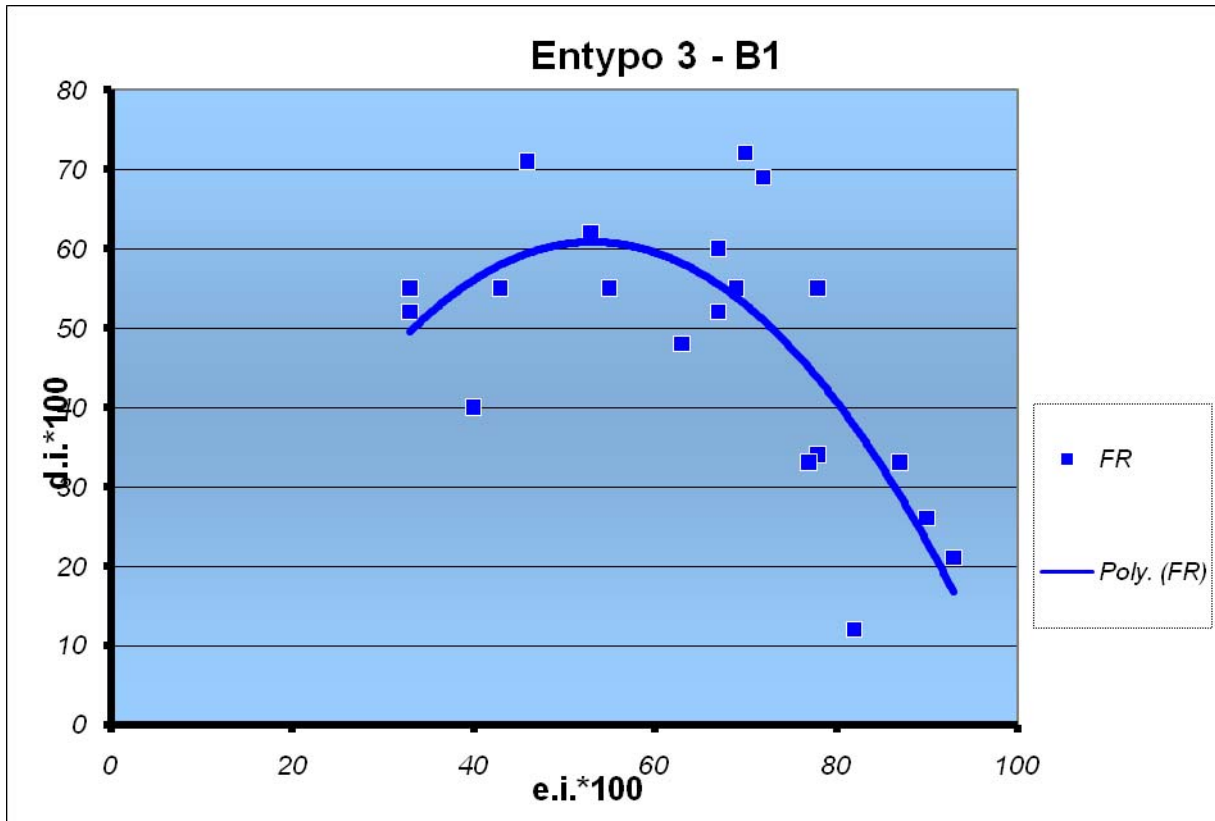
(εξεταστική περίοδος Μαΐου 2007)

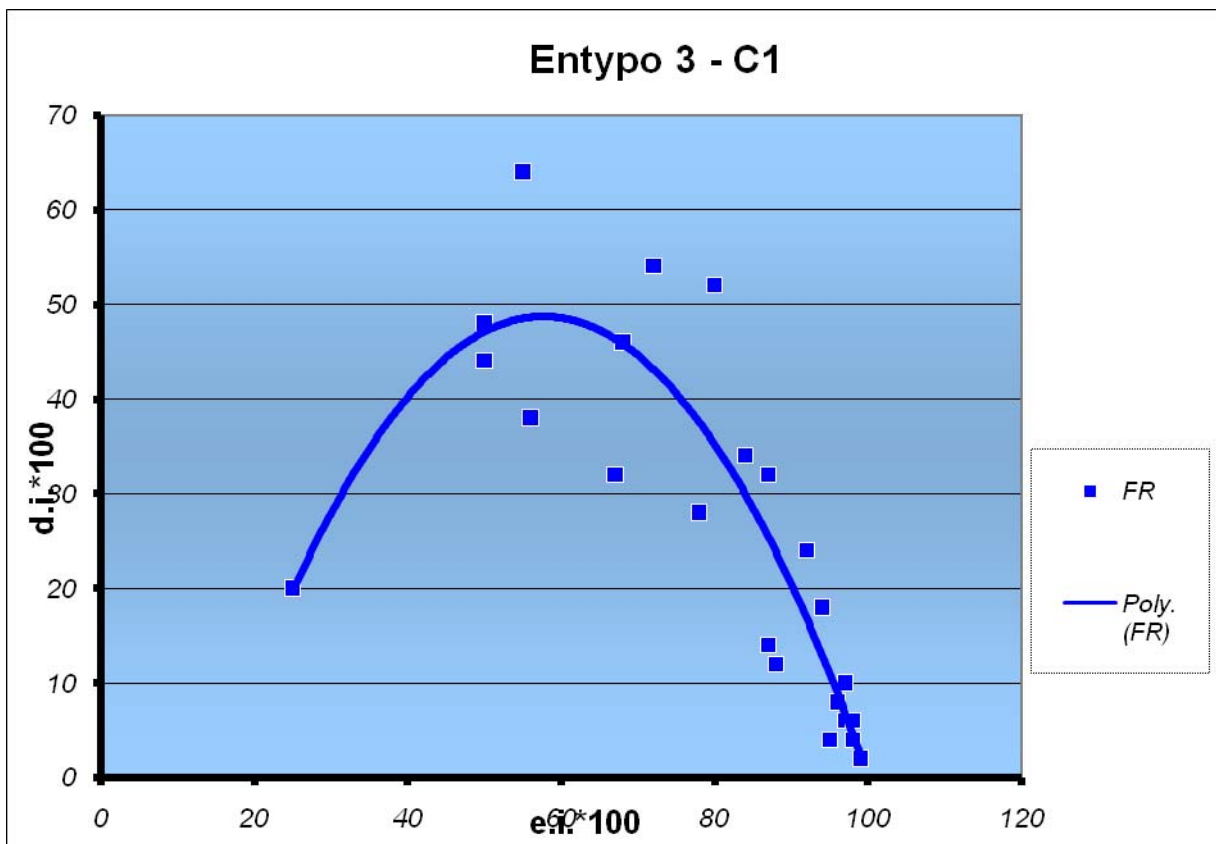
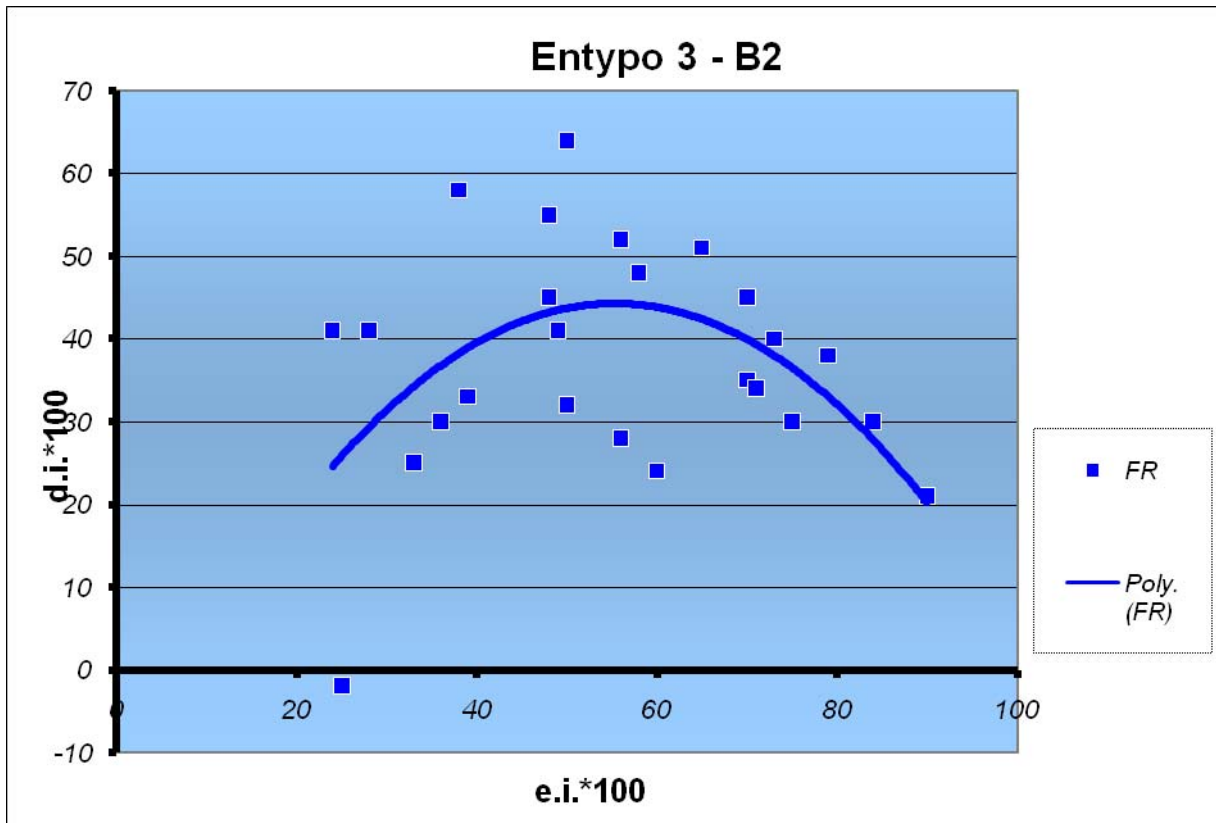




Συμπεριφορά των ερωτημάτων κατά τη δοκιμασία 3

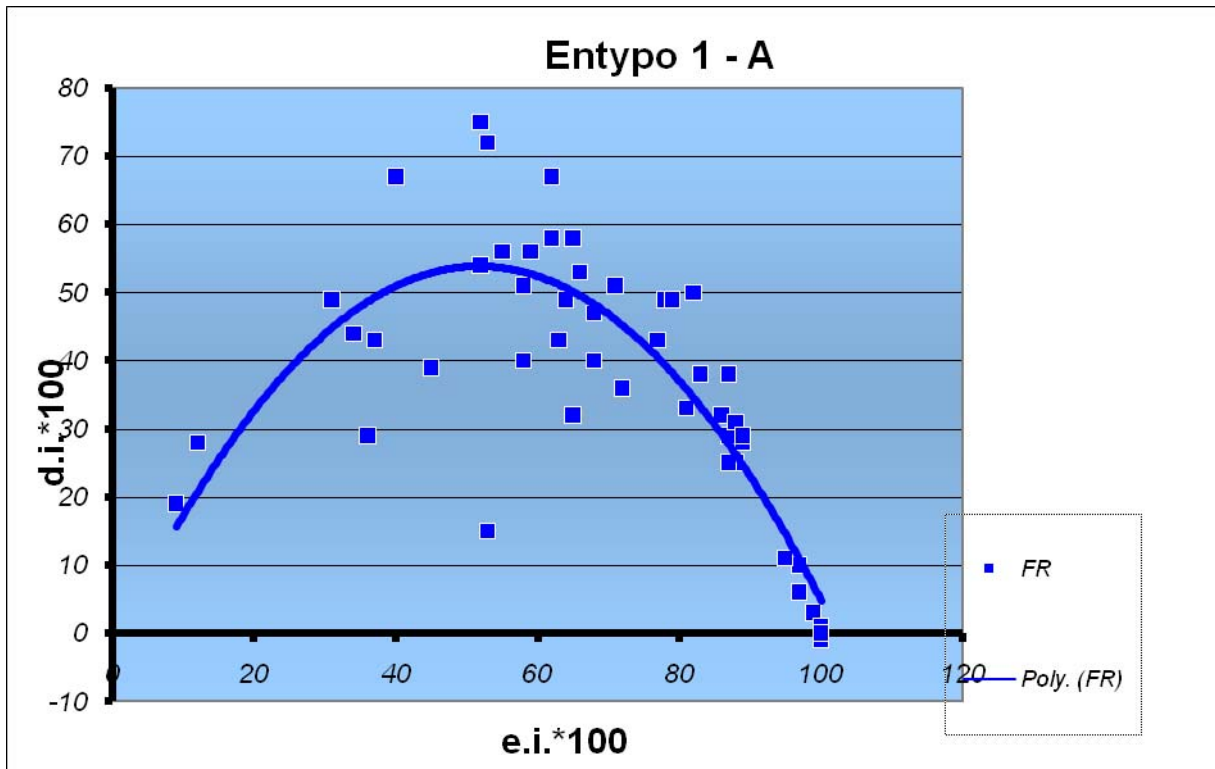
(εξεταστική περίοδος Μαΐου 2007)

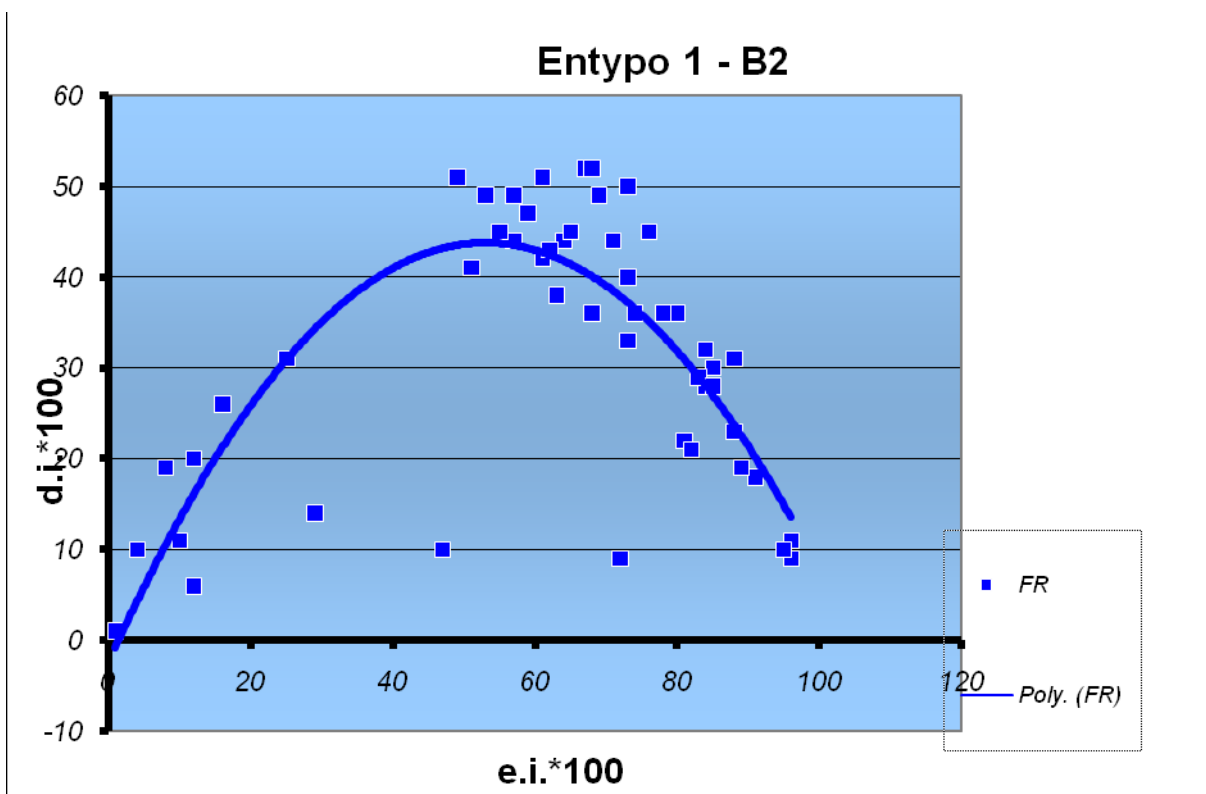
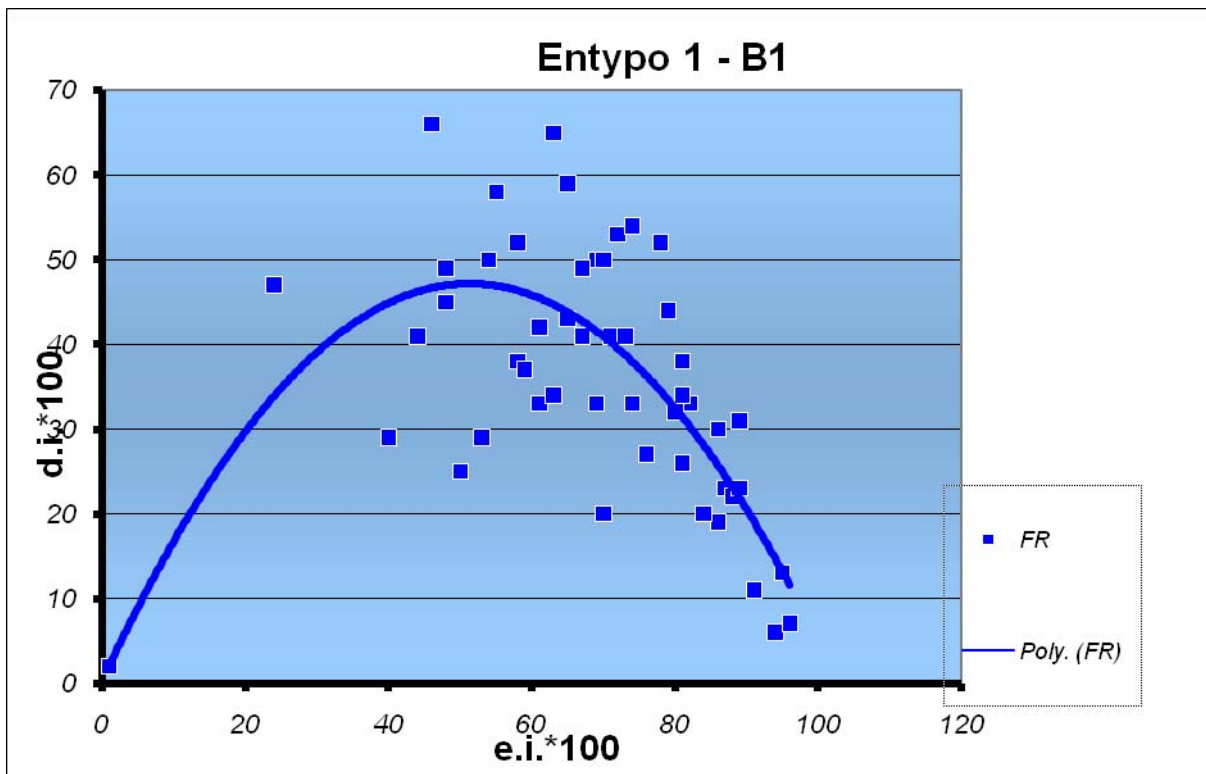


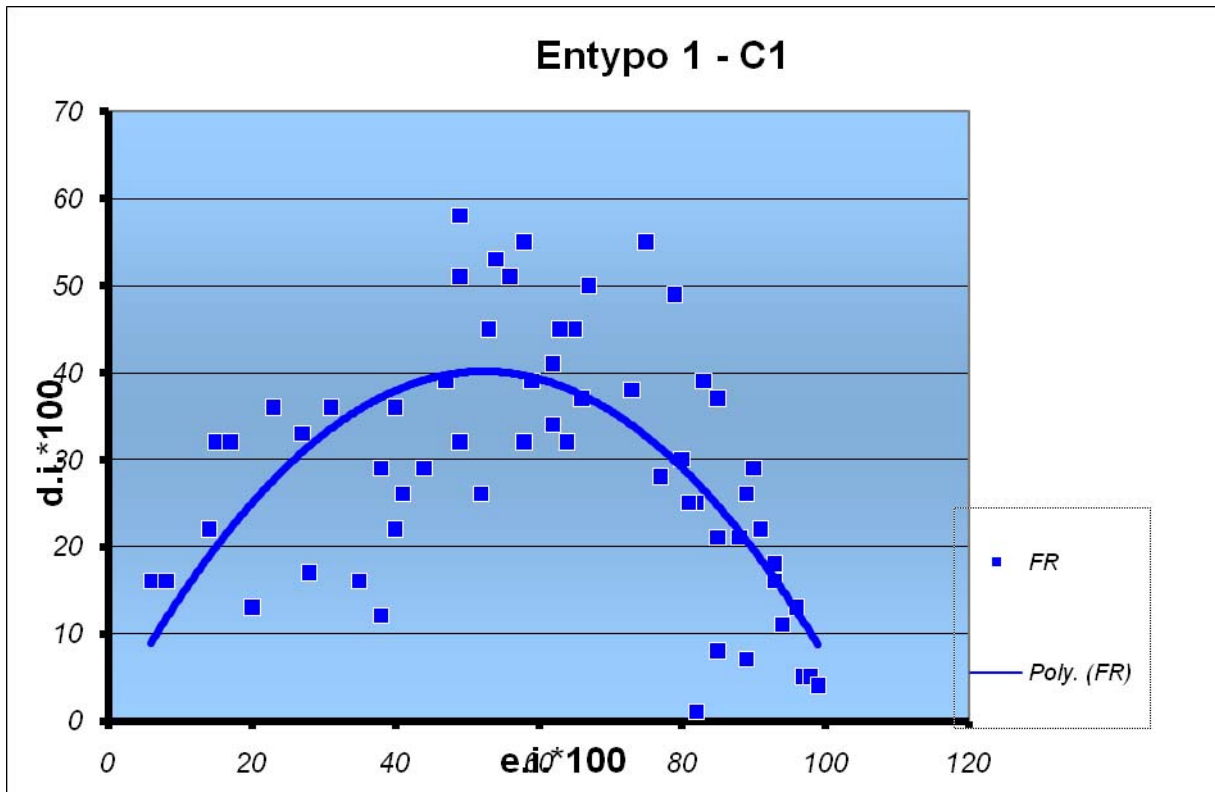


Συμπεριφορά των ερωτημάτων κατά τη δοκιμασία 1

(εξεταστική περίοδος Μαΐου 2008)

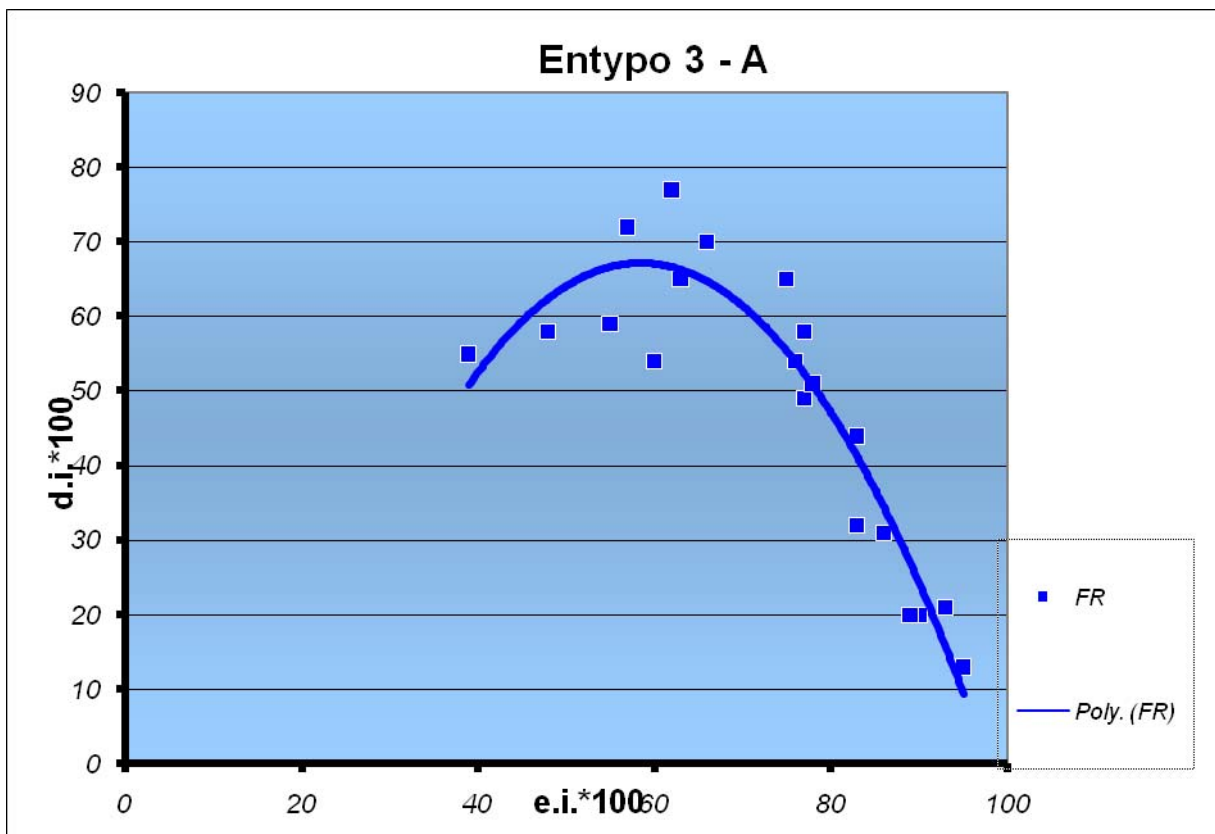


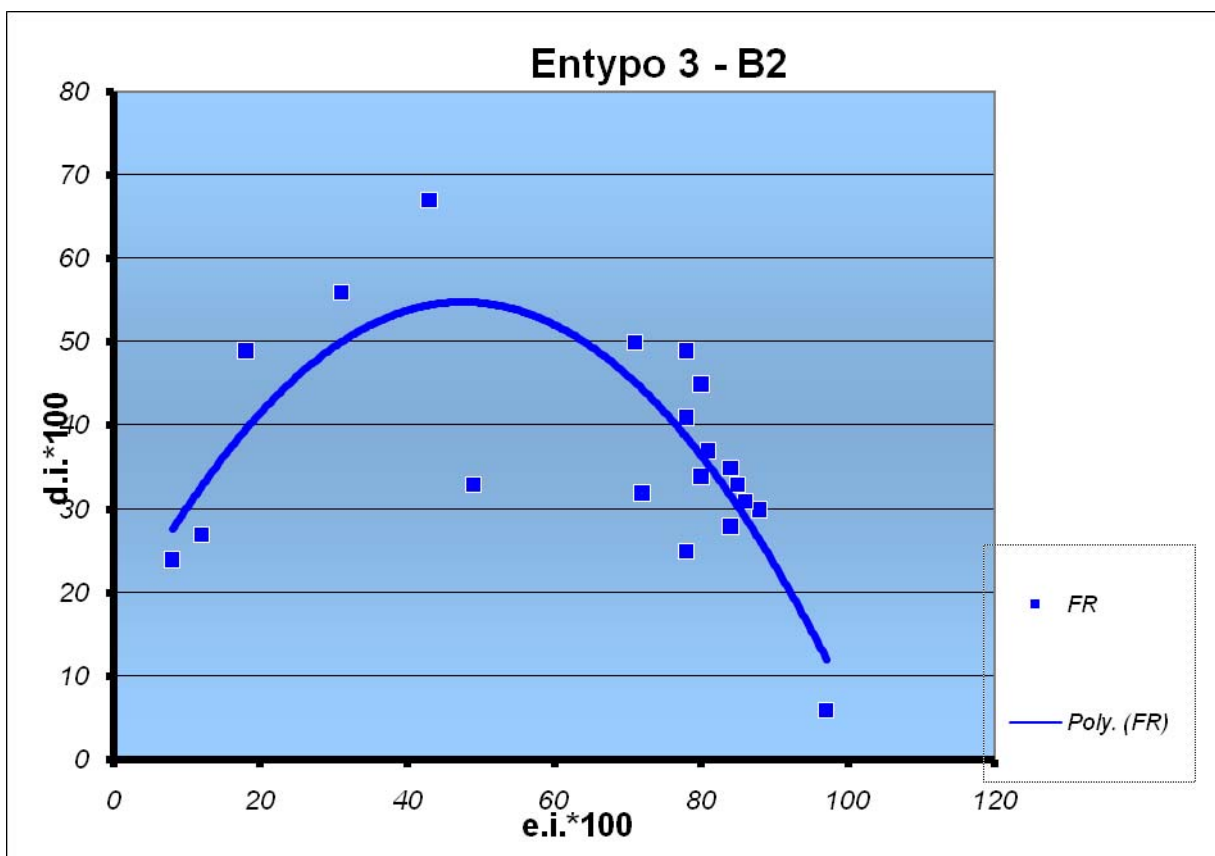
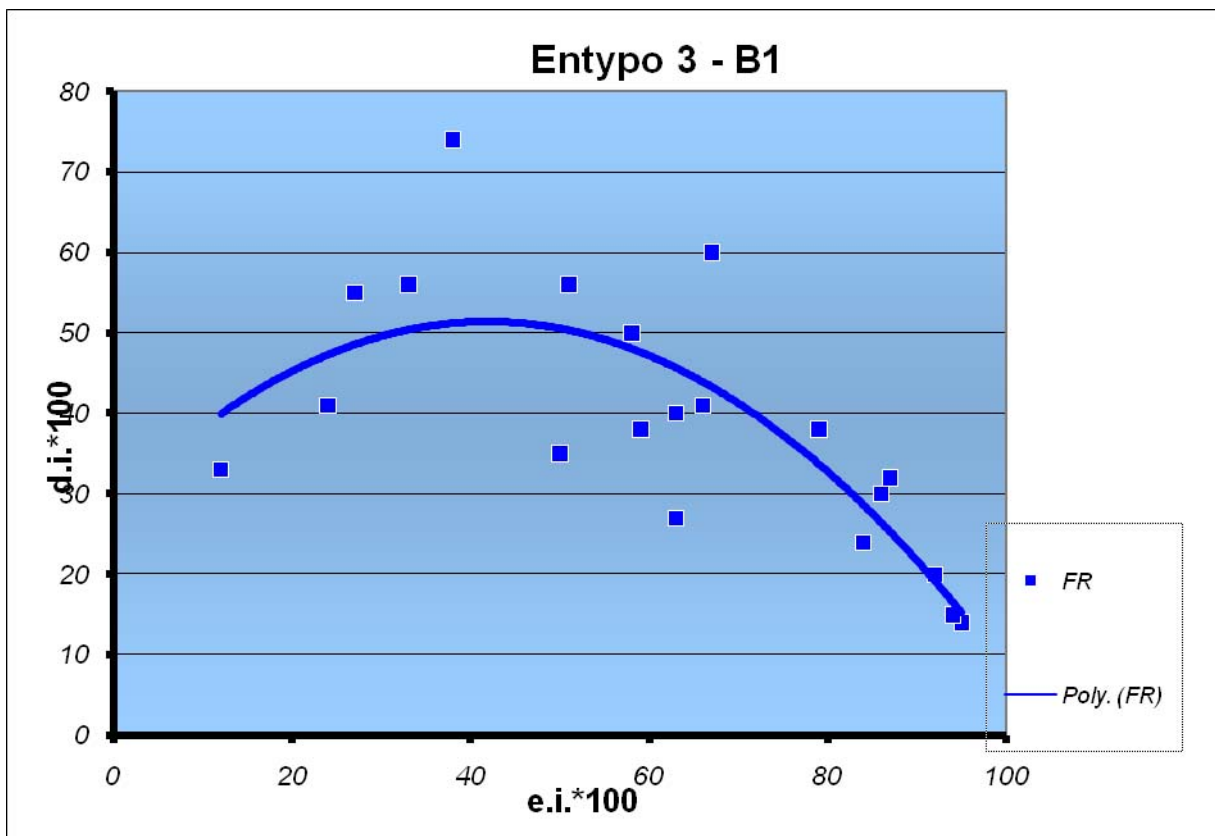


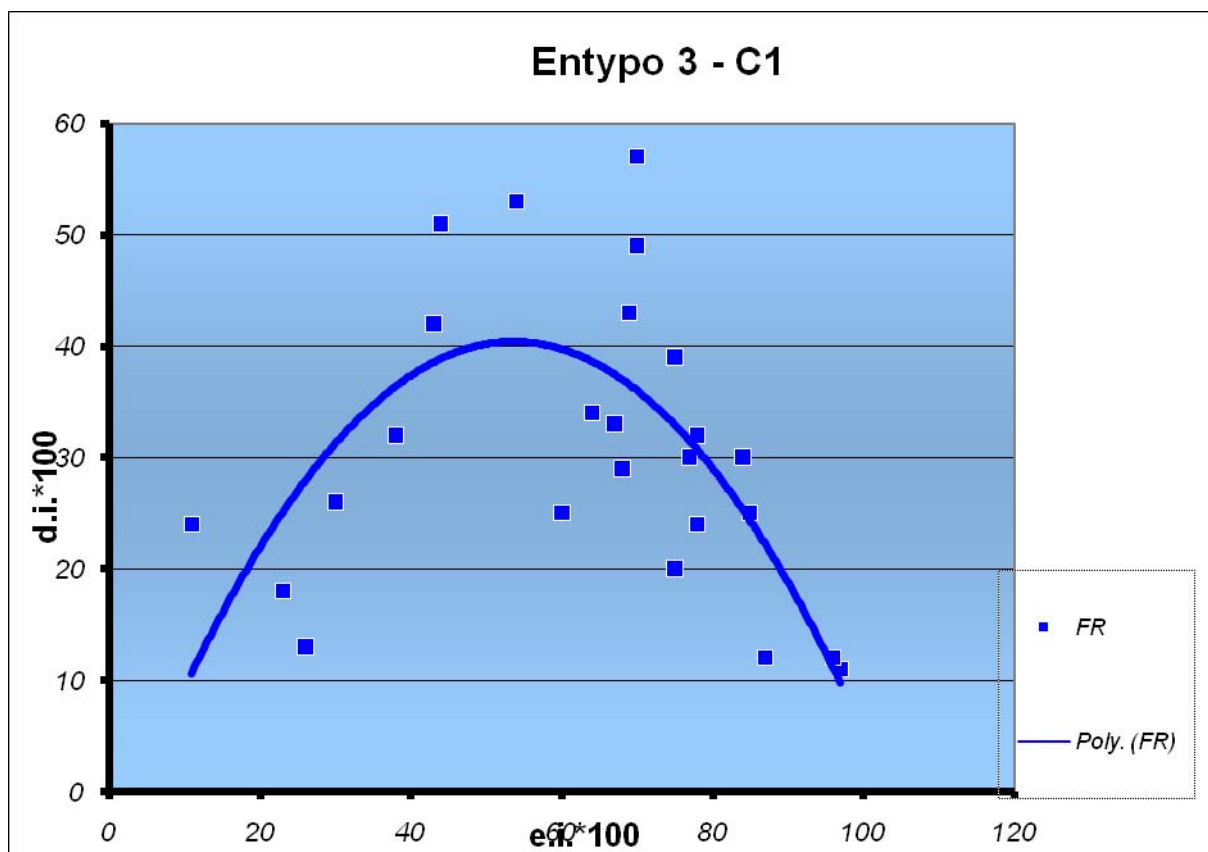


Συμπεριφορά των ερωτημάτων κατά τη δοκιμασία 3

(εξεταστική περίοδος Μαΐου
2007)







Η απογραφή των ερωτημάτων των οποίων η λειτουργία δεν ήταν ικανοποιητική κατά τις εξεταστικές περιόδους Μαΐου 2007 και Μαΐου 2008 αποκαλύπτουν ότι στη δοκιμασία 1, 85% των ερωτημάτων είχαν ικανοποιητική ποιότητα έναντι 90% το Μάιο 2008. Για τη δοκιμασία 3, 89% των ερωτημάτων είχαν ικανοποιητική ποιότητα τόσο το Μάιο 2007 όσο και το Μάιο 2008.

Οι παρεμβάσεις εσωτερικού ελέγχου φαίνεται ότι συμβάλλουν στη διατήρηση και βελτίωση της ποιότητας των ερωτημάτων.

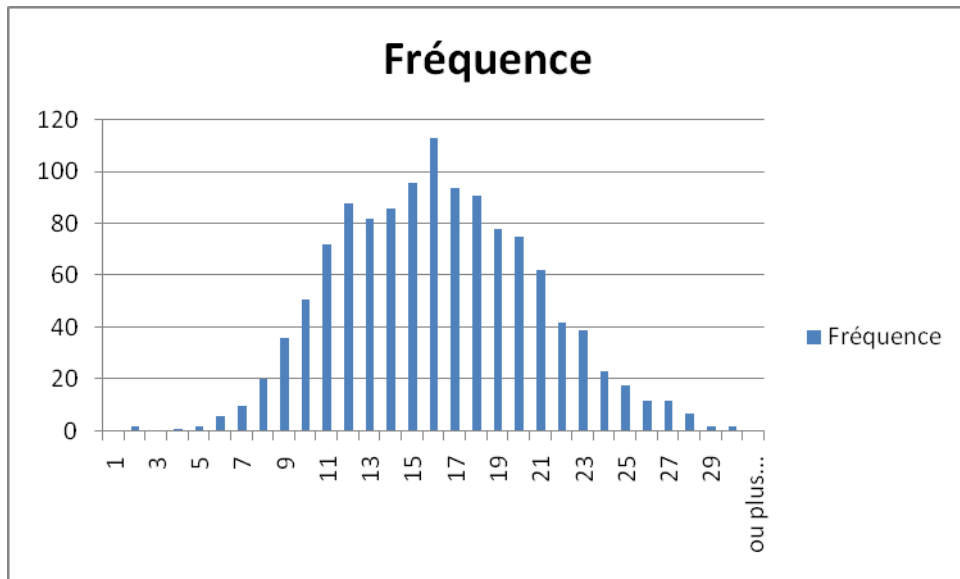
1.2 Ανάλυση βαθμολογιών - Δοκιμασίες 2 και 4

Οι αναλύσεις αφορούν τους βαθμούς μόνο των υποψηφίων της εξεταστικής περιόδου Μαΐου 2007.

1.2.1 Παρουσίαση των μετρήσεων

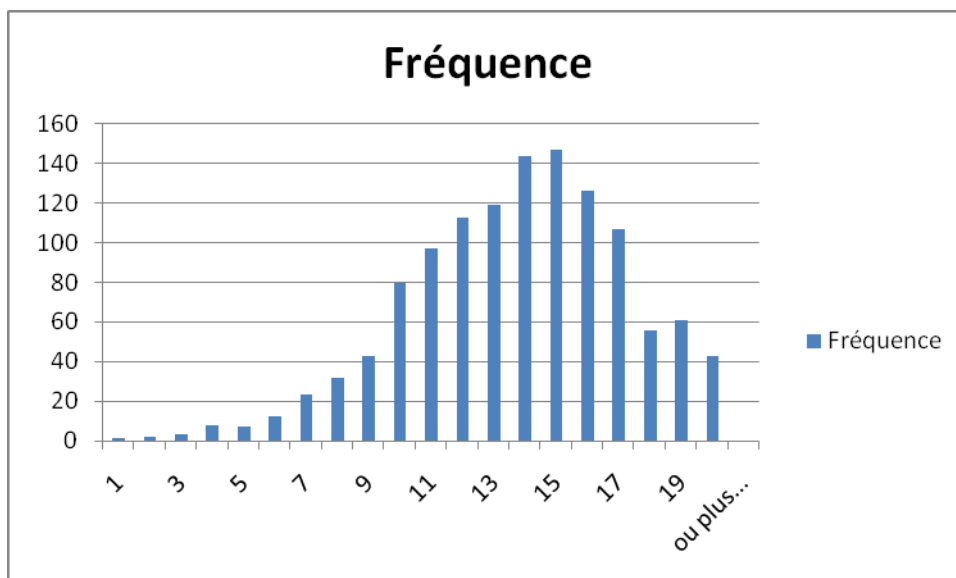
Αποτελέσματα της δοκιμασίας 2 :

	<i>Συχνότητα</i>
1	0
2	2
3	0
4	1
5	2
6	6
7	10
8	20
9	36
10	51
11	72
12	88
13	82
14	86
15	96
16	113
17	94
18	91
19	78
20	75
21	62
22	42
23	39
24	23
25	18
26	12
27	12
28	7
29	2
30	2



Αποτελέσματα της δοκιμασίας 4 :

	Συχνότητα
1	1
2	2
3	3
4	8
5	7
6	12
7	23
8	32
9	43
10	80
11	97
12	113
13	119
14	144
15	147
16	126
17	107
18	56
19	61
20	43



Τα σχήματα των καμπυλών διανομής είναι φυσιολογικά και σχετικά παρόμοια το ένα επίπεδο στο άλλο. Αυτό που είναι θετικό στην κατανομή των βαθμολογιών στη δοκιμασία 4 απορρέει από το γεγονός ότι η βεντάλια των βαθμολογιών που δόθηκαν (1 για την παρουσία, 2 για κάθε κείμενο που μη αναγνώσιμο, 3 για κάθε γραπτό εκτός θέματος κτλ.) είναι η ίδια, ανεξάρτητα από το μέγιστο βαθμό, και η βεντάλια των βαθμών που δόθηκαν είναι πιο περιορισμένη (με άριστα το 20).

1.2.2 Συνάφεια μεταξύ των βαθμολογητών

Ο υπολογισμός του μέσου όρου των βαθμών που δόθηκαν από τον πρώτο και το δεύτερο βαθμολογητή μας οδηγεί στο συμπέρασμα ότι οι βαθμολογητές δεν παρουσιάζουν ιδιαίτερη απόκλιση στο τρόπο που βαθμολογούν και ότι η ευαισθησία της βαθμολόγησής τους είναι συγκρίσιμη.

Μέσος όρος αποτελεσμάτων υποψηφίων στις δοκιμασίες γαλλικών

της εξεταστικής περιόδου Μαΐου 2007- Δοκιμασία 2

	Πρώτη Βαθμολόγηση	Δεύτερη Βαθμολόγηση	Συνολική Βαθμολογία
B1	15	14,6	14,8
	5,3	5,2	4,9
B2	16,1	16,1	16,1
	4,1	4,1	4,1
Γ1	15,8	17,5	16,6
	4,6	4,9	4,5

Η συνάφεια μεταξύ των βαθμών για το ίδιο γραπτό από τον κάθε βαθμολογητή είναι καλή, και γίνεται ακόμη και εξαιρετική (>.80) εάν τυποποιήσουμε τις βαθμολογίες επικεντρώνοντας στο μέσο όρο και μειώνοντας την απόκλισή τους στο 1.

Στην εξεταστική περίοδο Μαΐου 2007, εντοπίζουμε σε δύο περιπτώσεις βαθμολογίες για το ίδιο γραπτό σημαντικά αποκλίνουσες. Αυτές οι μη αποδεκτές αποκλίσεις μπορούν να εντοπισθούν με έναν ηλεκτρονικό έλεγχο.

Η αναφορά της ταυτότητας των βαθμολογητών στα δελτία βαθμολόγησης που δίνονται στη συντακτική ομάδα θα επέτρεπε μια πιο λεπτομερή ανάλυση των ενδεχόμενων αποκλίσεων, μια πιο ακριβή τυποποίηση της βαθμολογίας και μια αντικειμενική επιλογή των βαθμολογητών.

Μέσος όρος αποτελεσμάτων υποψηφίων στις δοκιμασίες γαλλικών

της εξεταστικής περιόδου Μαΐου 2007- Δοκιμασία 4

	Πρώτη Βαθμολόγηση	Δεύτερη Βαθμολόγηση	Συνολική Βαθμολογία
B1	13,8	13,5	13,6
	3,6	3,5	3,5
2B2	13,4	13,4	13,4
	3,4	3,3	3,3
Γ1	13,7	13,7	13,7
	3,9	3,8	3,87

Η συνάφεια μεταξύ των βαθμών που δίνονται στο ίδιο υποψήφιο για το ίδιο κριτήριο από κάθε αξιολογητή είναι αντίθετα ανεπαρκής. Ωστόσο, τα σύνολα των βαθμολογιών που δίνονται σε κάθε υποψήφιο ελάχιστα επηρεάζονται στο μέτρο που οι αποκλίσεις αλληλοεξουδετερώνονται. Είναι αδύνατο να εξηγήσουμε τις αποκλίσεις όσο η ταυτότητα των βαθμολογητών δεν κοινοποιείται στα δελτία των αποτελεσμάτων που δίνονται στους συντάκτες των θεμάτων. Αυτή η δυνατότητα τυποποίησης θα μπορούσε να επιτρέψει μελλοντικά τον αποκλεισμό των βαθμολογητών των οποίων οι βαθμοί παρουσιάζουν σημαντικές αποκλίσεις.

2 Έλεγχος εγκυρότητας

Με τον όρο εγκυρότητα ορίζουμε την καταλληλότητα που έχει ένα εργαλείο να καλύπτει από κάθε άποψη τους λόγους για τους οποίους δημιουργήθηκε. Υπάρχουν δύο τρόποι για να επαληθεύσουμε σε πιο βαθμό ισχύει αυτό.

- Ο υπολογισμός του συντελεστή συνάφειας (συντελεστής r του Bravais-Pearson) μεταξύ, για παράδειγμα,
 - ο των αποτελεσμάτων που δίνονται από το εργαλείο και
 - ο αυτών ενός άλλου εργαλείου που εφαρμόζεται στο ίδιο δείγμα, και του οποίου η εγκυρότητα είναι ήδη αναγνωρισμένη

Μπορούμε επίσης να λάβουμε υπόψη τους συντελεστές ομοιογένειας² ως δείκτες εγκυρότητας.

- Κατά την ανάπτυξη του εργαλείου, μπορούμε να δημιουργήσουμε ένα πίνακα εξειδίκευσης που να επιτρέπει
 - ο την αναγνώριση των συντελεστών του τομέα που μελετάμε,
 - ο την ανάδειξη της υπολανθάνουσας δομής και
 - ο την επιβεβαίωση μιας όσο πιο ικανοποιητικής γίνεται συσχέτισης μεταξύ των συνιστωσών με αυτές του εργαλείου που χρησιμοποιήθηκε.

Ζητούμε στη συνέχεια από ειδικούς να πουν τη γνώμη τους για την καταλληλότητα των στοιχείων σε σχέση με το αξιολογούμενο αντικείμενο.

2.1 Εγκυρότητα της δομής

Ο υπολογισμός των συντελεστών ομοιογένειας βασίζεται στη συνάφεια μεταξύ του κάθε ερωτήματος του τεστ και της συνολικής βαθμολογίας στο τεστ που θεωρείται ως η καλύτερη μέτρηση της δομής που υποτίθεται ότι μετρά το τεστ. Επιλέγουμε λοιπόν τα ερωτήματα βάσει της συνεισφοράς τους στη διατήρηση της ομοιογένειας της μέτρησης. Αυτή η ίδια τεχνική μπορεί να χρησιμοποιηθεί για να εξασφαλίσει την «εγκυρότητα» των υποδοκιμασιών των οποίων τα αποτελέσματα θα έπρεπε

- Να είναι σε υψηλή συνάφεια (συγκλίνουσα συνάφεια) με τη συνολική βαθμολογία αλλά
- Να παρουσιάζουν χαμηλή συνάφεια (αποκλίνουσα συνάφεια) μεταξύ των υποδοκιμασιών.

² Είναι σημαντικό να σημειώσουμε ότι οι συντελεστές ομοιογένειας είναι πριν απ' όλα δείκτες συνεκτικότητας της μέτρησης, και κατά συνέπεια, πρέπει να θεωρούνται ως έμμεσοι δείκτες που επιτρέπουν την διατύπωση μιας υπόθεσης εγκυρότητας η οποία θα πρέπει οπωσδήποτε να επαληθευτεί κατά τις διαδικασίες εγκυροποίησης του περιεχομένου για τις οποίες θα γίνει λόγος παρακάτω.

Παρακάτω παρουσιάζονται τα αποτελέσματα του υπολογισμού του συντελεστή ομοιογένειας r του Bravais-Pearson στο σύνολο του πληθυσμού των εξεταζομένων της εξεταστικής περιόδου Μαΐου 2007 (1225 υποψήφιοι μετά τον αποκλεισμό των μη κατάλληλων εγγραφών).

	<i>δοκ1</i>	<i>δοκ2</i>	<i>δοκ3</i>	<i>δοκ4</i>	<i>συν</i>
<i>δοκ1</i>	1				
<i>δοκ2</i>	0,55367075	1			
<i>δοκ3</i>	0,51888612	0,37877027	1		
<i>δοκ4</i>	0,39058116	0,35436232	0,36203062	1	
<i>συν</i>	0,92507016	0,72791334	0,69961487	0,58614429	1

Σε μια καλύτερα αναγνώσιμη μορφή :

	<i>δοκ1</i>	<i>δοκ2</i>	<i>δοκ3</i>	<i>δοκ4</i>	<i>συν</i>
<i>δοκ1</i>	1				
<i>δοκ2</i>	.55	1			
<i>δοκ3</i>	.51	.37	1		
<i>δοκ4</i>	.39	.35	.36	1	
<i>συν</i>	.92	.72	.69	.58	1

Ο πίνακας των διακυμάνσεων που παρουσιάζεται παραπάνω επιβεβαιώνει την ύπαρξη μιας θετικής και συγκλίνουσας συνάφειας των βαθμολογιών που δόθηκαν σε κάθε δοκιμασία με την τελική βαθμολογία. Οι συνάφειες μεταξύ των βαθμολογιών κάθε δοκιμασίας είναι σαφώς πιο ασθενείς και παρουσιάζουν μια σημαντική απόκλιση.

Παρατηρούμε ότι η συνάφεια μεταξύ των βαθμολογιών που δόθηκαν στη δοκιμασία 1 και της τελικής βαθμολογίας είναι ιδιαίτερα αποκλίνουσα (.92) και ότι οι βαθμολογίες που δόθηκαν στη δοκιμασία 4 είναι σε ασθενέστερη συνάφεια με τις βαθμολογίες που δόθηκαν στις άλλες δοκιμασίες.

Οι πίνακες που πραγματοποιήθηκαν ανά επίπεδο και ανά εξεταστική περίοδο είναι αισθητά οι ίδιες. Ο υπολογισμός των συντελεστών ομοιογένειας μοιάζει να μας επιτρέπει την επιστημονική διατύπωση της υπόθεσης της εγκυρότητας της δομής του εργαλείου αξιολόγησης.

2.2 Εγκυρότητα του περιεχομένου

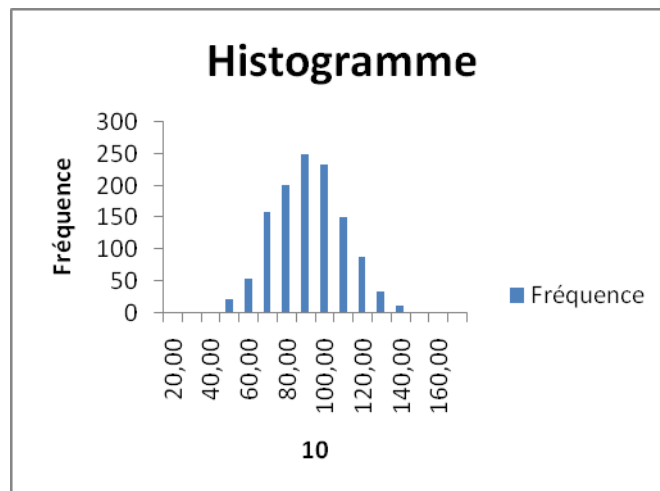
Η εγκυρότητα του περιεχομένου συνίσταται στην ανάλυση του περιεχομένου των δοκιμασιών της εξέτασης, με σκοπό να εξασφαλίσει ότι είναι αντιπροσωπευτικό, ως δείγμα της δεξιότητας που αξιολογείται.

Αυτή η ουσιαστική ανάλυση έχει ήδη πραγματοποιηθεί κατά τη συσχέτιση του κοινού πλαισίου του ΚΠΓ και των προτεινόμενων δραστηριοτήτων των εξετάσεων του ΚΠΓ με το Κοινό Ευρωπαϊκό Πλαίσιο Αναφοράς (βλ. παραδοτέα του ΠΕ6) . Αυτή η συσχέτιση πραγματοποιήθηκε βάσει του σχετικού «Οδηγού» (*Manual for Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching and Assessment*) που έχει εκδώσει το Τμήμα Σύγχρονων Γλωσσών του Συμβουλίου της Ευρώπης.

3 Έλεγχος συνάφειας αποτελεσμάτων εξέτασης και προσδοκιών

Η καταλληλότητα των αποτελεσμάτων των εξετάσεων σε σχέση με τις προσδοκίες των υποψηφίων και των εκπαιδευτών τους μπορεί να υπολογιστεί με ένα λιγότερο υποκειμενικό τρόπο από την διοργάνωση μιας σφυγμομέτρησης μεταξύ των εμπλεκόμενων ατόμων, με μια μελέτη της κατανομής των βαθμολογιών και των αποφάσεων επιτυχίας ανά επίπεδο, ανά δοκιμασία, ανά εξεταζόμενη γλώσσα³, ανά έτος κτλ..

Κατανομή των βαθμολογιών κατά την εξεταστική περίοδο Μαΐου 2007



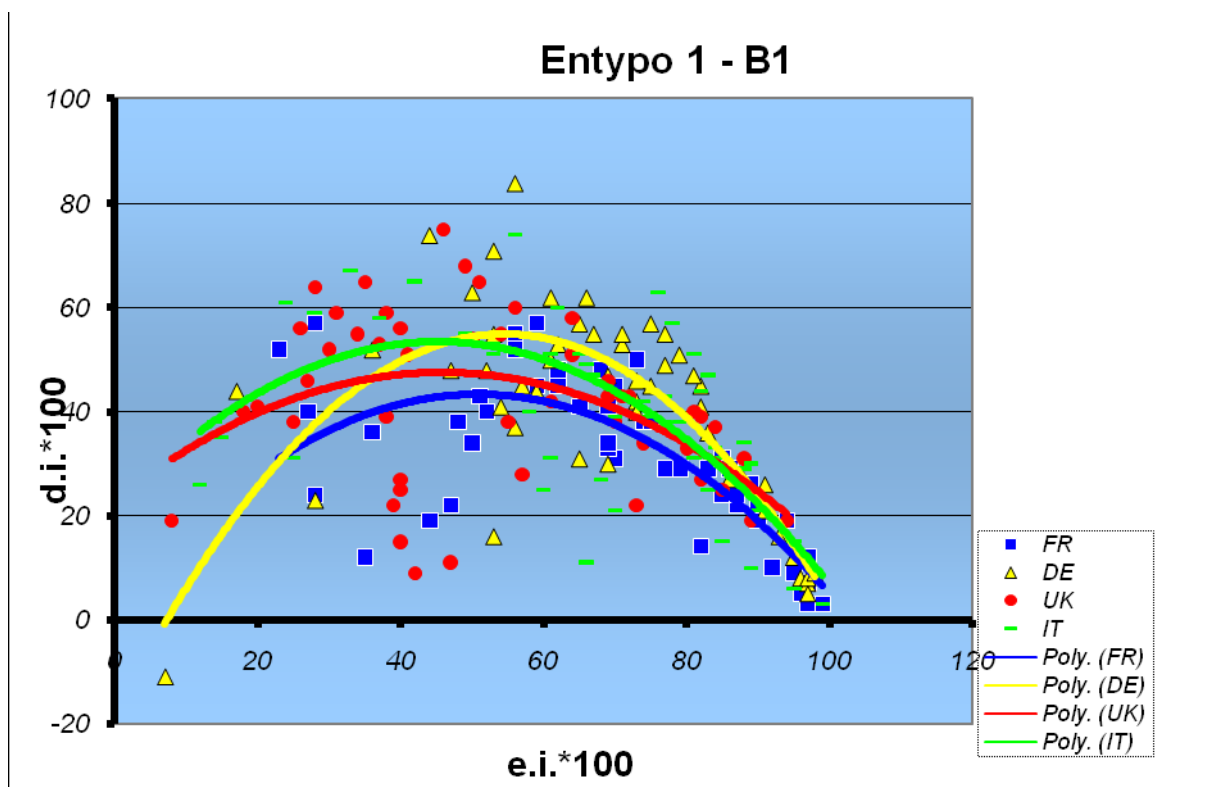
Ανεξάρτητα από το κριτήριο ομαδοποίησης που υιοθετήθηκε, η κατανομή που παρουσιάζεται είναι φυσιολογική όπως φαίνεται ενδεικτικά από το ιστόγραμμα την κατανομής των συνολικών βαθμολογιών του συνόλου των υποψηφίων που παρουσιάζεται παραπάνω.

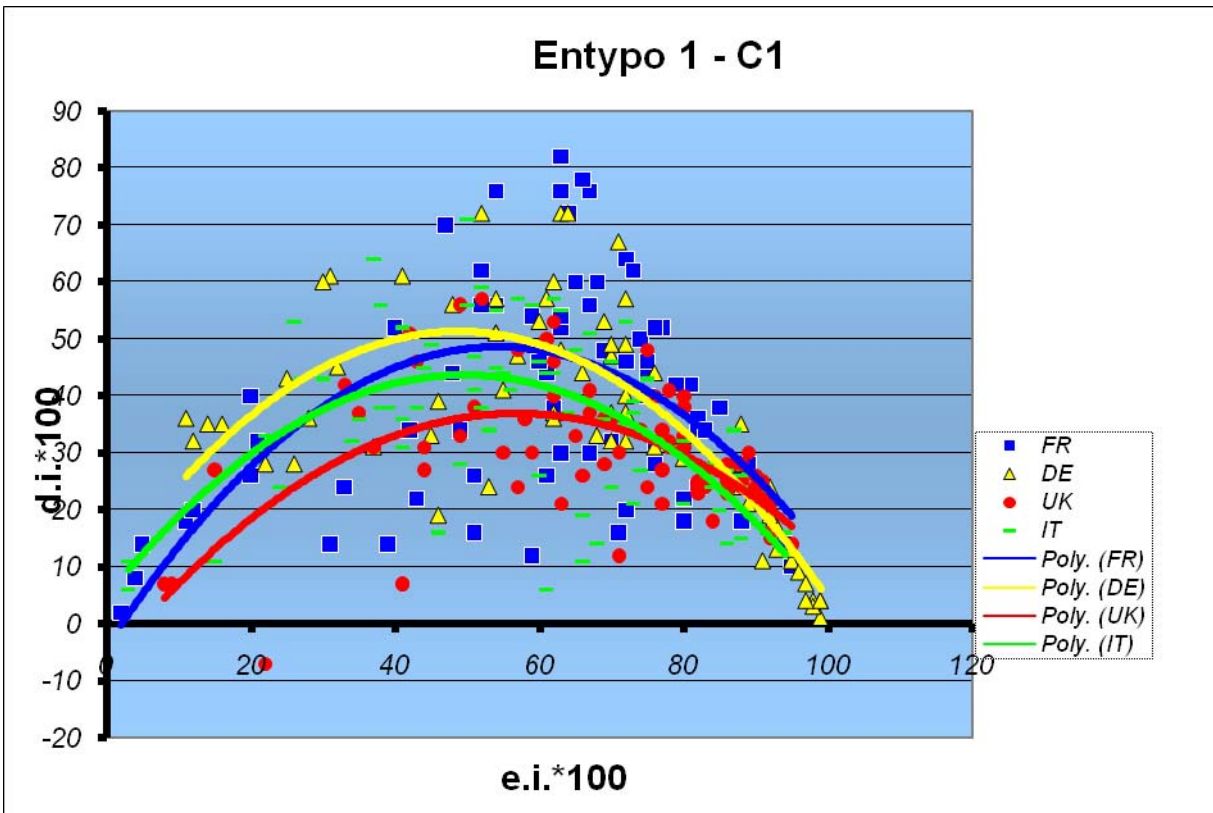
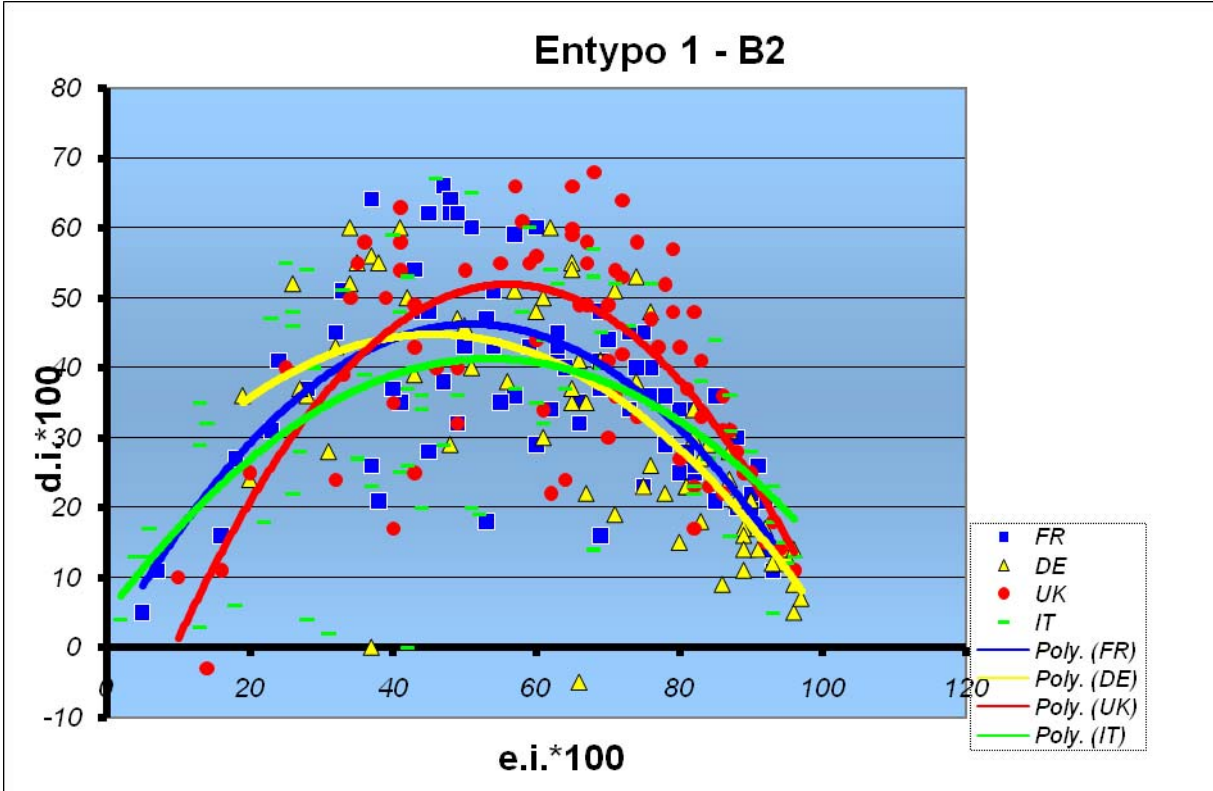
³ Προκειμένου να καταστεί δυνατή η σύγκριση, να σχετικοποιηθεί η σημαντικότητα των παρατηρήσεων, να υπάρξει ένα σημείο αναφοράς

4 Ερμηνεία των αποτελεσμάτων των ελέγχων

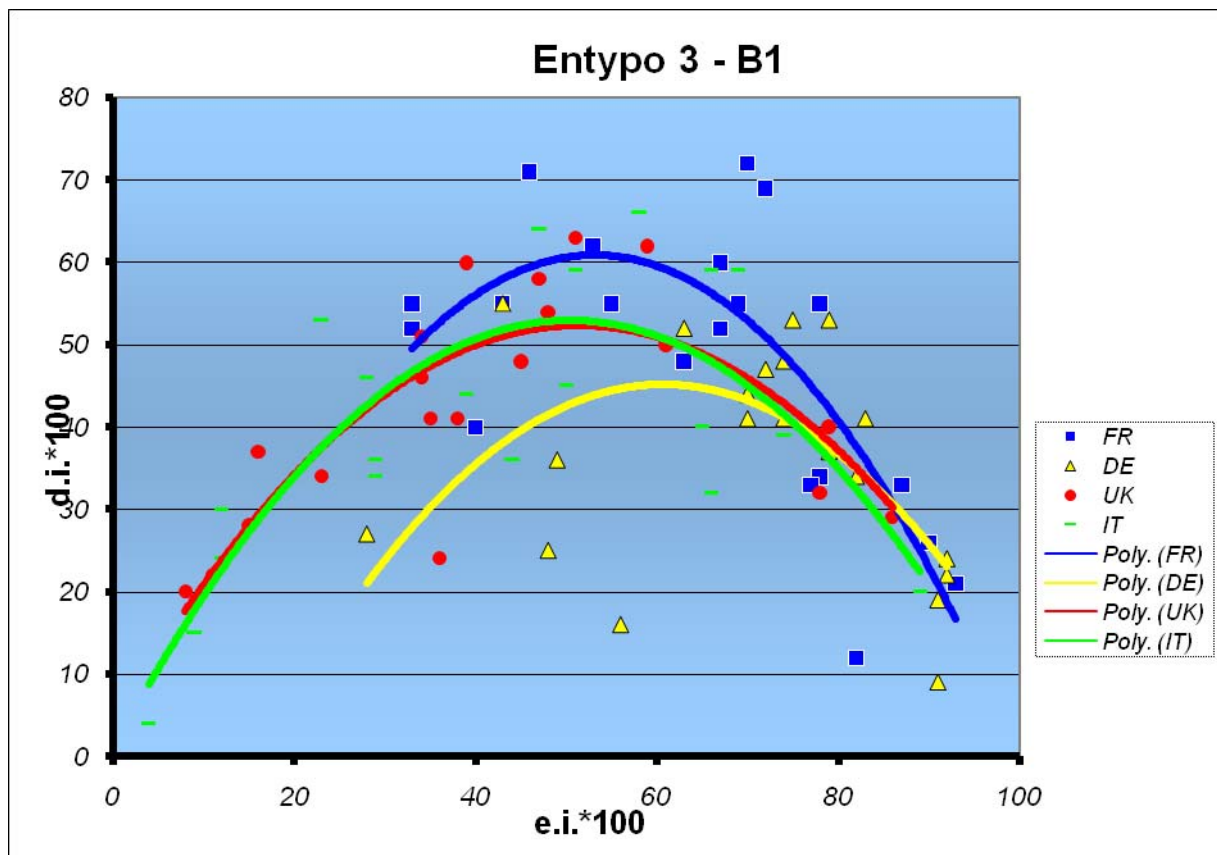
Η επιβεβαίωση των αποτελεσμάτων των διαφορετικών ελέγχων ποιότητας που παρουσιάστηκαν στις προηγούμενες σελίδες μπορεί να συσχετιστεί με αυτά των ελέγχων που πραγματοποιήθηκαν κατά τις επανειλημμένες έρευνες από τις ομάδες σύνταξης θεμάτων για τα ιταλικά, τα αγγλικά και τα γερμανικά. Έτσι, η παρουσίαση των ελλειπτικών καμπυλών που σχηματίζονται από τη συσχέτιση μεταξύ των δεικτών ευκολίας και διάκρισης για τα ερωτήματα των δοκιμασιών για την εξεταστική περίοδο Μαΐου 2007 για όλες τις εξεταζόμενες γλώσσες της συγκεκριμένης περιόδου μας επιτρέπουν να διαπιστώσουμε ότι το εργαλείο της πιστοποίησης του ΚΠΓ είναι στο σύνολό του ομοιογενές για όλες τις γλώσσες, όσον αφορά τουλάχιστον τα μετρολογικά χαρακτηριστικά των δοκιμασιών 1 και 3, καθώς και των ερωτημάτων που τις συνθέτουν.

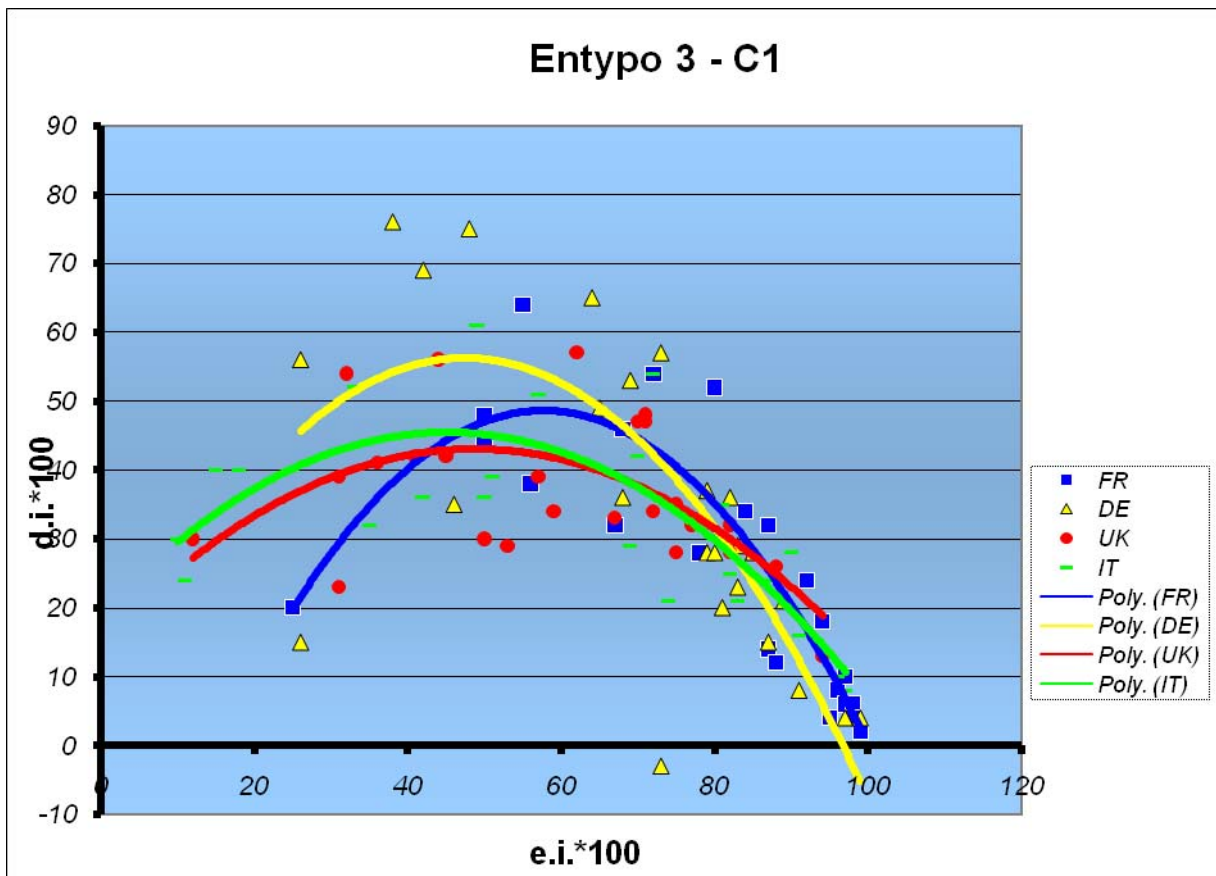
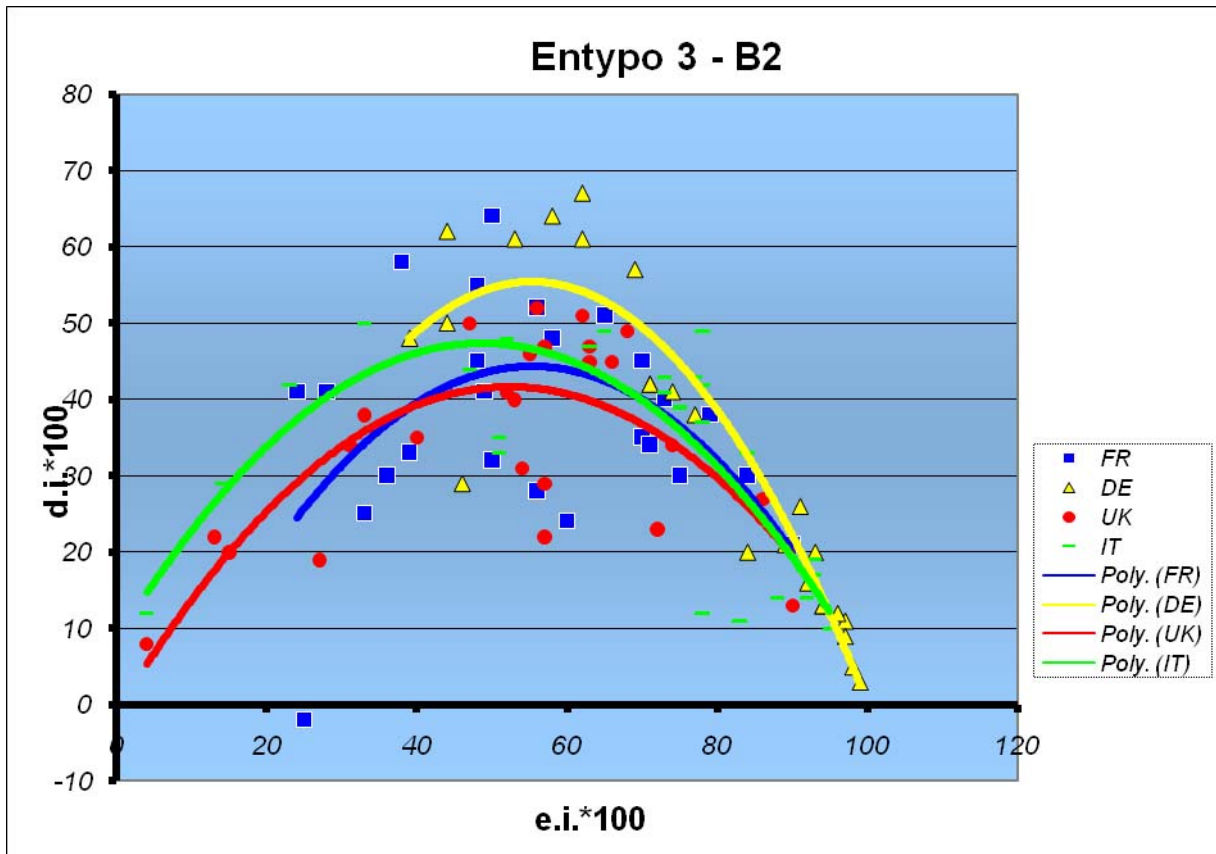
Συμπεριφορά των ερωτημάτων της δοκιμασίας 1 (εξεταστική περίοδος Μαΐου 2007)



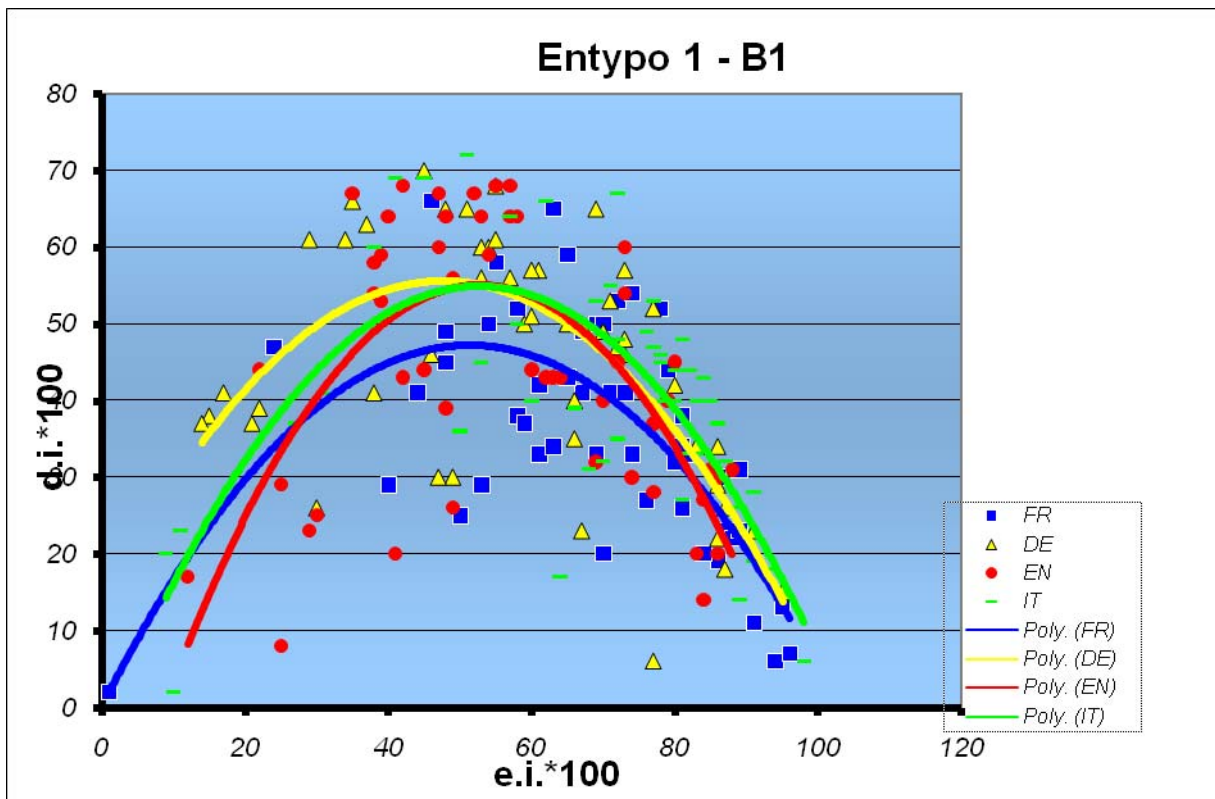
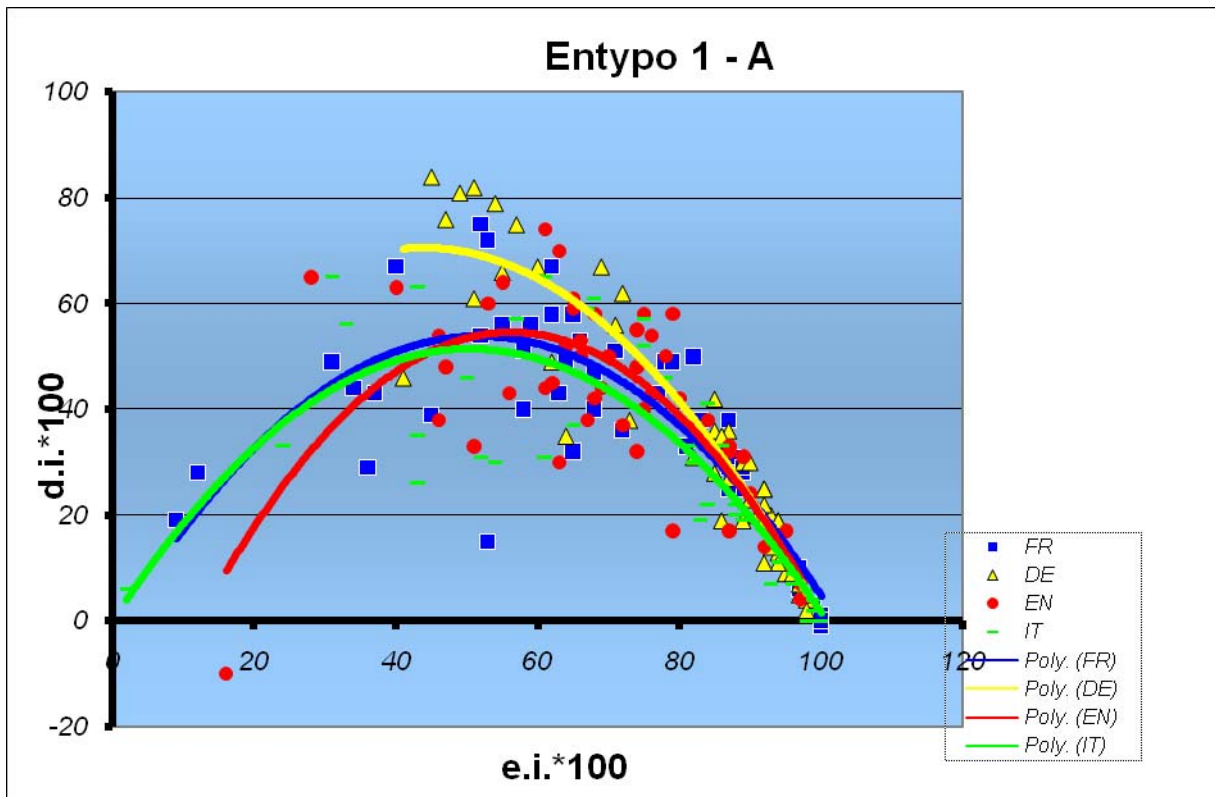


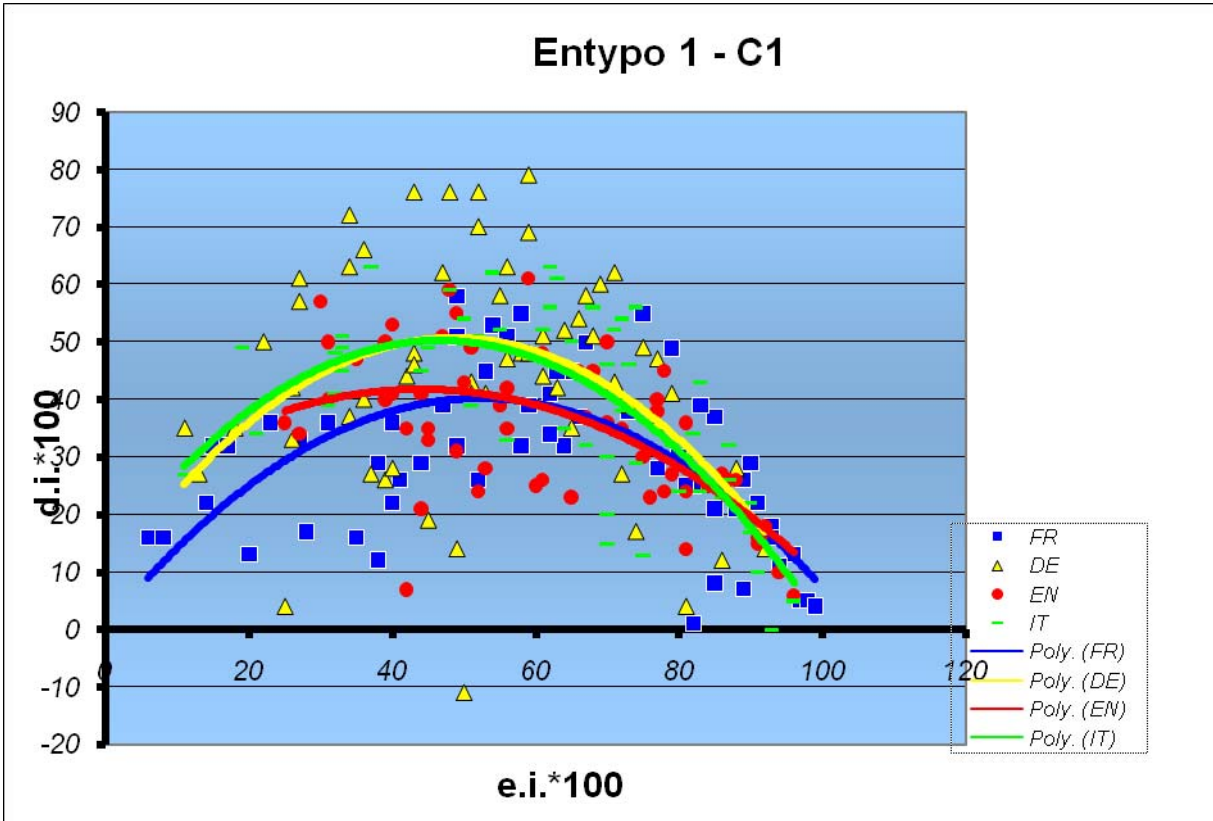
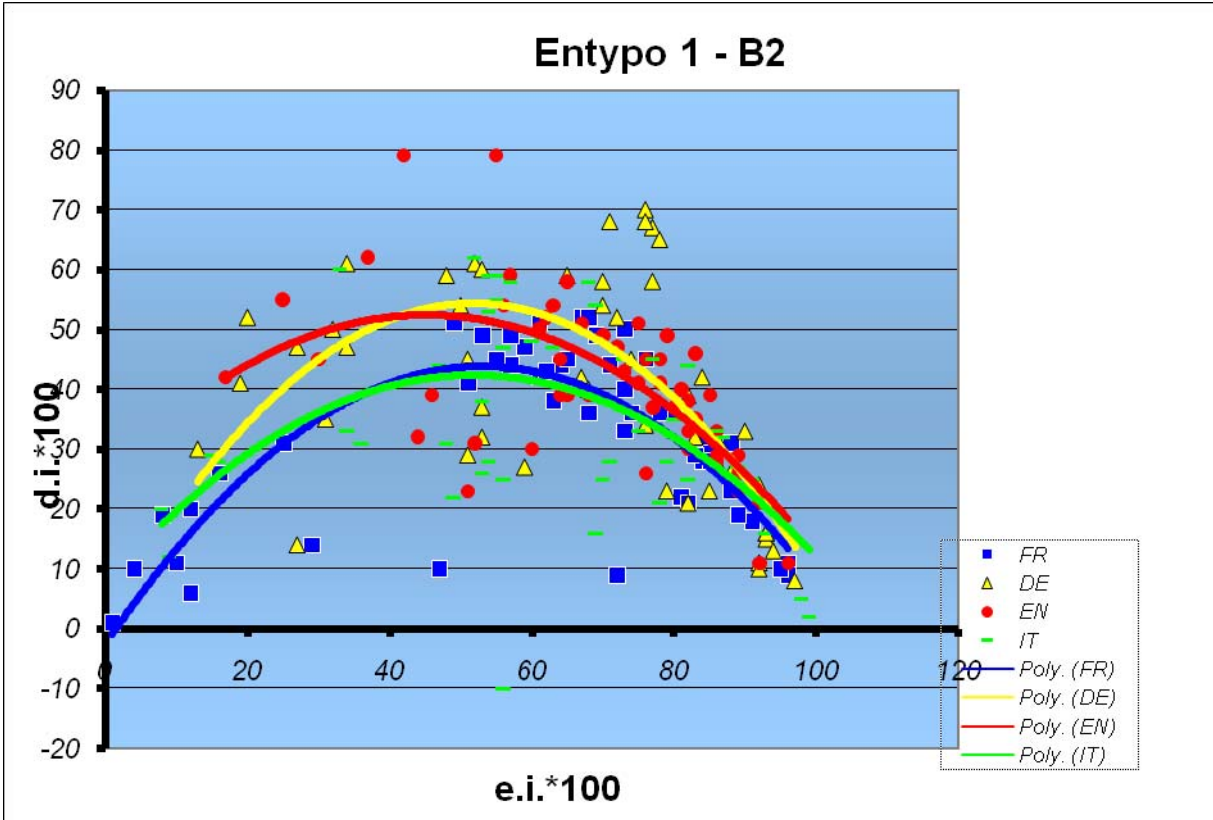
Συμπεριφορά των ερωτημάτων της δοκιμασίας 3 (εξεταστική περίοδος Μαΐου 2007)



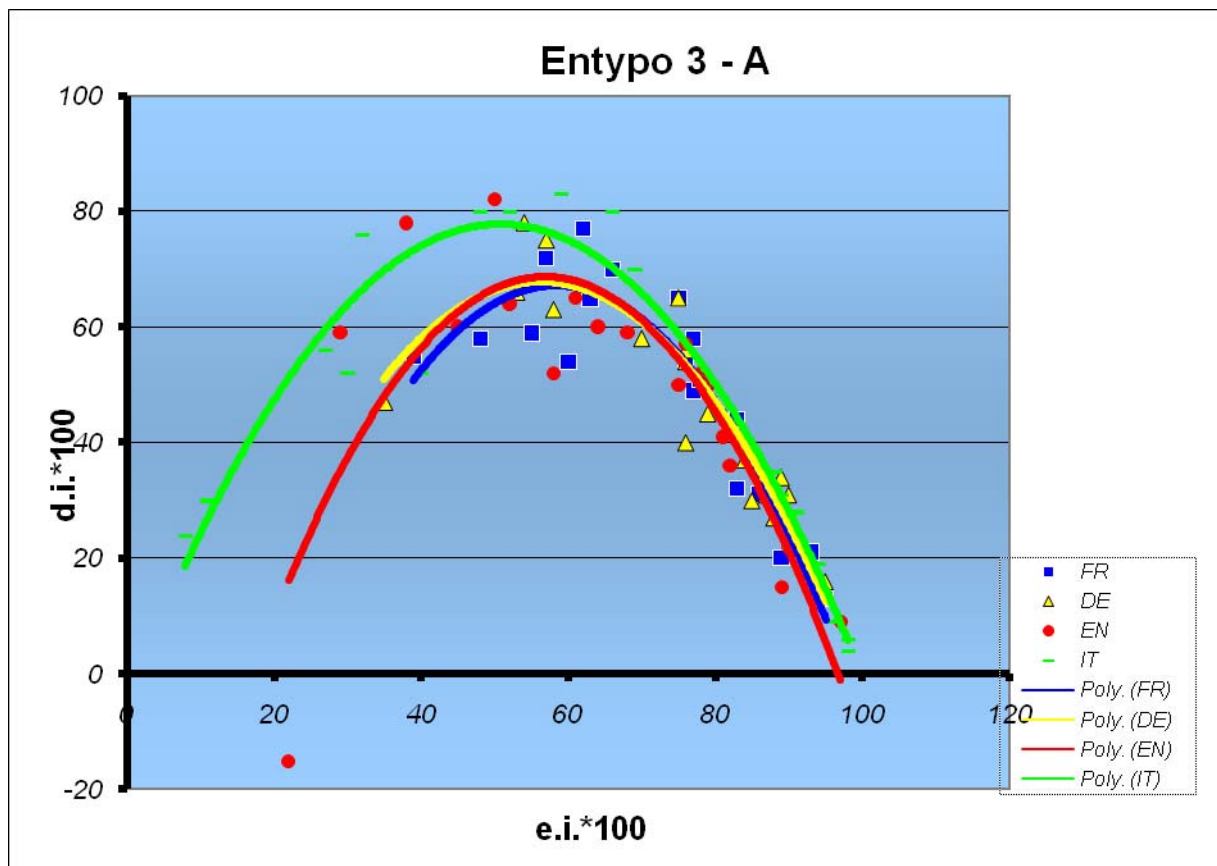


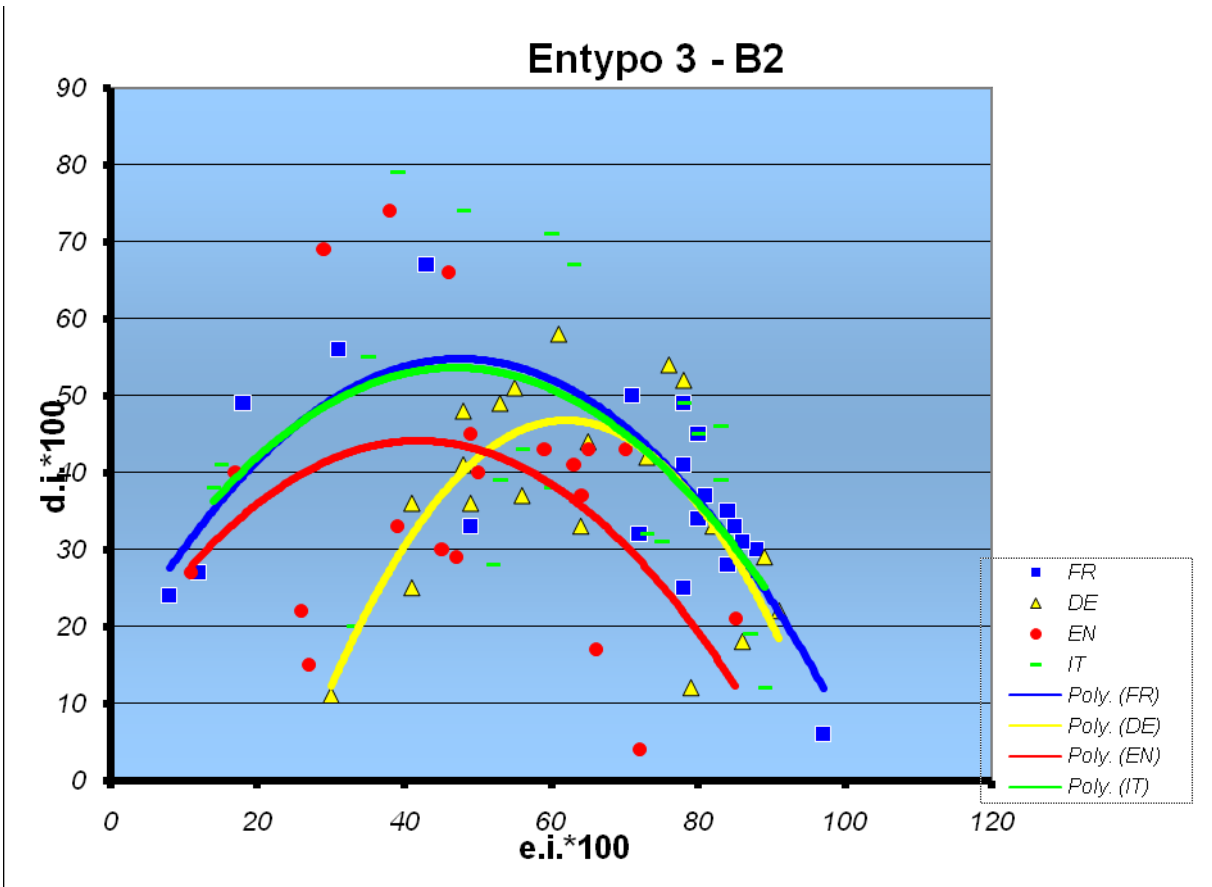
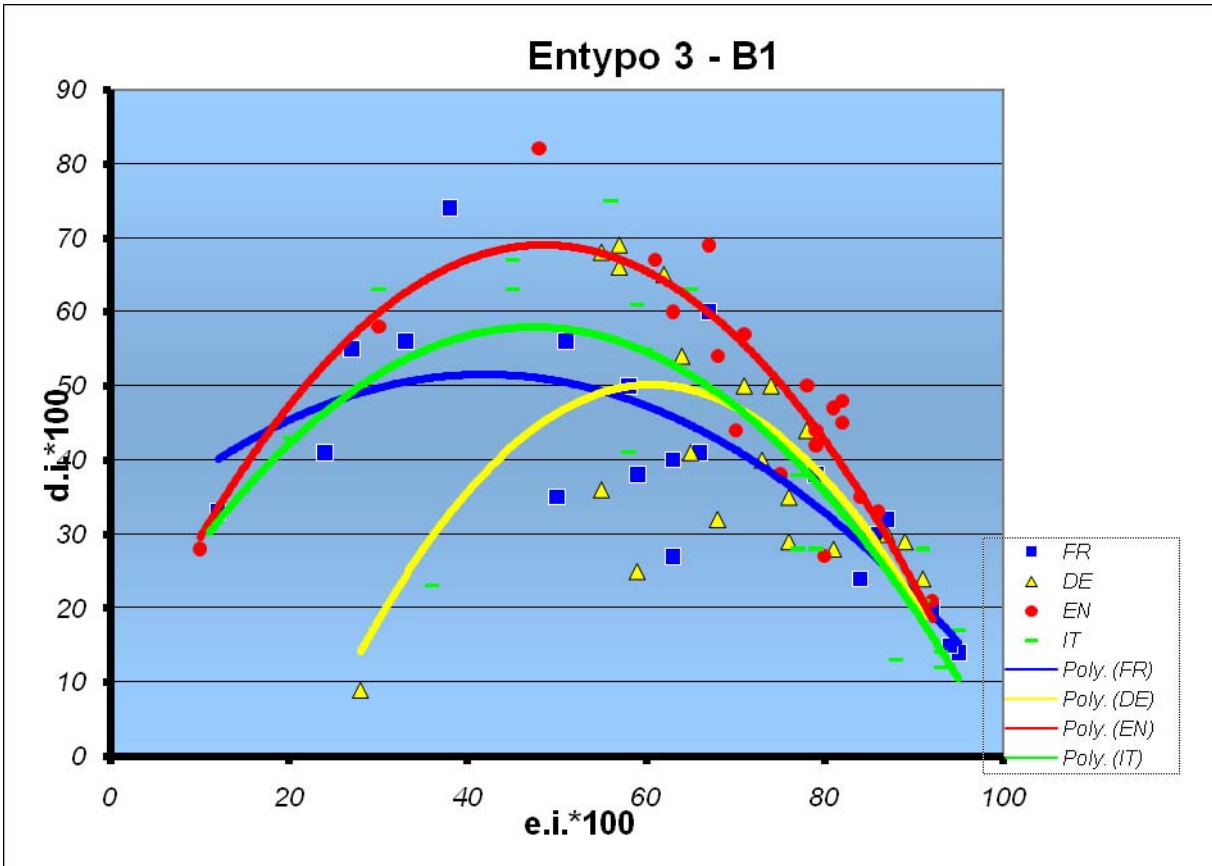
Συμπεριφορά των ερωτημάτων της δοκιμασίας 1 (εξεταστική περίοδος Μαΐου 2008)

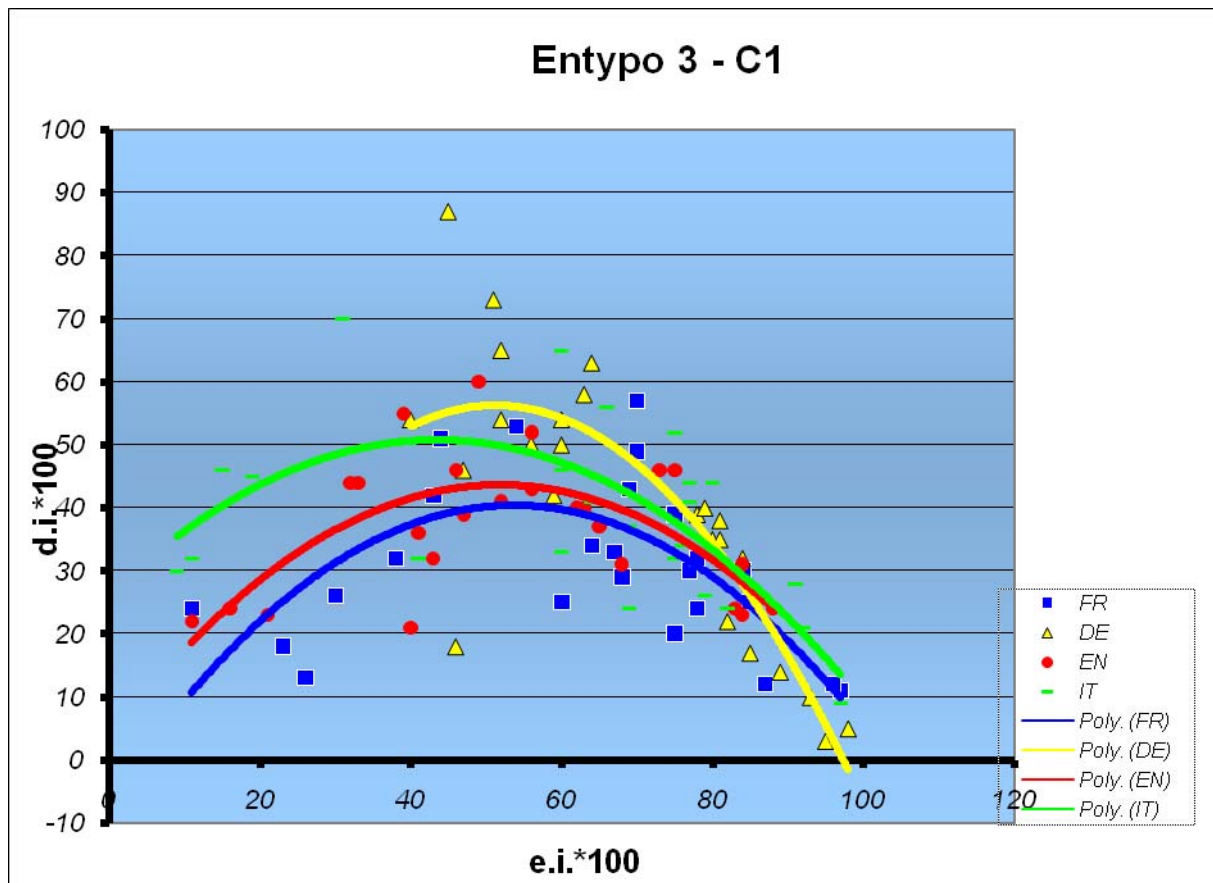




Συμπεριφορά των ερωτημάτων της δοκιμασίας 3 (εξεταστική περίοδος Μαΐου 2008)







Τα δεδομένα και οι ερμηνείες που προηγήθηκαν θα έπρεπε να επιτρέπουν τη συνέχιση της βελτίωσης των μετρολογικών και διδακτικών χαρακτηριστικών του εργαλείου του ΚΠγ και την ευκολότερη ανάλυση των συνιστωσών από ένα ενδεχόμενο εξωτερικό οργανισμό αξιολόγησης.

Τα δείγματα των αναλύσεων των ερωτημάτων παρατίθενται στο παράρτημα⁴.

⁴ Είναι προφανές ότι για τεχνικούς λόγους δεν είναι δυνατό να αναπαραχθεί στο παράρτημα ούτε το σύνολο των στατιστικών δεδομένων στα οποία βασίζονται τα συμπεράσματα ούτε οι σχετικές μελέτες.

Παράρτημα : Δείγματα των αναλύσεων των ερωτημάτων