

Η Βαθμολόγηση του Μαθήματος Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον στις Πανελλήνιες Εξετάσεις 2011

Ευάγγελος Κανίδης

Σχολικός Σύμβουλος Πληροφορικής Β' Αθήνας και Ανατολικής Αττικής
vkanidis@sch.gr

Περίληψη

Η παρούσα εργασία παρουσιάζει μια μελέτη περίπτωσης σχετική με την ποιότητα βαθμολόγησης των γραπτών δοκιμίων των μαθητών στο Μάθημα Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον (ΑΕΠΠ) το οποίο αποτελεί εξεταζόμενο μάθημα στις Πανελλήνιες Εξετάσεις για την Τεχνολογική Κατεύθυνση του κύκλου Πληροφορικής και Υπηρεσιών. Στόχος της εργασίας είναι η μελέτη των προσωπικών χαρακτηριστικών των βαθμολογητών του μαθήματος μέσα από τις βαθμολογίες ενός μεγάλου αριθμού γραπτών δοκιμίων. Για την επίτευξη του στόχου εξετάστηκαν οι βαθμολογίες σε 1638 γραπτά μαθητών ενός κεντρικού Κέντρου Βαθμολόγησης. Η μελέτη της ποιότητας βαθμολόγησης επιτυγχάνεται μέσα από την σύγκριση της βαθμολογίας κάθε βαθμολογητή με τους βαθμολογητές που διόρθωσαν τα ίδια γραπτά με αυτόν. Για την πληρέστερη αξιολόγηση συγκρίνεται και η βαθμολογία του με τον αναβαθμολογητή σε όσα γραπτά οδηγήθηκαν σε αναβαθμολόγηση. Διαπιστώνεται ότι η ποιότητα βαθμολόγησης είναι καλή αλλά υπάρχουν περιθώρια περαιτέρω βελτίωσης ορισμένων βαθμολογητών. Προτείνεται η επιμόρφωση των βαθμολογητών και η δημιουργία σώματος βαθμολογητών.

Λέξεις κλειδιά: Πανελλήνιες εξετάσεις, Βαθμολόγηση, ΑΕΠΠ.

1. Εισαγωγή

Οι Πανελλήνιες Εξετάσεις διεξάγονται στο τέλος της Γ' τάξης Λυκείου και το αποτέλεσμα αυτών των εξετάσεων καθορίζει την είσοδο του υποψηφίου στην τριτοβάθμια εκπαίδευση. Η αυξημένη ζήτηση των υποψηφίων, αλλά και της κοινωνίας για εισαγωγή στην τριτοβάθμια εκπαίδευση οδήγησε τις Πανελλήνιες Εξετάσεις στο κέντρο του ενδιαφέροντος της σχολικής εκπαιδευτικής διαδικασίας. Το αποτέλεσμα είναι ένα υπέρμετρο συχνά ενδιαφέρον της εκπαιδευτικής κοινότητας και των γονέων, σχετικά με τον τρόπο που διεξάγονται οι εξετάσεις, τα θέματα στα οποία καλούνται να απαντήσουν οι μαθητές, αλλά και η απαίτηση από όλους για μια αξιόπιστη, αντικειμενική και αδιάβλητη αποτίμηση των γραπτών.

Το μάθημα Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον (ΑΕΠΠ) καθιερώθηκε με το νόμο 2525/1997 και διδάσκεται στους μαθητές της Τεχνολογικής Κατεύθυνσης του κύκλου Πληροφορικής & Υπηρεσιών της Γ' Λυκείου από το 2000 μέχρι σήμερα. Η αξιολόγηση των θεμάτων καθώς και του τρόπου βαθμολόγησης, που

έχουν τεθεί στις εξετάσεις του μαθήματος ΑΕΠΠ, παρουσιάζει ερευνητικό ενδιαφέρον, γιατί η μορφή και η δυσκολία των θεμάτων και οι απαντήσεις των μαθητών δίνουν πληροφορίες για τον τρόπο που αντιμετωπίζουν οι μαθητές το γνωστικό αντικείμενο στο οποίο αξιολογήθηκαν (Dietel, Herman & Knuth, 1991). Επίσης οι πληροφορίες που προκύπτουν από τη μελέτη των χαρακτηριστικών αυτής της αξιολόγησης, μπορούν να βοηθήσουν τους καθηγητές και την πολιτεία να πάρουν σωστές αποφάσεις για τη βελτίωση του τρόπου αποτίμησης των απαντήσεων.

Στη βιβλιογραφία αναφέρεται ένα πλήθος παραγόντων που μπορούν να επηρεάσουν τα αποτελέσματα μιας εξέτασης καθώς και την αξιολόγησή της (Δημητρόπουλος, 1998). Ένα από τα χαρακτηριστικά αυτά είναι η *αξιοπιστία*, της εξέτασης.

Η έννοια της αξιοπιστίας αναφέρεται στο βαθμό εμπιστοσύνης που μπορούμε να έχουμε στα αποτελέσματα της εξέτασης, δηλαδή πόσο σταθερά και ακριβή είναι τα αποτελέσματα της μέτρησης της επίδοσης (Kirk & Miller, 1986).

Οι Πανελλήνιες Εξετάσεις έχουν καθιερωθεί στη συνείδηση των εκπαιδευτικών, των μαθητών και των γονέων ως αντικειμενικές και αξιόπιστες. Το Υπουργείο Παιδείας εξασφαλίζει την αντικειμενικότητα και την αξιοπιστία των εξετάσεων με την τήρηση μιας ελεγχόμενης, αυστηρής και δομημένης διαδικασίας, η οποία ξεκινά από την επιλογή των θεμάτων και καταλήγει στη βαθμολόγηση των γραπτών. Αρκεί όμως αυτή η διαδικασία για να εξασφαλιστεί και η αξιοπιστία των εξετάσεων;

Η αξιοπιστία μιας εξέτασης είναι δύσκολο να οριστεί με ακρίβεια και συνήθως οι ορισμοί που δίνονται αντανakλούν τη μεθοδολογία που χρησιμοποιείται για τη διαπίστωσή της (Κανίδης, 2010). Χαρακτηριστικά οι Cook and Campbell (1979) την ορίζουν ως «Η βέλτιστη διαθέσιμη προσέγγιση ως προς την ορθότητα ή το λάθος μιας συγκεκριμένης ερμηνείας, πρότασης ή συμπεράσματος». Ουσιαστικά δηλαδή τονίζουν ότι η μέτρηση της απόδοσης (βαθμολόγηση) δεν είναι απόλυτη αλλά πάντα έχει ένα βαθμό υποκειμενικότητας. Η αποτίμηση επίδοσης σε μια εξέταση επηρεάζεται από αυτό που οι ερευνητές ονομάζουν "αξιοπιστία αξιολογητών", δηλαδή υποκειμενικές διαφορές εκτίμησης της επίδοσης από τους αξιολογητές (Ebel & Friesbie, 1990)

Η παρούσα εργασία εξετάζει κυρίως τη βαθμολογική συμπεριφορά 30 καθηγητών Πληροφορικής που εργάστηκαν ως βαθμολογητές του μαθήματος Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον σε μεγάλο κεντρικό βαθμολογικό κέντρο. Ο αριθμός των γραπτών ΑΕΠΠ που διόρθωσαν ανέρχεται σε 1637. Συνεπώς εξετάζονται 3274 βαθμολογίες. Παράλληλα εξετάζονται οι βαθμοί αναβαθμολόγησης σε 43 γραπτά τα οποία αναβαθμολογήθηκαν, επειδή η διαφορά βαθμολογίας τους υπερέβαινε τις 12 μονάδες. Τα γραπτά προέρχονται από μια ευρεία περιφέρεια η οποία περιλαμβάνει αστικές ημιαστικές, αγροτικές περιοχές, καθώς και νησιά του Αιγαίου.

Η αναλυτική καταγραφή της βαθμολογίας των καθηγητών σε κάθε ερώτημα των θεμάτων, η οποία εφαρμόστηκε για πρώτη φορά στις Πανελλήνιες Εξετάσεις του 2010, επέτρεψε την αναλυτική μελέτη της βαθμολογικής συμπεριφοράς των

καθηγητών. Η βαθμολογική συμπεριφορά των καθηγητών ανά γραπτό και ανά θέμα μελετήθηκε συγκριτικά τόσο ως προς την αντίστοιχη συμπεριφορά του δεύτερου βαθμολογητή του γραπτού, όσο και ως προς την βαθμολογική συμπεριφορά του αναβαθμολογητή (όπου υπήρξε αναβαθμολόγηση).

2. Ταυτότητα αξιολόγησης

Στο μάθημα ΑΕΠΠ στις Πανελλήνιες Εξετάσεις του 2011 τέθηκαν 4 θέματα. Ο πίνακας 1 εμφανίζει τα ερωτήματα κάθε θέματος καθώς και την αντίστοιχη βαθμολογία του.

Πίνακας 1: Ερωτήματα και η βαθμολογία τους ανά θέμα

ΘΕΜΑΤΑ	ΘΕΜΑ Α					ΘΕΜΑ Β		ΘΕΜΑ Γ					ΘΕΜΑ Δ			
ΕΡΩΤΗΜΑΤΑ	A1	A2	A3	A4	A5	B1	B2	Γ1	Γ2	Γ3	Γ4	Γ5	Δ1	Δ2	Δ3	Δ4
ΒΑΘΜΟΛΟΓΙΑ	10	10	5	3	12	10	10	2	5	4	4	5	4	4	4	8

2.1 Διαφορές βαθμολόγησης

Αρχικά μελετήθηκε ο βαθμός ταύτισης της βαθμολογίας των δύο βαθμολογητών. Δηλαδή σε πόσες περιπτώσεις ο βαθμός του Α βαθμολογητή συμπίπτει με το βαθμό του Β βαθμολογητή σε κάθε θέμα καθώς και συνολικά στο γραπτό. Ο πίνακας 2 παρουσιάζει τον αριθμό αυτών των περιπτώσεων.

Πίνακας 2: Ταύτιση βαθμολογίας των δύο βαθμολογητών

	ΘΕΜΑ Α	ΘΕΜΑ Β	ΘΕΜΑ Γ	ΘΕΜΑ Δ	ΓΡΑΠΤΟ
Πλήθος	586	791	277	794	276
Ποσοστό	35,8%	48,3%	16,9%	48,5%	16,9%

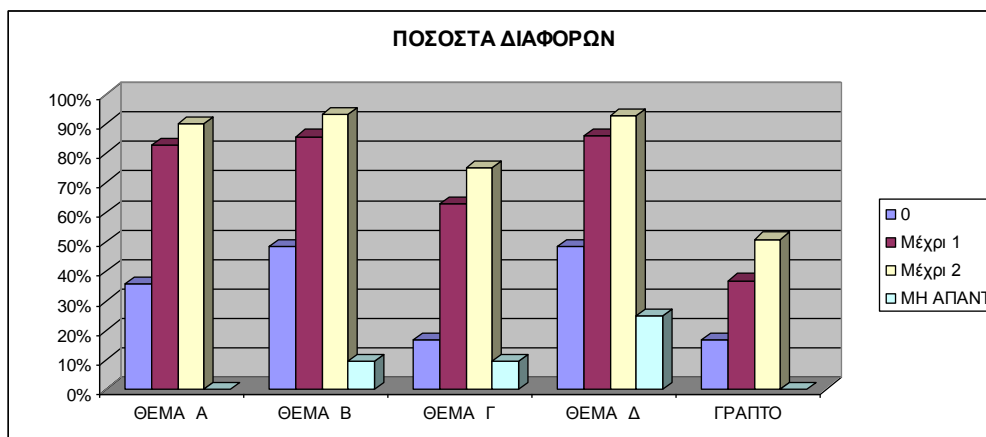
Παρατηρούμε ότι η μεγαλύτερη σύμπτωση βαθμολογιών παρουσιάζεται στο δεύτερο και στο τέταρτο θέμα ενώ η μικρότερη στο τρίτο θέμα. Βέβαια η ταύτιση της βαθμολογίας στατιστικά περιλαμβάνει και τους μαθητές που δεν απάντησαν στα επιμέρους θέματα και βαθμολογήθηκαν και από τους δύο βαθμολογητές με μηδέν. Ο πίνακας 3 εμφανίζει τον αριθμό και το ποσοστό των μαθητών που δεν απάντησαν σε κάθε θέμα και συνολικά.

Πίνακας 3: Μαθητές που δεν απάντησαν σε κάποιο θέμα

	ΘΕΜΑ Α	ΘΕΜΑ Β	ΘΕΜΑ Γ	ΘΕΜΑ Δ	ΓΡΑΠΤΟ
Πλήθος	2	157	157	408	2
Ποσοστό	0,1%	9,6%	9,6%	24,9%	0,1%

Παρατηρούμε ότι το ένα τέταρτο των μαθητών δεν απαντά καθόλου το 4^ο θέμα, ενώ τα ποσοστά αυτών που δεν απαντούν στο πρώτο θέμα συμπίπτουν με τα ποσοστά αυτών που δεν απάντησαν σε κανένα θέμα και είναι αμελητέα.

Το σχήμα 1 αναφέρεται στα ποσοστά σύγκλισης των δύο βαθμολογιών ενός γραπτού και παρουσιάζει τις διαφορές βαθμολογίας ανά θέμα και συνολικά μέχρι και 2 μονάδες.



Σχήμα 1: Ποσοστιαίες διαφορές βαθμολογίας ανά θέμα και συνολικά

Παρατηρούμε ότι σε όλα τα θέματα το ποσοστό των γραπτών που έχουν διαφορά μέχρι και 2 μονάδες είναι πάνω από το 85%. Το αντίστοιχο ποσοστό για τη συνολική βαθμολογία, επειδή οι διαφορές αθροίζονται, είναι 51%.

Το ποσοστό αναβαθμολογήσεων πανελλαδικά στο μάθημα ΑΕΠΠ το 2011 ήταν 3,15% και στο αναφερόμενο εξεταστικό κέντρο ήταν 2,63%.

3. Διαφοροποίηση βαθμολογίας συνολικά και ανά θέμα

Οι διαφορές της βαθμολογίας συνολικά στο κάθε γραπτό πριν από τις αναβαθμολογήσεις δεν παρουσίασαν μεγάλες αποκλίσεις. Φυσικά μετά τις αναβαθμολογήσεις οι διαφορές στις ακραίες περιπτώσεις μειώθηκαν. Ο πίνακας 4 παρουσιάζει τις διαφορές συνολικά.

Πίνακας 4. Διαφορές βαθμολογίας συνολικά

Μονάδες	0	1	2	3	4	5	6	7	8	9	10	11	12	>12
Αρ. γραπτών	276	328	229	189	146	98	101	63	50	34	41	21	18	43
Ποσοστο %	16,9	20	14	11,5	8,9	6	6,2	3,8	3,1	2,1	2,5	1,3	1,1	2,6

Παρατηρούμε ότι και στον πίνακα 3 οι διαφορές πάνω από 10 μονάδες είναι σχετικά σπάνιες. Το μικρό ποσοστό στις μεγάλες διαφορές δείχνει ότι υπάρχει ένα σημαντικά μεγάλο ποσοστό αξιοπιστίας στη βαθμολόγηση του μαθήματος. Σε κάθε περίπτωση όμως δεν θα πρέπει να ξεχνάμε ότι το κάθε γραπτό αντιπροσωπεύει ένα μαθητή και η διαφορά 10 μονάδων σε ένα μαθητή μπορεί να του στερήσει την εισαγωγή του στη σχολή που επιθυμεί.

3.1 Θέμα Α

Το πρώτο θέμα των Πανελληνίων Εξετάσεων περιείχε 5 υποενότητες η μορφή των οποίων ήταν:

- Η Α1 και Α2 ενότητες περιείχαν από πέντε ερωτήσεις τύπου Σωστό–Λάθος.
- Η Α3 ενότητα έδινε ένα τμήμα αλγορίθμου και ζητούσε τις τιμές που θα εμφανιστούν μετά την εκτέλεση του αλγορίθμου
- Η Α4 έδινε πάλι ένα τμήμα αλγορίθμου και ζητούσε τη συμπλήρωση μιας εντολής εκχώρησης έτσι ώστε να έχει το ίδιο ακριβώς αποτέλεσμα με το τμήμα του αλγορίθμου.
- Η Α5 ενότητα περιείχε τρεις ερωτήσεις θεωρίας.

Οι διαφορές βαθμολογίας στο πρώτο θέμα εμφανίζονται στον πίνακα 5 .

Πίνακας 5. Διαφορές βαθμολογίας στο πρώτο θέμα.

Μονάδες	0	1	2	3	4	5	6	7	8	9	10
Αρ.γραπτών	586	419	269	179	89	47	28	9	4	5	2
Ποσοστό	35,8	25,6	16,4	10,9	5,4	2,9	1,7	0,5	0,2	0,3	0,1

Οι διαφορές βαθμολογίας εντοπίζονται κυρίως στην ενότητα Α5(12 μονάδες) και δευτερευόντως στην Α4. Οι διαφορές στην Α5 προκαλούν εντύπωση γιατί το θέμα αφορούσε αξιολόγηση ερωτήσεων θεωρίας. Ενώ στην ίδια ενότητα 250 μαθητές (15,3%) δεν απάντησαν καθόλου.

Παράδειγμα: Η παρακάτω απάντηση μαθητή βαθμολογήθηκε από ένα βαθμολογητή με 12 (άριστα) και από τον άλλο βαθμολογητή με 4!

Α) Υποπρόγραμμα λέγεται ένας αλγόριθμος ο οποίος χρησιμοποιείται από το κυρίως πρόγραμμα για την εκτέλεση κάποιας συγκεκριμένης λειτουργίας. Δεν αποτελεί ξεχωριστό πρόγραμμα δηλαδή δεν μπορεί να χρησιμοποιηθεί μόνο του. Υποπρογράμματα είναι οι διαδικασίες και οι συναρτήσεις.

Β) Τμηματικός προγραμματισμός ονομάζεται ο προγραμματισμός που γίνεται με τη χρήση υποπρογράμματος. Κάθε πρόβλημα αναλύεται σε υποπροβλήματα τα οποία λύνονται το καθένα με ξεχωριστό αλγόριθμο. Έπειτα αυτά συνθέτονται και αποτελούν τη συνολική λύση. Έτσι η δημιουργία του προγράμματος μπορεί να ανατεθεί σε περισσότερους προγραμματιστές που ο καθένας φτιάχνει ένα συγκεκριμένο τμήμα του προγράμματος μειώνοντας το χρόνο που χρειάζεται για να δημιουργηθεί.

Γ) Παράμετρος ονομάζεται η μεταβλητή που λαμβάνει το υποπρόγραμμα από το κυρίως πρόγραμμα ή ένα άλλο υποπρόγραμμα από το καλείται.

Αντίστοιχα στην ενότητα Α4 εντοπίστηκαν γραπτά όπου η απάντηση $K \leftarrow K \cdot H$ (ΟΧΙ K) και η $K \leftarrow X-1$ βαθμολογήθηκαν με 3 μονάδες (με άριστα το 4) από τον έναν βαθμολογητή και μηδέν από τον άλλον.

3.2 Θέμα Β

Το δεύτερο θέμα των Πανελληνίων Εξετάσεων περιείχε δύο ενότητες B1 και B2. Η πρώτη αφορούσε τη μετατροπή ενός διαγράμματος ροής σε αλγόριθμο. Η δεύτερη ενότητα έδινε ένα κύριο πρόγραμμα και ένα υποπρόγραμμα και ζητούσε τον υπολογισμό των τιμών που θα εμφανιστούν με την εκτέλεση του προγράμματος. Οι διαφορές βαθμολογίας στο δεύτερο θέμα εμφανίζονται στον πίνακα 6.

Πίνακας 6. Διαφορές βαθμολογίας στο δεύτερο θέμα.

Μονάδες	0	1	2	3	4	5	6	7	8	9	10
Αρ.γραπτών	791	356	247	117	68	37	6	6	5	2	2
Ποσοστό	48,3	21,7	15,1	7,1	4,2	2,3	0,4	0,4	0,3	0,1	0,1

Ακραίες περιπτώσεις διαφορών είχαμε στην ενότητα B1, όπου δύο γραπτά βαθμολογήθηκαν με άριστα από τον ένα βαθμολογητή και μηδέν από τον άλλον. Ενώ υπήρξαν και δύο γραπτά στην ενότητα B2 που οι δύο βαθμολογητές είχαν 9 μονάδες διαφορά. Η λύση της B1 που εμφανίζεται στο πλαίσιο δεξιά αξιολογήθηκε από τον ένα βαθμολογητή με 10 (άριστα) και από τον άλλον με 4.

```

Σ ← 0
Κ ← 1
Αρχή_επανάληψης
Αν Κ<=100 τότε
    Διάβασε Χ
    Σ ← Σ + Χ
    Κ ← Κ + 1
Τέλος Αν
Μέχρις_ότου Σ>1000
    
```

3.3 Θέμα Γ

Το τρίτο θέμα των Πανελληνίων εξετάσεων ζητούσε τη δημιουργία ενός αλγορίθμου σχετικού με τις εξετάσεις του ΑΣΕΠ. Το θέμα περιείχε πέντε ενότητες με αντίστοιχες βαθμολογίες 2, 5, 4, 4, 5 μονάδες. Οι διαφορές βαθμολογίας στο δεύτερο θέμα εμφανίζονται στον πίνακα 7.

Πίνακας 7. Διαφορές βαθμολογίας στο τρίτο θέμα.

Μονάδες	0	1	2	3	4	5	6	7	8	9	10	11	12
Αρ.γραπτών	277	301	344	223	222	130	108	8	9	3	5	5	2
Ποσοστό %	16,9	18,4	21	13,6	13,6	7,9	6,6	0,5	0,5	0,2	0,3	0,3	0,1

Παρατηρούνται διαφορές βαθμολογίας και στο θέμα Γ όπου υπάρχουν γραπτά και με 12 μονάδες διαφορά (στις 20). Στην ενότητα Γ2 υπήρξαν περισσότερα από 120 γραπτά με διαφορές 3 και 4 μονάδων. Για παράδειγμα το τμήμα του αλγορίθμου δεξιά βαθμολογήθηκε από τον ένα βαθμολογητή με 4 (άριστα) και από τον άλλο με μηδέν. Εντυπωσιακό επίσης είναι το

```

ΔΙΑΒΑΣΕ ονομ, βαθ1, βαθ2, βαθ3
Αρχή_επανάληψης
Για Ι από 1 μέχρι 3
    Αν βαθ1 > βαθ2 και βαθ1 > βαθ3 τότε
        ΓΡΑΨΕ « Ο Μεγαλύτερος Βαθμός είναι ο :» βαθ1
    ΑΛΛΙΩΣ_ΑΝ βαθ2 > βαθ1 και βαθ2 > βαθ3 τότε
        ΓΡΑΨΕ «Ο μεγαλύτερος βαθμός είναι ο :» βαθ2
    ΑΛΛΙΩΣ
        ΓΡΑΨΕ «Ο μεγαλύτερος βαθμός είναι ο :» βαθ3
Τέλος_ΑΝ
.....
    
```

στοιχείο ότι στην ενότητα Γ4 (να τερματίζει ο αλγόριθμος με τη λέξη "ΤΕΛΟΣ") υπάρχουν 458 γραπτά όπου υπάρχει απόλυτη διαφορά βαθμολογίας, δηλαδή ο ένας βαθμολογητής τα έχει αξιολογήσει με 4 (άριστα) και ο άλλος με μηδέν!

Θέμα Δ.

Το τέταρτο θέμα των Πανελληνίων εξετάσεων ζητούσε τη δημιουργία ενός αλγορίθμου σχετικού με την εκλογή τριών αρχηγών μιας ποδοσφαιρικής ομάδας 22 ατόμων. Το θέμα περιείχε τέσσερις υποενότητες με αντίστοιχες βαθμολογίες 4, 4, 4 και 8 μονάδες. Οι διαφορές βαθμολογίας στο τέταρτο θέμα εμφανίζονται στον πίνακα 8.

Πίνακας 8. Διαφορές βαθμολογίας στο τέταρτο θέμα.

Μονάδες	0	1	2	3	4	5	6	7	8	9	10	11	12
Αρ.γραπτών	794	385	215	117	65	35	16	5	1	3	0	1	0
Ποσοστό %	48,5	23,5	13,1	7,1	4	2,1	1	0,3	0,1	0,2	0	0,1	0

Στο τέταρτο θέμα δεν εντοπίζονται μεγάλες διαφορές λόγω του ότι ένα 25% των μαθητών δεν το έχει απαντήσει καθόλου. Οι μεγαλύτερες διαφορές εντοπίζονται στην ενότητα Δ4 ή οποία ήταν απαιτητική και αντιστοιχούσε σε 8 μονάδες.

4. Βαθμολογική συμπεριφορά βαθμολογητών

Στην ενότητα αυτή εξετάζεται κυρίως η βαθμολογική συμπεριφορά 16 βαθμολογητών από ένα σύνολο 30 βαθμολογητών. Η επιλογή έγινε με κριτήριο τον αριθμό των γραπτών που διόρθωσε ο κάθε βαθμολογητής. Επιλέχθηκαν οι καθηγητές που διόρθωσαν τουλάχιστον 75 γραπτά (τρία πακέτα). Αξιολογήθηκαν οι βαθμολογίες σε σύγκριση με τον έτερο βαθμολογητή του κάθε γραπτού καθώς και με τον αναβαθμολογητή του γραπτού στην περίπτωση που υπάρχει αναβαθμολόγηση. Το ερευνητικό ερώτημα που μας απασχόλησε είναι: "Υπάρχουν καθηγητές που συστηματικά βαθμολογούν πάνω ή κάτω από τον έτερο βαθμολογητή του κάθε γραπτού". Για το λόγο αυτό εξετάστηκε για κάθε βαθμολογητή το ποσοστό των γραπτών που παρουσιάζει ίση, μεγαλύτερη ή μικρότερη βαθμολογία από τον άλλο βαθμολογητή του γραπτού. Ένα ποσοστό μεγαλύτερο του 50% με υψηλότερη βαθμολογία υποδεικνύει πιθανά έναν "επιεική" βαθμολογητή, ενώ το αντίθετο υποδεικνύει πιθανά έναν "αυστηρό" βαθμολογητή. Ο διπλός πίνακας 9 εξετάζει τη βαθμολογική συμπεριφορά των 16 καθηγητών.

Για μεγαλύτερη ακρίβεια ορίσαμε ένα "Δείκτη Βαθμολόγησης" για κάθε βαθμολογητή. Ο δείκτης βαθμολόγησης υπολογίζεται από το μέσο όρο της διαφοράς των βαθμολογιών που έχει κάθε βαθμολογητής με τον άλλο βαθμολογητή σε κάθε γραπτό. Ο δείκτης βαθμολόγησης μπορεί να πάρει θετικές τιμές, αν κάποιος βαθμολογητής συστηματικά (κατά μέσο όρο) βαθμολογεί πάνω από το βαθμό των άλλων βαθμολογητών. Επίσης μπορεί να πάρει αρνητικές τιμές, αν ο συγκεκριμένος βαθμολογητής βαθμολογεί συστηματικά κάτω από το βαθμό των άλλων

βαθμολογητών. Ουδέτερος και τυπικά "ιδανικός" βαθμολογητής θα ήταν ένας που έχει δείκτη βαθμολόγησης μηδέν.

Πίνακας 9: Δείκτες βαθμολόγησης καθηγητών

Βαθμολογητής	4	5	6	12	14	15	16	18
ΓΡΑΠΤΑ	>150	>150	>250	<150	<150	<150	>250	>150
Αναβαθμολογήσεις	2,0%	5,4%	0,4%	3,0%	0,8%	1,8%	2,0%	3,2%
Σύμπτωση βαθμολογίας	16,8%	10,9%	19,7%	16,3%	18,2%	14,5%	14,1%	13,0%
Μεγαλύτερη βαθμολογία	27,9%	22,9%	46,8%	24,5%	27,3%	42,7%	44,3%	51,1%
Μικρότερη βαθμολογία	55,3%	66,3%	33,5%	59,2%	54,5%	43,7%	41,6%	35,9%
Δείκτης Αναβαθμολόγησης	-7	0,4	X	-11	X	0	0,8	3,8
Δείκτης Βαθμολόγησης	-1,9	-3,2	0,9	-2,8	-1,6	0	0,1	1,4
Βαθμολογητής	19	20	21	23	24	27	28	29
ΓΡΑΠΤΑ	>250	<100	<100	<150	<150	>250	>250	<100
Αναβαθμολογήσεις	3,1%	2,7%	7,0%	4,3%	2,1%	2,8%	2,2%	1,4%
Σύμπτωση βαθμολογίας	15,7%	15,1%	8,5%	8,5%	21,1%	18,5%	18,9%	24,3%
Μεγαλύτερη βαθμολογία	54,7%	50,7%	83,1%	16,0%	35,8%	46,9%	43,3%	43,2%
Μικρότερη βαθμολογία	29,6%	34,2%	8,5%	75,5%	43,2%	34,6%	37,8%	32,4%
Δείκτης Αναβαθμολόγησης	7,9	2,5	7,6	-8,3	-8	-0,2	9,2	X
Δείκτης Βαθμολόγησης	1,5	1,1	5,3	-3,3	-0,7	0,6	0,5	1,3

Μεγάλος θετικός βαθμός του δείκτη αξιολόγησης δείχνει έναν "επιεική" βαθμολογητή, ενώ αντίστοιχα μεγάλος (απόλυτα) αρνητικός βαθμός δείχνει έναν "αυστηρό" βαθμολογητή. Ο δείκτης βαθμολόγησης έχει μεγαλύτερη αξιοπιστία, όταν υπολογίζεται από μεγάλο αριθμό διορθώσεων και για το λόγο αυτό επελέγησαν βαθμολογητές που είχαν διορθώσει τουλάχιστον 75 γραπτά. Για έλεγχο του δείκτη βαθμολόγησης εξετάστηκε και ο δείκτης αναβαθμολόγησης. Ο δείκτης αναβαθμολόγησης εξετάζει το μέσο όρο των διαφορών ενός βαθμολογητή από τους βαθμούς του αναβαθμολογητή. Επειδή ορισμένοι βαθμολογητές είχαν πολύ λίγα αναβαθμολογούμενα γραπτά ο δείκτης αναβαθμολόγησης εξετάζεται μόνο συμπληρωματικά ως προς το δείκτη βαθμολόγησης για επιβεβαίωσή του. Οι βαθμολογητές που έχουν στον δείκτη αναβαθμολόγησης τον χαρακτήρα "X" έχουν μία ή καμία αναβαθμολόγηση.

Από τον πίνακα αυτό συμπεραίνουμε ότι την καλύτερη βαθμολογική συμπεριφορά παρουσιάζουν οι βαθμολογητές με κωδικό 15, 16 και 27. Οι τρεις αυτοί βαθμολογητές έχουν λίγες αναβαθμολογήσεις, ισορροπημένη βαθμολογία και μικρούς δείκτες αναβαθμολόγησης και βαθμολόγησης. Στο άλλο άκρο βρίσκονται οι βαθμολογητές με κωδικούς 21 και 23. Ο βαθμολογητής 21 παρατηρούμε ότι είναι υπερβολικά "επιεικής", αφού κατά 83,1% βαθμολογεί υψηλότερα από τον άλλο

βαθμολογητή. Ο δείκτης βαθμολόγησης μας δείχνει ότι θα έπρεπε κατά μέσο όρο οι βαθμοί του να είναι κατά 5,3 μονάδες μικρότεροι. Αυτό επιβεβαιώνεται από τον δείκτη αναβαθμολόγησης ο οποίος δείχνει ότι ο συγκεκριμένος βαθμολογητής έχει βαθμολογήσει κατά μέσο όρο 7,6 μονάδες υψηλότερα κάθε γραπτό που οδηγήθηκε σε αναβαθμολόγηση. Τα χαρακτηριστικά αυτά συμπληρώνονται από ένα μεγάλο αριθμό αναβαθμολογήσεων. Ο βαθμολογητής 23 έχει αντίθετα χαρακτηριστικά από τον 21. Είναι "αυστηρός", βαθμολογεί κατά μέσο όρο 3,3 μονάδες χαμηλότερα κάθε γραπτό. Η συμπεριφορά αυτή επιβεβαιώνεται από το σχετικά υψηλό αριθμό αναβαθμολογήσεων, όπου ο δείκτης αναβαθμολόγησης δείχνει ότι ο συγκεκριμένος βαθμολογητής έχει βαθμολογήσει κατά μέσο όρο 8,3 μονάδες χαμηλότερα κάθε γραπτό που οδηγήθηκε σε αναβαθμολόγηση.

Καλή βαθμολογική συμπεριφορά παρουσιάζουν επίσης οι βαθμολογητές με κωδικούς 6 και 28 που έχουν διορθώσει ένα μεγάλο αριθμό γραπτών, έχουν λίγες αναβαθμολογήσεις και παρουσιάζουν ένα καλό δείκτη βαθμολόγησης.

Ενδιαφέρον παρουσιάζει ο βαθμολογητής με κωδικό 4, διότι η μικρότερη βαθμολογία που παρουσιάζει κατά μέσο όρο διαπιστώθηκε μετά από αναλυτική μελέτη των βαθμολογιών του ότι οφείλεται στο θέμα θεωρίας (Α5) όπου η βαθμολογία του ήταν ιδιαίτερα αυστηρή. Επίσης αξίζει να σημειωθεί ότι υπήρξαν βαθμολογητές οι οποίοι είχαν 16% αναβαθμολογήσεις (4 γραπτά στα 25) αλλά λόγω μικρού αριθμού γραπτών που διόρθωσαν δεν έχουν αξιολογηθεί.

Μια τελευταία παρατήρηση είναι ότι οι βαθμολογητές που διορθώνουν πολλά γραπτά έχουν καλύτερη βαθμολογική συμπεριφορά από αυτούς που έχουν βαθμολογήσει λίγα γραπτά.

3. Συμπεράσματα – προτάσεις

Η καταγραφή της αναλυτικής βαθμολογίας σε κάθε ερώτημα κάθε μαθήματος στις Πανελλήνιες Εξετάσεις που καθιερώθηκε το 2010 επιτρέπει τον καθορισμό της βαθμολογικής συμπεριφοράς κάθε βαθμολογητή. Από την έρευνα διαπιστώθηκε ότι η αξιοπιστία βαθμολόγησης στο μάθημα ΑΕΠΠ, στο εξεταζόμενο βαθμολογικό κέντρο, βρίσκεται σε υψηλό επίπεδο. Διαπιστώθηκαν σημαντικές διαφορές στο τρόπο βαθμολόγησης των γραπτών και εντοπίστηκαν καθηγητές που βαθμολογούν συστηματικά με "επιείκεια" ή με "αυστηρότητα". Ενδιαφέρον παρουσιάζει η διαχρονική σύγκριση της βαθμολογικής συμπεριφοράς των εκπαιδευτικών για την εξαγωγή περισσότερο έγκυρων συμπερασμάτων την οποία σκοπεύουμε να πραγματοποιήσουμε σε επόμενη εργασία μας. Τέλος θα αναφέρουμε δύο προτάσεις:

- **Επιμόρφωση βαθμολογητών:** Η επιμόρφωση των βαθμολογητών με συγκεκριμένα παραδείγματα γραπτών που παρουσιάζουν μεγάλες διαφορές βαθμολόγησης θα μπορούσε να δώσει ακόμα μεγαλύτερη αξιοπιστία στη βαθμολόγηση των γραπτών.

- Δημιουργία σώματος βαθμολογητών με μέλη καθηγητές που επιμορφώνονται και ελέγχεται διαχρονικά η βαθμολογική συμπεριφορά τους.

Η εργασία αυτή θα μπορούσε ένα πραγματοποιηθεί σε πανελλαδικό επίπεδο αν διατεθούν σταθεροί κωδικοί διόρθωσης σε κάθε εκπαιδευτικό. Επίσης θα ήταν ιδιαίτερα ενδιαφέρονσα η εφαρμογή της σε θεωρητικά μαθήματα τα οποία παρουσιάζουν μεγάλα ποσοστά αναβαθμολογήσεων.

Αναφορές

- Cook, T.D. and Campbell, D.T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Houghton Mifflin, Boston.
- Dietael, J. R. Herman, L. J. and R.A. Knuth A. R., *What Does Research Say About Assessment?*, North Central Regional Educational Laboratory (NCREL), Oak Brook, 1991 Διαθέσιμο στο Διαδίκτυο στο http://www.ncrel.org/sdrs/areas/stw_esyss/4assess.htm, πρόσβαση Δεκέμβριος 2009.
- Ebel, R. L., and Frisbie, D. A. (1990). *Essentials of Educational Measurement*. (5th ed.) Englewood Cliffs, N.J.: Prentice-Hall.
- Kirk, J., & Miller, M. L. (1986). *Reliability and validity in qualitative research*. Beverly Hills:Sage Publications.
- Stiggins, R. J. (1991). *Assessment literacy*. Phi Delta Kappan. 72 (7)
- Δημητρόπουλος Γ. Ε.,(1998), Εκπαιδευτική αξιολόγηση η αξιολόγηση του μαθητή, εκδόσεις Γρηγόρη, Αθήνα.
- Κανίδης Ε., (2010). " Αξιολόγηση του μαθήματος 'Ανάπτυξη Εφαρμογών σε Προγραμματιστικό περιβάλλον' στις Πανελλήνιες εξετάσεις των ετών 2002-2009" 5ο Συνέδριο Διδακτική της Πληροφορικής, Αθήνα, Απρίλιος. Πρακτικά σε CD, Διαθέσιμα on-line στην διεύθυνση: <http://www.etpe.gr>