

ΜΕΡΟΣ ΕΚΤΟ

Εκπαιδευτική πολιτική και απόψεις σχετικά με τις εφαρμογές των ΤΠΕ στην Εκπαίδευση

ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΘΕΜΑΤΩΝ ΤΟΥ ΜΑΘΗΜΑΤΟΣ “ΑΝΑΠΤΥΞΗ ΕΦΑΡΜΟΓΩΝ ΣΕ ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΟ ΠΕΡΙΒΑΛΛΟΝ” ΣΤΙΣ ΓΕΝΙΚΕΣ ΕΞΕΤΑΣΕΙΣ ΤΩΝ ΕΤΩΝ 2002-2005

Ευάγγελος Κανίδης

Μεταπτυχιακό πρόγραμμα στη Βασική και Εφαρμοσμένη Γνωσιακή Επιστήμη
vkanidis@di.uoa.gr

Εμμανουήλ Μόρμορης

2ο Ε.Λ Βριλησίων
emormoris@sch.gr

Περίληψη

Η εργασία παρουσιάζει μια αξιολόγηση των θεμάτων του μαθήματος “Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον (ΑΕΠΠ)” που έχουν τεθεί στις Γενικές Εξετάσεις των ετών 2002 έως 2005. Ως κριτήρια αξιολόγησης των θεμάτων και κατ’ επέκταση των εξετάσεων χρησιμοποιήθηκαν οι παράγοντες αξιοπιστία, αντικειμενικότητα, εγκυρότητα και διακριτικότητα. Διαπιστώθηκε ότι η αξιοπιστία και η αντικειμενικότητα των εξετάσεων είναι σε υψηλό επίπεδο, ενώ υπάρχουν περιθώρια βελτίωσης της αντικειμενικότητας της βαθμολόγησης. Επίσης η εγκυρότητα των θεμάτων είναι σε υψηλό επίπεδο με τα θέματα να καλύπτουν ευρέως και ικανοποιητικά την εξεταστέα ύλη. Η διακριτικότητα εμφανίζει αμφιλεγόμενη εικόνα με τον υπολογισμό του δείκτη διακριτικότητας να βρίσκεται σε υψηλά επίπεδα αλλά η κατανομή της βαθμολογίας απέχει πολύ από την “ιδανική” εικόνα της κανονικής κατανομής. Προτείνονται βελτιώσεις στο σύστημα βαθμολόγησης των γραπτών δοκιμών καθώς και στην επιλογή των θεμάτων των εξετάσεων.

Λέξεις Κλειδιά

Αξιολόγηση θεμάτων, Γενικές εξετάσεις .

ΕΙΣΑΓΩΓΗ

Η εργασία αυτή αποσκοπεί στην αξιολόγηση των θεμάτων του μαθήματος “Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον (ΑΕΠΠ)” που έχουν τεθεί στις Γενικές Εξετάσεις των ετών 2002 έως 2005. Το μάθημα ΑΕΠΠ διδάσκεται στους μαθητές της Τεχνολογικής Κατεύθυνσης του κύκλου Πληροφορικής & Υπηρεσιών της Γ΄ Λυκείου. Το μάθημα άρχισε να διδάσκεται και να εξετάζεται πανελλαδικά από το έτος 2000 (Ε.Λ. 1998) . Τα δύο πρώτα χρόνια (2000 & 2001) υπήρχαν αρκετά προβλήματα τόσο σε επίπεδο διδασκαλίας όσο και σε επίπεδο γενικών εξετάσεων. Για το λόγο αυτό τα θέματα των

αντίστοιχων γενικών εξετάσεων δεν προσφέρονται για την εξαγωγή γενικότερων συμπερασμάτων και στην εργασία αυτή εξετάζονται μόνο τα θέματα των ημερήσιων Ενιαίων Λυκείων κατά τις γενικές εξετάσεις Ιουνίου των ετών 2002-2005.

Η αξιολόγηση των θεμάτων στο μάθημα ΑΕΠΠ παρουσιάζει ερευνητικό ενδιαφέρον για δύο λόγους. Πρώτον είναι χρήσιμη στους καθηγητές που διδάσκουν το μάθημα, γιατί τα αποτελέσματα μιας “σωστής” εξέτασης, αντιπροσωπεύουν κάτι περισσότερο από το βαθμό που οι μαθητές πήραν στη συγκεκριμένη εξέταση. Τα αποτελέσματα δίνουν γενικότερες πληροφορίες για τον τρόπο που αντιμετωπίζουν οι μαθητές το γνωστικό αντικείμενο στο οποίο αξιολογήθηκαν (Dietel, Herman & Knuth, 1991). Δεύτερον οι πληροφορίες που προκύπτουν από την αξιολόγηση των μαθητών βοηθούν τους καθηγητές και την πολιτεία να πάρουν σωστές αποφάσεις για τη βελτίωση του μαθήματος αλλά και της παιδείας γενικότερα.

Η αξιολόγηση των θεμάτων θα γίνει με μελέτη των χαρακτηριστικών και των παραγόντων που πρέπει να πληρούν οι εξετάσεις και τα αντίστοιχα θέματα για να εξασφαλίζουν μια “σωστή” αξιολόγηση των μαθητών (Geiger & Cooper, 1995).

Οι ερευνητές της αξιολόγησης συμφωνούν ότι η εγκυρότητα ενός τεστ είναι στενά συνδεδεμένη με τους σκοπούς για τους οποίους χρησιμοποιείται το συγκεκριμένο τεστ. (Dietel, Herman & Knuth, 1991). Η εισαγωγή των μαθητών στην τριτοβάθμια εκπαίδευση βασίζεται στις γενικές εξετάσεις και για το λόγο αυτό οι εξετάσεις είναι στενά συνδεδεμένες με την εκάστοτε εκπαιδευτική πολιτική (ίδρυση σχολών, συνολικός αριθμός φοιτητών, κατανομή φοιτητών σε τμήματα). Η σύνδεση εκπαιδευτικής πολιτικής και γενικών εξετάσεων αναμένεται να ενταθεί από τις γενικές εξετάσεις του 2006. Από τις εξετάσεις αυτές θα εισάγονται στην τριτοβάθμια εκπαίδευση μόνο οι μαθητές που έχουν επιτύχει βαθμολογία πάνω από τη βαθμολογική βάση (Μέσος όρος βαθμών 10). Συνεπώς η κατανομή της βαθμολογίας θα επηρεάζει άμεσα τον αριθμό των φοιτητών. Στην εργασία αυτή ως “σωστή” αξιολόγηση των μαθητών θεωρούμε εκείνη την αξιολόγηση η οποία αποτυπώνει το γνωστικό επίπεδο των μαθητών, όπως αποτυπώνεται από τις απαντήσεις τους και παράλληλα είναι σύμφωνη με την εκπαιδευτική πολιτική (Ε.Λ. 1998).

ΠΑΡΑΓΟΝΤΕΣ ΠΟΥ ΕΠΗΡΕΑΖΟΥΝ ΜΙΑ ΕΞΕΤΑΣΗ

Η αξία των αποτελεσμάτων μιας εξέτασης εξαρτάται από ορισμένους παράγοντες που πρέπει να πληρεί η σχετική διαδικασία. Στη βιβλιογραφία αναφέρεται ένα μεγάλο πλήθος παραγόντων που μπορούν να επηρεάσουν τα αποτελέσματα της εξέτασης (Stiggins, 1991). Ένας ευρύτερα αποδεκτός πυρήνας από τους παράγοντες αυτούς είναι:

Η αξιοπιστία: Η αξιοπιστία μιας εξέτασης είναι δύσκολο να οριστεί με ακρίβεια. Οι ορισμοί που δίνονται συνήθως αντανακλούν τη μεθοδολογία που χρησιμοποιείται για τη διαπίστωσή της. Γενικά σαν αξιοπιστία θεωρείται ο βαθμός εμπιστοσύνης που μπορούμε να έχουμε στα αποτελέσματά της, δηλαδή πόσο σταθερά και ακριβή είναι τα αποτελέσματα της μέτρησης (Kerlinger, 1979). Κάθε επίδοση σε μια εξέταση ουσιαστικά αποτελείται από δύο παράγοντες: Από την πραγματική επίδοση και το σφάλμα εκτίμησης της επίδοσης. Το σφάλμα εκτίμησης οφείλεται κυρίως σε αυτό που ο Ebel (1979) ονομάζει

“αξιοπιστία αξιολογητών”, δηλαδή υποκειμενικές διαφορές εκτίμησης της επίδοσης. Επίσης αξιόπιστη επιστημονικά θεωρείται μια εξέταση η οποία όταν επαναληφθεί κάτω από μια παρόμοια μεθοδολογία θα δώσει όμοια αποτελέσματα (Kirk & Miller, 1986; Charles, 1995; Jorpe, 2000).

Η αντικειμενικότητα: Αντικειμενική είναι μια εξέταση, όταν η τελική βαθμολογία δεν επηρεάζεται από άλλους παράγοντες (όπως η συμπάθεια, η αντιπάθεια, η εντοπιότητα κλπ) πέρα από τη επίδοση του κρινόμενου. Η αντικειμενικότητα προϋποθέτει ομοιόμορφο και καθορισμένο τρόπο εκτίμησης της επίδοσης των μαθητών.

Η εγκυρότητα ή κύρος: Η έννοια της εγκυρότητας αναφέρεται στο κατά πόσο η εξέταση επιτελεί σωστά το σκοπό της, δηλαδή αν εξετάζει πραγματικά το αντικείμενο για το οποίο πραγματοποιείται. (Jorpe, 2000; Wainer & Braun, 1998)

Μια εξέταση θεωρείται έγκυρη, όταν τα θέματα που περιέχει:

- Καλύπτουν ένα ευρύ φάσμα της ύλης και δεν επικεντρώνονται μόνο σε μερικά σημεία της.
- Εξετάζουν όχι μόνο τη δυνατότητα απομνημόνευσης των μαθητών αλλά και άλλες γνωστικές δυνατότητές τους όπως η κατανόηση, η ικανότητα εφαρμογής και ανάλυσης (Κανίδης, Ραχωβίτσας, 2005).
- Αντικατοπτρίζουν το πραγματικό επίπεδο γνώσεων ή άλλων δεξιοτήτων των οποίων η ανάπτυξη επιδιώχθηκε κατά την εκπαιδευτική διαδικασία.

Η διακριτικότητα: Σε μια σωστή εξέταση αναμένεται να υπάρχουν κακοί βαθμοί, μέτριοι, σχεδόν καλοί, καλοί, πολύ καλοί και άριστοι. Η εξέταση για να έχει διακριτικότητα, θα πρέπει να κατατάσσει τους μαθητές ανάλογα με τη γνώση του καθενός σε διαφορετικές κατηγορίες αξιολόγησης. Στενά συνδεδεμένη με τη διακριτικότητα είναι και η έννοια της **ευκολίας** μιας εξέτασης. Η ευκολία της εξέτασης υπολογίζεται από το ποσοστό των μαθητών που απάντησαν με επιτυχία στα θέματα. Όσο πιο πολλοί μαθητές απαντούν με επιτυχία τόσο πιο εύκολη θεωρείται η εξέταση. Η μέγιστη διακριτικότητα επιτυγχάνεται, όταν το 50% των μαθητών απαντήσει με επιτυχία, ενώ υπάρχει ελάχιστη διακριτικότητα, όταν απαντούν με επιτυχία όλοι οι μαθητές ή όταν αποτυγχάνουν όλοι οι μαθητές. Θεωρητικά η διακριτικότητα μιας εξέτασης είναι άριστη, όταν δίνει αποτελέσματα που κατανέμονται σύμφωνα με την κανονική κατανομή (Κασσωτάκης, 1981).

Η ΑΞΙΟΠΙΣΤΙΑ ΚΑΙ Η ΑΝΤΙΚΕΙΜΕΝΙΚΟΤΗΤΑ ΤΩΝ ΘΕΜΑΤΩΝ ΣΤΙΣ ΓΕΝΙΚΕΣ ΕΞΕΤΑΣΕΙΣ

Οι γενικές εξετάσεις έχουν καθιερωθεί στη συνείδηση των εκπαιδευτικών, των μαθητών και των γονέων ως αντικειμενικές και αξιόπιστες εξετάσεις. Η αντικειμενικότητα εξασφαλίζεται από το Υπουργείο Παιδείας με την τήρηση μιας ελεγχόμενης και δομημένης διαδικασίας η οποία εξασφαλίζει την αντικειμενικότητα και την αξιοπιστία των εξετάσεων. Η όλη διαδικασία επιλογής των θεμάτων, η μετάδοση τους, η επιτήρηση των μαθητών, η διασπορά των γραπτών σε άλλες περιφέρειες, η ανώνυμη βαθμολόγηση των γραπτών, η διαδικασία αναβαθμολόγησης και η επαλήθευση της βαθμολογίας, ακολουθούν αυστηρούς κανόνες. Κάθε γραπτό βαθμολογείται σε κλίμακα από 0 έως 100 από δύο καθηγητές που έχουν διδάξει το αντίστοιχο μάθημα. Σε περίπτωση διαφοράς βαθμολογίας μεγαλύτερης από 12 μόρια βαθμολογείται και από

τρίτο βαθμολογητή (αναβαθμολόγηση). Η αξιοπιστία των βαθμολογήσεων είναι γενικά υψηλή με τα μαθήματα θετικής κατεύθυνσης να παρουσιάζουν ένα ποσοστό αναβαθμολόγησης 1% έως 5% ενώ στα μαθήματα που υπεισέρχεται περισσότερο ο προσωπικός παράγοντας του βαθμολογητή, όπως η έκθεση, το ποσοστό αναβαθμολόγησης μπορεί να φτάσει και το 30%. Υπάρχουν βέβαια περιθώρια βελτίωσης της αξιοπιστίας της βαθμολογησης. Μερικές προτάσεις προς αυτή την κατεύθυνση είναι και οι ακόλουθες: α) Η συγκρότηση ειδικού σώματος “βαθμολογητών” που είχε προταθεί παλαιότερα από το Παιδαγωγικό Ινστιτούτο. β) Η πιλοτική υποδειγματική βαθμολόγηση γραπτών σε Εθνικό επίπεδο από τους συντονιστές, που είχε προταθεί από την ηγεσία του ΥΠΕΠΘ. γ) Ειδική επιμόρφωση των βαθμολογητών δ) Έλεγχος, ενημέρωση και πιθανά απομάκρυνση των βαθμολογητών που παρουσιάζουν συστηματικά πολλές, ακραίες αναβαθμολογήσεις.

Η ΕΓΚΥΡΟΤΗΤΑ ΤΩΝ ΘΕΜΑΤΩΝ

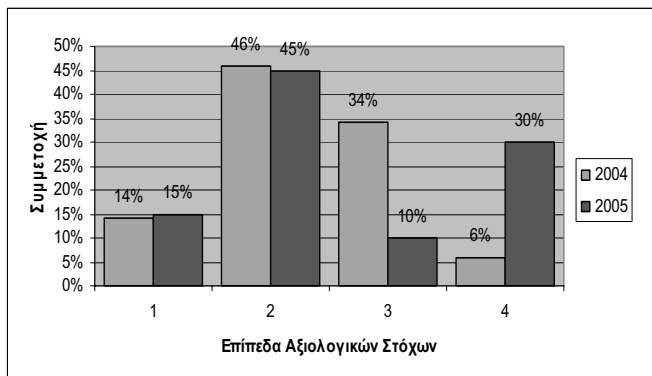
Ο Δημητρόπουλος (1998) τονίζει ότι η εγκυρότητα (ή κύρος) μια εξέτασης είναι σπουδαιότερος παράγοντας από την αξιοπιστία, γιατί “είναι προτιμότερο να έχουμε μέτρηση έστω και μέτριας ακρίβειας (χαμηλή αξιοπιστία) αλλά να μετράμε το σωστό μέγεθος (υψηλό κύρος) παρά πολύ αξιόπιστη μέτρηση ενός λανθασμένου μεγέθους”. Η τοποθέτηση αυτή δείχνει την προσπάθεια που πρέπει να καταβάλει η γενική επιτροπή εξετάσεων για την επιλογή έγκυρων θεμάτων.

Στις γενικές Εξετάσεις του μαθήματος ΑΕΠΠ η μορφή των θεμάτων καθορίζεται από το ΠΔ 246/98. Το διάταγμα καθορίζει ότι τα θέματα πρέπει να είναι τέσσερα και να περιέχουν 40 % θεωρία και 60 % ασκήσεις. Σχετικά με τη μορφή των θεμάτων το ΠΔ δεν είναι πολύ δεσμευτικό σε αντίθεση με άλλα θετικά μαθήματα όπως η Φυσική και τα Μαθηματικά των οποίων η μορφή καθορίζεται με μεγαλύτερη ακρίβεια. (Κανίδης, Ραχωβίτσας, 2005). Στους μαθητές που εξετάζονται στο μάθημα ΑΕΠΠ δίνονται προς διαπραγμάτευση τέσσερα θέματα με βαρύτητα 40, 20, 20 και 20 μόρια αντίστοιχα. Το πρώτο θέμα αποσκοπεί στη διερεύνηση του γνωστικού επιπέδου “ανάκληση” και “κατανόηση” εννοιών. Περιλαμβάνει ορισμούς, ερωτήσεις σύντομης ανάπτυξης, πολλαπλών επιλογών, αντιστοίχισης, επιλογής σωστού – λάθους κλπ. Το δεύτερο θέμα είναι άσκηση και αφορά κυρίως τον έλεγχο του γνωστικού επιπέδου “κατανόηση” του περιεχομένου των μαθημάτων. Συνήθως στο θέμα αυτό δίδεται τμήμα προγράμματος (σε Γλώσσα) ή αλγόριθμου και ζητείται από τους μαθητές να παρακολουθήσουν τις τιμές των μεταβλητών κατά τη διάρκεια της εκτέλεσης. Το τρίτο και το τέταρτο θέμα είναι ασκήσεις που ελέγχουν το γνωστικό επίπεδο “εφαρμογή” και “ανάλυση” των μαθητών. Για την επίλυση των θεμάτων αυτών οι μαθητές πρέπει να εφαρμόσουν αλγοριθμικές δομές οι οποίες συχνά δεν είναι προφανείς αλλά προκύπτουν μετά από ανάλυση των απαιτήσεων του προβλήματος.

Τα θέματα των εξετάσεων, όλα τα εξεταζόμενα χρόνια είχαν μια ευρύτητα και κάλυπταν το μεγαλύτερο μέρος της ύλης. Το θεωρητικό σκέλος των θεμάτων σε όλες τις εξετάσεις καλύπτει πάνω από το 70% της θεωρίας του μαθήματος.

Σχετικά με την εξέταση των γνωστικών ικανοτήτων (ανάκληση, κατανόηση, εφαρμογή και ανάλυση) τα θέματα όλων των ετών κάλυψαν πλήρως τα

τρία πρώτα επίπεδα με θεωρία και ασκήσεις. Σχετικά με το τέταρτο επίπεδο (ανάλυση) διαπιστώνεται ότι τα έτη 2002 και 2003 ουσιαστικά απουσιάζει, ενώ παρουσιάζεται μια αυξημένη προοδευτικά συμμετοχή της στα θέματα των ετών 2004 και 2005 (Κανίδης, Φανίκος, 2005) (βλέπε σχήμα 1).



Σχήμα 1. Συγκριτική Κατανομή Μονάδων σε Επίπεδα Αξιολογικών Στόχων.

ΔΙΑΚΡΙΤΙΚΟΤΗΤΑ ΤΩΝ ΘΕΜΑΤΩΝ

Ένα ισχυρό και διαδεδομένο εργαλείο μελέτης και αξιολόγησης των αποτελεσμάτων μιας εξεταστικής διαδικασίας είναι η στατιστική συμπεριφορά των βαθμολογικών κατανομών των εξεταζομένων. Από τους περισσότερους ερευνητές γίνεται αποδεκτό, ότι μια βαθμολογική κατανομή είναι τόσο πιο “κανονική” όσο πιο ομαλή στατιστικά είναι. Η επιστήμη της στατιστικής θεωρεί ότι δύο πλέον ομαλές κατανομές είναι οι κατανομές Gauss και Poisson. Ειδικότερα η κατανομή Gauss ονομάζεται “κανονική κατανομή” και αποτελεί ένα πρότυπο για τη διακριτικότητα ενός μεγέθους. Σύμφωνα με αυτό, άριστη διακριτικότητα έχει μια εξεταστική διαδικασία, όταν οι βαθμοί των εξετασθέντων κατανέμονται σύμφωνα με μια τυπική κατανομή Gauss. Αν η βαθμολογία των μαθητών ακολουθούσε την κατανομή αυτή θα πρέπει ο μέσος όρος της βαθμολογίας θα πρέπει να είναι το 10, οι περισσότεροι μαθητές να έχουν βαθμολογηθεί με 10 και το 76% των μαθητών να έχει βαθμολογία από 7 έως 13.

Αν η μορφή της καμπύλης βρίσκεται μετατοπισμένη προς το δεξιό μέρος της βαθμολογίας (μεγάλοι βαθμοί), τότε η εξέταση χαρακτηρίζεται ως εύκολη. Αντίθετα αν η καμπύλη είναι μετατοπισμένη προς το αριστερό μέρος της βαθμολογίας τότε η εξέταση χαρακτηρίζεται δύσκολη.

Ο Δημητρόπουλος (1998) προτείνει μια εναλλακτική προσέγγιση της εύρεσης της διακριτικότητας μια εξέτασης παραδοσιακού τύπου όπως είναι οι Γενικές Εξετάσεις.

Η μέθοδος που προτείνει είναι: Η βαθμολογία των μαθητών χωρίζεται σε τρεις ομάδες. α) Εκείνη που αποτελεί το άνω 1/3 των επιδόσεων που ονομάζεται Άνω Ομάδα (ΑΟ). β) Εκείνη που αποτελεί το κάτω 1/3 των επιδόσεων που ονομάζεται Κάτω Ομάδα (ΚΟ). Ο δείκτης διακριτικότητας της εξέτασης υπολογίζεται με δύο τύπους χρησιμοποιώντας τα αθροίσματα ή τους μέσους όρους των βαθμών της κάθε ομάδας. Για την περίπτωση της χρήσης των μέσων όρων επίδοσης ο τύπος είναι:

$$\Delta\Delta T = \frac{\bar{X}AO - \bar{X}KO}{X_{\max}}$$

όπου: $\bar{X}AO$ = Μέσος όρος βαθμών που πήρε η Άνω Ομάδα $\bar{X}KO$ = Μέσος όρος βαθμών που πήρε η Κάτω Ομάδα X_{\max} = Μέγιστος βαθμός που μπορεί να πάρει ένας μαθητής.

Ο δείκτης διακριτικότητας μπορεί να πάρει τιμές από 0 έως 1 και θεωρείται ικανοποιητικός, όταν είναι πάνω από το 0,40 (Ebel, 1979).

Για τον υπολογισμό της διακριτικότητας των θεμάτων των ετών 2002-2005 θα χρησιμοποιηθεί για την Άνω και Κάτω Ομάδα το ποσοστό 25% (δηλαδή οι βαθμοί 0 – 4.99 & 15.01 – 20) γιατί όπως τονίζει και ο ίδιος ο Δημητρόπουλος (1991 σελ 228): “Στατιστικές αναλύσεις έχουν δείξει πως καλύτερα αποτελέσματα προκύπτουν αν χρησιμοποιηθεί το 22-27%...”

Ο δείκτης διακριτικότητας των γενικών εξετάσεων με τη χρήση του τύπου (1) εμφανίζεται στον πίνακα 1.

Πίνακας 1. Δείκτες διακριτικότητας των Γενικών Εξετάσεων 2002-2005.

Έτος	2002	2003	2004	2005
Δείκτης Διακριτικότητας	0,68	0,62	0,74	0,72

Παρατηρούμε ότι ο δείκτης διακριτικότητας ($\Delta\Delta$) όλων των εξετάσεων είναι ικανοποιητικός και η μεγαλύτερη διακριτικότητα παρουσιάζεται στις εξετάσεις των ετών 2004-2005. Από έρευνα που πραγματοποιήθηκε στο 33^ο Βαθμολογικό Κέντρο Αθηνών σε δείγμα 276 γραπτών το έτος 2004 και 650 γραπτών το έτος 2005 προκύπτει μια διαφοροποίηση του Δείκτης Διακριτικότητας των ετών αυτών σε σχέση με τον πίνακα.

Με βάση τα δείγματα που αναφέρθηκαν οι Δείκτες Διακριτικότητας είναι: $\Delta\Delta_{2004} = 0,70$ και $\Delta\Delta_{2005} = 0,80$ δηλαδή προκύπτει μια σαφής υπεροχή του $\Delta\Delta$ του 2005.

Ενδιαφέροντα συμπέρασμα επίσης προκύπτουν από την κατανομή της βαθμολογίας τα έτη 2002 έως 2005.

Στο σχήμα 2 παρατηρούμε ότι η κατανομή βαθμολογίας το έτος 2002 ήταν ακριβώς αντίθετη από αυτή του έτους 2003. Το έτος 2004 παρουσιάζει μια εξομοίωση των τριών πρώτων κλάσεων με ταυτόχρονη αύξηση της υψηλής βαθμολογίας. Το έτος 2005 η κατανομή δείχνει να πλησιάζει περισσότερο προς την κανονική κατανομή.

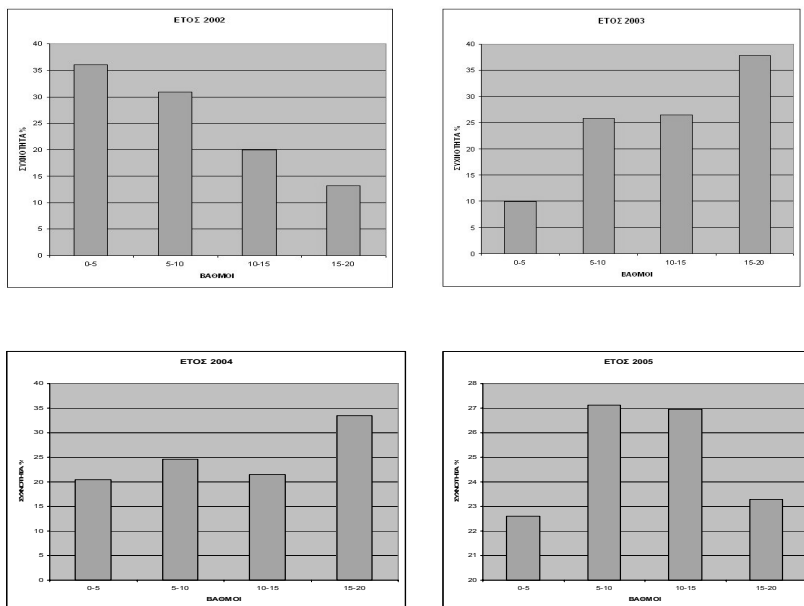
ΣΥΜΠΕΡΑΣΜΑΤΑ – ΠΡΟΤΑΣΕΙΣ

Η αξιοπιστία και η αντικειμενικότητα των εξετάσεων, λόγω του μηχανισμού που έχει αναπτυχθεί από το ΥΠΕΠΘ, βρίσκονται σε υψηλό επίπεδο. Η αντικειμενικότητα της βαθμολόγησης είναι ικανοποιητική αλλά έχει περιθώρια βελτίωσης. Η υιοθέτηση προτάσεων όπως η ύπαρξη σώματος βαθμολογητών, ενιαία πιλοτική βαθμολόγηση, υποδειγματικές “βαθμολογήσεις” και ειδική επιμόρφωση βαθμολογητών μπορούν να βελτιώσουν ακόμα περισσότερο την αντικειμενικότητα της βαθμολόγησης.

Η εγκυρότητα των θεμάτων είναι σε υψηλό επίπεδο με τα θέματα να πα-

ρουσιάζουν ικανοποιητικό ποσοστό κάλυψης της ύλης. Η σταδιακή εισαγωγή θεμάτων που απαιτούν ανάλυση για την αντιμετώπισή τους αποτελούν εν δυνάμει στοιχεία βελτίωσης.

Η διακριτικότητα των θεμάτων εμφανίζει αμφιλεγόμενη εικόνα. Ο υπολογισμός του δείκτη διακριτικότητας βρίσκεται σε υψηλά επίπεδα, αλλά η κατανομή της βαθμολογίας απέχει πολύ από την εικόνα της κανονικής κατανομής. Διαφορές στην κατανομή της βαθμολογίας υπάρχουν από έτος σε έτος με χαρακτηριστικότερες εκείνες του 2002-2003 όπου οι κατανομές είναι ακριβώς αντίθετες. Η κατανομή του 2005 παρουσιάζει μια καλύτερη εικόνα χωρίς να είναι άριστη. Μια πιο προσεκτική επιλογή του γνωστικού επιπέδου των θεμάτων πιστεύουμε ότι θα βελτιώσει περισσότερο την εικόνα της κατανομής βαθμολογίας.



Σχήμα 2. Κατανομές Βαθμολογίας τα έτη 2002 έως 2005.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Δημητρόπουλος Γ. Ε.,(1998), Εκπαιδευτική αξιολόγηση η αξιολόγηση του μαθητή, εκδόσεις Γρηγόρη, Αθήνα.
- Ενιαίο Λύκειο (1998): Ενιαίο Πλαίσιο Προγραμμάτων Σπουδών, Η εισήγηση του Παιδαγωγικού Ινστιτούτου. ΥΠΕΠΘ-ΠΙ
- Κανίδης Ε., Ραχωβίτσας Η., (2005), Αξιολόγηση των Θεμάτων του Μαθήματος “Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον” στις Πανελλαδικές Εξετάσεις 2003-2004, 3^ο Πανελλήνιο Συνέδριο Εκπαιδευτικών για τις ΤΠΕ με θέμα “Αξιοποίηση των Τεχνολογιών της Πληροφορίας και της Επικοινωνίας στη Διδακτική Πράξη”. Σύρος 13-15 Μαΐου 2005. Πρακτικά του συνεδρίου σε CD.
- Κανίδης Ε., Φανίκος Α., (2005) “Αξιολόγηση των Θεμάτων του Μαθήματος “Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον” στις Γενικές Εξετάσεις

- Ενιαίων Λυκείων 2004-2005” 3ο Πανελλήνιο Συνέδριο με θέμα “Διδακτική της Πληροφορικής”, Πρακτικά του συνεδρίου σε CD.
- Κασσωτάκης Ι. Μ., (1981), Η αξιολόγηση της επιδόσεως των μαθητών, εκδόσεις Γρηγόρη, Αθήνα
- Charles, C. M. (1995). *Introduction to educational research* (2nd ed.). San Diego, Longman.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Toronto: Holt, Rinehart, and Winston, Inc.
- Dietel, 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
- Ebel, R. L., (1979), *Essentials of education Measurement*, Prentice Hall, N.J.
- Ediger, Marlow (1994), “Measurement and Evaluation,” *Studies in Educational Evaluation* 20 (2). 169-174.
- Geiger M.A & Cooper E.A (1995). Predicting academic performance: the impact of Expectancy and Needs Theory. *Journal of Experimental Education*, 63(3):251-262.
- Golafshani Nahid, (2003), *Understanding Reliability and Validity in Qualitative Research*, *The Qualitative Report* Volume 8 Number 4 597-607
- Joppe, M. (2000). *The Research Process*. Retrieved February 25, 1998, from <http://www.ryerson.ca/~mjoppe/rp.htm>
- Kerlinger, R. J., (1979), *Foundations of behavioural Research*, Holt, Rinehart & Winston, N.Y
- Kirk, J., & Miller, M. L. (1986). *Reliability and validity in qualitative research*. Beverly Hills: Sage Publications.
- Li, H. (2003). The resolution of some paradoxes related to reliability and validity. *Journal of Educational and Behavioral Statistics*, 28, 89-95.
- Mislevy, R. (2004). Can there be reliability without reliability? *Journal of Educational and Behavioral Statistics*, 29, 241-244.
- Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, 23, 5-12.
- Stiggins, R. J. (1991). *Assessment literacy*. *Phi Delta Kappan*. 72 (7)
- Wainer, H., & Braun, H. I. (1988). *Test validity*. Hilldale, NJ: Lawrence Earlbaum Associates.