



ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

1ο Πανελλήνιο Συνέδριο
για την Ανοικτή & εξ Αποστάσεως Εκπαίδευση

Πάτρα, 25 - 27 Μαΐου 2001

ΠΡΑΚΤΙΚΑ ΕΙΣΗΓΗΣΕΩΝ

Επιμέλεια: Αντώνης Λιοναράκης

ΤΟΜΟΣ Β'

Η Αξιολόγηση των Γραπτών Εργασιών στην «Αεξαε»: Οι Αξιολογητές και τα Κριτήρια Βαθμολόγησης

Χρήστος Παναγιωτακόπουλος^[1], Γεράσιμος Κουστουράκης^[2].

[1]Ελληνικό Ανοικτό Πανεπιστήμιο (cranag@eap.gr)

[2]Ελληνικό Ανοικτό Πανεπιστήμιο (koustger@hol.gr)

ΠΕΡΙΛΗΨΗ

Η αξιολόγηση των φοιτητών στο ΕΑΠ γίνεται και μέσα από τις γραπτές εργασίες τύπου δοκιμίου. Σε μερικές Θεματικές Ενότητες (όπως η Ανοικτή και εξ Αποστάσεως Εκπαίδευση – Αεξαε), τα μέλη ΣΕΠ έχουν διαφορετικό γνωστικό υπόβαθρο και η αξιολόγηση των εργασιών μπορεί να αποτελέσει πρόβλημα, αφού συναρτάται και με το πεδίο γνώσης του ΣΕΠ.

Στην παρούσα μελέτη δόθηκε για αξιολόγηση μια κοινότυπη εργασία της Αεξαε σε δύο ομάδες μελών ΣΕΠ από διάφορες Θ.Ε., εκ των οποίων η μία χρησιμοποίησε για να την αξιολογήσει συγκεκριμένα κριτήρια, ενώ η άλλη στηρίχτηκε στην ακαδημαϊκή της εμπειρία. Τα αποτελέσματα έδειξαν στατιστικά σημαντικές διαφορές στη βαθμολογία μεταξύ των δύο προηγούμενων ομάδων. Αντίθετα δεν παρατηρήθηκαν στατιστικά σημαντικές διαφορές: (α) στη βαθμολογία μεταξύ φύλου, (β) στη βαθμολογία μεταξύ αυτών που είναι μέλη ΣΕΠ στη Θ.Ε. Αεξαε και αυτών που είναι σε άλλες Θ.Ε., και (γ) στη βαθμολογία μεταξύ αυτών που κατά το παρελθόν είχαν παρακολουθήσει επιτυχώς (ως φοιτητές) τη Θ.Ε. της Αεξαε και αυτών που δεν είχαν.

ABSTRACT

The students' evaluation in Humanities Faculty of the Hellenic Open University (H.O.U.) is mostly achieving by essays Tutor Marked Assignment (T.M.A.). Also, in several Modules, as Open Distance Education, the teachers have different cognitive background and the evaluation of T.M.A. may be problematic.

In this research paper a common homework about Open Distance Education is given for evaluation in two groups of H.O.U. teachers, which work in different cognitive units. The first group has used concrete and specific criteria for the evaluation of the T.M.A. and the second has evaluated this by rely upon their own academic experience and sufficiency. The outcomes of our research have shown statistical significant differences between the two groups, watching the grades that have put in the essay T.M.A. Doesn't exist statistical significant differences in the grades of the essay depending of the: (a) sex of evaluators, (b) module that the evaluators is teaching, and (c) previous attendance of the Open Distance Education Module.

1. Εισαγωγή

Οι φοιτητές των ιδρυμάτων τριτοβάθμιας εκπαίδευσης και ειδικότερα εκείνοι των Ανοικτών Συστημάτων Εκπαίδευσης (Ανοικτών Πανεπιστημίων) αξιολογούνται συνήθως είτε μέσα από γραπτές εργασίες ή/και από τη συμμετοχή τους στις τελικές γραπτές εξετάσεις των διαφόρων

γνωστικών αντικειμένων. Και στις δύο περιπτώσεις οι φοιτητές καλούνται ν' αξιοποιήσουν την ικανότητα έκφρασης των σκέψεών τους με το γραπτό λόγο για ν' αποδείξουν ότι κατέκτησαν τις απαραίτητες επιστημονικές γνώσεις.

Ο ρόλος των γραπτών εργασιών είναι ιδιαίτερα κρίσιμος και ουσιώδης στην περίπτωση των Ανοικτών Πανεπιστημίων. Στο Ελληνικό Ανοικτό Πανεπιστήμιο (ΕΑΠ), ένα μέρος του τελικού βαθμού (30%) προέρχεται από τη βαθμολογία των γραπτών εργασιών¹¹. Επιπλέον, η χρησιμότητα των συγκεκριμένων εργασιών στο ΕΑΠ είναι μεγάλη διότι, αφενός, μέσα από αυτές επιτυγχάνεται η ανατροφοδότηση κατά τη μαθησιακή πορεία των φοιτητών και αφετέρου, γίνεται, ουσιαστικά, η εκπλήρωση από τα μέλη ΣΕΠ του διδακτικού τους έργου (Ματραλής, 1998). Στις σχολές Ανθρωπιστικών Επιστημών η συχνότερα χρησιμοποιούμενη μορφή γραπτών εργασιών είναι αυτή του τύπου δοκιμίου. Μάλιστα, η αξιολόγηση μιας γραπτής εργασίας χαρακτηρίζεται ως (Ματραλής, 1998):

- Έγκυρη, όταν αποτιμά πράγματι τις γνώσεις και τις ικανότητες των φοιτητών,
- Αξίопιστη, όταν δίνει όμοια αποτελέσματα διεξαγόμενη κάτω από παρόμοιες συνθήκες και
- έντιμη, όταν στοχεύει στον έλεγχο της κατανόησης των προσδοκώμενων αποτελεσμάτων.

2. Αξιολόγηση Γραπτών Εργασιών: Θεωρητική Προσέγγιση

Σύμφωνα με τους ερευνητές του χώρου ο προβληματισμός που αναπτύσσεται σχετικά με τον τρόπο αξιολόγησης των εργασιών των φοιτητών εστιάζεται σε δύο σημεία (Graham & Weintraub, 1996 – Dimakos, 1998):

- 1) στη χρησιμοποίηση διαφόρων τύπων τεστ αξιολόγησης των γραπτών εργασιών,
- 2) στην αναγνώριση της ικανότητας των εκπαιδευτικών, που εργάζονται επαγγελματικά σε ένα συγκεκριμένο επιστημονικό πεδίο και ασχολούνται διδακτικά με ένα συγκεκριμένο τμήμα σπουδαστών, να μπορούν ν' αξιολογούν σωστά τους τελευταίους.

Αναφορικά με τον τρόπο βαθμολόγησης των γραπτών εργασιών διακρίνουμε τις παρακάτω τρεις προτάσεις αξιολόγησης (Perkins, 1983 – Shohamy, Gordon & Kraemer, 1992):

- Ολική βαθμολόγηση (holistic scoring): Αναφέρεται στην αξιολόγηση που βασίζεται στη συνολική εντύπωση που έχουν οι αξιολογητές από το γραπτό κείμενο. Χρησιμοποιείται κυρίως για τους σκοπούς της συνολικής – τελικής αξιολόγησης (summative).
- Αναλυτική βαθμολόγηση (analytic scoring): Αναφέρεται στην αξιολόγηση που προϋποθέτει τη διαίρεση του γραπτού κειμένου σε συγκεκριμένες και εξειδικευμένες περιπτώσεις – κατηγορίες όπως π.χ. γραμματικά, πραγματολογικά ή λεξιλογικά στοιχεία.
- Βαθμολόγηση των βασικών – πρωταρχικών γνωρισμάτων – προϋποθέσεων (primary trait scoring): Ο όρος αναφέρεται στη διαδικασία κατά την οποία δίνεται ένας συνολικός βαθμός (holistic score) στα διάφορα χαρακτηριστικά του γραπτού. Η διαδικασία αυτή απαιτεί την εξακρίβωση του εάν ένα γραπτό διαθέτει κάποια πολύ σημαντικά χαρακτηριστικά στοιχεία, τα οποία και ο γράφων προσπαθεί να εγυτοπίσει.

Οι παράγοντες, τέλος, που επηρεάζουν την αξιοπιστία της αξιολόγησης γραπτών εργασιών εντοπίζονται στα εξής (Breland, 1983 – Shohamy, Gordon & Kraemer, 1992):

- Στην επαγγελματική θέση του αξιολογητή (professional background): Για παράδειγμα όσοι

εμπλέκονται διδακτικά με ένα συγκεκριμένο γνωστικό τομέα μπορεί να επηρεαστούν από τους υπάρχοντες εκπαιδευτικούς στόχους και να δώσουν έμφαση σε εξειδικευμένα στοιχεία του γραπτού, που αξιολογούν. Από την άλλη πλευρά όσοι διδάσκουν σε διαφορετικά γνωστικά αντικείμενα μπορεί να προσεγγίσουν το ίδιο γραπτό με διαφορετικό τρόπο, βασιζόμενοι στις προσωπικές τους επιστημονικές θεωρήσεις. Έτσι το ερώτημα που προκύπτει είναι: Κατά πόσον οι αξιολογητές που απασχολούνται σε διαφορετικά γνωστικά αντικείμενα διαφέρουν ως προς τον τρόπο βαθμολόγησης ενός συγκεκριμένου γραπτού;

- Στην εκπαίδευση (training). Η εντατική εκπαίδευση στις διαδικασίες και τις προϋποθέσεις αξιολόγησης των γραπτών εργασιών στην Ανοικτή Εκπαίδευση μπορεί να οδηγήσει στη συναίνεση των αξιολογητών αναφορικά με την επιστημονική ποιότητα ενός γραπτού. Εδώ το ερώτημα που προκύπτει είναι: Κατά πόσον οι βαθμολογητές που έχουν εκπαιδευτεί σχετικά με τις προϋποθέσεις και τον τρόπο παροχής σπουδών από απόσταση εμφανίζουν μεγαλύτερο βαθμό αξιοπιστίας, κατά την αξιολόγηση ενός γραπτού, απ' ό,τι εκείνοι που δεν έχουν;
- Στην κλίμακα - τρόπο βαθμολόγησης (rating scale): Ο τρόπος βαθμολόγησης (με ή χωρίς κριτήρια) είναι το εργαλείο που επηρεάζει το προϊόν της αξιολόγησης ενός γραπτού. Το ερώτημα που προκύπτει είναι: Εάν χρησιμοποιήσουμε μια δεδομένη κλίμακα αξιολόγησης, η οποία στηρίζεται σε συγκεκριμένα κριτήρια, μπορούμε να αυξήσουμε την αξιοπιστία της βαθμολόγησης ενός γραπτού τύπου δοκιμίου;

Εστιάζοντας, στη συνέχεια, την προσοχή μας ενδεικτικά σε σχετικές έρευνες, που ασχολούνται με την αξιολόγηση γραπτών εργασιών επισημαίνουμε τα εξής^[2]:

- Τα αποτελέσματα της έρευνας των Shohamy, Gordon & Kraemer (1992), στην οποία έγινε η αξιολόγηση γραπτών εργασιών επιπέδου proficiency από βαθμολογητές, που διέφεραν μεταξύ τους ως προς την επαγγελματική απασχόληση και την εκπαίδευσή τους στο συγκεκριμένο γνωστικό τομέα, φανέρωσαν ότι: α) ανεξάρτητα από την κλίμακα βαθμολόγησης, που χρησιμοποιήθηκε, υπήρξε σημαντική διαφοροποίηση στις βαθμολογήσεις ανάλογα με το επίπεδο εκπαίδευσης των αξιολογητών, και β) η επαγγελματική εξειδίκευση δεν έπαιξε σημαντικό ρόλο στον τρόπο αξιολόγησης.
- Στην έρευνα της Freedman (1979) αξιολογήθηκαν με τη χρήση μιας ολικής κλίμακας τεσσάρων περιπτώσεων (4-point holistic rating scale) εργασίες φοιτητών κολεγίου από έμπειρους αξιολογητές. Τα τέσσερα σημεία - περιπτώσεις της κλίμακας αυτής ήταν: α) περιεχόμενο (content), β) οργάνωση - δομή της εργασίας (organization), γ) σύνταξη προτάσεων - παραγράφων, ορθή εκφραστική χρήση της γλώσσας (sentence structure) και δ) μηχανισμός γραφής - ορθογραφία (mechanics). Τα αποτελέσματα της συγκεκριμένης έρευνας έδειξαν ότι οι αξιολογητές έδωσαν μεγαλύτερη βαρύτητα στο περιεχόμενο και τη δομή των εργασιών που προσέγγισαν.
- Σημαντικός αριθμός ερευνών, τέλος, έχει εστιάσει το ενδιαφέρον του στην ανίχνευση του κατά πόσον οι εξειδικευμένοι στο αντικείμενό τους εκπαιδευτικοί μπορούν να κρίνουν και να αξιολογήσουν με επιτυχία (βασίζόμενοι στην επαγγελματική τους επάρκεια κι εμπειρία) τους μαθητές τους και μάλιστα σε βαθμό τόσο αξιόπιστο όσο με τη χρήση αντικειμενικά σταθμισμένων τεστ. Τα αποτελέσματα των ερευνών αυτών φανέρωσαν την ικανότητα των συγκεκριμένων εκπαιδευτικών στο να κρίνουν και να βαθμολογούν με μεγάλη αξιοπιστία

την επίδοση και τις δυνατότητες των μαθητών τους, όπως συμβαίνει και με τη χρήση κατάλληλων κλιμάκων αξιολόγησης της ακαδημαϊκής επάρκειας των φοιτητών. (Hoge, 1983 – Neeper & Lahey, 1986 – Gresham, Reschly & Carey, 1987 – DuPaul, Rapport & Perriello, 1991 – Salvensen & Undheim, 1994 – Gresham, MacMillan & Bocian, 1997).

Αναφορικά με την παροχή Ανοικτής κι Εξ Αποστάσεως Εκπαίδευσης οι Race (1993) και Rowntree (1997) υποστηρίζουν την αναγκαιότητα της σωστής βαθμολόγησης των γραπτών εργασιών των φοιτητών. Για το λόγο αυτό θεωρούν ως απαραίτητο το σχεδιασμό από τους συντάκτες των διαφόρων πακέτων παροχής σπουδών ΑεξΑΕ ενός μοντέλου αξιολόγησης, με το οποίο να φανερώνονται με ακρίβεια τα σημεία, που μετρούν θετικά ή αρνητικά σε μια γραπτή εργασία, όπως και το πόσοι βαθμοί χάνονται ή κερδίζονται σε καθένα από αυτά. Μάλιστα τα συγκεκριμένα μοντέλα εκτός από τη χρηστική τους αξία, από την πλευρά των διδασκόντων, διαδραματίζουν ταυτόχρονα και διπλό θετικό παιδαγωγικό ρόλο. Δηλαδή παρέχουν, αφενός, ανατροφοδότηση στο φοιτητή και τον πληροφορούν για τα αδύνατα σημεία της προσπάθειάς του, όπως και τον τρόπο υπέρβασής τους στη συνέχεια. Και αφετέρου μαθαίνουν στους φοιτητές τις απαιτήσεις – προϋποθέσεις και τις τεχνικές σύνταξης των διαφόρων επιστημονικών εργασιών.

3. Ερευνητικά Ερωτήματα

Η έρευνα που ακολουθεί έχει ως στόχο να εξετάσει την σταθερότητα – εγκυρότητα της βαθμολόγησης γραπτών εργασιών τύπου δοκιμίου από ΣΕΠ του ΕΑΠ στους οποίους έχουν δοθεί ή όχι συγκεκριμένα κριτήρια βαθμολόγησης και που διαφέρουν μεταξύ τους ως προς την εμπλοκή τους με τη Θεματική Ενότητα της ΑεξΑΕ. Πιο συγκεκριμένα, η εμπλοκή αυτή των ΣΕΠ εστιάζεται σε δύο επίπεδα: α) στο κατά πόσον εργάζονται ή όχι στη Θ.Ε. της ΑεξΑΕ, από το επιστημονικό πεδίο της οποίας είναι παρμένη η υπό αξιολόγηση εργασία και β) στο κατά πόσον έχουν παρακολουθήσει επιτυχώς τη Θεματική αυτή Ενότητα στο ΕΑΠ ή έχουν κάνει σχετικές σπουδές σε άλλο εκπαιδευτικό ίδρυμα.

Το κύριο ερευνητικό μας ερώτημα είναι το κατά πόσον επιδρά στην αξιολόγηση γραπτών εργασιών τύπου δοκιμίου η παροχή ή όχι συγκεκριμένων κριτηρίων αξιολόγησης. Επίσης, εξετάζονται και τα εξής:

- 1) Ποια είναι η επίδραση του παράγοντα της επαγγελματικής εξειδίκευσης στο ΕΑΠ κατά τη διόρθωση συγκεκριμένης γραπτής εργασίας;
- 2) Ποια είναι η επίδραση του παράγοντα «εκπαίδευση στην ΑεξΑΕ» κατά τη διόρθωση γραπτής εργασίας τύπου δοκιμίου στο ΕΑΠ;
- 3) Ποια κριτήρια θεωρούν οι ΣΕΠ ως σημαντικά όταν τους ζητηθεί να αξιολογήσουν συγκεκριμένη γραπτή εργασία τύπου δοκιμίου βασιζόμενοι μόνο και μόνο στην επιστημονική τους επάρκεια;

4. Μεθοδολογία

Από τον πληθυσμό των 248 διδασκόντων (ΣΕΠ) σε όλες τις Θεματικές Ενότητες του ΕΑΠ, επελέγη τυχαία δείγμα 120 ΣΕΠ (οι 25 ήσαν κατά το τρέχον ακαδημαϊκό έτος ΣΕΠ στην ΑεξΑΕ) και στους οποίους απεστάλη προς αξιολόγηση γραπτή εργασία με τίτλο: «Επισημάνετε και ανα-

λύστε τις ομοιότητες και τις διαφορές ενός τριτοβάθμιου ιδρύματος ανοικτής και εξ αποστάσεως εκπαίδευσης, όπως είναι το ΕΑΠ, σε σχέση με κάποιο άλλο Ελληνικό τριτοβάθμιο εκπαιδευτικό ίδρυμα».

Όπως φαίνεται και από τον τίτλο της εργασίας, το θέμα είναι κοινότυπο. Έτσι, ακόμα και αν ο βαθμολογητής δεν εργάζεται ως ΣΕΠ στην ΑεΞΑΕ, από την εμπλοκή του μόνο και μόνο στις διαδικασίες παροχής σπουδών στο ΕΑΠ έχει αποκτήσει το απαραίτητο γνωστικό υπόβαθρο ώστε να μπορέσει να την αξιολογήσει.

Το δείγμα μας χωρίστηκε τυχαία σε δύο ομάδες από 60 ΣΕΠ η καθεμία. Στην πρώτη ομάδα, εστάλησαν μαζί με την εργασία συγκεκριμένα κριτήρια βαθμολόγησης και ζητήθηκε από τους ΣΕΠ να τη βαθμολογήσουν ανώνυμα βασιζόμενοι στα κριτήρια αυτά. Επίσης κλήθηκαν να συμπληρώσουν και τα εξής πεδία σ' ένα σύντομο ερωτηματολόγιο:

- βαθμός εργασίας,
- επιτυχής παρακολούθηση, κατά το παρελθόν, της Θεματικής Ενότητας ΑεΞΑΕ,
- εργασία στη Θ.Ε. ΑεΞΑΕ ή αλλού,
- φύλο, και
- μερικά γενικά στοιχεία αναφορικά με το γνωστικό υπόβαθρο και την εμπλοκή τους στο ΕΑΠ.

Στη δεύτερη ομάδα των υπόλοιπων 60 ΣΕΠ, στάλθηκε η γραπτή εργασία χωρίς κριτήρια αξιολόγησης με σκοπό να την βαθμολογήσουν ανώνυμα στηριζόμενοι στην ακαδημαϊκή τους εμπειρία κι επάρκεια. Μάλιστα, κλήθηκαν και αυτοί να συμπληρώσουν τα ίδια πεδία συντόμου ερωτηματολογίου, όπως και οι συνάδελφοί τους της πρώτης ομάδας. Επιπλέον, τους ζητήθηκε ν' αναφέρουν με σαφήνεια τα κριτήρια αξιολόγησης, τα οποία θεώρησαν ως χρήσιμα και στα οποία στηρίχθηκαν για τη διόρθωση και βαθμολόγηση της γραπτής εργασίας, που τους δόθηκε.

Για την τυχαία επιλογή των ΣΕΠ όσο και για τον τυχαίο διαχωρισμό τους σε δύο ομάδες χρησιμοποιήθηκαν οι αλγόριθμοι *Shift Register Sequence (R250)* (Kirkpatrick & Stoll, 1981) και *Prime Modulus Multiplicative Linear Congruential algorithm* (Park & Miller, 1988). Οι προηγούμενοι αλγόριθμοι χρησιμοποιούνται και για την τυχαία επιλογή των υποψηφίων φοιτητών στο ΕΑΠ.

Τελικά απάντησαν συνολικά 63 ΣΕΠ (52,5% του συνόλου που επελέγησαν), από τους οποίους:

- 33 χρησιμοποίησαν τα κριτήρια που τους είχαν αποσταλεί προκειμένου να βαθμολογήσουν την εργασία και
- 30 τη βαθμολόγησαν στηριζόμενοι στην εμπειρία τους.

Μερικά βασικά χαρακτηριστικά του δείγματος που απάντησε έμπρακτα στην πρόσκλησή μας να βαθμολογήσει την εργασία, έχουν ως εξής:

- Ως προς το φύλο, 44 ήσαν άνδρες και 19 ήσαν γυναίκες.
- Αυτοί που είχαν παρακολουθήσει επιτυχώς τη Θ.Ε. ΑεΞΑΕ ήσαν 19.
- Αυτοί που ήσαν κατά το τρέχον ακαδημαϊκό έτος ΣΕΠ στην ΑεΞΑΕ ήταν 16. Οι υπόλοιποι ήταν ΣΕΠ σε άλλες Θ.Ε.

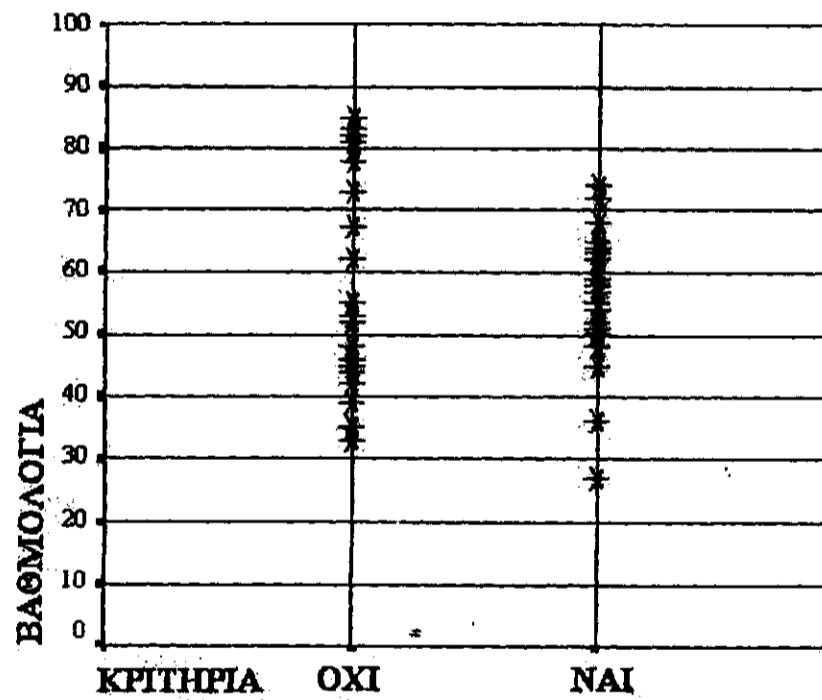
Τα χαρακτηριστικά της βαθμολογίας της ομάδας που χρησιμοποίησε τα κριτήρια που της είχαν αποσταλεί, είναι τα εξής:

Με κριτήρια βαθμολόγησης	
Μέγιστο βαθμολογίας	74,00
Ελάχιστο βαθμολογίας	27,00
Μέσος Όρος	57,15
Τυπική απόκλιση	9,86

Τα χαρακτηριστικά της βαθμολογίας της ομάδας που χρησιμοποίησε την εμπειρία της για την εργασία, είναι τα εξής:

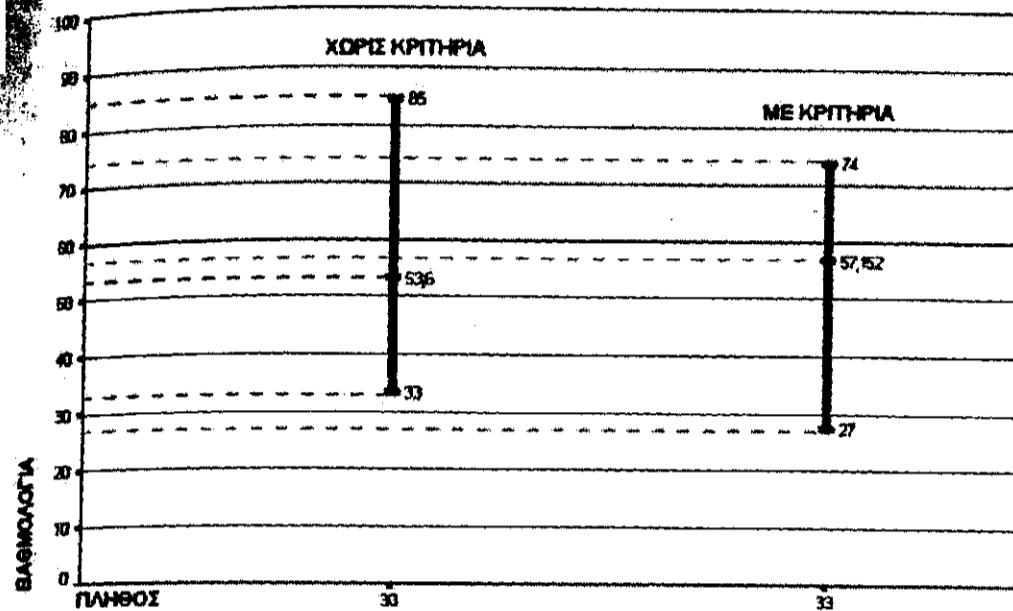
Χωρίς κριτήρια βαθμολόγησης	
Μέγιστο βαθμολογίας	85,00
Ελάχιστο βαθμολογίας	33,00
Μέσος Όρος	53,60
Τυπική απόκλιση	16,07

Στο Διάγραμμα 1 δίνεται με γραφικό τρόπο η διασπορά της βαθμολογίας και στις δύο ομάδες αξιολογητών.



Διάγραμμα 1: Η διασπορά της βαθμολογίας στις δύο ομάδες αξιολογητών

Το Διάγραμμα 2 παρουσιάζεται, επίσης, με γραφικό τρόπο η διασπορά της βαθμολογίας και ο μέσος όρος στις δύο ομάδες αξιολογητών της γραπτής εργασίας.



Διάγραμμα 2: Η διασπορά της βαθμολογίας και οι μέσοι όροι στις δύο ομάδες αξιολογητών

Για τη στατιστική επεξεργασία των ερευνητικών μας δεδομένων και προκειμένου να διαπιστώσουμε την ύπαρξη ή όχι διαφορών μεταξύ της βαθμολογίας της γραπτής εργασίας και των ομάδων που την βαθμολόγησαν με ή χωρίς κριτήρια (ΚΑ), όπως και των εξής παραγόντων: φύλο (ΦΥ), επιτυχής παρακολούθηση της Θ.Ε. ΑεΞΑΕ από το βαθμολογητή (ΑΕ) και διδακτική απασχόληση του βαθμολογητή ως ΣΕΠ στην ΑεΞΑΕ ή αλλού (ΣΑ) χρησιμοποιήθηκε το *t*-κριτήριο⁽³⁾. Από τη στατιστική επεξεργασία των μέσων όρων της βαθμολογίας των δύο ομάδων αξιολόγησης (με ή χωρίς κριτήρια βαθμολόγησης) με το *t*-κριτήριο, δεν ανιχνεύθηκαν στατιστικά σημαντικές διαφορές. Το γεγονός αυτό ερμηνεύεται από το ότι οι μέσοι όροι της βαθμολογίας των δύο συγκεκριμένων ομάδων βρίσκονται αριθμητικά πολύ κοντά μεταξύ τους. Και τούτο μολονότι η διασπορά των βαθμών τους, όπως προκύπτει από την παρατήρηση των δύο παραπάνω Διαγραμμάτων, είναι πολύ διαφορετική. Για τη μελέτη του φαινομένου αυτού αντί του μέσου όρου των βαθμολογιών της κάθε ομάδας αξιολόγησης του γραπτού δοκιμίου επικεντρώσαμε το ενδιαφέρον μας στην ανάλυση των εξής περιπτώσεων:

- της απόστασης της κάθε βαθμολογίας του από το μέσο όρο της ίδιας ομάδας αξιολόγησης (ΜΟ),
- της απόσταση της κάθε βαθμολογίας από τη μέγιστη της βαθμολογία της ίδιας ομάδας (ΜΓ), και
- της απόσταση της κάθε βαθμολογίας από την ελάχιστη βαθμολογία της ίδιας ομάδας αξιολόγησης (ΕΛ).

Έτσι οδηγηθήκαμε στη χρησιμοποίηση του t-κριτηρίου μεταξύ των διαφορών: ΚΑ – ΜΟ, ΚΑ – ΜΓ και ΚΑ – ΕΛ, προκειμένου να ελέγξουμε την ύπαρξη στατιστικά σημαντικών διαφοροποιήσεων μεταξύ των ομάδων που βαθμολόγησαν τη συγκεκριμένη γραπτή εργασία (με ή χωρίς κριτήρια).

Επιπλέον χρησιμοποιήσαμε το t-κριτήριο και στις ακόλουθες περιπτώσεις:

- ΦΥ – ΜΟ, ΦΥ – ΜΓ και ΦΥ – ΕΛ, προκειμένου να ελέγξουμε την ύπαρξη στατιστικά σημαντικών διαφορών στη βαθμολογία μεταξύ ανδρών και γυναικών.
- ΑΕ – ΜΟ, ΑΕ – ΜΓ και ΑΕ – ΕΛ, προκειμένου να ελέγξουμε την ύπαρξη στατιστικά σημαντικών διαφορών στη βαθμολογία μεταξύ αυτών που είχαν παρακολουθήσει επιτυχώς το πρόγραμμα της ΑεξΑΕ και αυτών που δεν το είχαν παρακολουθήσει.
- ΣΑ – ΜΟ, ΣΑ – ΜΓ και ΣΑ – ΕΛ, προκειμένου να ελέγξουμε την ύπαρξη στατιστικά σημαντικών διαφορών στη βαθμολογία μεταξύ αυτών που είναι τώρα ΣΕΠ στην ΑεξΑΕ και αυτών που είναι ΣΕΠ σε άλλες Θ.Ε.

5. Αποτελέσματα – Συζήτηση

Τα αποτελέσματα της στατιστικής ανάλυσης σε κάθε περίπτωση έδειξαν τα εξής:

1. Δεν υπάρχουν στατιστικά σημαντικές διαφορές στη βαθμολογία μεταξύ ανδρών και γυναικών (κάπι άλλωστε, που ήταν αναμενόμενο), σε όλες τις περιπτώσεις.
2. Δεν υπάρχουν στατιστικά σημαντικές διαφορές στη βαθμολογία μεταξύ αυτών που είχαν παρακολουθήσει στο παρελθόν την ΑεξΑΕ και εκείνων που δεν είχαν, σε όλες τις περιπτώσεις.
3. Δεν υπάρχουν στατιστικά σημαντικές διαφορές στη βαθμολογία μεταξύ των ΣΕΠ της ΑεξΑΕ και των ΣΕΠ άλλων Θ.Ε., σε όλες τις περιπτώσεις.
4. Υπάρχουν στατιστικά σημαντικές διαφορές στη βαθμολογία μεταξύ της ομάδας στην οποία δόθηκαν κριτήρια βαθμολόγησης και εκείνης στην οποία δεν δόθηκαν. Η στατιστική αυτή διαφορά ευρέθηκε μεταξύ των δύο ομάδων σε σχέση με:
 - τις αποστάσεις της βαθμολογίας τους από το μέσο όρο κάθε ομάδας ($p < 0,05$),
 - τις αποστάσεις της βαθμολογίας τους από το μέγιστο της βαθμολογίας κάθε ομάδας ($p < 0,05$), και
 - τις αποστάσεις της βαθμολογίας τους από τον ελάχιστο της βαθμολογίας κάθε ομάδας ($p < 0,05$).

Από τα προηγούμενα προκύπτει ότι ο μόνος παράγων που βρέθηκε να διάφοροποιεί τη βαθμολόγηση των δοθείσας γραπτής εργασίας είναι η ύπαρξη ή όχι συγκεκριμένων κριτηρίων.

Παρατηρώντας προσεκτικά το Διάγραμμα 1, μπορεί κανείς να διαπιστώσει τη σχετική ομοιογένεια και συσώρευση των βαθμολογιών αυτών (εκτός από δύο) που αξιολόγησαν τη γραπτή εργασία με κριτήρια. Δεν παρατηρείται το ίδιο στην ομάδα όπου στηρίχθηκε στην ακαδημαϊκή εμπειρία για τη διόρθωσή της. Όπως μάλιστα τονίζει ο Ματραλής (1998), η χρήση συγκεκριμένων κριτηρίων αξιολόγησης αυξάνει την αξιοπιστία, την εγκυρότητα και την αντικειμενικότητα στην αξιολόγηση των γραπτών εργασιών. Θα μπορούσε βέβαια να αντιτάξει κανείς ότι έτσι αποδυναμώνεται το προσωπικό στοιχείο του μέλους ΣΕΠ ή η έμφαση σε κάποιο ζήτημα που είχε δοθεί από το μέλος ΣΕΠ κατά τη διάρκεια των ΟΣΣ ή μέσα από τα σχόλια κατά την

αξιολόγηση των προηγούμενων γραπτών εργασιών. Όμως, η διερεύνηση των ζητημάτων αυτών δεν εντάσσεται στα πλαίσια της παρούσας μελέτης.

Αξίζει να αναφερθεί πως δεν εντοπίστηκαν στατιστικά σημαντικές διαφορές στη βαθμολογία μεταξύ: α) αυτών που εργάζονται ως ΣΕΠ στην ΑεΞΑΕ και εκείνων που διδάσκουν σε άλλες Θ.Ε., και β) αυτών που είχαν παρακολουθήσει επιτυχώς τη Θ.Ε. της ΑεΞΑΕ και εκείνων που δεν την είχαν παρακολουθήσει. Τα ευρήματά μας αυτά μπορούν να ερμηνευτούν από το ότι το θέμα της γραπτής εργασίας, που δόθηκε για βαθμολόγηση ήταν αρκετά κοινότυπο και δεν απαιτούσε ιδιαίτερα εξειδικευμένες γνώσεις για την προσέγγισή του. Μάλιστα, όπως αναφέρθηκε και παραπάνω, η εμπλοκή του ΣΕΠ στο διδακτικό έργο και τις διαδικασίες του ΕΑΠ, όπως λ.χ. και η προαπαιτούμενη παρακολούθηση σχετικών σεμιναρίων παρείχε στον κάθε βαθμολογητή το απαραίτητο γνωστικό υπόβαθρο για την αξιολόγηση της συγκεκριμένης εργασίας.

Έπειτα, από την ποιοτική ανάλυση των ερωτηματολογίων σχετικά με τα κριτήρια αξιολόγησης, που προτάθηκαν από τους ΣΕΠ της ομάδας, που κλήθηκε να βαθμολογήσει τη γραπτή εργασία στηριζόμενη στην ακαδημαϊκή της εμπειρία, προέκυψαν τα ευρήματα που συνοψίζονται στις στήλες του Πίνακα 1. Σημειωτέον, ότι για τη σύνταξη του εντοπίστηκαν τα κοινά κριτήρια αξιολόγησης, που αναφέρθηκαν από τους διαφόρους ΣΕΠ και τα οποία στη συνέχεια κωδικοποιήθηκαν σε μικρού πλήθους όμοιες, από άποψη γνωστικού περιεχομένου, κατηγορίες (Bogdan & Biklen, 1982).

Πίνακας 1: Κριτήρια βαθμολόγησης της γραπτής εργασίας ανάλογα με την επαγγελματική απασχόληση των ΣΕΠ στο ΕΑΠ (διόρθωση χωρίς τη χρήση συγκεκριμένων κριτηρίων αξιολόγησης)

Περιπτώσεις - Κριτήρια αξιολόγησης	ΣΕΠ στην ΑεΞΑΕ	ΣΕΠ σε άλλη Θ.Ε.
Εστίαση στο ερώτημα - Κατανόηση των ζητούμενων από την εργασία		0
Γενική εικόνα		0
Περιεχόμενο εργασίας		0
Δομή - χωρισμός σε ενότητες με σαφείς τίτλους - υπότιτλους (οργάνωση κειμένου)	0	0
Γλώσσα - έκφραση - σαφήνεια	0	0
Επιχειρηματολογία - εγκυρότητα - κριτική ικανότητα (προσέγγιση)	0	0
Παράθεση επιχειρημάτων - όχι μηχανιστική αντιμετώπιση		0
Βιβλιογραφική τεκμηρίωση εκφραζόμενων απόψεων	0	
Ομοιότητες	0	
Διαφορές	0	
Αξιολογική κατάταξη ομοιοτήτων - διαφορών κι εντοπισμός της βαρύτητας του κάθε κριτηρίου		0
Πληρότητα κάλυψης του θέματος		0

Σχολιάζοντας τα δεδομένα του Πίνακα 1 και προβαίνοντας, παράλληλα, στην εις βάθος ανάλυση των απαντήσεων των ΣΕΠ αναφορικά με τα κριτήρια, που θεωρούν ως σημαντικά κατά την αξιολόγηση της συγκεκριμένης γραπτής εργασίας σημειώνουμε τα εξής:

- Σε γενικές γραμμές συμπίπτουν οι απόψεις των ΣΕΠ σχετικά με τον τρόπο που βαθμολόγησαν τη συγκεκριμένη γραπτή εργασία όσον αφορά τη μορφή και τους παράγοντες που κατά τεκμήριο συνιστούν μια επιστημονική γραπτή εργασία.
- Οι ΣΕΠ της ΑεΞΑΕ είναι περισσότερο σαφείς και συγκεκριμένοι στις αξιολογήσεις τους καθώς, ενδεχομένως, εμφανίζουν μεγαλύτερη γνωστική επάρκεια σχετικά με το υπό αξιολόγηση αντικείμενο. Έτσι εστιάζουν με ακρίβεια στα ζητούμενα της εργασίας που αξιολογούν.
- Οι ΣΕΠ των άλλων Θεματικών Ενοτήτων τείνουν να προτείνουν περισσότερο γενικότερης υφής κριτήρια αξιολόγησης.

Η μικρή διαφοροποίηση ως προς τη διόρθωση της γραπτής εργασίας χωρίς την εξαρχής εφαρμογή δεσμευτικών κριτηρίων αξιολόγησης μεταξύ των ΣΕΠ της ΑεΞΑΕ και αυτών που εργάζονται σε άλλες Θ.Ε. ενδεχομένως να ήταν αναμενόμενη. Κι αυτό λόγω του ότι οι ΣΕΠ της ΑεΞΑΕ διακρίνονται για τη μεγαλύτερη επαγγελματική τους εξειδίκευση στο συγκεκριμένο γνωστικό πεδίο σε σύγκριση με τους λοιπούς συναδέλφους τους.

6. Συμπεράσματα

Από την ανάλυση κι επεξεργασία των ερευνητικών μας δεδομένων καταλήγουμε στο συμπέρασμα ότι με τη χρήση συγκεκριμένων και λεπτομερών κριτηρίων βαθμολόγησης μπορεί να αντιμετωπιστεί το πρόβλημα της υποκειμενικότητας κατά την αξιολόγηση γραπτών εργασιών. Επίσης, η εφαρμογή της πρακτικής αυτής στο πλαίσιο λειτουργίας του ΕΑΠ ενισχύει την αξιοπιστία, την αντικειμενικότητα και την εγκυρότητα της αξιολόγησης των γραπτών εργασιών που εκπονούνται από τους φοιτητές κατά τη διάρκεια των σπουδών τους στις διάφορες Θ.Ε. Παράλληλα μεταβιβάζεται σε μεγάλο βαθμό την ευθύνη αντιμετώπισης των σχετικών προβλημάτων στο συντάκτη των κριτηρίων αξιολόγησης, ο οποίος είναι συνήθως και ο συντονιστής της αντίστοιχης Θ.Ε.

Εξάλλου, σε συνάρτηση με τα προηγούμενα και λαμβάνοντας υπόψη ότι κατά το ακαδημαϊκό έτος 2000-2001 στη Θ.Ε. της ΑεΞΑΕ εργάζονται 32 μέλη ΣΕΠ με σημαντικές διαφοροποιήσεις στα πεδία των σπουδών τους, χωρίς δηλαδή να έχουν απόλυτη ή καθόλου γνωστική συνάφεια με τις Επιστήμες της Αγωγής, κρίνεται ως απαραίτητη η εφαρμογή σαφών κριτηρίων στη διαδικασία της αξιολόγησης των πάσης φύσεως γραπτών εργασιών ή γραπτών κειμένων (π.χ. τελικές εξετάσεις). Και τούτο διότι η έλλειψη κριτηρίων αξιολόγησης, όπως άλλωστε αποδείχτηκε, μπορεί να δημιουργήσει σημαντικά προβλήματα αξιοπιστίας κατά τη διόρθωση, το σχολιασμό και τη γνωστική αξιοποίηση των διαφόρων γραπτών εργασιών στο πλαίσιο της παροχής σπουδών στο ΕΑΠ.

Σημειώσεις

[1] Πληροφοριακό Δελτίο: «Οι Σπουδές στο ΕΑΠ», Εκδόσεις ΕΑΠ, Πάτρα, 2000.

[2] Σημειωτέον, ότι τα ευρήματα αυτά αφορούν τη συμβατική και όχι την Ανοικτή Εκπαίδευση.

[3] Για τη στατιστική ανάλυση χρησιμοποιήθηκε το πακέτο SPSS v 9.0 με επίπεδο στατιστικής σημαντικότητας $\alpha=0,05$

7. Βιβλιογραφικές Αναφορές

- Bogdan, R. & Bilken, S., *Qualitative Research for Education: An Introduction to Theory and Methods*. Boston: Allyn and Bacon Inc, 1982.
- Breland H., *The Direct Assessment of Writing Skill: A Measurement Review*, College Board Report 83-6; ETS Research Report 83-32, New York: CEEB, 1983.
- Dimakos I. C., *Teacher assessment of students' writing skills*, Doctoral Dissertation, Syracuse University, 1998.
- DuPaul G. J., Rapport M. D. & Perriello L. M., «Teacher Ratings of Academic Skills: The Development of the Academic Performance Rating Scale», *School Psychology Review*, 20, 2, 1991, pp. 284-300.
- Ματραλής Χ., «Εργασίες Αξιολόγησης», στο βιβλίο: Κόκκος Α., Λιοναράκης Α., Ματραλής Χ., Παναγιωτακόπουλος Χ., «Ανοικτή και εξ Αποστάσεως Εκπαίδευση – Το Εκπαιδευτικό Υλικό και οι Νέες Τεχνολογίες» – Τόμος Γ', ΕΑΠ, Πάτρα, 1998, σελ. 105-149.
- Freedman S. W., «How Characteristics of Students Essays Influence Teachers' Evaluation», *Journal of Educational Psychology*, 71, 3, 1979, pp. 328-338.
- Graham S. & Weintraub N., «A review of handwriting research: Progress and prospects from 1980 to 1994», *Educational Psychology Review*, 8, 1996, pp. 7-87.
- Gresham F. M., MacMillan D. L. & Bocian K. M., «Teachers as «Tests»: Differential Validity of Teacher Judgments in Identifying Students At-Risk for Learning Difficulties», *School Psychology Review*, 26, 1, 1997, pp. 47-60.
- Gresham F. M., Reschly D. J. & Carey M.P., «Teachers as «Tests»: Classification accuracy and concurrent validation in the identification of learning disabled children», *School Psychology Review*, 16, 4, 1987, pp. 543-553.
- Hoge R. D., «Psychometric properties of teacher-judgment measures of pupil aptitudes, classroom behaviors, and achievement levels», *Journal of Special Education*, 17, 1983, pp. 401-429.
- Kirkpatrick, S. and Stoll E., A Very Fast Shift Register Sequence Random Number Generator, *Journal of Computational Physics*, V. 40, 1981, pp 517-526.
- Neeper R. & Lahey B., «The Children's Behavior Rating Scale: A factor analytic developmental study», *School Psychology Review*, 15, 1986, pp. 277-288.
- Park, S.K. and Miller K. W., Random Number Generators: Good Ones are Hard to Find, *Comm. of the ACM*, V. 31, No 10, 1988, pp 1192-1201.
- Perkins K., «On the Use of Composition Scoring Techniques: Objective Measures and Objective Tests to Evaluate ESL Writing Ability», *TESOL Quarterly*, 17, 1983, pp. 651-671.
- Race P., *The Open Learning Handbook*, London: Kogan Page, 1993.
- Rowntree D., *Preparing Materials for Open, Distance and Flexible Learning. An Action Guide for Teachers and Learners*, London: Kogan Page, 1997.
- Salvensen K. A. & Undheim J. O., «Screening for Learning Disabilities with Teacher Rating Scales», *Journal of Learning Disabilities*, 27, 1, 1994, pp. 60-66.
- Shohamy E., Gordon C.M. & Kraemer R., «The Effect of Raters' Background and Training on the Reliability of Direct Writing Tests», *The Modern Language Journal*, 76, 1992, pp. 27-33.