

ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΕΚΠΑΙΔΕΥΣΗΣ ΚΑΙ ΑΓΩΓΗΣ ΣΤΗΝ ΠΡΟΣΧΟΛΙΚΗ ΗΛΙΚΙΑ

**ΑΝΑΠΤΥΞΗ ΕΡΕΥΝΗΤΙΚΟΥ ΕΡΓΟΥ
ΣΤΙΣ ΕΠΙΣΤΗΜΕΣ ΤΗΣ ΕΚΠΑΙΔΕΥΣΗΣ ΚΑΙ ΤΗΣ
ΑΓΩΓΗΣ
ΤΕΑΠΗ, 2010-11 & 2011-12
ΧΡΗΜΑΤΟΔΟΤΗΣΗ: ΕΛΚΕ-ΕΚΠΑ**

Επιστημονικός Υπεύθυνος: Βασίλης Τσελφές

ΕΡΕΥΝΗΤΙΚΟΣ ΑΞΟΝΑΣ 7

**Ανάπτυξη εργαλείων αξιολόγησης και δημοσιότητας
ακαδημαϊκών μονάδων: η περίπτωση του ΤΕΑΠΗ**

**Έκθεση
για τις διεθνείς πρακτικές αξιολόγησης
του εκπαιδευτικού έργου που πραγματοποιείται
στα ανώτατα εκπαιδευτικά ιδρύματα**

Μαρία Γιαλλούση

Αθήνα
Απρίλιος 2012

Έκθεση για τις διεθνείς πρακτικές αξιολόγησης του εκπαιδευτικού έργου που πραγματοποιείται στα ανώτατα εκπαιδευτικά ιδρύματα

Μαρία Γιαλλούση

1. Εισαγωγή

Η παρούσα Έκθεση επιχειρεί μία επισκόπηση της υπάρχουσας έρευνας για την αξιολόγηση του εκπαιδευτικού έργου στην ανώτατη εκπαίδευση με βάση τη διεθνή βιβλιογραφία. Αυτή αποτελεί μία πρώτη απόπειρα εντοπισμού, σύνθεσης και παρουσίασης θεμάτων και ευρημάτων που αφορούν στην *εγκυρότητα* και στη *χρησιμότητα* των αξιολογήσεων.

Η Έκθεση περιλαμβάνει τα εξής κεφάλαια:

Κεφάλαιο 1: Εισαγωγή – παρουσιάζει συνοπτικά το πεδίο εφαρμογής, τη μεθοδολογία και τους περιορισμούς αυτής της έκθεσης.

Κεφάλαιο 2: Πλαίσιο – παρουσιάζει σε γενικές γραμμές τα ισχύοντα σχετικά με την αξιολόγηση του εκπαιδευτικού έργου στην ανώτατη εκπαίδευση. Αναλυτικότερα, παρουσιάζει τις αντιλήψεις και τις στάσεις όσων εμπλέκονται σε αυτή, τα συνήθη χαρακτηριστικά των διαδικασιών με τις οποίες γίνεται μέτρηση της αποτελεσματικότητας της διδασκαλίας καθώς και τις συνήθεις χρήσεις αυτής της αξιολόγησης.

Κεφάλαιο 3: Συστήματα αξιολόγησης & Χρηματοδότηση – αναφέρεται στο ρόλο των εθνικών προτύπων/ δεικτών (standards) αξιολόγησης, παρουσιάζει σύντομα τα τρέχοντα συστήματα ανάπτυξης των δεικτών επίδοσης, τη μέχρι τώρα εμπειρία από τη χρήση τους καθώς και τις πολιτικές σύνδεσης των δεικτών επίδοσης με τη χρηματοδότηση. Τέλος, γίνεται μία σύντομη αναφορά στο σχετικό τοπίο της ανώτατης εκπαίδευσης στην Ευρώπη.

Κεφάλαιο 4: Στρατηγικές Μέτρησης Αποτελεσματικότητας Διδασκαλίας – τεκμηριώνεται η αναγκαιότητα της χρήσης πολλαπλών πηγών αξιολόγησης και στη συνέχεια παρουσιάζονται και περιγράφονται μία προς μία οι δώδεκα επικρατέστερες στρατηγικές μέτρησης.

Κεφάλαιο 5: Ζητήματα Μεθοδολογίας: Εξωτερική Εγκυρότητα, Αξιοπιστία και Ερμηνεία αποτελεσμάτων – γίνεται αναφορά σε κατ' επανάληψη ερευνητικά ευρήματα που σχετίζονται με τη βελτίωση της διαχείρισης και της ερμηνείας των αξιολογήσεων των μαθημάτων στη ανώτατη εκπαίδευση.

Κεφάλαιο 6: Παρουσίαση Μοντέλων Αξιολόγησης – παρουσιάζονται το μοντέλο *360^ο Πολλαπλών Πηγών Ανατροφοδότησης (360^ο ΠΠΑ)*, το μοντέλο της *Περιβάλλουσας Αξιολόγησης Δεδομένων (ΠΑΔ)*, το μοντέλο *Πολλαπλών Διαστάσεων* του Fink, και τέλος ένα μοντέλο *έρευνας μικτών μεθόδων* για διαμορφωτική αξιολόγηση του εκπαιδευτικού έργου.

Κεφάλαιο 7: Τελικές Επισημάνσεις – σταχυολογούνται και παρουσιάζονται στοιχεία της Έκθεσης σχετικά με θέματα εγκυρότητας και αξιοπιστίας κατά την αξιολόγηση του εκπαιδευτικού έργου στην ανώτατη εκπαίδευση καθώς και ερμηνείας των αποτελεσμάτων της.

Παράρτημα - περιλαμβάνει ενδεικτικά ερωτηματολόγια για τις διάφορες πηγές αξιολόγησης και για τα μοντέλα αξιολόγησης που περιγράφονται στο Τμήμα 6.

Οι πληροφορίες που εμπεριέχει η παρούσα Έκθεση προέρχονται από μία εκτενή επισκόπηση αρκετών από τις δημοσιευμένες μελέτες πάνω στην αξιολόγηση των μαθημάτων και του διδακτικού έργου στην τριτοβάθμια εκπαίδευση σε Ευρώπη, Ηνωμένες Πολιτείες Αμερικής και Κίνα. Ο βασικός κορμός της περιλαμβάνει δημοσιεύσεις των τελευταίων 20 χρόνων κυρίως αλλά και αρκετές από τις προηγούμενες οι οποίες είτε επαναλήφθηκαν είτε τα ευρήματα τους είχαν αμφισβητηθεί. Ομολογουμένως το πεδίο είναι τεράστιο και ποικίλης ποιότητας ως προς την εφαρμογή της προαναφερθείσας αξιολόγησης και σίγουρα θα υπάρχουν περιοχές οι οποίες δεν αναδεικνύονται κατάλληλα στο παρόν κείμενο (όπως προγενέστερες μελέτες στις οποίες θα μπορούσε να γίνει μία συνοπτική παρουσίαση).

Επίσης, η παρούσα Έκθεση δεν αναφέρεται στη πολιτική και στις πρακτικές αξιολόγησης που ισχύουν στην ανώτατη εκπαίδευση στην Ελλάδα. Οι πηγές που μπορούν να δώσουν την αντίστοιχη πληροφόρηση (νόμος για την τριτοβάθμια εκπαίδευση, οι εκθέσεις αυτό-αξιολόγησης ανά τμήμα) δεν περιελήφθησαν στο σκεπτικό της οργάνωσης της παρούσας έκθεσης η οποία εστίασε την αναζήτησή της στη διεθνή βιβλιογραφία.

Σε κάθε περίπτωση, το τελικό αποτέλεσμα της επισκόπησης είναι σχετικό με την κατανόηση και τη διαχείριση αυτού του υλικού από την πλευρά της συγγραφέως.

ΠΕΡΙΕΧΟΜΕΝΑ

1. Εισαγωγή

2. Πλαίσιο

- 2.1 Αξιολογώντας τη διδασκαλία στη τριτοβάθμια εκπαίδευση
- 2.2 Αντιλήψεις των μελών ΔΕΠ, διοίκησης και φοιτητών για την αξιολόγηση των μαθημάτων
 - 2.2.1 *Αντιλήψεις των μελών ΔΕΠ*
 - 2.2.2 *Αντιλήψεις των διοικούντων*
 - 2.2.3 *Αντιλήψεις των φοιτητών*
- 2.3 Ορίζοντας την αξιολόγηση
 - 2.3.1 *Τα εθνικά πρότυπα (standards)*
- 2.4 Τα συνήθη χαρακτηριστικά των αξιολογήσεων των μαθημάτων
 - 2.4.1 *Συνήθεις μετρήσεις της αποτελεσματικότητας της διδασκαλίας*
 - 2.4.2 *Συλλέγοντας και ερμηνεύοντας ποιοτική ανατροφοδότηση*
 - 2.4.3 *Συλλέγοντας και ερμηνεύοντας διαμορφωτική ανατροφοδότηση*
- 2.5 Συνήθεις χρήσεις των στοιχείων αξιολόγησης των μαθημάτων
 - 2.5.1 *Βελτίωση της διδασκαλίας*
 - 2.5.2 *Αποφάσεις για το προσωπικό*
 - 2.5.3 *Επιλογή μαθημάτων από τους φοιτητές*
 - 2.4.4 *Βραβεία διδασκαλίας*

3. Συστήματα Αξιολόγησης & Χρηματοδότηση

- 3.1 Αξιολόγηση και Εθνικά Πρότυπα (*standards*)
- 3.2 Προσεγγίσεις στην ανάπτυξη των δεικτών επίδοσης
- 3.3 Η μέχρι τώρα εμπειρία από τη χρήση των δεικτών επίδοσης
- 3.4 Δείκτες επίδοσης στην Ανώτατη Εκπαίδευση: περιορισμοί & κίνδυνοι
- 3.5 Συνδέοντας επίδοση με χρηματοδότηση
 - 3.5.1 *Χαρακτηριστικά του προϋπολογισμού με βάση την επίδοση*
 - 3.5.2 *Προϋποθέσεις για τη χρηματοδότηση με βάση την επίδοση*

- 3.6 Προβλήματα στη χρηματοδότηση με βάση την επίδοση
- 3.7 Το τοπίο της τριτοβάθμιας εκπαίδευσης στην Ευρώπη

4. Στρατηγικές Μέτρησης Αποτελεσματικότητας Διδασκαλίας

- 4.1 Οι πολλαπλές πηγές αξιολόγησης
- 4.2 «Οι 12 Πηγές Απόδειξης»
 - 4.2.1 Βαθμολογίες από τους φοιτητές
 - 4.2.2 Βαθμολογίες από τους ομοβάθμιους
 - 4.2.3 Αυτό-αξιολόγηση
 - 4.2.4 Μαγνητοσκοπήσεις
 - 4.2.5 Συνεντεύξεις φοιτητών
 - 4.2.6 Βαθμολογίες εξόδου και από τους αποφοίτους
 - 4.2.7 Βαθμολογίες από τον εργοδότη
 - 4.2.8 Βαθμολογίες από τα μέλη της διοίκησης
 - 4.2.9 Κατάρτιση στη διδασκαλία
 - 4.2.10 Βραβεία διδασκαλίας
 - 4.2.11 Μετρήσεις αποτελεσμάτων μάθησης
 - 4.2.12 Χαρτοφυλάκιο διδασκαλίας

5. Ζητήματα Μεθοδολογίας

Α΄ Ζητήματα Εγκυρότητας

- 5.1 Αποτελεσματικότητα διδασκαλίας και εγκυρότητα αξιολογήσεων
- 5.2 Γενικές έναντι πολλαπλών διαστάσεων αξιολογήσεις διδασκαλίας
- 5.3 Το ζήτημα της μέτρησης της μάθησης φοιτητών
 - 5.3.1 Βαθμοί φοιτητών & αξιολογήσεις διδασκαλίας από φοιτητές
 - 5.3.2 Αντιλήψεις φοιτητών για τη μάθηση
 - 5.3.3 Πραγματικοί βαθμοί και μάθηση
- 5.4 Μεθοδολογικού τύπου επισημάνσεις για τη μέτρηση της μάθησης
- 5.5 Το παράδοξο της αυστηρότητας
 - 5.5.1 Εξήγηση του παράδοξου της αυστηρότητας

5.5.2 *Συνοψίζοντας τα πορίσματα για τη σχέση μάθησης και αξιολογήσεων*

5.6 *Ανάπτυξη μέσων αξιολόγησης*

5.7 *Έκθεση και ερμηνεία αποτελεσμάτων αξιολόγησης*

5.7.1 *Έκθεση αποτελεσμάτων αξιολόγησης*

5.7.2 *Προκλήσεις στην ερμηνεία και χρήση για συγκεντρωτικούς σκοπούς*

5.7.3 *Χρησιμοποιώντας τις αξιολογήσεις για συγκριτικούς σκοπούς*

5.8 *Διαβούλευση ή συμβουλευτική ανατροφοδότηση*

Β΄ Ζητήματα Αξιοπιστίας

5.9 *Χαρακτηριστικά υποβάθρου & αξιολόγηση*

5.9.1 *Συνθήκες διαχείρισης της έρευνας*

5.9.2 *Χαρακτηριστικά των μαθημάτων*

5.9.3 *Ένας τελικός αναστοχασμός στα χαρακτηριστικά των μαθημάτων*

5.9.4 *Χαρακτηριστικά του διδάσκοντα*

5.9.5 *Χαρακτηριστικά των φοιτητών*

6. Κατευθύνσεις και Μοντέλα Αξιολόγησης

6.1 *Κατευθύνσεις αξιολόγησης*

6.1.1 *Τα ηλεκτρονικά εργαλεία αξιολόγησης*

6.1.2 *Η σύνδεση της αξιολόγησης με τις μετρήσεις λογοδοσίας και τα αποτελέσματα μάθησης με βάση την επάρκεια*

6.1.3 *Αξιολογήσεις εκπαιδευτικού έργου για διαμορφωτικούς σκοπούς*

6.1.4 *Αξιολογήσεις εκπαιδευτικού έργου για συγκεντρωτικούς σκοπούς*

6.2 *Η αναγκαιότητα των πολλαπλών πηγών αξιολόγησης*

6.3 *Παραδείγματα μοντέλων αξιολόγησης*

6.4 *Μοντέλο Πολλαπλών Διαστάσεων του Fink*

6.4.1 *Τα πέντε γενικά ερωτήματα*

6.4.2 *Αξιολογώντας την ποιότητα της διδασκαλίας*

6.4.3 *Προτεινόμενες διαδικασίες*

6.5 *Το μοντέλο 360^ο Πολλαπλών Πηγών Ανατροφοδότησης (ΠΠΑ)*

6.5.1 *Διαμορφωτικές αποφάσεις σχετικές με τις συμπεριφορές διδασκαλίας*

6.5.2 *Συγκεντρωτικές αποφάσεις σχετικές με τις συμπεριφορές διδασκαλίας.*

6.5.3 *Διαμορφωτικές αποφάσεις σχετικές με επαγγελματικές συμπεριφορές.*

6.5.4 *Τα 10 κορυφαία προβλήματα*

6.6 Το μοντέλο της Περιβάλλουσας Αξιολόγησης Δεδομένων (ΠΑΔ)

6.6.1 *Βασικά προτερήματα της χρήση του ΠΑΔ*

6.6.2 *Επικρίσεις για το μοντέλο ΠΑΔ*

6.7 Ένα μοντέλο έρευνας μικτών μεθόδων για την αξιολόγηση του περιβάλλοντος μάθησης με βάση την αλλαγή των πεποιθήσεων αυτό-αποτελεσματικότητας των φοιτητών

7. Τελικές Επισημάνσεις

Παράρτημα

Βιβλιογραφία

2. Πλαίσιο

2.1 Αξιολογώντας τη διδασκαλία στη τριτοβάθμια εκπαίδευση

Στα προηγούμενα 30 χρόνια οι αξιολογήσεις από τους φοιτητές είχαν επικρατήσει ως η κυρίαρχη μέτρηση για την αποτελεσματικότητα της διδασκαλίας (Seldin, 1999a). Ωστόσο, κατά τις τελευταίες δύο δεκαετίες έχει παρατηρηθεί μία τάση αύξησης εκείνων των αξιολογήσεων που χρησιμοποιούν δεδομένα και από άλλες πηγές. Τέτοιες πηγές μπορούν να λειτουργήσουν για να διευρύνουν και να βαθύνουν τη βάση απόδειξης που χρησιμοποιείται στην αποτίμηση μιας σειράς μαθημάτων και στην αξιολόγηση της ποιότητας της διδασκαλίας (Arreola, 2000; Braskamp & Ory, 1994; Knapper & Cranton, 2001; Seldin & Associates, 1999).

Διάφορα περιεκτικά μοντέλα για την αξιολόγηση των διδασκόντων έχουν προταθεί (Arreola, 2000; Braskamp & Ory, 1994; Centra, 1999; Keig & Waggoner, 1994; Romberg, 1985; Soderberg, 1986). Αυτά περιλαμβάνουν πολλαπλές πηγές απόδειξης με το μεγαλύτερο βάρος να αποδίδεται στις εισροές των φοιτητών και των ομοβάθμιων και λιγότερο στην αυτό-αξιολόγηση, στους απόφοιτους, στους διοικητικούς και σε άλλους. Όλα αυτά τα μοντέλα χρησιμοποιούνται για να καταλήξουν σε αποφάσεις διαμορφωτικού και συγκριτικού χαρακτήρα.

Η ποσοτικοποίηση και η συγκρισιμότητα των περισσότερων αξιολογήσεων των μαθημάτων καθιστά την ασαφή τέχνη της αξιολόγησης της διδασκαλίας να φαίνεται περισσότερο αντικειμενική και διαχειρίσιμη.

Τα συνήθη μέσα της αξιολόγησης της διδασκαλίας περιλαμβάνουν τις αξιολογήσεις των μαθημάτων, γράμματα αλληλογραφίας από τους φοιτητές και τους συναδέλφους, αξιολογήσεις στην-τάξη/από ομοβάθμιους, την παραλαβή βραβείων διδασκαλίας, τα υλικά μαθημάτων και τα βιβλία, καθώς και απόδειξη καινοτόμων στρατηγικών και πρακτικών. Κάθε μία από αυτές τις μετρήσεις φέρνει μαζί της τους δικούς της περιορισμούς και όρια. Αυτός είναι ο λόγος που τα περισσότερα ιδρύματα βασίζονται σε περισσότερες από μία φόρμες απόδειξης για να αναπτύξουν μία ολοκληρωμένη κατανόηση των συνιστωσών της διδασκαλίας ενός διδάσκοντα. Ωστόσο, οι αξιολογήσεις των μαθημάτων ή οι βαθμολογίες από τους φοιτητές αποτελούν ένα από τα πλέον συνήθη εργαλεία που χρησιμοποιούνται για να αξιολογήσουν τη διδασκαλία στην τάξη (Wright, 2006; Seldin, 1999; Centra, 1979) και σύμφωνα με τις

πεποιθήσεις κάποιων είναι τα πιο σημαντικά σταθμισμένα εργαλεία για τη λήψη αποφάσεων για το προσωπικό (personnel decision) (Franklin, 2001).

Σε γενικές γραμμές, οι αξιολογήσεις των μαθημάτων χρησιμοποιούνται καταρχήν με συγκεντρωτική/ αθροιστική λογική, σε αντιδιαστολή με τη διαμορφωτική, που σημαίνει, ως ένα μέσο για τη λήψη αποφάσεων για το προσωπικό (δηλαδή πρόσληψη, παράταση θητείας, προαγωγή, και ετήσια επανεξέταση) βασιζόμενη εν μέρει στις βαθμολογίες από τους φοιτητές για την αποτελεσματικότητα της διδασκαλίας του διδάσκοντα. Τα στοιχεία που συλλέγονται, ιδίως τα ποιοτικά, χρησιμοποιούνται επίσης από τους διδάσκοντες και τα γραφεία υποστήριξης της διδασκαλίας για να παρέχουν διαμορφωτική ανατροφοδότηση που έχει την πρόθεση να διευκολύνει τη βελτίωση της διδασκαλίας και την ανάπτυξη των μαθημάτων. Ο Wright (2006) προειδοποιεί για τον κίνδυνο από τη χρήση των εργαλείων που δεν είναι ειδικά σχεδιασμένα να παρέχουν διαμορφωτική ανατροφοδότηση και ότι γι' αυτό το λόγο θα **πρέπει να σχεδιαστούν ξεχωριστά εργαλεία για να παρέχουν συγκεντρωτική και διαμορφωτική ανατροφοδότηση** αντίστοιχα.

Πολλά έχουν γραφτεί σχετικά με τα προβλήματα στις αξιολογήσεις των μαθημάτων. Οι μελετητές στο χώρο της εκπαίδευσης έχουν εξετάσει θέματα προκατάληψης/μεροληψίας (bias), έχουν αναγνωρίσει ανησυχίες που σχετίζονται με τη στατιστική τους αξιοπιστία και έχουν θέσει υπό αμφισβήτηση την ικανότητά τους να εκτιμούν με ακρίβεια την αποτελεσματικότητα της διδασκαλίας των μελών ΔΕΠ. Επιπρόσθετα, κάποιοι έχουν υποστηρίξει ότι η ανατροφοδότηση που παρέχεται από την αξιολόγηση των μαθημάτων δεν προωθεί αποτελεσματικά την αλλαγή στη συμπεριφορά των μελών ΔΕΠ. Ωστόσο, μία σημαντική πλειοψηφία των ερευνητών θεωρεί ότι οι αξιολογήσεις της συμπεριφοράς των διδασκόντων από τους φοιτητές συνιστούν ένα χρήσιμο εργαλείο που συμβάλλει στην αποτελεσματικότητα της διδασκαλίας (ανάμεσά τους Beran, Violato & Kline, 2007; Abrami, 2001; Schmelkin, Spencer & Gellman, 1997; Marsh, 1987).

2.2 Αντιλήψεις των μελών ΔΕΠ, διοίκησης και φοιτητών για την αξιολόγηση των μαθημάτων

Υπάρχουν αμέτρητοι μύθοι και παρανοήσεις που αφορούν στις αξιολογήσεις των μαθημάτων και αναπόφευκτα επηρεάζουν τα μέλη ΔΕΠ, τα μέλη της διοίκησης (administrators) του πανεπιστημίου και τις αντιλήψεις των φοιτητών. Παρά την εμπειριστατωμένη έρευνα για την αντιμετώπιση αυτών των υποθέσεων, τέτοιες

πεποιθήσεις υπάρχουν και συνεχίζουν να διαχέονται. Αρκεί κάποιος να θέσει το θέμα σε μία συνάντηση του τμήματος, σε ένα γεύμα της σχολής ή σε μία εκδήλωση της πανεπιστημιούπολης για να αποσπάσει μία σειρά από «ανεπίσημα στοιχεία» από διάφορα μέλη της πανεπιστημιακής κοινότητας. Όπως επισημαίνουν οι Nasser και Fresko (2002), λίγες εκτενείς μελέτες έχουν διεξαχθεί πάνω στις στάσεις και στις αντιλήψεις σχετικά με τα συστήματα αξιολόγησης των μαθημάτων από εκείνους που τα χρησιμοποιούν και που επηρεάζονται από αυτά (κυρίως μέλη ΔΕΠ, φοιτητές και διοίκηση). Μελέτες που συζητούν τέτοια θέματα είναι τυπικά μικρές και προκύπτουν από απαντήσεις ενός περιορισμένου αριθμού ατόμων, ωστόσο, υπάρχει κάτι το συνεπές στις καθεστηκυίες στάσεις και έτσι είναι δυνατό να προκύψουν κάποιες γενικεύσεις.

2.2.1 Αντιλήψεις των μελών ΔΕΠ

Οι αξιολογήσεις των μαθημάτων από τους φοιτητές έχουν καθιερωθεί ως μία πηγή άγχους για τα μέλη ΔΕΠ (Hodges & Stanton, 2007; Ryan, Anderson & Birchler, 1980) και για ορισμένα μέλη υποδαυλίζουν απερίφραστη εχθρότητα (Franklin & Theall, 1989). Τέτοιες στάσεις προκύπτουν από ισχύουσες πεποιθήσεις ότι οι αξιολογήσεις μεροληπτούν (are biased) (Eiszler, 2002; Feldman, 1976), ότι οι φοιτητές δεν είναι αρμόδιοι αξιολογητές (Nasser & Fresko, 2002; Ryan, Anderson & Birchler, 1980) και ότι οι αξιολογήσεις επηρεάζονται από τις προσδοκίες των φοιτητών για τους βαθμούς τους (Baldwin & Blattner, 2003). Αυτά κι άλλα θέματα έχουν οδηγήσει τόσο τα μέλη ΔΕΠ όσο και τα μέλη της διοίκησης να αμφισβητούν την όλη εγκυρότητα των αξιολογήσεων από τους φοιτητές και την χρήση τους (Beran, Violato & Kline, 2007; Ory, 2001), ειδικότερα σε σχέση με τις αποφάσεις για το προσωπικό (Nasser & Fresko, 2002; Sproule, 2000; Ryan, Anderson & Birchler, 1980). Παρ' όλα αυτά, οι Theall και Franklin (2000) έχουν παρατηρήσει ότι «η δυσφορία των μελών ΔΕΠ για τις αξιολογήσεις και τις ελλείψεις σε καλές πρακτικές είναι σημάδια της επίμονης κακής συνύπαρξης των κόσμων της έρευνας και της πρακτικής» (σελ. 95). Ακόμα οι Aleamoni (1999) και Ory (2001) έχουν υποστηρίξει ότι τόσο τα μέλη ΔΕΠ όσο και η διοίκηση των Πανεπιστημίων συνεχίζουν να δημιουργούν και να διαιωνίζουν τη μυθολογία και τις παρανοήσεις σχετικά με τις αξιολογήσεις των μαθημάτων.

Κάποιες έρευνες έχουν δείξει ότι είναι υψηλότερο το ποσοστό των μελών ΔΕΠ που κατέχεται από αρνητικές παρά από θετικές ή ουδέτερες στάσεις απέναντι στα εργαλεία αξιολόγησης (Nasser & Fresko, 2002; Abrami, 2001; Theall & Franklin,

2000; Wachtel, 1998; Centra, 1993). Μελέτες επίσης έχουν βρει ότι η χρήση ενός συστήματος αξιολόγησης των μαθημάτων μπορεί να αποδυναμώσει το ηθικό των μελών ΔΕΠ (Ryan, Anderson & Birchler, 1980). Ωστόσο, τα ευρήματα είναι μικτά και οι απόψεις των μελών ΔΕΠ ποικίλουν ευρέως (Wachtel, 1998; Schmelkin, Spencer & Gellman, 1997; Newport, 1996).

2.2.2 Αντιλήψεις των διοικούντων

Οι περισσότερες μελέτες έχουν δείξει ότι οι άνθρωποι της διοίκησης, γενικά, έχουν μία θετική στάση απέναντι στα στοιχεία αξιολόγησης και την αποτιμούν ως μία χρήσιμη πηγή πληροφοριών για τη λήψη αποφάσεων για το προσωπικό - αποφάσεις σχετικές με την προαγωγή, την παράταση της θητείας και το μισθό (Campbell & Bozeman, 2008; Beran et al., 2005). Παρ' ότι οι διοικούντες μπορεί να συμφωνούν ότι αυτά τα εργαλεία είναι αποτελεσματικά, οι στάσεις τους υπόκεινται στις διάχυτες παρανοήσεις γύρω από τις ανησυχίες για την εγκυρότητα. Οι Theall και Franklin (2001), ο Abrami (2001) και άλλοι έχουν τονίσει ότι τέτοιες παρανοήσεις επικρατούν εξαιτίας της γενικής έλλειψης οικειότητας με την έρευνα πάνω στην εγκυρότητα των αξιολογήσεων ή την απροθυμία να γίνουν δεκτά τα ευρήματα. Επιπλέον, η βιβλιογραφία έχει δείξει ότι οι διοικούντες επηρεάζονται από τις προσεγγίσεις και τις στάσεις τους απέναντι στη διδασκαλία και απέναντι στους μεμονωμένους διδάσκοντες. Ο McKeachie (1997) σημειώνει ότι τα στερεότυπα των διοικούντων σχετικά με τη διδασκαλία επηρεάζουν τις κρίσεις τους γύρω από την αποτελεσματικότητα της διδασκαλίας: εάν ένας υποψήφιος δεν ανταποκρίνεται στα υπάρχοντα στερεότυπά τους για το τι κάνει έναν καλό διδάσκοντα, αυτοί τον αντιμετωπίζουν αρνητικά.

2.2.3 Αντιλήψεις φοιτητών

Η έρευνα πάνω στις αντιλήψεις των φοιτητών για τις αξιολογήσεις των μαθημάτων και τη χρήση των στοιχείων αξιολόγησης είναι περιορισμένη. Οι περισσότερες μελέτες έχουν υπάρξει μικρές, σχεδιασμένες σε δείγματα από ένα ίδρυμα (Campbell & Bozeman, 2008; Beran et al., 2005). Από αυτές, θα μπορούσε να βγει ότι οι φοιτητές αντιλαμβάνονται τη διαδικασία της συλλογής των σχετικών στοιχείων και της ανατροφοδότησης από τους φοιτητές ως έγκυρη και χρήσιμη. Επιπλέον, οι φοιτητές πιστεύουν ότι μπορούν να είναι και είναι αποτελεσματικοί αξιολογητές της διδασκαλίας. Ωστόσο, κάποιες μελέτες δείχνουν ότι οι φοιτητές δεν είναι πάντα ενήμεροι με ποιον τρόπο τα ιδρύματα χρησιμοποιούν τα στοιχεία που συλλέγονται (Campbell & Bozeman, 2008; Beran et al., 2005; Wachtel, 1998), ούτε

καταλαβαίνουν πάντα την επίδραση που έχουν οι αξιολογήσεις στις αποφάσεις για το προσωπικό. Κάποιες μελέτες βρίσκουν ότι οι φοιτητές στέκονται σκεπτικά απέναντι στο γεγονός ότι οι εισροές τους χρησιμοποιούνται, αξιολογούνται και αναλύονται (Wachtel, 1998). Επιπλέον, πολλοί φοιτητές κάνουν μικρή χρήση των στοιχείων της αξιολόγησης. Για παράδειγμα, σε μία μελέτη που συμμετείχαν 1194 φοιτητές και 35 απόφοιτοι, σε ένα ίδρυμα στον Καναδά, ο Beran και οι συνεργάτες του (2005) βρήκαν ότι 56% των φοιτητών δεν χρησιμοποιούσε τα στοιχεία των αξιολογήσεων. Από το 43% που έδειξε ότι τα συμβουλευτήκε, λιγότεροι από το ένα τρίτο (31%) τα χρησιμοποίησαν για να επιλέξουν μάθημα με βάση το περιεχόμενο και τη δομή του (δηλαδή αναθέσεις, φόρτος εργασίας, θέματα) και σχεδόν τα δύο τρίτα τα χρησιμοποίησαν για να επιλέξουν μάθημα με βάση τον διδάσκοντα (64%).

2.3 Ορίζοντας την αξιολόγηση

Σύμφωνα με τον Scriven (1991) αξιολόγηση (evaluation) είναι «η διαδικασία, της οποίας το καθήκον είναι ο συστηματικός και αντικειμενικός καθορισμός του πλεονεκτήματος (merit), της αξίας (worth) ή της τιμής (value). Χωρίς μία τέτοια διαδικασία, δεν υπάρχει τρόπος να διαχωριστεί αυτό που αξίζει από αυτό που δεν αξίζει» (σελ. 4). Αυτή διαδικασία περιλαμβάνει δύο διαστάσεις: α) τη συλλογή δεδομένων και β) τη χρήση αυτών των δεδομένων για διαμόρφωση κρίσης και απόφασης σε σχέση με συμφωνημένα standards. Τα εργαλεία μέτρησης όπως είναι οι έλεγχοι, οι κλίμακες, και τα ερωτηματολόγια χρησιμοποιούνται για να συλλέξουν αυτά τα δεδομένα. Τα κριτήρια για την αποτελεσματικότητα της διδασκαλίας είναι ενσωματωμένα στο περιεχόμενο αυτών των μετρήσεων. Οι περισσότερο κοινές μετρήσεις που χρησιμοποιούνται για τη συλλογή δεδομένων για τη αποτίμηση των διδασκόντων είναι οι κλίμακες αξιολόγησης.

2.3.2 Τα εθνικά πρότυπα (standards)

Τα εθνικά standards – εφόσον υπάρχουν- καθορίζουν το πώς πρέπει να γίνεται η μέτρηση της αποτελεσματικότητας της διδασκαλίας ή της επίδοσης (για παράδειγμα βλέπε the Standards for Educational and Psychological Testing AERA, APA, & NCME Joint Committee on Standards, 1999).

Αυτά μπορούν να καθοδηγήσουν:

- την ανάπτυξη των εργαλείων μέτρησης,
- την τεχνική της ανάλυσης των αποτελεσμάτων, και
- την αναφορά και την ερμηνεία της απόδειξης για τη λήψη αποφάσεων.

Τα standards καταπιάνονται με το ΤΙ μετριέται και ως εκ τούτου με το ΠΩΣ μετριέται. Για παράδειγμα, σύμφωνα με τα προαναφερθέντα standards στα πλαίσια του τι και πώς ισχύουν τα εξής:

- **ΤΙ**- Το περιεχόμενο κάθε εργαλείου (τι μετριέται;), για παράδειγμα μίας κλίμακας αξιολόγησης από φοιτητή ή από ομοβάθμιο (συνάδελφο). Αυτό απαιτεί έναν εφ' όλης της ύλης και κατηγορηματικό ορισμό των γνώσεων, των δεξιοτήτων και των ικανοτήτων καθώς και άλλων χαρακτηριστικών και συμπεριφορών που περιγράφουν το έργο της «αποτελεσματικής διδασκαλίας».
- **ΠΩΣ**- Τον τρόπο που αξιοποιούνται τα δεδομένα από μία κλίμακα αξιολόγησης ή άλλου εργαλείου που βασίζεται στη συστηματική συλλογή των γνωμοδοτήσεων ή των αποφάσεων από τους βαθμολογητές, τους παρατηρητές, ή τους κριτές εξαρτώμενα από την τεχνογνωσία, τα προσόντα και την εμπειρία τους.

Οι άμεσες παρατηρήσεις των φοιτητών και των ομοβάθμιων του ΤΙ αυτοί βλέπουν στην τάξη αποτελούν τα θεμέλια των αξιολογήσεών τους. Ωστόσο, άλλες πηγές, όπως είναι τα αποτελέσματα (βαθμοί) των φοιτητών και οι προτάσεις πάνω σε καινοτόμες στρατηγικές διδασκαλίας, αποτελούν έμμεσες πηγές, από τις οποίες συνάγεται η αποτελεσματικότητα της διδασκαλίας. Αυτές οι διαφορετικές πηγές δεδομένων διαφέρουν αξιοσημείωτα στο ΠΩΣ (τρόπο) μετρούν το ΤΙ. Απαιτείται επομένως ιδιαίτερη προσοχή στον εντοπισμό των διακρίσεων ανάμεσα σε όλες τις διαθέσιμες πηγές.

2.4 Τα συνήθη χαρακτηριστικά των αξιολογήσεων των μαθημάτων

Στη διαδικασία διερεύνησης των αξιολογήσεων των μαθημάτων, πολλοί μελετητές έχουν αναγνωρίσει μια σειρά από συνήθη χαρακτηριστικά (Algozzine et al. 2004).

Οι φόρμες αξιολόγησης των μαθημάτων συνήθως διανέμονται στη λήξη μιας συγκεκριμένης ενότητας της διδασκαλίας. Σχεδόν πάντα είναι ανώνυμες (ή, λιγότερο συχνά, εμπιστευτικές) και τις περισσότερες φορές συμπεριλαμβάνουν ποιοτικές και ποσοτικές μετρήσεις. Οι ποσοτικές ερωτήσεις ζητούν οι φοιτητές να απαντήσουν σε μία κλίμακα Likert και σε άλλες κλίμακες βαθμολόγησης, τις περισσότερες φορές με πέντε ή επτά σημεία επιλογών.

Οι αξιολογήσεις γενικά απαιτούν συγκεκριμένη ανατροφοδότηση στις μετρήσεις της αποτελεσματικότητας της διδασκαλίας και στις ειδικές πλευρές των μαθημάτων, καθώς και γενικές ερωτήσεις (global question) αξιολόγησης και συχνά έναν περιορισμένο αριθμό ερωτήσεων ανοικτού-κλειστού τύπου οι οποίες απαιτούν ποιοτικές γραπτές απαντήσεις. Τα περισσότερα εργαλεία αξιολόγησης είναι σχεδιασμένα για να χρησιμοποιούνται για συγκεντρωτική αξιολόγηση της διδασκαλίας, αλλά η διαμορφωτική αξιολόγηση είναι πιθανή διαμέσου εναλλακτικών μοντέλων των ερωτηματολογίων αξιολόγησης από τους φοιτητές και διαμέσου της διαγνωστικής ερμηνείας των αποτελεσμάτων από τα είδη των αξιολογήσεων που περιγράφονται παραπάνω.

2.4.1 Συνήθεις μετρήσεις της αποτελεσματικότητας της διδασκαλίας

Οι δηλώσεις στις αξιολογήσεις των μαθημάτων ψάχνουν πληροφορία σχετικά με το σχεδιασμό του μαθήματος και την παράδοση καθώς και τη συμπεριφορά του διδάσκοντα. Η Cashin (1995) επισημαίνει έξι στοιχεία τα οποία συνήθως εμφανίζονται στις αξιολογήσεις:

- ερωτήσεις σχετικές με το περιεχόμενο των μαθημάτων,
- ερωτήσεις σχετικές με τις δεξιότητες επικοινωνίας του διδάσκοντα,
- ερωτήσεις σχετικές με την αλληλεπίδραση καθηγητή-φοιτητή,
- ερωτήσεις σχετικές με τη δυσκολία των μαθημάτων και το φόρτο εργασίας,
- ερωτήσεις σχετικές με τις πρακτικές αξιολόγησης των φοιτητών στα μαθήματα,
- ερωτήσεις αυτό-αξιολόγησης των φοιτητών.

Τα διαφορετικά συστατικά των αξιολογήσεων των μαθημάτων προέρχονται επίσης από τη έρευνα για την μάθηση των φοιτητών και για την επιτυχημένη συμπεριφορά του διδάσκοντα κατά τη διδασκαλία. Η εκτιμώμενη ανάγκη να διατυπωθούν ερωτήσεις για τις πολλαπλές κατηγορίες της συμπεριφοράς στη διδασκαλία προκύπτει από την πεποίθηση ότι η αποτελεσματικότητα της διδασκαλίας είναι πολλαπλών διαστάσεων. Αυτό σημαίνει ότι οι διδάσκοντες μπορεί να υπερέχουν σε κάποια σημεία της διδασκαλίας και σε κάποια άλλα όχι (Algozzine et al., 2004; Marsh & Roche, 1997; Marsh 1987). Το 1987, ο Marsh ανέπτυξε το εργαλείο αξιολόγησης Student Evaluations of Education Quality (SEEQ), το οποίο περιλαμβάνει εννέα κατηγορίες ερωτήσεων για τη συμπεριφορά διδασκαλίας. Αυτός ισχυρίστηκε ότι θα πρέπει όλες να είναι παρούσες με σκοπό να διασφαλιστεί ότι η αξιολόγηση είναι αντιπροσωπευτική της αποτελεσματικότητας της διδασκαλίας:

1. μάθηση/ αξία
2. ενθουσιασμός διδάσκοντα,
3. οργάνωση,
4. ατομική επικοινωνία/συμπάθεια,
5. αλληλεπίδραση της ομάδας,
6. εύρος της κάλυψης,
7. εξέταση/ βαθμοί,
8. αναθέσεις/ μελέτες, και
9. φόρτος εργασίας/ δυσκολία.

Παρόμοιες μετρήσεις της αποτελεσματικότητας της διδασκαλίας έχουν αναγνωριστεί από τους Braskamp και Ory (1994) και τον Centra (1993) και στο Individual Development and Educational Assessment (IDEA), σύστημα αξιολόγησης που ανέπτυξε η Cashin στο πανεπιστήμιο του Κάνσας των ΗΠΑ. Αυτές περιλαμβάνουν την οργάνωση και τον σχεδιασμό των μαθημάτων, τις δεξιότητες αποσαφήνισης/ επικοινωνίας, την αλληλεπίδραση/ συμπάθεια καθηγητή φοιτητών, τη δυσκολία/ φόρτο εργασίας μαθημάτων, τη βαθμολόγηση στις εξετάσεις και την αυτό-αξιολόγηση της μάθησης των φοιτητών. Άλλες μελέτες όπως είναι εκείνη του Feldman (1989) έχουν αναγνωρίσει γύρω στις 28 κατηγορίες συμπεριφορών διδασκαλίας. Η πρόκληση της ανεύρεσης και του ορισμού εκείνων των δραστηριοτήτων διδασκαλίας και των στρατηγικών που περισσότερο συμβάλλουν στην εκμάθηση των φοιτητών συζητούνται στο τμήμα της έκθεσης: *Ορίζοντας την Αποτελεσματική Διδασκαλία*. Ο Harry Murray (1987) στο πανεπιστήμιο του Δυτικού Οντάριο ανέπτυξε το Teaching Behavior Inventory, το οποίο μπορεί να χρησιμοποιηθεί για να συλλέγει πληροφορία από τους φοιτητές για 60 συμπεριφορές καθηγητή και δραστηριότητες διδασκαλίας. Αυτές οι συμπεριφορές μετράνε δραστηριότητες διδασκαλίας που πιστεύεται ότι βελτιώνουν την εκμάθηση των φοιτητών και ομαδοποιούνται σε εννέα κατηγορίες: 1. σαφήνεια, 2. έκφραση, 3. αλληλεπίδραση, 4. οργάνωση, 5. υπομονή, 6. κοινοποίηση, 7. ομιλία, 8. συμπάθεια και 9. στόχοι διδασκαλίας. Παραδείγματα διαφόρων τέτοιων ερωτηματολογίων μπορούν να βρεθούν στο Παράρτημα.

2.4.2 Συλλέγοντας και ερμηνεύοντας ποιοτική ανατροφοδότηση

Πολλοί είναι οι μελετητές που ισχυρίζονται ότι η αξιολόγηση της αποτελεσματικότητας της διδασκαλίας είναι καλύτερα να συμπεριλαμβάνει εκτός από τις βαθμολογίες από τους φοιτητές και πολλαπλού τύπου ποιοτικά στοιχεία (Lattuca

& Domagal-Goldman, 2007; Ory, 2001; Arreola, 1983). Πολλά από αυτά τα ποιοτικού τύπου μέσα αξιολόγησης της διδασκαλίας (όπως το χαρτοφυλάκιο, αυτό-αξιολόγηση, αξιολόγηση από τους ομοβάθμιους και γραπτές διηγήσεις διδασκαλίας) δεν αφορούν στους φοιτητές. Πράγματι, **ενώ οι φοιτητές είναι αποτελεσματικοί στη μέτρηση των συμπεριφορών διδασκαλίας και των δραστηριοτήτων μέσα στην τάξη, δεν έχουν τα απαραίτητα προσόντα να αποτιμήσουν το περιεχόμενο των μαθημάτων ή τους στόχους και γι' αυτό χρειάζεται να αξιοποιηθούν άλλες πηγές πληροφόρησης.**

Παρ' όλα αυτά είναι γενικά αποδεκτό ότι τα ποιοτικά μέσα αξιολόγησης παρέχουν γνώση που δεν μπορεί να εξασφαλιστεί με τα ποσοτικά (βλέπε Harper & Kuh, 2007) και η ανατροφοδότηση από τα ποιοτικά στοιχεία χρησιμεύει περισσότερο στη διαμορφωτική παρά στη συγκεντρωτική αξιολόγηση (Franklin, 2001; Lewis, 2001). Συνήθως μία ποιοτικού τύπου έρευνα λαμβάνει χώρα στα μέσα και στο τέλος του μαθήματος (Lewis, 2001) ή ως μία αξιολόγηση μέσα στην τάξη (Diamond, 2004).

Σύμφωνα με την Cashin (1995) αρκετά από τα μέλη ΔΕΠ εμπιστεύονται περισσότερο τις απαντήσεις στις ανοικτού-κλειστού τύπου ερωτήσεις από τις ποσοτικές βαθμολογίες από τους φοιτητές, και αρκετά μέλη υποστηρίζουν το αντίθετο. Ωστόσο, οι μελέτες έδειξαν ότι υπάρχει ένας συνδυασμός/ σχέση μεταξύ των ποιοτικών εισροών και των ποσοτικών βαθμολογιών.

Η μη αξιοποίηση των ποιοτικών στοιχείων στη συγκεντρωτική αξιολόγηση της διδασκαλίας αποδίδεται στην δυσκολία να αξιολογηθούν η εγκυρότητα και η αξιοπιστία τους (Abrami, 2001). Οι Harper και Kuh (2007) όμως υποστηρίζουν ότι τέτοιου τύπου ανησυχίες είναι άσχετες με τον τρόπο που τα ποιοτικά στοιχεία είναι δυνατό και πρέπει να χρησιμοποιούνται στις συγκεντρωτικές αξιολογήσεις.

2.4.3 Συλλέγοντας και ερμηνεύοντας διαμορφωτική ανατροφοδότηση

Οι Lewis (2001) και Ory (2001) τονίζουν ότι για να είναι αποτελεσματικές στην βελτίωση της διδασκαλίας οι αξιολογήσεις πρέπει να είναι συνεχείς και διαμορφωτικές και να λαμβάνουν χώρα στο πλαίσιο των προσωπικών στόχων ενός διδάσκοντα για τη βελτίωση της διδασκαλίας του.

Η ανατροφοδότηση από τη διαμορφωτική αξιολόγηση μπορεί να διεξαχθεί με τη χρήση παραδοσιακών αξιολογήσεων στο τέλος του εξαμήνου ή διαμέσου εναλλακτικών μορφών αξιολόγησης. Οι Aultman (2006) και Lewis (2001)

υπερασπίζονται την χρήση των αξιολογήσεων στις αρχές και στα μέσα του εξαμήνου οι οποίες μπορούν να παρέχουν ανατροφοδότηση διαμορφωτικού τύπου η οποία μπορεί να φανεί άμεσα χρήσιμη. Οι Hodges και Stanton (2007) περιγράφουν τον τρόπο με τον οποίο τα γραπτά σχόλια των φοιτητών μπορούν να αποκαλύψουν πληροφορίες σχετικές με τις πλευρές της διαδικασίας και μάθησης τις οποίες οι φοιτητές δεν κατανοούν και μπορούν να λειτουργήσουν ως ένα σημαντικό εργαλείο βελτίωσης των μαθημάτων.

Ακόμα ένα άλλο είδος διαμορφωτικού τύπου ανατροφοδότησης μπορεί να προκύψει από συγκεντρωτικού τύπου αξιολογήσεις όπως είναι η διαγνωστική αξιολόγηση. Εάν οι αποτιμήσεις που χρησιμοποιούνται είναι πολλαπλών διαστάσεων, μία έκθεση μπορεί να παρέχει στους διδάσκοντες πληροφορία η οποία να τους προσδιορίζει τα δυνατά σημεία και αυτά που χρειάζονται βελτίωση. Μια τέτοια έκθεση μπορεί να διευκολύνει την αυτό-καθοδηγούμενη αλλά και την υποστηριζόμενη από το ίδρυμα βελτίωση της διδασκαλίας (Crosson et al., 2006; Marsh & Roche, 1997).

2.5 Συνήθειες χρήσεις των στοιχείων αξιολόγησης των μαθημάτων

Τα στοιχεία αξιολόγησης χρησιμοποιούνται συνήθως στη διαμόρφωση δύο τύπων αποφάσεων (Berk, 2006, σελ. 11):

1. Οι **διαμορφωτικές** αποφάσεις (formative decisions). Αυτές είναι αποφάσεις τις οποίες λαμβάνουν οι διδάσκοντες για να βελτιώσουν και να διαμορφώσουν την ποιότητα της διδασκαλίας τους. Αυτές βασίζονται σε στοιχεία που αφορούν στην αποτελεσματικότητα της διδασκαλίας τους, τα οποία συλλέγουν για να την σχεδιάσουν και να την αναθεωρήσουν από εξάμηνο σε εξάμηνο. Η συλλογή της απόδειξης και οι συνεπαγόμενες κρίσεις για τη διδασκαλία μπορούν να προκύψουν κάθε στιγμή κατά τη διάρκεια των μαθημάτων, έτσι οι φοιτητές μπορούν να ωφεληθούν από αυτές τις αλλαγές, ή μετά την ολοκλήρωση των μαθημάτων κατά την προετοιμασία της επόμενης σειράς μαθημάτων.
2. **Συγκεντρωτικές** αποφάσεις (summative decisions): Αυτές οι αποφάσεις αποδίδονται δια μέσου του «*da boss*», του ατόμου με διοικητικού τύπου αρμοδιότητες, το οποίο ελέγχει τη μοίρα (destiny) και το μέλλον του διδάσκοντα στην τριτοβάθμια εκπαίδευση. Αυτό το άτομο συνήθως είναι ο κοσμήτορας, ο αναπληρωτής κοσμήτορας, ο επιμελητής του προγράμματος, ο πρόεδρος του τμήματος. Ο διοικητικός χρησιμοποιεί αποδεικτικά στοιχεία για την αποτελεσματικότητα της διδασκαλίας του διδάσκοντα για να

«συναθροίσει» τη συνολική του απόδοση ή την επίδοση ή την κατάσταση και να μπορέσει να αποφασίσει σχετικά με την ανανέωση της σύμβασης ή την απόλυση, το μισθό, την προαγωγή και τη μονιμότητα. Αν και η προαγωγή και η μονιμότητα αποφασίζονται συνήθως από την επιτροπή των διδασκόντων, μία συστατική επιστολή από τον κοσμήτορα τυπικά απαιτείται για την επίτευξη της επανεξέτασης από την επιτροπή. Οι συγκεντρωτικού τύπου αποφάσεις είναι τελικές αποφάσεις απασχόλησης που καθορίζουν την πρόοδο ενός διδάσκοντα δια μέσου των βαθμών ιεραρχίας και την επιτυχία του ως ενός ακαδημαϊκού ή «ενοχλητικού». Αυτές οι αποφάσεις έχουν επίδραση στην ποιότητα της επαγγελματικής και προσωπικής ζωής ενός διδάσκοντα.

Οι διάφορες πηγές απόδειξης για την αποτελεσματικότητα της διδασκαλίας μπορούν να χρησιμοποιηθούν είτε για τις διαμορφωτικού τύπου αποφάσεις είτε για τις συγκριτικού ή και για τις δύο (Berk, 2009).

Διάφορες πηγές μπορούν επίσης να χρησιμοποιηθούν για προγραμματικού τύπου αποφάσεις. Τέτοιες αποφάσεις σχετίζονται με τα προγράμματα σπουδών, τις εισαγωγές και τις απαιτήσεις της φοίτησης και την αποτελεσματικότητα των προγραμμάτων. Αυτές οι αποφάσεις εστιάζονται σε διαδικασίες και προϊόντα. Η απόδειξη συνήθως έρχεται από τις εισροές (input) των διδασκόντων και των φοιτητών και από την αξιολόγηση των εργαζομένων από τους φοιτητές. Αυτή επίσης συλλέγεται για να παρέχει τεκμηρίωση που να ικανοποιεί τα κριτήρια σε μία επισκόπηση πιστοποίησης (accreditation review).

Οι Moore και Kuol (2005) έχουν βρει ότι τα συστήματα αξιολόγησης από τους φοιτητές βοηθούν να αντιμετωπισθεί η ανεπίσημη πληροφορία για τις συμπεριφορές διδασκαλίας και την αποτελεσματικότητα. Αυτοί επίσης υποστηρίζουν ότι τέτοια εργαλεία παρέχουν ένα άλλο μέσο αξιολόγησης της διδασκαλίας και έτσι βοηθάνε να συρρικνώνονται τα υπάρχοντα κενά μεταξύ των μεθόδων αξιολόγησης για τη διδασκαλία και την έρευνα.

Υπάρχουν πολλές κοινές χρήσεις για τα στοιχεία αξιολόγησης των μαθημάτων: βελτίωση διδασκαλίας, αποφάσεις για το προσωπικό, επιλογή μαθήματος από τους φοιτητές και όλο και περισσότερο, αρχεία υποψηφιότητας για βραβεία κατάρτισης επί της διδασκαλίας.

2.5.1 Βελτίωση της διδασκαλίας

Από τότε που διαδόθηκε η χρήση της αξιολόγησης οι ερευνητές έχουν συμφωνήσει ότι τα στοιχεία αξιολόγησης των μαθημάτων μπορούν αποτελεσματικά να

χρησιμοποιηθούν για το σκοπό της βελτίωσης της διδασκαλίας και ως εκ τούτου και της μάθησης των φοιτητών (Goldschmid, 1978). Ωστόσο, όταν κάποιος εστιάζει την προσοχή του στο ερώτημα «Με ποιο τρόπο θα πρέπει να αξιολογείται η διδασκαλία;» τότε είναι σχεδόν βέβαιο ότι μπορεί να εργαστεί στη βάση ότι κάθε βελτίωσή της απαιτεί δύο διαδικασίες (Fink, 2008):

- την εξέταση των πολλαπλών διαστάσεων της διδασκαλίας, κάτι που είναι περισσότερο από αυτό που κάνει ο καθηγητής στην τάξη, και
- τη χρήση των πολλαπλών πηγών πληροφόρησης, κάτι που είναι περισσότερο από μία αποκλειστική προσφυγή στις βαθμολογίες των φοιτητών για τους καθηγητές τους.

Ο March (2007) και ο Goldschmid (1978) έχουν βρει ότι τα στοιχεία αξιολόγησης των μαθημάτων μόνα τους σπάνια γίνονται φορείς αλλαγών για τις συμπεριφορές διδασκαλίας αφού πολλά μέλη ΔΕΠ δεν εκπαιδεύονται στην ανάλυση των στοιχείων και είναι επομένως λιγότερο πιθανό να έχουν τις αναγκαίες δεξιότητες για να ερμηνεύσουν τις αξιολογήσεις τους. Επιπλέον, σε πολλά μέλη ΔΕΠ δεν δίνεται η ευκαιρία (εθελοντικά ή υποχρεωτικά) να συζητήσουν τα αποτελέσματά τους με τους προέδρους των τμημάτων ή τους πρυτάνεις και επωφελούνται μόνο εν μέρει από τις υπηρεσίες και τους πόρους που προσφέρονται από τα γραφεία υποστήριξης του Πανεπιστημίου για τη διδασκαλία και τη μάθηση. Ως αποτέλεσμα, τα περισσότερα μέλη ΔΕΠ απλά διεξάγουν μία βιαστική επισκόπηση των στοιχείων που προκύπτουν και σπάνια προσπαθούν να κάνουν εξειδικευμένες αλλαγές που να βασίζονται στην ανατροφοδότηση από τους φοιτητές.

Η έρευνα έχει δείξει ότι όταν στα μέλη ΔΕΠ παρέχεται εκπαίδευση ή ενίσχυση και διαβουλεύσεις με τους ομοβάθμιους ή με τους ΔΕΠ/ εκπαιδευτικούς προγραμματιστές, αυτοί κάνουν αλλαγές στη διδασκαλία τους (Penny & Coe, 2004). Για την ενθάρρυνση των αλλαγών και των θετικών επιδράσεων στις συμπεριφορές διδασκαλίας, ο Abrami (2001) έχει προτείνει περισσότερο ανοικτή επικοινωνία που αφορά στα στοιχεία που συλλέγονται και στην ερμηνεία των αποτελεσμάτων. Οι Beran, Violato και Kline (2007) προτείνουν οι αξιολογήσεις «να συμπληρώνονται από συμπληρωματικές πηγές πληροφοριών που αφορούν στην αποτελεσματικότητα της διδασκαλίας» και υποστηρίζουν ότι «όλες οι ομάδες που είναι χρήστες, συμπεριλαμβανομένων των μελών της διοίκησης, των μελών ΔΕΠ και των φοιτητών

θα πρέπει να είναι ενήμεροι» (σελ. 37) της ανάγκης αυτής της συμπληρωματικής πληροφορίας όταν χρησιμοποιούν τις αξιολογήσεις από τους φοιτητές.

2.5.2 Αποφάσεις για το προσωπικό

Οι ερευνητές τις δεκαετίες του '80 και '90 ανελλιπώς αμφισβητούσαν τη χρήση των αξιολογήσεων των μαθημάτων για συγκεντρωτικού τύπου αποφάσεις. Εν μέρει, αυτές οι ανησυχίες προέρχονταν από πεποιθήσεις ότι τα στοιχεία από τις βαθμολογίες δεν χρησιμοποιούνται αποτελεσματικά ή ισότιμα. Ωστόσο, ο διάλογος σχετικά με την αποτελεσματική χρήση των στοιχείων αξιολόγησης για συγκεντρωτικούς (και για διαμορφωτικούς) σκοπούς σχετίζεται επίσης με τις ερωτήσεις που καθοδηγούν αυτές τις αποφάσεις προσωπικού (personnel decisions). Στην τελευταία δεκαετία, οι στάσεις έχουν αναστραφεί και οι περισσότεροι μελετητές, μεταξύ των οποίων ο Abrami (2001) και ο Algozzine με τους συνεργάτες του (2004), αποδέχονται γενικά – και / ή βεβαιώνουν – την εγκυρότητα των αξιολογήσεων των μαθημάτων για αυτές τις αποφάσεις προσωπικού, που περιλαμβάνουν την πρόσληψη, τη θητεία και την προαγωγή.

Πριν τριάντα χρόνια, η έρευνα έδειχνε ότι ενώ τα μέλη ΔΕΠ ήταν ευνοϊκά απέναντι στην χρήση των αξιολογήσεων από τους φοιτητές σε αποφάσεις για την προαγωγή και τη θητεία (Rich, 1976), η διοίκηση του πανεπιστημίου δεν στηρίζονταν τακτικά σε αυτές για τέτοιους σκοπούς (McKeachie & Lin, 1975). Πιο πρόσφατα, μερικές μελέτες έχουν υποστηρίξει ότι οι διοικούντες κάνουν χρήση των στοιχείων αξιολόγησης των μαθημάτων πιο συχνά απ' ό,τι κάνουν οι διδάσκοντες (Bernar et al., 2005), ειδικότερα για αποφάσεις προσωπικού (Nasser & Fresko, 2002; Haskell, 1997; Schmelkin, Spencer & Gellman, 1997). Μερικές μελέτες έχουν επισημάνει ότι είναι ασαφές κατά πόσο οι διοικούντες χρησιμοποιούν τις πληροφορίες που συλλέγονται κατάλληλα (Abrami, 2001), ή εάν αυτές παρερμηνεύεται ή γίνεται κακή χρήση τους, ως μόνης πηγής στοιχείων σχετικά με τη διδασκαλία (Franklin & Theall, 1989). Επιπλέον, αρκετές μελέτες (βλέπε d' Apollonia & Abrami, 1997; Cashin & Downey, 1992) έχουν δείξει κατ' επανάληψη ότι τα «γενικά στοιχεία» (global items) τα οποία παρείχαν πληροφορίες για τη συνολική ποιότητα των μαθημάτων ή του διδάσκοντα αποτελούσαν τους πιο χρήσιμους δείκτες για την αξιολόγηση της αποτελεσματικότητας της διδασκαλίας.

Παρά τη χρησιμότητα τους στις συγκεντρωτικές αξιολογήσεις και αποφάσεις προσωπικού υπάρχει μία γενική ομοφωνία ότι τα στοιχεία αξιολόγησης των μαθημάτων δεν θα πρέπει να χρησιμοποιούνται μεμονωμένα αλλά μάλλον θα πρέπει

να είναι ένας από τους πολλούς δείκτες που χρησιμοποιούνται για να αξιολογήσουν τη διδασκαλία (Beran, Violato & Kline, 2007; Ory, 2001).

2.5.3 Επιλογή μαθημάτων από τους φοιτητές

Σε κάποια ιδρύματα οι φοιτητές έχουν πρόσβαση στα στοιχεία αξιολόγησης των μαθημάτων μέσα από εκδόσεις, τα «αντί-ημερολόγια». Αυτά τυπικά παρέχουν περιλήψεις των στοιχείων αξιολόγησης, μαζί με σχόλια από τους φοιτητές. Τα τεκμήρια αυτά είναι σχεδιασμένα για να χρησιμοποιούνται από τους φοιτητές με σκοπό την επιλογή των μαθημάτων, άλλα όπως φαίνεται αυτή η χρήση τους είναι περιορισμένη (Beran et al., 2005).

2.5.4 Βραβεία διδασκαλίας

Τα στοιχεία αξιολόγησης των μαθημάτων αποτελούν συχνά ένα στοιχείο βράβευσης της διδασκαλίας μέσα στους φακέλους υποψηφιότητας τόσο σε εσωτερικό επίπεδο (τμήματος, νομαρχίας ή ιδρύματος) όσο και σε εξωτερικό. Σε αυτή την περίπτωση η προσδοκία είναι ότι, οι υποψήφιοι θα επιδείξουν υπεροχή στη διδασκαλία στα επιστημονικά τους πεδία, για τα οποία οι αξιολογήσεις των μαθημάτων λειτουργούν ως ένας δείκτης. Επιπλέον, από τη στιγμή που αυτά τα δεδομένα συλλέγονται τακτικά, οι υποψήφιοι μπορούν να επιδείξουν κανονικά ισχύουσα υπεροχή ή να παρέχουν συγκρίσιμα στοιχεία που να δείχνουν τις σχετικές τους επιδόσεις μέσα στο τμήμα τους, στη νομαρχία ή στο ίδρυμα.

3. Συστήματα Αξιολόγησης & Χρηματοδότηση

3.1 Αξιολόγηση και Εθνικά Πρότυπα (*standards*)

Τα εθνικά πρότυπα καθοδηγούν το πώς πρέπει να γίνεται η μέτρηση της αποτελεσματικότητας της διδασκαλίας ή της επίδοσης επομένως καταπιάνονται με το ΤΙ μετριέται και με το ΠΩΣ μετριέται. Συγκεκριμένα, αυτά καθοδηγούν την ανάπτυξη των εργαλείων μέτρησης, τις τεχνικές της ανάλυσης καθώς και την έκθεση και την ερμηνεία της απόδειξης που μπορεί να χρησιμοποιηθεί στη λήψη αποφάσεων (Berk, 2006). Αναλυτικότερα, καθορίζουν:

- το περιεχόμενο κάθε εργαλείου (τι μετριέται;), για παράδειγμα μίας κλίμακας αξιολόγησης από φοιτητή ή από ομοβάθμιο (συνάδελφο). Αυτό απαιτεί έναν εφ' όλης της ύλης και κατηγορηματικό ορισμό των γνώσεων, των δεξιοτήτων και των ικανοτήτων καθώς και άλλων χαρακτηριστικών και συμπεριφορών που περιγράφουν το έργο της «αποτελεσματικής διδασκαλίας».
- Τον τρόπο που αξιοποιούνται τα δεδομένα από μία κλίμακα αξιολόγησης ή άλλου εργαλείου που βασίζεται στη συστηματική συλλογή των γνωμοδοτήσεων ή των αποφάσεων από τους βαθμολογητές, τους παρατηρητές, ή τους κριτές εξαρτώμενα από την τεχνογνωσία, τα προσόντα και την εμπειρία τους.

Οι άμεσες παρατηρήσεις των φοιτητών και των ομοβάθμιων σχετικές με το ΤΙ βλέπουν στην τάξη αποτελούν τα θεμέλια των αξιολογήσεων. Άλλες έμμεσες πηγές από τις οποίες συνάγεται η αποτελεσματικότητα της διδασκαλίας είναι τα αποτελέσματα των φοιτητών και οι δημοσιεύσεις με αντικείμενο καινοτόμες στρατηγικές διδασκαλίας. Αυτές οι διαφορετικές πηγές δεδομένων διαφέρουν αξιολογούμενα στο ΠΩΣ μετρούν το ΤΙ.

Κατά τη διάρκεια της τελευταίας δεκαετίας έχει αλλάξει ο τρόπος που οι κυβερνήσεις αλληλεπιδρούν με τα ιδρύματα στο χώρο της τριτοβάθμιας εκπαίδευσης. Οι κυβερνητικές αρχές δεν είναι πλέον τόσο δεκτικές στις παραδοσιακές διαδικασίες αυτό-ρύθμισης που έχουν επικρατήσει στους χώρους αυτούς για αιώνες. Ένα νέο οικονομικό κίνητρο τις οδηγεί να αναθεωρήσουν τις σχέσεις τους πιέζοντας τα ιδρύματα να γίνουν περισσότερο ικανά να λογοδοτούν, περισσότερο αποδοτικά, και περισσότερο αποτελεσματικά στη χρήση δημόσιων πόρων. Προγενέστερες

προσπάθειες μέτρησης της θεσμικής αποδοτικότητας (efficiency) και επίδοσης (performance) έχουν γενικά συναντήσει την παθητική αντίσταση ή απερρίφθησαν από τους ακαδημαϊκούς κύκλους. Πίσω από τη φαινομενική ανησυχία για τη βελτιωμένη επίδοση του δημόσιου τομέα, τα αποτελέσματα των δεικτών έχουν αναδειχθεί ως ένα όργανο οικονομικού ορθολογισμού ώστε να βελτιωθεί η αποδοτικότητα και η αποτελεσματικότητα (effectiveness) του ιδρύματος (King, 2000, σελ. 419) Αν και αυτή η στάση επικρατεί, αρκετοί είναι οι ηγέτες του χώρου της τριτοβάθμιας εκπαίδευσης που εκτιμούν ότι η ισχύουσα κατάσταση (status quo) δεν αποτελεί μια βιώσιμη επιλογή για την τριτοβάθμια εκπαίδευση (Barnett, 1992).

3.2 Προσεγγίσεις στην ανάπτυξη των δεικτών επίδοσης

Η ανάπτυξη των δεικτών επίδοσης εν κατακλείδι αφορά στην επιθυμία για λογοδοσία (accountability). Στην τριτοβάθμια εκπαίδευση η λογοδοσία υλοποιείται με τον καθορισμό σκοπών και στόχων (μακροπρόθεσμων) και την περιοδική αξιολόγηση της επίτευξής τους με τη χρήση αποδεκτών δεικτών (Layzell, 1998). Τα συστήματα αξιολόγησης της επίδοσης συμπεριλαμβάνουν όλο και περισσότερο δημοσιονομικού τύπου κίνητρα στο νέο διαχειριστικό πλαίσιο, παροτρύνοντας έτσι τα ιδρύματα να γίνουν περισσότερο παραγωγικά στην επίτευξη των προκαθορισμένων στόχων ή αλλιώς να διακινδυνεύουν μειώσεις στις ετήσιες πιστώσεις. Ο σχεδιασμός και η χρηματοδότηση με βάση την επίδοση έχει επίσης γίνει ένα βολικό μέσο για τις κυβερνήσεις να συγκρίνουν και να ταξινομούν την παραγωγικότητα κάθε ιδρύματος έναντι της παραγωγικότητας των άλλων. Πίνακες επιτυχόντων, κάρτες βαθμολογίας, εκθέσεις είναι κάποιες από τα δημοφιλή εργαλεία που αναπτύχθηκαν για να συγκρίνονται οι μετρήσεις επίδοσης των ιδρυμάτων (King, 2000 σελ. 419).

Σύμφωνα με τους Ewell και Jones (1994), οι επικρατέστερες τέσσερις προσεγγίσεις που συνήθως χρησιμοποιούνται στη μέτρηση της προόδου που λογοδοτεί για την επίτευξη των σκοπών και των στόχων είναι οι εξής:

- *Εισροές, διαδικασίες, αποτελέσματα*: ένα μοντέλο της διαδικασίας ‘παραγωγής’ που έχει στόχο να μετρά την προστιθέμενη αξία για τους φοιτητές που αποχωρούν, ίσως διαμέσου μιας προ- και μετά- αξιολόγησης.
- *Αποδοτικότητα πόρων και αποτελεσματικότητα*: μία προσέγγιση που έχει στόχο να μετρά την αποδοτική χρήση πόρων-κλειδιά όπως είναι τα μέλη ΔΕΠ, ο

χώρος και ο εξοπλισμός, χρησιμοποιώντας αναλύσεις αναλογίας ή παρόμοιες τεχνικές.

- *Ανάγκες κράτους και απόδοση επενδύσεων*: μία προσέγγιση που χτίζεται πάνω στην υπόθεση ότι η τριτοβάθμια εκπαίδευση αποτελεί μία στρατηγικού τύπου επένδυση για την πολιτεία: είναι σχεδιασμένη για να μετρά την προσαρμογή της τριτοβάθμιας εκπαίδευσης στις ανάγκες της πολιτείας (π. χ προετοιμασία εργατικού δυναμικού).
- *Ανάγκες «καταναλωτή» και απόδοση επενδύσεων*: μία προσέγγιση που χτίζεται πάνω στην έννοια του «καταναλωτισμού» (consumerism) η οποία είναι σχεδιασμένη να μετρά τον αντίκτυπο της τριτοβάθμιας εκπαίδευσης στην αντιμετώπιση των ατομικών αναγκών (π. χ διατήρηση και ποσοστά αποφοίτησης, απασχολησιμότητα αποφοίτων).

Οι τέσσερις προσεγγίσεις δεν είναι ανεξάρτητες, ως εκ τούτου, η χρησιμοποίηση των δεικτών επίδοσης είναι δυνατό να αντλεί στοιχεία από μία ή περισσότερες περιοχές. Επιπλέον οι Ewell και Jones (1994) τονίζουν ότι οι πολιτικοί στόχοι και σκοποί πρέπει να καθοδηγούν την επιλογή των δεικτών επίδοσης και όχι το αντίστροφο.

3.3 Η μέχρι τώρα εμπειρία από τη χρήση των δεικτών επίδοσης

Ο Gaither (1994) παραθέτει διάφορα «μαθήματα» που απεκόμισε από την προγενέστερη εμπειρία με τους δείκτες επίδοσης:

- Οι δείκτες επίδοσης πρέπει να περιορίζονται στο ελάχιστο (< 20).
- Οι δείκτες επίδοσης δεν πρέπει να αναπτύσσονται με τον από πάνω προς τα κάτω (top-down) τρόπο.
- Τόσο η σχολή (ΔΕΠ) όσο και το νομοθετικό σώμα πρέπει να συμμετέχουν στην ανάπτυξη των δεικτών για μακρόχρονη επιτυχία.
- Ένα μοντέλο δεικτών δεν μπορεί να εφαρμόζεται σε όλους τους τύπους των ιδρυμάτων αποτελεσματικά χωρίς μείωση διαφόρων αποστολών (δηλαδή ένα μέγεθος δεν ταιριάζει σε όλους).
- Οι υπεύθυνοι χάραξης πολιτικής τείνουν να προτιμούν την ποσοτική από την ποιοτική μέτρηση.
- Οι δείκτες πρέπει να έχουν οικονομικά κίνητρα για τα ιδρύματα.

- Τα αποτελέσματα της επίδοσης πρέπει να κοινοποιούνται έγκαιρα και με κατανοητό τρόπο για τους υπεύθυνους χάραξης πολιτικής και για το ευρύ κοινό.

Η σχετική σημασία καθενός από αυτά τα «μαθήματα» ποικίλει. Τέλος, όπως επισημαίνει ο Layzel (1998) «Κάθε ένα από αυτά (τα «μαθήματα») υπήρξε ένα πικρό μάθημα από μία τουλάχιστο Πολιτεία (των ΗΠΑ) κατά τη διάρκεια εφαρμογής των δεικτών επίδοσης» (σελ.105).

3.4 Δείκτες επίδοσης στην Ανώτατη Εκπαίδευση: περιορισμοί & κίνδυνοι

Η χρήση των δεικτών επίδοσης εμπεριέχει ένα αριθμό περιορισμών και κινδύνων που πρέπει να έχουν υπόψη τους όσοι ασχολούνται με την ανάπτυξή τους. Αυτοί περιγράφονται στη συνέχεια (Layzell, 1998, σελ. 105-6).

- *Περιορισμοί Δεδομένων:* Ένα σημαντικό πρακτικό θέμα το οποίο ασκεί επίδραση στην ανάπτυξη και στην εφαρμογή των δεικτών επίδοσης είναι η διαθεσιμότητα των δεδομένων. Η δυνατότητα αξιοποίησης υπαρχόντων συστημάτων συλλογής δεδομένων μειώνει το χρόνο εκκίνησης και το κόστος εφαρμογής ενός συστήματος δεικτών επίδοσης. Αυτό επίσης βελτιώνει το «επίπεδο άνεσης» εκείνων που εμπλέκονται και συνεπώς την αξιοπιστία της διαδικασίας. Από την άλλη, η διατήρηση μόνο εκείνων τους δεικτών για τους οποίους υπάρχουν διαθέσιμα πρόσφατα δεδομένα μπορεί να μη καταλήξει στα πιο χρήσιμα και κατάλληλα σύνολα επίδοσης.
- *Το περισσότερο δεν είναι απαραίτητα το καλύτερο:* (η κατοχή τόσο πολλών δεικτών). Ένας κοινός κίνδυνος στον οποίο πολλοί φορείς που χαράσσουν πολιτική πέφτουν θύμα είναι η επιθυμία να χρησιμοποιούν διάφορους δείκτες επίδοσης, σκεπτόμενοι ότι αυτό κατά κάποιο τρόπο παρέχει μία πιο πλήρη εικόνα της επίδοσης των ιδρυμάτων. Αυτό τυπικά συνιστά ένα σύμπτωμα της μη κατοχής κάποιου σαφούς πλαισίου (standards) που να καθοδηγεί την ανάπτυξη των δεικτών επίδοσης. Ακόμη συνιστά το αποτέλεσμα μιας πολιτικής διαδικασίας ανάπτυξης κατά την οποία οι γνώμες διαφέρουν σχετικά με το τι είναι σημαντικό να μετρηθεί. Το αποτέλεσμα από την κατοχή τόσων πολλών δεικτών είναι διττό. Πρώτο, με όσους περισσότερους δείκτες ένα ίδρυμα μετριέται, τόσο λιγότερο σημαντικός γίνεται κάθε ένας από αυτούς τους δείκτες, και το αντίθετο. Αυτό οφείλεται στο γεγονός ότι η διαδικασία της ελαχιστοποίησης του αριθμού των δεικτών αποτελεί μία

διαδικασία προτεραιότητας η οποία επιβεβαιώνει ότι οι δείκτες επίδοσης που τελικά επιλέγονται αποτιμώνται ως σημαντικοί. Δεύτερο, καθώς προστίθενται οι δείκτες και οι στόχοι, το ίδρυμα διατρέχει τον κίνδυνο των αντικρουόμενων στόχων και αποτελεσμάτων. Για παράδειγμα, εάν οι υπεύθυνοι χάραξης πολιτικής αποφασίσουν ότι τα ιδρύματα πρέπει να δείξουν ένα υψηλό επίπεδο πρόσβασης και εισαγωγής για τους πρωτοετείς (δηλαδή μία πολιτική ανοιχτών θυρών) και παράλληλα υψηλές βαθμολογίες αποφοίτων, αυτοί είναι πιθανό να διαπιστώσουν ότι κάποιος από τους δείκτες δεν είναι απαραίτητα συμβατός με τους άλλους.

- *Πλαίσιο Πολιτικής και Δείκτες Επίδοσης:* Μία παγίδα που πρέπει να αποφεύγεται είναι η βιασύνη για την ανάπτυξη των δεικτών επίδοσης χωρίς να υπάρχει ένα σαφές πολιτικό πλαίσιο που να καθοδηγεί αυτή την ανάπτυξη. Οι πρώτες δύο ερωτήσεις που πρέπει να απαντηθούν είναι:
 - ποιοι στόχοι στην τριτοβάθμια εκπαίδευση θεωρούνται ως οι πιο σημαντικοί για να επιτευχθούν;
 - με ποιο τρόπο πρέπει να μετρηθεί αυτή η επίτευξη;

Κάθε υπεύθυνος χάραξης πολιτικής διαθέτει τους δικούς του /ή της «αγαπημένους» δείκτες οι οποίοι μπορεί να ταιριάζουν, μπορεί και όχι, σε αυτούς τους ευρύτερους στρατηγικούς στόχους. Χωρίς αυτούς τους στρατηγικούς στόχους, οι δείκτες επίδοσης καταλήγουν απλά σε μία συλλογή ενοχλητικών δεδομένων και δραστηριοτήτων αναφοράς χωρίς τέλος.

- *Ποσοτικοί εναντίον ποιοτικών δεικτών.* Τα συστήματα των δεικτών επίδοσης που επικεντρώνονται μόνο σε ποσοτικές μετρήσεις μπορεί να είναι πιο άνετα και οικεία στους υπεύθυνους χάραξης πολιτικής αλλά μπορεί επίσης να παρέχουν μία μονοδιάστατη εικόνα του ιδρύματος. Η ανάπτυξη έγκυρων ποιοτικών δεικτών της οργανωτικής επίδοσης απαιτεί αυστηρότητα και πειθαρχία, οι οποίες όμως δύσκολα επιτυγχάνονται. Ένα καλά-ισορροπημένο σύστημα θα συμπεριλαμβάνει και τους δύο τύπους μετρήσεων.
- *Συγχέοντας «Εισροές», «Διαδικασίες» και «Αποτελέσματα».* Μία κοινή παγίδα στα συστήματα των δεικτών επίδοσης είναι η εξασθένιση των εισροών, διαδικασιών και αποτελεσμάτων στην ανάπτυξη των μετρήσεων. Για παράδειγμα, μία δημοφιλής «μέτρηση επίδοσης» η οποία χρησιμοποιείται από πολλές χώρες είναι ο *φόρτος διδασκαλίας* των μελών ΔΕΠ. Ενώ αυτή είναι μία

σημαντική μέτρηση θεσμικής χρήσης των πόρων και λέει κάτι για την εσωτερική διαδικασία προϋπολογισμού του ιδρύματος, δεν λέει κάτι σχετικά με το αποτέλεσμα της διδασκαλίας. Αλλά τα θεσμικά *αποτελέσματα* δεν συμβαίνουν εν κενό: αυτά σχετίζονται άμεσα με τις *εισροές* και τις *διαδικασίες* του ιδρύματος. Συνοψίζοντας, ένα καλά-στρογγυλεμένο σύνολο δεικτών επίδοσης πρέπει ρητά να συμπεριλαμβάνει πλευρές των εισροών, διαδικασιών και αποτελεσμάτων αλλά και να προηγείται των διαφορών τους.

- *Έλλειψη γενικού προσανατολισμού εκ των προτέρων* (Lack of Broad “Buy in” up front): Ο πιο γρήγορος τρόπος για την καταδίκη κάθε νέας πολιτικής πρωτοβουλίας, στην οποία περιλαμβάνεται και η καθιέρωση των δεικτών επίδοσης, είναι η ανάπτυξη τους χωρίς τις επαρκείς εισροές από τα βασικά ενδιαφερόμενα μέλη του οργανισμού. Μολονότι μία περισσότερο συμβουλευτική διαδικασία απαιτεί χρόνο, πειθαρχία και υπομονή, φαίνεται επίσης να είναι πιο επιτυχημένη και βιώσιμη σε βάθος χρόνου.

3.5 Συνδέοντας επίδοση με χρηματοδότηση

Η χρηματοδότηση με βάση την επίδοση (*performance-based funding*) αποτελεί την επέκταση της λογικής ενός συστήματος δεικτών επίδοσης και διασυνδέεται άμεσα με την απόδοση ευθυνών (λογοδοσία), την επίδοση, και τα επίπεδα χρηματοδότησης. Η κατανόηση της μοναδικής φύσης αυτών των μηχανισμών απαιτεί να εξεταστούν οι κυρίαρχοι τρόποι κατανομής των πόρων στα δημόσια ιδρύματα της τριτοβάθμιας εκπαίδευσης που επί του παρόντος χρησιμοποιούνται: ο στοιχειώδης προϋπολογισμός (*incremental budgeting*) και ο τύπος προϋπολογισμού (*formula budgeting*).

Στοιχειώδης προϋπολογισμός είναι η παραδοσιακή και κυρίαρχη φόρμα του κρατικού προϋπολογισμού. Αυτή ξεκινά με τον βασικό προϋπολογισμό του προηγούμενου έτους του ιδρύματος και χορηγεί αυξήσεις (ή και μειώσεις) σύμφωνα με ένα σύνολο καθιερωμένων κανόνων απόφασης ή κατευθυντήριων γραμμών προϋπολογισμού. Παραδείγματα περιλαμβάνουν, πληθωριστικές αυξήσεις για τις προμήθειες και τις υπηρεσίες κοινής ωφέλειας, ή αναπροσαρμογές των μισθών των εργαζομένων με βάση το κόστος ζωής. Οι υπεύθυνοι για τη χάραξη πολιτικής επικεντρώνονται για παράδειγμα σε αυξήσεις / μειώσεις που οφείλονται σε πληθωριστικές αυξήσεις ή σε αλλαγές φόρτου εργασίας και μερικές φορές, σε νέες πρωτοβουλίες του προγράμματος.

Ο *τύπος προϋπολογισμού*, από την άλλη, αναφέρεται σε μία μαθηματική βάση για την απαίτηση και / ή την κατανομή των κεφαλαίων στα ιδρύματα της τριτοβάθμιας

εκπαίδευσης, χρησιμοποιώντας ένα σύνολο κόστους και παραγόντων προσωπικού (δηλαδή κόστος ανά πίστωση (cost per credit), αναλογίες φοιτητών/μέλη ΔΕΠ) σε σχέση με τις καθορισμένες εισροές (δηλαδή διδακτικές ώρες φοιτητή, επίπεδα εγγραφών).

Η χρηματοδότηση με βάση την επίδοση είναι διαφορετική από τους δύο προαναφερθέντες τρόπους που χρησιμοποιούνται σήμερα για την κατανομή των πόρων στα δημόσια πανεπιστήμια. Σύμφωνα με αυτή «οι πόροι ρέουν μόνο μετά την απόδειξη από την πλευρά του αποδέκτη ότι ένα καθορισμένο αποτέλεσμα πράγματι έχει παραχθεί με την αξιοποίηση των κεφαλαίων» (Jones, 1997).

Συνοψίζοντας, οι μέθοδοι του *στοιχειώδους* και με βάση τύπου προϋπολογισμού έχουν μία προσέγγιση «*με βάση τις ανάγκες*» (“needs based”) για την κατανομή των πόρων, ενώ η *χρηματοδότηση / προϋπολογισμός με βάση την επίδοση* υιοθετεί μία περισσότερο «*με βάση τα προσόντα*» (“merit based”) προσέγγιση.

3.5.1 Χαρακτηριστικά του προϋπολογισμού με βάση την επίδοση

Σύμφωνα με τον Carter (1994), ένας *προϋπολογισμός με βάση την επίδοση* έχει τα ακόλουθα τέσσερα χαρακτηριστικά:

- Παρουσιάζει τον κύριο σκοπό για τον οποίο διατίθενται οι χρηματοδοτήσεις και θέτει μετρήσιμους στόχους.
- Αναφέρεται σε προηγούμενες επιδόσεις και χρησιμοποιεί κοινές ταξινομήσεις κόστους που επιτρέπουν στα προγράμματα να συγκρίνονται από το να επικεντρώνεται σε γραμμικές συγκρίσεις.
- Προσφέρει ευελιξία στη διαχείριση της ανακατανομής των χρημάτων που απαιτούνται και παρέχει αμοιβές για την επίτευξη ή τιμωρίες για την αποτυχία.
- Συμπεριλαμβάνει αποτελέσματα από περιοδικές αξιολογήσεις προγραμμάτων οι οποίες υποστηρίζονται από αξιόπιστες πληροφορίες που μπορούν να ελέγχονται ανεξάρτητα.

3.5.2 Προϋποθέσεις για τη χρηματοδότηση με βάση την επίδοση

Ο Jones (1997) επισημαίνει επιπλέον ότι για να εφαρμοστεί μία χρηματοδότηση πρωτοβουλίας με βάση την επίδοση, πρέπει να λαμβάνουν χώρα τα ακόλουθα τέσσερα στοιχεία:

- Οι προς επίτευξη στόχοι – είτε αποτελέσματα είτε επίδειξη καλής πρακτικής.

- Οι «μετρήσεις της επιτυχίας» - ειδικές μετρήσεις ή ορισμοί με τα οποία υπολογίζεται η επίδοση.
- Η βάση της ανταμοιβής – τα κριτήρια αξιολόγησης της επιτυχίας.
- Η μέθοδος για την κατανομή των πόρων.

Συγκεκριμένα, ο Hölttä (1998) αναφέρει ότι, πέντε δείκτες μέτρησης της θεσμικής αποτελεσματικότητας και αποδοτικότητας των ακαδημαϊκών δραστηριοτήτων καθορίζουν τη χρηματοδότηση με βάση την επίδοση. Αυτοί αφορούν:

- στην ικανότητα του ιδρύματος να προσελκύει εξωτερική χρηματοδότηση από τις εγχώριες και διεθνείς πηγές,
- στη σχετική θέση του ιδρύματος στις ταξινομήσεις των ιδρυμάτων για τις διεθνείς ανταλλαγές φοιτητών και διδασκόντων,
- στην αποτελεσματικότητα των υπηρεσιών στην εκπαίδευση ενηλίκων,
- στις θέσεις εργασίας των αποφοίτων στην αγορά εργασίας,
- στην επιτυχία του ιδρύματος να δημιουργεί ακαδημαϊκή ποιότητα στη διδασκαλία και στην έρευνα (σελ. 59).

3.6 Προβλήματα στη χρηματοδότηση με βάση την επίδοση

Ο Ashworth (1994) προειδοποιεί ότι η πλήρης εφαρμογή της χρηματοδότησης με βάση την επίδοση έχει δύο θεμελιώδη προβλήματα. Πρώτον, «δεν υπάρχει ενιαία συμφωνία για τις αξίες που πρέπει «να μπουν μπρος να δουλεύουν» σε ένα τύπο (formula) και δεν υπάρχουν διαθέσιμα στοιχεία για λογικά ή ανεκτά κόστη που να τροφοδοτούν ένα τέτοιο σύστημα (formula system)» (σελ. 11). Δεύτερο, εάν η κατανομή όλων των χρηματοδοτήσεων γίνει σε μία βάση επίδοσης είναι πιθανό να υπάρξει σημαντική ανακατανομή των χορηγήσεων από χρόνο σε χρόνο. Αυτό θα επηρέαζε δυσμενώς την ικανότητα του ιδρύματος να σχεδιάζει και να εκτελεί, προκαλώντας τελικά την ήττα του σκοπού για τον οποίο καταρτίστηκε ο προϋπολογισμός με βάση την επίδοση.

Οι Burke και Serban (1997) εξερεύνησαν μεταξύ άλλων τις δυσκολίες που αντιμετώπισαν οι Πολιτείες που εφήρμοσαν τη χρηματοδότηση των πρωτοβουλιών με βάση την επίδοση. Στην έρευνά τους συμπεριέλαβαν διάφορους ανθρώπους εμπλεκόμενους στη διαδικασία εφαρμογής, μεταξύ αυτών κυβερνώντες, νομοθέτες, υψηλά στελέχη της εκπαίδευσης σε κρατικό επίπεδο, θεσμικά συμβούλια, μέλη της διοίκησης των ιδρυμάτων και ηγέτες σχολών.

Τα αποτελέσματα έδειξαν ότι αυτοί οι άνθρωποι ανέφεραν τρεις βασικούς παράγοντες που προκαλούσαν δυσκολίες στην εφαρμογή της χρηματοδότησης με βάση την επίδοση: (1) την επιλογή των δεικτών επίδοσης, (2) την επιλογή των κριτηρίων «επιτυχίας» (τα σημεία αναφοράς της επιτυχίας), (3) τη μικρή χορηγούμενη χρηματοδότηση για την πρωτοβουλία. Σαφώς, οι δύο πρώτοι συνιστούν παράγοντες κλειδιά όσον αφορά στην επιτυχία κάθε χρηματοδότησης με βάση την επίδοση. Ο τρίτος παράγοντας αντανακλά την ανάγκη να παρέχεται ένα σημαντικό κίνητρο στα ιδρύματα για να πάρουν στα σοβαρά αυτή τη χρηματοδότηση.

3.7 Το τοπίο της Ανώτατης Εκπαίδευσης στην Ευρώπη

Η Ανώτατη Εκπαίδευση στην Ευρώπη έχει βιώσει πρωτοφανείς αλλαγές κατά τη διάρκεια των δύο τελευταίων δεκαετιών. Η επεκτατική και, σε κάποιο βαθμό, η φάση της μη-επέμβασης (“laissez-faire”) που παρατηρήθηκαν τις δεκαετίες του '80 και του '90, παρήκμασαν στο ξεκίνημα του 21^{ου} αιώνα. Βασικοί λόγοι υπήρξαν το δημόσιο έλλειμμα και χρέος που επέφερε η συνθήκη του Μάαστριχτ καθώς και δημογραφικοί λόγοι που σχετίζονταν με την προβλέψιμη μείωση του πληθυσμού σε κάποιες χώρες της Δυτικής Ευρώπης (ΟΟΣΑ, 2005).

Ο προκύπτων εξορθολογισμός στις δημόσιες δαπάνες έχει εγείρει ανησυχίες για το κατά πόσον τα πανεπιστήμια λειτουργούν στο μέγιστο δυνατό επίπεδο της οικονομικής αποδοτικότητας (cost-efficiency) (Abbot & Doucouliagos, 2003). Κατά την περίοδο 1980-1995 οι χώρες Αυστρία, Δανία, Γαλλία, Γερμανία, Βέλγιο, Νορβηγία, Σουηδία, Ελβετία παρουσίασαν πτώση στις δαπάνες ανά φοιτητή, ως ποσοστό του εθνικού ΑΕΠ ανά κάτοικο, αναγκάζοντας πολλά ιδρύματα να υιοθετήσουν νέες διαχειριστικές στρατηγικές (Eicher, 1998).

Η μέχρι τώρα αυτονομία των ιδρυμάτων Ανώτατης Εκπαίδευσης έχει οδηγήσει τις κυβερνήσεις να αλλάξουν τις διαδικασίες ελέγχου από τον παραδοσιακό εκ των προτέρων (ex ante) έλεγχο στην εκ των υστέρων (ex post) αιτιολόγηση χρησιμοποιώντας τη διασφάλιση της ποιότητας και τις μετρήσεις λογοδοσίας/απόδοσης ευθυνών (Harvey & Knight, 1996; Neave, 1998). Στις περισσότερες χώρες του ΟΟΣΑ η πολιτική ατζέντα της Ανώτατης Εκπαίδευσης περιλαμβάνει θέματα όπως είναι η λογοδοσία, η αποτελεσματικότητα και η αξία χρήματος. Ιδιαίτερα για τα Ευρωπαϊκά συστήματα Ανώτατης Εκπαίδευσης η διαδικασία της σύγκλισης άρχισε μετά τη Διακήρυξη της Μπολόνια (1999). Η

δημιουργία έως το 2010 του Ευρωπαϊκού Χώρου Ανώτατης Εκπαίδευσης (EAEU) (European Higher Education Area, EHEA) στόχευε στην αύξηση της απασχόλησης, της κινητικότητας, της διαφάνειας και της συγκρισιμότητας αυτών των συστημάτων (Mora, 2003). Ακόμα, στόχευε να δώσει έμφαση στον ανταγωνισμό μεταξύ των ιδρυμάτων με κριτήριο την ποιότητα της παροχής των υπηρεσιών αναφορικά με έναν υψηλότερο βαθμό αποτελεσματικότητας, ο οποίος θα συνέβαλε στην αύξηση των εγγραφών των νέων φοιτητών.

Σύμφωνα με την Επιτροπή των Ευρωπαϊκών Κοινοτήτων (2003), υπάρχουν τρεις κατευθύνσεις για τις προσπάθειες που στοχεύουν στη βελτίωση της ποιότητας της Ανώτατης Εκπαίδευσης. Αυτές είναι, η κατεύθυνση:

1. της εξασφάλισης ότι τα Ευρωπαϊκά Πανεπιστήμια έχουν επαρκείς και βιώσιμες πηγές και τις χρησιμοποιούν αποτελεσματικά,
2. της εδραίωσης της αριστείας στη διδασκαλία και στην έρευνα,
3. του ανοίγματος των πανεπιστημίων σε μεγαλύτερη έκταση προς τα έξω και αύξηση της διεθνούς τους ελκυστικότητας.

Στις περισσότερες Ευρωπαϊκές χώρες έχουν δημιουργηθεί Εθνικά Συστήματα αξιολόγησης με αποστολή την αξιολόγηση της Ανώτατης Εκπαίδευσης (Thune, 1998). Έτσι, η αξιολόγηση των δραστηριοτήτων που διεξάγουν τα ιδρύματα έχει καταστεί η τρέχουσα πρακτική, η οποία συνιστά μέρος της στρατηγικής για βελτίωση της ποιότητας της Ευρωπαϊκής Ανώτατης Εκπαίδευσης. Από την άλλη, τα ιδρύματα αυτά ενδιαφέρονται για την καθιέρωση των δικών τους εσωτερικών μηχανισμών αξιολόγησης με σκοπό την καθοδήγηση της ολοκλήρωσης των δικών τους αντικειμένων. Η αυτό-αξιολόγηση παίζει ένα σημαντικό ρόλο στις στρατηγικές που εφαρμόζονται για τη μέτρηση της επίδοσης και τη βελτίωση της ποιότητας των ιδρυμάτων της Ανώτατης Εκπαίδευσης (Thune, 1998). Αυτή παρέχει κρίσιμη πληροφόρηση σχετικά με τις επιδόσεις των μονάδων που απαρτίζουν το ίδρυμα και έτσι επιτρέπει τη λήψη αποφάσεων που στοχεύουν στη βελτίωση της αποτελεσματικότητας και της ποιότητας των διαδικασιών.

Επισημαίνεται ότι, η απουσία τιμών αγοράς για τους παράγοντες και τα αποτελέσματα στα πανεπιστήμια καθιστούν την αξιολόγηση ιδιαίτερα πολύπλοκη και οδηγούν στην ανάπτυξη εναλλακτικών μετρήσεων επίδοσης (Johnes & Johnes, 1993). Υπάρχει ένας αυξανόμενος αριθμός περιοδικών και εφημερίδων που δημοσιεύουν

δείκτες επίδοσης και πίνακες με τις θέσεις των πανεπιστημίων που ασκούν μία σημαντική επίδραση στις αποφάσεις των φοιτητών και των εργοδοτών (Colbert et al., 2000). Επομένως, η ανάγκη για την παρακολούθηση και την αξιολόγηση των θεσμικών επιδόσεων στην Ανώτατη Εκπαίδευση είναι ένα θέμα που επανέρχεται συνεχώς μεταξύ των ευρωπαϊκών κρατών. Την τελευταία δεκαετία, στις περισσότερες περιπτώσεις η δυναμική του περιλαμβάνει μεταρρυθμιστικές προσπάθειες σαν κι αυτές των ΗΠΑ και της Βρετανίας, τις οποίες εκείνοι που χαράσσουν πολιτική προσπαθούν να εισηγηθούν. Για παράδειγμα, στη Φινλανδία ένα νέο μοντέλο χρηματοδότησης τέθηκε σε εφαρμογή από το 1997, που ενσωματώνει το στοιχείο της επίδοσης και το οποίο καλύπτει το 3% του συνολικού προϋπολογισμού του πανεπιστημίου (Hämäläinen & Moitus, 1999; Hölttä, 1998, σ. 59). Στην Ολλανδία, για τη χάραξη πολιτικής απαιτείται η υλοποίηση ενός μηχανισμού χρηματοδότησης της έρευνας, που είναι παρόμοιος με τον βρετανικό, σύμφωνα με τον οποίο η παραγωγικότητα της έρευνας αποτιμάται με βάση μια κυβερνητική κλίμακα αξιολόγησης (El-Khawas & Massey, 1996). Τα πανεπιστήμια στις χώρες Γερμανία, Σουηδία, Δανία και Αυστρία εξασφαλίζουν οικονομικούς πόρους μέσω πιλοτικών προγραμμάτων χρηματοδότησης και μέσω των περιορισμένων κατανομών επιχορήγησης, στα πλαίσια μιας συναλλαγής που απαιτεί από τα ιδρύματα την αύξηση των επιδόσεών τους (Cave, Hanney, Henkel, & Kogan, 1997).

Είναι προφανές ότι, αυτές οι διεθνείς πρωτοβουλίες που βασίζονται στην επίδοση αποτελούν ένα κύμα αλλαγών στη σχέση μεταξύ των κυβερνήσεων και της τριτοβάθμιας εκπαίδευσης. Το κράτος στην προσπάθειά του να παράγει μεγαλύτερη λογοδοσία, αναθέτει στους θεσμικούς ηγέτες των πανεπιστημίων εκπαιδευτικές αποστολές που σκοπό έχουν να αντιμετωπίζουν τις ανάγκες της εθνικής οικονομίας σε μια ανταγωνιστική παγκόσμια αγορά. Σε αυτό το πιο ανταγωνιστικό μοντέλο οι κυβερνήσεις εξετάζουν προσεκτικά τη συνολική πρόοδο και τις επιδόσεις των πανεπιστημίων για την αντιμετώπιση των οικονομικών αναγκών του κράτους μέσω της χρήσης των δημοσιονομικών (fiscal) κινήτρων και αντικινήτρων. Τα πανεπιστήμια, από την άλλη πλευρά, αναμένεται να ανταγωνίζονται έντονα για να εξασφαλίζουν πρόσθετους πόρους με βάση τους προκαθορισμένους στόχους επιδόσεων που τίθενται από το κράτος. Σε αυτό το είδος του συστήματος με βάση την επίδοση η ευθύνη για τη βελτίωση βαρύνει τα ίδια τα ιδρύματα και τα τμήματά τους (King, 2000, σελ. 426).

4. Στρατηγικές Μέτρησης Αποτελεσματικότητας Διδασκαλίας

Στα προηγούμενα 30 χρόνια, η κυρίαρχη μέτρηση όσον αφορά στην αποτελεσματικότητα της διδασκαλίας στην τριτοβάθμια εκπαίδευση είχε επικρατήσει να είναι οι βαθμολογίες από τους φοιτητές (Seldin, 1999a). Ωστόσο, την τελευταία δεκαετία παρατηρείται η τάση να πληθαίνουν οι αξιολογήσεις της διδασκαλίας με δεδομένα από πολλαπλές πηγές. Με την τακτική αυτή διευρύνεται και βαθαίνει η βάση απόδειξης για την αποτίμηση των μαθημάτων και την αξιολόγηση της ποιότητας της διδασκαλίας (Arreola, 2000; Braskamp & Ory, 1994; Knapper & Cranton, 2001; Seldin & Associates, 1999). Ακόμα έχουν προταθεί διάφορα περιεκτικά μοντέλα για την αξιολόγηση των διδασκόντων (Arreola, 2000; Braskamp & Ory, 1994; Centra, 1999; Keig & Waggoner, 1994; Romberg, 1985; Soderberg, 1986). Αυτά περιλαμβάνουν πολλαπλές πηγές απόδειξης με το μεγαλύτερο βάρος να αποδίδεται στις εισροές των φοιτητών και των ομοβάθμιων και λιγότερο στην αυτό-αξιολόγηση, στους απόφοιτους, στα μέλη της διοίκησης και σε άλλους. Όλα αυτά τα μοντέλα χρησιμοποιούνται για να καταλήξουν σε αποφάσεις διαμορφωτικού και συγκριτικού τύπου.

4.1 Η αναγκαιότητα πολλαπλών πηγών αξιολόγησης

Κάθε πηγή μπορεί να υποστηρίξει μοναδική πληροφόρηση αλλά επίσης είναι επιδεκτική πλάνης ή σφάλματος συνήθως με ένα τρόπο διαφορετικό από τις άλλες πηγές. Για παράδειγμα, η αναξιοπιστία ή οι προκαταλήψεις (bias) των αξιολογήσεων από τους ομοβάθμιους δεν είναι ίδιες με εκείνες από τους φοιτητές οι οποίες έχουν άλλες αδυναμίες. Η απόδειξη από τρεις ή περισσότερες διαφορετικές πηγές έχει ως αποτέλεσμα οι αρετές κάθε πηγής να μπορούν να αντισταθμίζουν τις αδυναμίες των άλλων και ως εκ τούτου να συμβάλλουν στη διαμόρφωση μίας απόφασης σχετικά με την αποτελεσματικότητα της διδασκαλίας η οποία να είναι πιο ακριβής από ό, τι μία που βασίζεται σε μία μόνο πηγή (Appling, Nauman & Berk, 2001).

4.2 «Οι 12 Πηγές Απόδειξης»

Σύμφωνα με τον Scriven (1991) αξιολόγηση είναι «η διαδικασία, της οποίας το καθήκον είναι ο συστηματικός και αντικειμενικός καθορισμός του πλεονεκτήματος (merit), της αξίας (worth) ή της τιμής (value). Χωρίς μία τέτοια διαδικασία, δεν υπάρχει τρόπος να διαχωριστεί αυτό που αξίζει από αυτό που δεν αξίζει» (σελ. 4). Η διαδικασία αυτή περιλαμβάνει: 1) τη συλλογή δεδομένων και 2) τη χρήση αυτών των δεδομένων για διαμόρφωση κρίσης και απόφασης σε σχέση με τα συμφωνημένα

πρότυπα. Εργαλεία μέτρησης χρειάζονται για να συλλέξουν αυτά τα δεδομένα, όπως είναι οι έλεγχοι, οι κλίμακες, και τα ερωτηματολόγια. Τα κριτήρια για την αποτελεσματικότητα της διδασκαλίας είναι ενσωματωμένα στο περιεχόμενο αυτών των μετρήσεων. Οι πιο συνήθεις μετρήσεις για την αποτίμηση των διδασκόντων γίνονται με τις κλίμακες αξιολόγησης.

Στο βιβλίο του Berk (2006, σελ. 14-34) περιγράφονται 13 δυναμικές πηγές απόδειξης της αποτελεσματικότητας της διδασκαλίας, οι κάτωθι:

- (1) βαθμολογίες από τους φοιτητές,
- (2) βαθμολογίες από τους ομοβάθμιους,
- (3) βαθμολογίες από εξωτερικούς ειδικούς,
- (4) αυτό-αξιολογήσεις,
- (5) μαγνητοσκοπήσεις,
- (6) συνεντεύξεις φοιτητών,
- (7) οι βαθμολογίες από τους απόφοιτους,
- (8) βαθμολογίες από τους εργοδότες,
- (9) βαθμολογίες των διοικούντων
- (10) κατάρτιση διδασκαλίας,
- (11) βραβεία διδασκαλίας,
- (12) μετρήσεις των αποτελεσμάτων μάθησης,
- (13) χαρτοφυλάκιο διδασκαλίας.

Ένα περίγραμμα αυτών των πηγών φαίνεται στον Πίνακα 1 μαζί με διάφορα εξέχοντα χαρακτηριστικά και πληροφορίες όπως είναι:

- ο τύπος μέτρησης που απαιτείται για τη συλλογή της απόδειξης,
- το υπεύθυνο πρόσωπο / -α που θα παρέχει την απόδειξη (φοιτητές, ομοβάθμιοι, διδάσκων, ή μέλος διοίκησης),
- το πρόσωπο ή η επιτροπή που χρησιμοποιεί την απόδειξη, και
- η απόφαση /-σεις που αποδίδονται τυπικά με βάση τα δεδομένα

(Δ = Διαμορφωτική, Σ = Συγκριτική, Π = Πρόγραμμα).

Πίνακας 1: Εξέχοντα χαρακτηριστικά των 13 πηγών απόδειξης για την αποτελεσματικότητα της διδασκαλίας

Πηγή Απόδειξης	Τύπος Μέτρησης	Ποιος παρέχει την απόδειξη	Ποιος χρησιμοποιεί την απόδειξη	Τύπος της απόφασης
Βαθμολογίες από φοιτητές	Κλίμακα Αξιολόγησης	Φοιτητές	Διδάσκοντες/ Διοικούντες	Δ/Σ/Π
Βαθμολογίες από τους ομοβάθμιους	Κλίμακα Αξιολόγησης	Ομοβάθμιοι	Διδάσκοντες	Δ/ Σ
Αυτό-αξιολογήσεις	Κλίμακα Αξιολόγησης	Διδάσκοντες/	Διδάσκοντες/ Διοικούντες	Δ/ Σ
Μαγνητοσκοπήσεις	Κλίμακα Αξιολόγησης	Διδάσκοντες/ Ομοβάθμιοι	Διδάσκοντες/ Ομοβάθμιοι	Δ/ Σ
Συνεντεύξεις φοιτητών	Ερωτηματολόγια	Φοιτητές	Διδάσκοντες/ Διοικούντες	Δ/ Σ
Βαθμολογίες από τους απόφοιτους	Κλίμακα Αξιολόγησης	Πτυχιούχοι	Διδάσκοντες/ Διοικούντες	Δ/ Σ /Π
Βαθμολογίες από τους εργοδότες	Κλίμακα Αξιολόγησης	Εργοδότες πτυχιούχων	Διδάσκοντες/ Διοικούντες	Π
Βαθμολογίες από διοικούντες	Κλίμακα Αξιολόγησης	Διοικούντες	Διοικούντες	Σ
Κατάρτιση διδασκαλίας	Κριτική επισκόπηση	Διδάσκοντες	Διοικητικοί	Σ
Βραβεία διδασκαλίας	Κριτική επισκόπηση	Διδάσκοντες	Επιτροπή Διδασκόντων/ Διοικούντων	Σ
Μετρήσεις των αποτελεσμάτων μάθησης	Έλεγχοι, ερευνητικές εργασίες, προσομοιώσεις	Φοιτητές	Διδάσκοντες/ Επιτροπή Προγραμμάτων	Δ/ Σ
Χαρτοφυλάκιο διδασκαλίας	Τα περισσότερα από τα παραπάνω	Διδάσκοντες, Φοιτητές, Ομοβάθμιοι	Επιτροπή Προαγωγών	Σ
Βαθμολογίες από εξωτερικούς ειδικούς	Κλίμακα Αξιολόγησης	Εξωτερικοί ειδικοί	Διδάσκοντες	Δ/Σ

Πηγή: Berk, A. R. (2006). *“Thirteen Strategies to Measure College Teaching”*, Sterling, Virginia: Stylus p: 13

4.2.1 Βαθμολογίες από τους φοιτητές

Οι βαθμολογίες από τους φοιτητές χρησιμοποιούνται στη λήψη αποφάσεων που δίνουν έμφαση στην αποτελεσματικότητα της διδασκαλίας και είναι σχετικές με την προαγωγή και τη μονιμότητα του εκπαιδευτικού στην τριτοβάθμια εκπαίδευση (Emery, Kramer & Tian, 2003). Προηγούμενες μετρήσεις δείχνουν ότι το 88% των liberal arts κολλεγίων (ΗΠΑ) χρησιμοποίησαν τις βαθμολογίες από τους φοιτητές στη διαμόρφωση αποφάσεων συγκριτικού (Σ) τύπου (Seldin, 1999a). Εν τούτοις αυτές έχουν δεχτεί την εχθρότητα και τον κυνισμό αρκετών διδασκόντων (Franklin &

Theall, 1989; Nasser & Fresko, 2002; Schmelkin-Pedhazur, Spencer, & Gellman, 1997). Ένα μεγάλο ποσοστό των διδασκόντων σε όλα τα επιστημονικά πεδία εκθέτουν μέτρια θετική στάση απέναντι στην εγκυρότητα των αξιολογήσεων από τους φοιτητές και τη χρησιμότητά τους στη βελτίωση της διδασκαλίας. Ωστόσο δεν υπάρχει ομοφωνία (Nasser & Fresko, 2002).

Στην τριτοβάθμια εκπαίδευση η έρευνα για τις βαθμολογίες από τους φοιτητές υπερέρχει ποσοτικά από οποιοδήποτε άλλο θέμα (Theall & Franklin, 1990). Περισσότερα από 2000 άρθρα έχουν εκδοθεί τα τελευταία 60 χρόνια (Cashin, 1999; McKeachie & Kaplan, 1996). Αν και υπάρχει ακόμη ένα ευρύ φάσμα απόψεων για την αξία τους ο McKeachie (1997) επισήμανε ότι «οι βαθμολογίες από τους φοιτητές αποτελούν τη μοναδική περισσότερο έγκυρη πηγή δεδομένων για την αποτελεσματικότητα της διδασκαλίας» (σελ. 1219). Στην πραγματικότητα υπάρχει λίγη απόδειξη για την εγκυρότητα οποιασδήποτε από τις άλλες πηγές δεδομένων (Marsh & Roche, 1997). Οι ειδικοί για την αξιολόγηση των διδασκόντων φαίνεται να συμφωνούν στο ότι οι βαθμολογίες από τους φοιτητές αποτελούν μία εξαιρετική πηγή απόδειξης τόσο για τις διαμορφωτικού όσο και για τις συγκριτικού τύπου αποφάσεις, στις οποίες (συγκριτικές) πρέπει να χρησιμοποιούνται κι άλλες πηγές (Arreola, 2000; Braskamp & Ory, 1994; Cashin, 1989, 1990; Centra, 1999; Seldin, 1999a). [Υπάρχουν αρκετές αναφορές για τις βαθμολογίες των φοιτητών που μπορούν να προκαλέσουν εξέγερση (Berk, 2006, σελ. 19)] (βλέπε Aleamoni, 1999; Cashin, 1999; d'Apollonia & Abrami, 1997; Eiszler, 2002; Emery et al., 2003; Greenwald, 1997; Greenwald & Gilmore, 1997; Greimel-Fuhrmann & Geyer, 2003; Havelka, Neal, & Beasley, 2003; Lewis, 2001; Millea & Grimes, 2002; Read, Rama, & Raghunandan, 2001; Shevlin, Banyard, Davies, & Griffiths, 2000; Sojka, Gupta, & Deeter-Schmelz, 2002; Sproule, 2002; Theall, Abrami, & Mets, 2001; Trinkaus, 2002; Wachtel, 1998).

Συμπέρασμα: Οι βαθμολογίες από τους φοιτητές αποτελούν μία αναγκαία πηγή απόδειξης για την αποτελεσματικότητα της διδασκαλίας τόσο για τις διαμορφωτικού όσο και για τις συγκριτικού τύπου αποφάσεις αλλά όχι όμως μία επαρκή πηγή για τις τελευταίες. Λαμβάνοντας υπόψη όλη την πολεμική ενάντια στην αξία τους, παραμένει μία σημαντική συνιστώσα κάθε συστήματος αξιολόγησης διδασκόντων στην τριτοβάθμια εκπαίδευση.

4.2.2 Βαθμολογίες από τους ομοβάθμιους

Σύμφωνα με τους Boyer (1990) και Rice (1991), διδασκαλία είναι ο τρόπος δια μέσου του οποίου η ανακαλυφθείσα, ενσωματωμένη και εφαρμοσμένη γνώση μεταφέρεται στην επόμενη γενιά μελετητών. Η διδασκαλία είναι μια ακαδημαϊκή δραστηριότητα.

Με σκοπό να προετοιμάσουν και να διδάξουν μία σειρά μαθημάτων, οι διδάσκοντες πρέπει να: ενημερώνονται από την βιβλιογραφία, αναπτύσσουν τους άξονες περιεχομένου, επιλέγουν τις πιο κατάλληλες πηγές για εκτύπωση και για μη, γράφουν και / ή να επιλέγουν σημειώσεις, ενσωματώνουν στη διδασκαλία τις σύγχρονες τεχνολογίες, σχεδιάζουν δραστηριότητες μάθησης, και κατασκευάζουν και να διαβαθμίζουν τις μετρήσεις αξιολόγησης.

Οι Webb και McEnerney (1995) υποστήριξαν ότι αυτά τα προϊόντα και οι δραστηριότητες μπορούν να είναι τόσο δημιουργικά και ακαδημαϊκά όσο και η πρωτότυπη έρευνα. Εάν η επίδοση στη διδασκαλία πρόκειται να αναγνωριστεί και να βραβευτεί πρέπει να υποβάλλεται στους ομοβάθμιους με την ίδια αυστηρή διαδικασία με την οποία υποβάλλεται ένα ερευνητικό χειρόγραφο πριν δημοσιευθεί σε ένα περιοδικό. Η αξιολόγηση από ομοβάθμιους ως μία εναλλακτική πηγή απόδειξης μοιάζει να ανεβαίνει τη σκάλα της αξιολόγησης, έτσι που περισσότερο από το 40% των liberal arts κολλεγίων (ΗΠΑ) να χρησιμοποιούν την παρατήρηση από αυτούς στη συγκριτική αξιολόγηση (Seldin, 1999a).

Η αξιολόγηση της διδασκαλίας από ομοβάθμιους περιλαμβάνει δύο δραστηριότητες:

- την παρατήρηση της επίδοσης της διδασκαλίας μέσα στην τάξη, και
- την αξιολόγηση των γραπτών ντοκουμέντων που χρησιμοποιούνται στο μάθημα.

Η παρατήρηση της επίδοσης της διδασκαλίας από τους ομοβάθμιους απαιτεί μία κλίμακα βαθμολόγησης η οποία να καλύπτει εκείνες τις πλευρές της διδασκαλίας που οι ομοβάθμιοι περισσότερο από τους φοιτητές, διαθέτουν τα προσόντα να τις αξιολογήσουν. Οι δηλώσεις της κλίμακας τυπικά εισάγουν στη συζήτηση (βλέπε Berk, Naumann, & Applling, 2004):

- τη γνώση περιεχομένου από τον διδάσκοντα,
- την παράδοση της διδασκαλίας,
- τις μεθόδους διδασκαλίας,
- τις δραστηριότητες μάθησης, και παρόμοια.

Οι βαθμολογίες μπορεί να καταγραφούν επί τόπου από ένα ή περισσότερους ομοβάθμιους για μία ή για περισσότερες περιπτώσεις μαθημάτων ή από μαγνητοσκοπημένες διδασκαλίες. Η αξιολόγηση

- των μέσων/υλικών διδασκαλίας

από ομοβάθμιους απαιτεί ένα διαφορετικό τύπο κλίμακας που να βαθμολογεί την ποιότητα των μαθημάτων, των σχεδίων διδασκαλίας, των εγχειριδίων, των εργασιών μελέτης / ανάθεσης, των φυλλαδίων, της εργασίας στο σπίτι και των tests / ερευνητικών εργασιών (projects). Μερικές φορές στις κλίμακες αξιολόγησης περιλαμβάνονται και

- οι συμπεριφορές διδασκαλίας, όπως είναι η αμεροληψία, οι πρακτικές βαθμολόγησης, η ηθική και ο επαγγελματισμός.

Ο Berk στο βιβλίο του (σελ. 20) χαρακτηρίζει αυτή την περίπτωση αξιολόγησης ως λιγότερο υποκειμενική αλλά και οικονομικά αποδοτική, επίσης περισσότερο αποτελεσματική και αξιόπιστη από ό, τι είναι οι παρατηρήσεις από τους ομοβάθμιους. Ωστόσο, οι παρατηρήσεις αποτελούν την πιο συχνή επιλογή επειδή αυτές παρέχουν άμεση αποτίμηση της πράξης της διδασκαλίας. Και οι δύο φόρμες της αξιολόγησης από τους ομοβάθμιους πρέπει να περιλαμβάνονται, όσο είναι δυνατό, σε ένα περιεκτικό σύστημα.

Παρά την ισχύουσα κατάσταση για την τεχνική αξιολόγησης από τους ομοβάθμιους υπάρχει αξιοπρόσεκτη αντίσταση από τους διδάσκοντες να την αποδεχτούν ως συμπλήρωμα της βαθμολογίας από τους φοιτητές. Η σχετική μη δημοτικότητα της οφείλεται στους παρακάτω 10 λόγους:

1. Οι παρατηρήσεις είναι προκατειλημμένες επειδή οι βαθμολογίες είναι προσωπικές και υποκειμενικές (*«η αξιολόγηση από ομοβάθμιους είναι τυφλού συστήματος και υποκειμενική»*).
2. Οι παρατηρήσεις είναι άδικες (*«οι αξιολογήσεις από ομοβάθμιους ενδέχεται επίσης να αποφέρει χαμηλή εσωτερική αξιοπιστία»*).
3. Ένας παρατηρητής είναι άδικός (*«οι αξιολογήσεις της προς δημοσίευση έρευνας γίνεται από δύο ή τρεις αξιολογητές»*).
4. Οι παρατηρήσεις στο μάθημα απαιτούν πολύ χρόνο (*«οι αξιολογήσεις από ομοβάθμιους μπορεί να είναι χρονοβόρες αλλά επαφίεται στη διακριτική ευχέρεια των αξιολογητών να συμβούν»*).
5. Μία ή δύο παρατηρήσεις δεν συνιστούν ένα αντιπροσωπευτικό δείγμα επίδοσης της διδασκαλίας για όλη τη σειρά των μαθημάτων.
6. Μόνο οι φοιτητές που παρακολουθούν ένα διδάσκοντα για πάνω από 40 ώρες επί του συνόλου των μαθημάτων μπορούν πραγματικά να αξιολογήσουν την επίδοση της διδασκαλίας.

7. Οι διαθέσιμες κλίμακες βαθμολογίας για ομοβάθμιους δεν μετρούν σημαντικά χαρακτηριστικά της αποτελεσματικότητας της διδασκαλίας.
8. Τα αποτελέσματα πιθανό να μην έχουν καμία επίδραση στη διδασκαλία.
9. Η διδασκαλία δεν εκτιμάται τόσο όσο η έρευνα, ιδιαίτερα σε μεγάλα, προσανατολισμένα στην έρευνα πανεπιστήμια.
10. Τα δεδομένα από την παρατήρηση είναι ακατάλληλα για αποφάσεις συγκριτικού τύπου από τα μέλη της διοίκησης.

Ο Berk (2006) επισημαίνει ότι οι περισσότεροι από αυτούς τους λόγους ή από αυτές τις αντιλήψεις είναι δικαιολογημένοι (legitimate) σύμφωνα με το πώς διάφορα ιδρύματα της τριτοβάθμιας εκπαίδευσης πραγματοποιούν ένα σύστημα αξιολόγησης από ομοβάθμιους. Ελάχιστα μπορούν να διορθωθούν για να ελαχιστοποιήσουν τις προκαταλήψεις και την αδικία και να βελτιώσουν την αντιπροσωπευτικότητα των παρατηρήσεων (σελ: 21).

Ωστόσο υφίσταται ομοφωνία από τους ειδικούς για το λόγο 10. Τα δεδομένα από τις παρατηρήσεις των ομοβάθμιων πρέπει να χρησιμοποιούνται περισσότερο για διαμορφωτικού τύπου αποφάσεις παρά για συγκριτικού τύπου (Aleamoni, 1982; Arreola, 2000; Centra, 1999; Cohen & McKeachie, 1980; Keig & Waggoner, 1995; Millis & Kaplan, 1995). Πράγματι, 60 χρόνια εμπειρίας από την αξιολόγηση από τους ομοβάθμιους στον στρατιωτικό και ιδιωτικό τομέα οδήγησαν στο ίδιο συμπέρασμα (Muchinsky, 1995). Οι εργαζόμενοι / απασχολούμενοι τείνουν να δέχονται τις παρατηρήσεις από τους ομοβάθμιους όταν τα αποτελέσματα χρησιμοποιούνται για *επικοινωνιακή διαγνωστική ανατροφοδότηση* αντί ως βάση για λήψη αποφάσεων διοικητικού τύπου (Cederblom & Lounsbury, 1980; Love, 1981).

Συμπέρασμα: Οι αξιολογήσεις από τους ομοβάθμιους σχετικά με την επίδοση της διδασκαλίας και των υλικών συνιστά την πλέον συμπληρωματική πηγή απόδειξης των βαθμολογήσεων από τους φοιτητές. Αυτή καλύπτει εκείνες τις πλευρές της διδασκαλίας που οι φοιτητές δεν είναι σε θέση να αποτιμήσουν. Οι βαθμολογίες των φοιτητών και των ομοβάθμιων, εξετάζοντάς τες μαζί, παρέχουν μια πολύ ολοκληρωμένη εικόνα της αποτελεσματικότητας της διδασκαλίας για τη βελτίωσή της. Οι βαθμολογίες από τους ομοβάθμιους δεν πρέπει να χρησιμοποιούνται για αποφάσεις προσωπικού (personnel decisions).

4.2.3 *Αυτό-αξιολόγηση*

Ελάχιστες μελέτες βρήκαν ότι οι διδάσκοντες βαθμολογούν τους εαυτούς τους υψηλότερα από (Centra, 1999), το ίδιο με (Bo-Linn, Gentry, Lowman, Pratt, and Zhu, 2004; Feldman, 1989), ή χαμηλότερα από (Bo – Linn et al., 2004) τις βαθμολογίες που δίνουν οι φοιτητές. Οι υψηλής βαθμίδας διδάσκοντες δίνουν στους εαυτούς τους υψηλούς βαθμούς από ό,τι οι χαμηλότερης βαθμίδας (Doyle & Crichton, 1978; Marsh, Overall, & Kesler, 1979). Οι ανώτερης βαθμίδας (superior) καθηγητές παρέχουν πιο ακριβείς αυτό-αξιολογήσεις από τους μέτριας βαθμίδας (mediocre) (Centra, 1973; Sorey, 1968). Ανάμεσα στους κοσμήτορες στα κολλέγια (ΗΠΑ) των liberal arts, οι 59% πάντοτε περιλαμβάνουν τις αυτό-αξιολογήσεις σε αποφάσεις συγκριτικού τύπου (Seldin, 1999a).

Η έκθεση δράσης της σχολής είναι ο πιο κοινός τύπος της αυτό-αξιολόγησης. Περιγράφει τη διδασκαλία, την κατάρτιση, την υπηρεσία και τις πρακτικές δραστηριότητες του προηγούμενου χρόνου. Αυτή η πληροφορία χρησιμοποιείται από τα μέλη της διοίκησης με σκοπό να δικαιολογούν τις αποφάσεις τους. Αυτή η ετήσια έκθεση ωστόσο δεν είναι μία αληθής αυτό-αξιολόγηση για την αποτελεσματικότητα της διδασκαλίας (Berk, 2006).

Ο Seldin (1999b) προτείνει μία δομημένη φόρμα που να εκθέτει τα αντικείμενα διδασκαλίας του καθηγητή, τις δραστηριότητες, τα επιτεύγματα αλλά και τις αποτυχίες. Κατευθυντήριες ερωτήσεις προτείνονται σε τομείς όπως είναι η προσέγγιση της τάξης, η επικοινωνία διδάσκοντα-φοιτητή, η γνώση του επιστημονικού πεδίου, η οργάνωση και ο σχεδιασμός των μαθημάτων και ερωτήσεις σχετικές με τη διδασκαλία. Ο Wergin (1992) και οι Braskamp και Ory (1994) προτείνουν τη συλλογή κι άλλων τύπων απόδειξης.

Ο διδάσκοντας μπορεί επίσης να συμπληρώσει την κλίμακα βαθμολόγησης που συμπληρώνουν οι φοιτητές σύμφωνα με δύο οπτικές που αφορούν: η πρώτη, στην άμεση μέτρηση της επίδοσης της διδασκαλίας του /ή της και η δεύτερη, στις προβλεπόμενες βαθμολογήσεις από τους φοιτητές. Ανακολουθίες ανάμεσα στις πηγές αυτής της τριάδας – βαθμολογίες από τους φοιτητές, αυτό-αξιολογήσεις διδάσκοντα, και αντιλήψεις διδάσκοντα για τις βαθμολογίες των φοιτητών – μπορούν να παρέχουν αξιοσημείωτες γνώσεις πάνω στην αποτελεσματικότητα της διδασκαλίας. Τα

αποτελέσματα μπορεί να είναι πολύ χρήσιμα για τη στόχευση συγκεκριμένων περιοχών που χρήζουν βελτίωσης. Επισημαίνεται ότι οι αξιολογήσεις από τους φοιτητές και οι αυτό - αξιολογήσεις τείνουν να αποδίδουν χαμηλές θετικές συσχετίσεις (Braskamp, Caulley, & Costin, 1979; Feldman, 1989).

Η τριάδα των βαθμολογήσεων μπορεί να αποδειχτεί χρήσιμη σε αποφάσεις διαμορφωτικού τύπου, αλλά μία μαγνητοσκόπηση με σκοπό την αξιολόγηση της επίδοσης της διδασκαλίας κάποιου μπορεί να είναι ακόμη πιο ενημερωτική ως μία πηγή αποδεικτικών στοιχείων αυτό – αξιολόγησης.

Η αυτό-αξιολόγηση ενός διδάσκοντα αποδεικνύει τη γνώση του ή της σχετικά με τη διδασκαλία και την αντιλαμβανόμενη αποτελεσματικότητα στην τάξη (Cranton, 2001). Αυτή η πληροφορία θα πρέπει να αξιολογηθεί κριτικά και να συγκριθεί με τις άλλες πηγές απόδειξης στις αποφάσεις προσωπικού. Το διαγνωστικό πορτραίτο μπορεί να χρησιμοποιηθεί για να καθοδηγήσει τη βελτίωση της διδασκαλίας.

Συμπέρασμα: Η αυτό – αξιολόγηση είναι μία σημαντική πηγή απόδειξης στη διαμόρφωση διαμορφωτικών και συγκριτικών αποφάσεων. Οι εισροές των διδασκόντων για τη διδασκαλία τους συμπληρώνουν τον τριγωνισμό (triangulation) των τριών άμεσων πηγών παρατήρησης της αποτελεσματικότητας της διδασκαλίας: τους φοιτητές, τους ομοβάθμιους και τους ίδιους.

4.2.4 Μαγνητοσκοπήσεις

Ποιος μπορεί να αξιολογήσει μία μαγνητοσκόπηση;

1. Ο ίδιος ο διδάσκων, ιδιωτικά στο γραφείο του.
2. Ο ίδιος συμπληρώνει την κλίμακα που δίνεται στους ομοβάθμιους, η οποία αφορά στην παρατήρηση των συμπεριφορών διδασκαλίας.
3. Ένας ομοβάθμιος συμπληρώνει την κλίμακα και παρέχει ανατροφοδότηση.
4. Δύο ή τρεις ομοβάθμιοι συμπληρώνουν την κλίμακα και παρέχουν ανατροφοδότηση.
5. MTV, VH-1, ή BET.

Οι Braskamp και Ory (1994) και ο Seldin (1999b) υποστηρίζουν τις αρετές της μαγνητοσκόπησης για τη βελτίωση της διδασκαλίας. Ωστόσο, υπάρχει μικρή απόδειξη για την αποτελεσματικότητά της. Εάν ο σκοπός της μαγνητοσκόπησης είναι να διαγνώσει τα δυνατά και τα αδύνατα σημεία σε μία ή περισσότερες περιπτώσεις διδασκαλίας τότε οι διδάσκοντες πρέπει να ενθαρρύνονται να αξιολογούν συστηματικά τις παρατηρούμενες συμπεριφορές χρησιμοποιώντας μία κλίμακα αξιολόγησης ή μία κατάσταση ελέγχου (Seldin, 1998). Οι καταστάσεις ελέγχου έχουν αναπτυχτεί από τον Brinko (1993) και τον Perlberg (1983). Αυτές μπορούν και

επικεντρώνουν την ανατροφοδότησή τους πάνω στο τι χρειάζεται να αλλάξει. Ωστόσο, ένας επιδέξιος ομοβάθμιος, σεβαστός μέντορας, ή σύμβουλος μπορεί να παρέχει ανατροφοδότηση εμπιστοσύνης, που θα μπορούσε να είναι ακόμα πιο χρήσιμη στο διδάσκοντα (Braskamp & Ory, 1994).

Οποιαδήποτε δυνατότητα επιλεγεί, το αποτέλεσμα της μαγνητοσκόπησης πρέπει να είναι ένα προφίλ θετικής και αρνητικής διδασκαλικής συμπεριφοράς συνοδευόμενης από ένα κατάλογο συγκεκριμένων στόχων που θα συζητηθούν για την αντιμετώπιση των ελλείψεων. Αυτή η άμεση απόδειξη της αποτελεσματικότητας της διδασκαλίας μπορεί να συμπεριληφθεί στην αυτό-αξιολόγηση του διδάσκοντα και στο χαρτοφυλάκιο διδασκαλίας του. Η μαγνητοσκόπηση αποτελεί ένα ισχυρό έγγραφο (documentary) της επίδοσης της διδασκαλίας.

Συμπέρασμα: Εάν οι διδάσκοντες είναι πράγματι συνειδητοποιημένοι απέναντι στη βελτίωση της διδασκαλίας τους, η μαγνητοσκόπηση αποτελεί μία από τις καλύτερες πηγές απόδειξης για αποφάσεις διαμορφωτικές, ερμηνευμένες από είτε μόνο τους διδάσκοντες ή, το προτιμότερο, με την εισροή ενός ομοβάθμιου. Εάν η μαγνητοσκόπηση χρησιμοποιείται με εμπιστοσύνη γι' αυτό το σκοπό, οι διδάσκοντες πρέπει να αποφασίσουν εάν θα την συμπεριλάβουν στην αυτό-αξιολόγησή τους ή στο χαρτοφυλάκιο ως ένα «δείγμα εργασίας» για τις συγκριτικές αποφάσεις.

4.2.5 Συνεντεύξεις φοιτητών

Οι συνεντεύξεις ομάδας φοιτητών χορηγούν απόδειξη την οποία οι διδάσκοντες βαθμολογούν ως περισσότερο ακριβή, αξιόπιστη, χρήσιμη, πλήρη και πιστευτή από τις βαθμολογίες που δίνουν οι φοιτητές και τα γραπτά σχόλιά τους (Braskamp & Ory, 1994), αν και η πληροφορία που συλλέγεται από τις τρεις άλλες πηγές είναι ιδιαίτερα συγκλίνουσα (Braskamp, Ory, & Pieper, 1981). Οι διδάσκοντες θεωρούν τα αποτελέσματα των συνεντεύξεων ως περισσότερο χρήσιμα στη βελτίωση της διδασκαλίας αλλά μπορεί επίσης να είναι άξια σημασίας στην προώθηση αποφάσεων (Ory & Braskamp, 1981).

Υπάρχουν τρεις τύποι συνεντεύξεων που οι Braskamp και Ory (1994) συνιστούν: 1) οι κύκλοι ελέγχου ποιότητας, 2) οι συνεντεύξεις ομάδας στην τάξη, και 3) οι συνεντεύξεις αποφοίτων. Ο πρώτος τύπος συνέντευξης προέρχεται από μία τεχνική διοίκησης (management) που χρησιμοποιείται στη βιομηχανία στην Ιαπωνία και ονομάζεται *κύκλοι ελέγχου ποιότητας* (*quality control circles*) (Shariff, 1999; Weimer, 1990), κατά τις οποίες δίνονται ευκαιρίες σε ομάδες απασχολούμενων να συμμετέχουν στη διαμόρφωση αποφάσεων για την εταιρεία.

Η εκδοχή του «κύκλου» για τη διδασκαλία περιλαμβάνει τη συγκρότηση μίας ομάδας εθελοντών φοιτητών οι οποίοι συναντώνται κανονικά (εβδομαδιαίως) με σκοπό να διατυπώνουν κριτική στη διδασκαλία να ελέγχουν τις στρατηγικές να επισημαίνουν περιοχές με προβλήματα και να διατυπώνουν προτάσεις για βελτίωση. Αυτές οι συναντήσεις στις οποίες ηγείται ο διδάσκων, ισχυροποιούν τη λογοδοσία (accountability) για καθετί που συμβαίνει στην τάξη. Ωστόσο, οι διδάσκοντες πρέπει να είναι ανοικτοί στα σχόλια των φοιτητών και να είναι πρόθυμοι να κάνουν τις αναγκαίες προσαρμογές για βελτίωση. Αυτή η τεχνική διαμορφωτικής αξιολόγησης επιτρέπει την ανατροφοδότηση από τους φοιτητές και τη συστηματική αλλαγή της διδασκαλίας κατά τη διάρκεια των μαθημάτων.

Οι συνεντεύξεις ομάδας στην τάξη περιλαμβάνουν όλη την τάξη, διεξάγονται από ένα άλλο άτομο, συνήθως ένα συνεργάτη στον ίδιο τμήμα, ένα μεταπτυχιακό τεχνικής βοήθειας (TA), ή ένα μέλος ΔΕΠ προγραμματιστή. Αυτός που κάνει τη συνέντευξη χρησιμοποιεί ένα δομημένο ερωτηματολόγιο για να «ψηλαφίσει» τα δυνατά και τα αδύνατα σημεία των μαθημάτων και των δραστηριοτήτων της διδασκαλίας. Μερικές από τις ερωτήσεις πρέπει να δίνουν τη δυνατότητα απόσπασης ενός ευρέος φάσματος απόψεων των φοιτητών για το μάθημα. Η πληροφορία που συλλέγεται κοινοποιείται στον διδάσκοντα για να χρησιμοποιηθεί στη βελτίωση της διδασκαλίας του αλλά και ως πηγή απόδειξης για συγκριτικού τύπου αποφάσεις.

Οι συνεντεύξεις από τον πτυχιούχο μπορούν να γίνουν είτε ατομικά ή σε ομάδα από τους διδάσκοντες, τα μέλη της διοίκησης, ή το προσωπικό που παρέχει υπηρεσίες. Δοσμένου του χρόνου που απαιτείται, οι ερωτήσεις πρέπει να επικεντρώνονται σε πληροφορίες που δεν συγκεντρώθηκαν με την κλίμακα βαθμολόγησης. Για παράδειγμα, οι δηλώσεις σε μία συνέντευξη ομάδας πρέπει να επικεντρώνονται στα πιο χρήσιμα μαθήματα, στα λιγότερο χρήσιμα μαθήματα, στους καλύτερους διδάσκοντες, στα κενά περιεχομένου, στην ποιότητα της διδασκαλίας, στην παροχή συμβουλών ποιότητας και στα σχέδια αποφοίτησης. Οι απαντήσεις των φοιτητών μπορεί να καταγράφονται κατά τη συνέντευξη ή μπορεί να ζητηθούν ως ανώνυμες γραπτές παρατηρήσεις για το πρόγραμμα. Τα αποτελέσματα πρέπει να προωθούνται στους κατάλληλους διδάσκοντες, στις επιτροπές αναλυτικών προγραμμάτων και στα μέλη της διοίκησης. Με βάση την ιδιαιτερότητα της πληροφορίας που συλλέχτηκε,

αυτή η απόδειξη μπορεί να χρησιμοποιηθεί για διαμορφωτική ανατροφοδότηση αλλά και για συγκριτικές αποφάσεις.

Συμπέρασμα: Ο κύκλος ελέγχου ποιότητας αποτελεί μία εξαιρετική τεχνική για να παρέχουν οι φοιτητές σταθερή ανατροφοδότηση στη βελτίωση της διδασκαλίας. Η *συνέντευξη ομάδας* ως μία ανεξάρτητη αξιολόγηση μπορεί να παρέχει αρκετή πληροφορία που στηρίζει τις βαθμολογίες από τους φοιτητές. Οι *συνεντεύξεις εξόδου* μπορεί να μην είναι πρακτικές στη διεξαγωγή ή περιττές με τις βαθμολογίες εξόδου, που περιγράφονται στη συνέχεια.

4.2.6 Βαθμολογίες εξόδου και από τους απόφοιτους

Μελέτες όπως εκείνες των Overall και Marsh (1980) και του Feldman (1989) δείχνουν ότι οι απόφοιτοι διατηρούν ένα υψηλό επίπεδο λεπτομερειών σχετικών με τις εμπειρίες τους από τα μαθήματα (Kulik, 2001).

Για ποιο πράγμα θα ρωτηθούν οι απόφοιτοι; Μία κλίμακα βαθμολόγησης ένα, πέντε, και δέκα χρόνια αργότερα μπορεί να παρέχει νέα πληροφορία για την ποιότητα της διδασκαλίας, τη χρησιμότητα των απαιτήσεων των μαθημάτων, την επίτευξη των αποτελεσμάτων του προγράμματος, την αποτελεσματικότητα των διαδικασιών εισδοχής, την προετοιμασία για μεταπτυχιακές σπουδές, την προετοιμασία για τον πραγματικό κόσμο και μία ποικιλία από άλλα θέματα που δεν μετρώνται με τις κλίμακες βαθμολόγησης από τους φοιτητές. Αυτή η στερεοσκοπική αξιολόγηση μπορεί να εκμαιεύσει αξιοσημείωτο αναστοχασμό για τις μεθόδους διδασκαλίας, τις απαιτήσεις των μαθημάτων, τις τεχνικές αξιολόγησης, την ενσωμάτωση των τεχνολογιών, το άνοιγμα στη διαφορετικότητα και για άλλα θέματα σχετικά με τα μαθήματα ή για το πρόγραμμα στο σύνολό του (Berk, 2006, σελ. 28). Οι μη δομημένες απαντήσεις μπορεί να φωτίσουν ιδιαίτερα δυνατά σημεία των διδασκόντων καθώς και να παρέχουν κατευθύνσεις για βελτίωση. Οι Hamilton, Smith, Heady, και Carson (1995) παρουσίασαν τα αποτελέσματα μίας μελέτης με ανοικτού – κλειστού τύπου ερωτήσεις σε αποφοίτους σε έρευνες εξόδου. Ο αναστοχασμός απεδείχθη χρήσιμος τόσο στους διδάσκοντες όσο και στα μέλη της διοίκησης. Αυτός ο τύπος έρευνας μπορεί επίσης να αξιοποιήσει τις πληροφορίες σε θέματα όπως είναι το περιεχόμενο του αναλυτικού προγράμματος σπουδών και τα επακόλουθα, ο προγραμματισμός των μαθημάτων, και τα μέσα διδασκαλίας. Ακόμα μπορεί σε ένα πιο γενικό επίπεδο να υπάρξει ιδιαίτερα χρήσιμος ως μία άλλη πηγή απόδειξης της ποιότητας της διδασκαλίας.

Συμπέρασμα: Αν και οι βαθμολογίες εξόδου και αποφοίτων είναι παρόμοιες με τις αυθεντικές βαθμολογίες των φοιτητών για την ίδια κλίμακα, διαφορετικές δηλώσεις της κλίμακας σχετικά με την ποιότητα της διδασκαλίας, τη σειρά των μαθημάτων, τις

εισαγωγές προγραμμάτων σπουδών και άλλα θέματα μπορούν να παρέχουν νέα πληροφορία. Οι βαθμολογίες των αποφοίτων πρέπει να θεωρηθούν ως μία άλλη σημαντική πηγή απόδειξης για την αποτελεσματικότητα της διδασκαλίας.

4.2.7 Βαθμολογίες από τον εργοδότη

Μετά από το πέρασμα το λιγότερο ενός χρόνου, μία αξιολόγηση, για παράδειγμα, της επίδοσης των πτυχιούχων στην εργασία τους μπορεί να παρέχει αναστοχασμό επί της συνολικής ποιότητας της διδασκαλίας, της σημασίας του προγράμματος σπουδών και του σχεδιασμού αυτού του προγράμματος. Τοιουτοτρόπως, μπορούν να αντληθούν συμπεράσματα που εξαρτώνται από την εξειδίκευση των αποτελεσμάτων και είναι σχετικά με την ατομική αποτελεσματικότητα του διδάσκοντα. Ωστόσο, αυτή η μέτρηση είναι περιορισμένη επειδή είναι έμμεση και βασίζεται στα αποτελέσματα του προγράμματος.

Το πρώτο βήμα είναι να εντοπιστούν οι απόφοιτοι. Το γραφείο εισαγωγής (εγγραφών) συνήθως διατηρεί εγγραφές εργασιακής απασχόλησης για λίγα χρόνια μετά την αποφοίτηση. Ο Seppanen (1995) προτείνει τη χρήση των βάσεων δεδομένων ανεργίας για την παρακολούθηση του ιστορικού της απασχόλησης των πτυχιούχων η οποία μπορεί να συνδεθεί άμεσα με τα συστήματα πληροφόρησης των σχολών.

Το επόμενο βήμα αφορά στην απόφαση για το ποιες συμπεριφορές θα μετρηθούν. Τα αποτελέσματα του προγράμματος μπορούν να χρησιμοποιηθούν όταν η σχολή προετοιμάζει τον πτυχιούχο για ένα ειδικό επάγγελμα, όπως είναι η διδασκαλία, η περίθαλψη, η λογιστική, η μηχανολογία, το ποδόσφαιρο, ή η κατασκοπεία. Αυτά τα αποτελέσματα μαζί με ερωτήσεις σχετικές με την ικανοποίηση του εργοδότη από την επίδοση του εργαζόμενου μπορούν να συναρμολογηθούν σε μία κλίμακα βαθμολόγησης για να καθορίσουν την ποιότητα της γνώσης του/ή της, των δεξιοτήτων και των ικανοτήτων με βάση την επίδοσή του. Οι βαθμολογίες μπορούν να επισημάνουν στους διδάσκοντες, τη σειρά των μαθημάτων και τα δυνατά σημεία του προγράμματος καθώς και τις αδυναμίες σε σχέση με την επίδοση στην εργασία.

Συμπέρασμα: Οι βαθμολογίες του εργοδότη παρέχουν μία έμμεση πηγή απόδειξης για αποφάσεις αξιολόγησης του προγράμματος σχετικές με την αποτελεσματικότητα της διδασκαλίας και την επίτευξη των αποτελεσμάτων του προγράμματος, ειδικά για τις επαγγελματικές σχολές. Τα δεδομένα από την επίδοση στην εργασία μπορούν να συνδεθούν με την ατομική επίδοση στη διδασκαλία αλλά σε μία πολύ περιορισμένη βάση.

4.2.8 Βαθμολογίες από τα μέλη της διοίκησης

Αναπληρωτές κοσμήτορες, υπεύθυνοι προγράμματος, ή πρόεδροι τμημάτων μπορούν να αξιολογήσουν τα μέλη ΔΕΠ στην ετήσια επανεξέταση προσόντων σύμφωνα με τα κριτήρια για τη διδασκαλία, την κατάρτιση, τις υπηρεσίες και / ή την πρακτική (Diamond, 2004). Ούτως ή άλλως, αυτοί υπήρξαν ή συνεχίζουν να είναι μέλη ΔΕΠ με εξειδίκευση στις μεθόδους διδασκαλίας, στην αξιολόγηση των τεχνικών που εφαρμόζονται στην τάξη και του περιεχομένου των επιστημονικών πεδίων. Το μέλος της διοίκησης μπορεί να παρακολουθεί την αποτελεσματικότητα της διδασκαλίας και να εξετάζει τα διαπιστευτήρια σε άλλες τρεις περιοχές που εκπονήθηκαν από κάθε μέλος ΔΕΠ. Οι βαθμολογίες στην επίδοση της διδασκαλίας από τα μέλη της διοίκησης απαιτούν μία πιο γενική κλίμακα από εκείνο τον τύπο κλίμακας που συμπληρώνουν οι φοιτητές.

Τυπικά, μία δομημένη έκθεση δραστηριοτήτων διαμοιράζεται σε όλα τα μέλη ΔΕΠ για να παρέχει μία πλήρη εικόνα των επιτευγμάτων του παρελθόντος έτους σε όλες τις περιοχές. Όσο πιο σαφείς οι κατηγορίες που ζητούνται στην έκθεση τόσο πιο εύκολο είναι για τα μέλη ΔΕΠ να την συμπληρώνουν και για τους διοικούντες να την αξιολογούν. Οι τελευταίοι τότε μπορούν να βαθμολογούν τη συνολική ποιότητα της επίδοσης σε κάθε κατηγορία. Η συνολική βαθμολογία κατά μήκος των κατηγοριών μπορεί να κλιμακωθεί για να καθορίσει τις αυξήσεις μισθών (Berk, 2006, σελ. 29-30).

Συμπέρασμα: Οι βαθμολογίες από τους διοικούντες τυπικά βασίζονται σε δεύτερου τύπου πηγές, όχι στην άμεση παρατήρηση της διδασκαλίας ή σε άλλες περιοχές επίδοσης. Αυτή η πηγή παρέχει μία προοπτική διαφορετική από τις άλλες η οποία αφορά στη λήψη αποφάσεων που σχετίζονται με μισθούς και προαγωγή.

4.2.9 Κατάρτιση στη διδασκαλία

Η κατάρτιση στη διδασκαλία και στη μάθηση σύμφωνα με την Carnegie Academy for the Scholarship of Teaching and Learning (CASTL), είναι «ένας δημόσιος απολογισμός για κάποιο μέρος ή για όλη την πράξη της διδασκαλίας – όραμα, σχεδιασμός, θέσπιση, αποτελέσματα, και ανάλυση – με ένα τρόπο που επιδέχεται την κριτική θεώρηση από τους επαγγελματίες συναδέλφους του εκπαιδευτικού και είναι επιδεκτικός παραγωγικής απασχόλησης σε μελλοντικές εργασίες από τα μέλη της ίδιας κοινότητας» (Shulman, 1998, p. 6). [Μετάφραση: Συμβάλλει σε ένα αυξανόμενο σώμα γνώσης σχετικά με τη διδασκαλία και τη μάθηση στην τριτοβάθμια εκπαίδευση παρουσιάζοντας την σε συνέδρια διδασκαλίας και μάθησης και δημοσιεύοντας την σε αντίστοιχα περιοδικά]. Παρουσιάσεις και εκδόσεις πάνω στη

διδασκαλία και στη μάθηση για καινοτόμες τεχνικές διδασκαλίας και σχετικά θέματα είναι δείκτες της διδασκαλικής εμπειρίας. Έρευνα πάνω σε σημαντικά ερωτήματα για τη διδασκαλία και τη μάθηση δεν μπορούν μόνο να βελτιώνουν την αποτελεσματικότητα ενός μέλους ΔΕΠ στο μάθημα του ή της αλλά να βελτιώνουν και την πρακτική συνέπειά της (Hutchings & Shulman, 1999). Η απόδειξη της κατάρτισης στη διδασκαλία μπορεί να συμπεριλαμβάνει παρουσιάσεις πάνω σε νέες μεθόδους διδασκαλίας, όπως είναι η έρευνα, τα εργαστήρια και κεντρικές ομιλίες σε ιδρύματα και συνέδρια.

Συμπέρασμα: Η κατάρτιση στη διδασκαλία είναι μία σημαντική πηγή απόδειξης που μπορεί να υποστηρίξει τις τρεις κυρίαρχες άμεσης παρατήρησης πηγές. Αυτή μπορεί εύκολα να διαχωρίσει τον «καθηγητή υπότροφο» και τα πολύ δημιουργικά μέλη ΔΕΠ από όλους τους άλλους κατά τη λήψη συγκεντρωτικών αποφάσεων.

4.2.10 Βραβεία διδασκαλίας

Το θέμα σε αυτή την περίπτωση είναι η διαδικασία με την οποία καθορίζεται το βραβείο. Ένα μέλος ΔΕΠ που προτάθηκε ως υποψήφιος/-α για οποιοδήποτε βραβείο πρέπει να υποστεί μία εξουθενωτική αξιολόγηση από ένα πάνελ κριτών σύμφωνα με τα κριτήρια της υποδειγματικής διδασκαλίας. Η απόδειξη για την αποτελεσματικότητα της διδασκαλίας θα ήταν περιορισμένη αν αξιοποιούντο μόνο τα κριτήρια βράβευσης, η έκθεση και οι υποψηφιότητες.

Αξιολογήσεις στη δεκαετία του '90 δείχνουν ότι περίπου το 70% των δύο-ετών κολλεγίων και των liberal arts ιδρυμάτων και το 96% των ερευνητικών πανεπιστημίων στις ΗΠΑ είχαν βραβεία ή προγράμματα με τα οποία τίμησαν την υποδειγματική διδασκαλία (Jenrette & Hayes, 1996; Zahorski, 1996).

Η βιβλιογραφία σχετικά με την αξία των βραβείων διδασκαλίας ως ενός κινήτρου για τη βελτίωση της διδασκαλίας είναι αραιή (Carusetta, 2001), αλλά διατρέχει μία κλίμακα από το ναι (Seldin & Associates, 1999; Wright & Associates, 1995) στο όχι (McNaught & Anwyl, 1993; Ruedrich, Cavey, Katz, & Grush, 1992; Zahorski, 1996). Έχει υπάρξει αξιοσημείωτη κριτική σχετικά με την επιλογή της πορείας αξιολόγησης που τείνει να είναι ακατάστατη, ασαφής, ύποπτη και υποκειμενική (Knapper, 1997; Menges, 1996; Weimer, 1990).

Συμπέρασμα: Ως μία πηγή απόδειξης για την αποτελεσματικότητα της διδασκαλίας, στην καλύτερη περίπτωση, τα βραβεία διδασκαλίας παρέχουν αξιοσημείωτη πληροφορία μόνο για τους υποψήφιους και στη χειρότερη περίπτωση, αυτά παρέχουν μη ακριβή και μη αξιόπιστη ανάδραση σε αμφισβητήσιμους υποψήφιους που μπορεί να έχουν εμφανιστεί δια του Νόμου και της Τάξης. Οι αρετές των βραβείων

διδασκαλίας πρέπει να αποτιμώνται στο πλαίσιο ενός δικτύου κινήτρων και βραβείων του ιδρύματος για τη διδασκαλία.

4.2.11 Μετρήσεις για τα αποτελέσματα μάθησης

Οι περισσότερες από τις προαναφερόμενες πηγές απόδειξης περιλαμβάνουν άμεσες βαθμολογίες της διδασκαλικής συμπεριφοράς. Οι μετρήσεις αποτελεσμάτων μάθησης είναι μία «κολλώδης» (sticky) πηγή επειδή δεν είναι άμεση. Η επίδοση της διδασκαλίας συμπεραίνεται από την επίδοση των φοιτητών – τι αυτοί μαθαίνουν κατά τη διάρκεια μιας σειράς μαθημάτων. Οι Theall και Franklin (2001) σημείωσαν σταθερά υψηλές συσχετίσεις μεταξύ των αξιολογήσεων από τους φοιτητές για το «πόσο έμαθαν» και της συνολική βαθμολογίας. Επιπλέον, υπάρχουν σαφείς συσχετίσεις μεταξύ των βαθμολογιών από τους φοιτητές και των επιδόσεών τους στις τελικές εξετάσεις (Cohen, 1981).

Παρά την ύπαρξη αυτών των σχέσεων, η θεμελίωση της επίδοσης των φοιτητών με βάση τα αποτελέσματα μάθησης ως μίας ανεξάρτητης, έγκυρης μέτρησης της αποτελεσματικότητας της διδασκαλίας είναι γεμάτη με πολλές δυσκολίες (Berk, 2006, σελ. 33). Η ουσία του προβλήματος είναι η απομόνωση της διδασκαλίας ως της μόνης εξήγησης για τη μάθηση του φοιτητή. Η επίδοση στα τεστ, στις ερευνητικές εργασίες (projects), στις εκθέσεις και σε άλλους δείκτες των μαθημάτων μπορεί να επηρεάζονται από τα χαρακτηριστικά των φοιτητών, το ίδρυμα, και από τις ίδιες τις μετρήσεις αποτελεσμάτων, πάνω στα οποία τα μέλη ΔΕΠ δεν έχουν κανένα έλεγχο (βλέπε αναλυτική παρουσίαση στο τμήμα 5: Ζητήματα Μεθοδολογίας).

Η αποτελεσματικότητα της διδασκαλίας αξιολογείται με όρους παραγωγικότητας του φοιτητή. Αυτό σημαίνει, ότι αυτή βασίζεται στα αποτελέσματα. Ως εκ τούτου, εάν η επίδοση ένας εργαζόμενου σε εργοστάσιο μπορεί να μετρηθεί από το πόσα μαραφέτια αυτός ή αυτή παράγει σε ένα συγκεκριμένο χρονικό διάστημα, γιατί να μην αξιολογηθεί το μέλος ΔΕΠ από την παραγωγικότητα ή την επιτυχία των φοιτητών του ή της με βάση τις μετρήσεις των αποτελεσμάτων; Τα επιχειρήματα για αυτή την αναλογία παραγωγικότητας εργοστασιακού εργάτη - καθηγητή αντλούνται από τις αρχές ενός συστήματος αποζημίωσης, της *αμοιβής με το τεμάχιο* (piece-rate compensation system) (Murnane & Cohen, 1986). Οι συμβάσεις *αμοιβής με το τεμάχιο* αποτελούν την πιο κοινή φόρμα «πληρωμής με βάση τα αποτελέσματα» (Pencavel, 1997). Αυτές οι συμβάσεις παρέχουν ένα ισχυρό κίνητρο για τους

εργαζόμενους να παράγουν, επειδή η υψηλή παραγωγικότητα καταλήγει σε άμεση αμοιβή. Όταν το σύστημα αυτό εφαρμόζεται στη διδασκαλία καταρρέει για δύο λόγους. Πρώτον, ένας εργαζόμενος σε εργοστάσιο χρησιμοποιεί τα ίδια υλικά για να κάνει κάθε προϊόν. Το μέλος ΔΕΠ εργάζεται με φοιτητές των οποίων τα χαρακτηριστικά ποικίλουν αξιοσημείωτα από μάθημα σε μάθημα. Δεύτερον, τα χαρακτηριστικά των υλικών που χρησιμοποιεί ο εργαζόμενος στο εργοστάσιο σπάνια επηρεάζουν τις δεξιότητές του/ ή της και το ρυθμό παραγωγής. Αυτό σημαίνει ότι η ποσότητα και η ποιότητα της παραγωγής μαραφετιών μπορεί να αποδοθεί εξολοκλήρου στον εργαζόμενο. Τα χαρακτηριστικά-κλειδιά των φοιτητών, όπως είναι η ικανότητα, οι στάσεις, τα κίνητρα, η ηλικία, το φύλο, και η ωριμότητα αλλά και του ιδρύματος, όπως είναι ο αριθμός φοιτητών στο μάθημα, οι εγκαταστάσεις της αίθουσας, η διαθέσιμη τεχνολογία και οι πηγές μάθησης, και το κλίμα της σχολής, μπορούν να επηρεάσουν την επίδοση του φοιτητή ανεξάρτητα από το τι κάνει ο διδάσκων στην αίθουσα διδασκαλίας (Berk, 2006, σελ. 33-34) (βλέπε αναλυτική παρουσίαση στο τμήμα 5: *Ζητήματα Μεθοδολογίας*).

Ο Fenwick (2001) προτείνει τα αποτελέσματα των σταθμισμένων μετρήσεων από τα τεστ, τις ασκήσεις επίλυσης –προβλημάτων, τις ερευνητικές εργασίες (projects), και τις προσομοιώσεις να συναθροίζονται ανά ομάδες φοιτητών εφόσον πρόκειται να αξιοποιηθούν στη λήψη αποφάσεων σχετικών με την αξιολόγηση προγράμματος για τις μεθόδους διδασκαλίας και στη βελτίωση του προγράμματος. Αυτό συνιστά ένα μη τετριμμένο κατόρθωμα από τη στιγμή που το περιεχόμενο των μαθημάτων, οι φοιτητές και οι μετρήσεις διαφέρουν από τάξη σε τάξη.

Συμπέρασμα: Οι μετρήσεις αποτελεσμάτων μάθησης πρέπει να χρησιμοποιούνται με μεγάλη προσοχή ως πηγή απόδειξης για την αξιολόγηση των διδασκόντων. Είναι πιο ασφαλές να χρησιμοποιούνται σε συνδυασμό με τις πηγές των άμεσων δεδομένων που περιγράφηκαν προηγούμενα για τη βελτίωση των προγραμμάτων.

4.2.12 Χαρτοφυλάκιο διδασκαλίας

Το χαρτοφυλάκιο διδασκαλίας είναι «ένα συνεκτικό σύνολο υλικών που περιλαμβάνει δείγματα εργασίας και σημειώσεις αναστοχασμού πάνω σε αυτά, το οποίο καταρτίζει ένα μέλος ΔΕΠ για να παρουσιάσει την πρακτική της διδασκαλίας του (ή της) τη σχετική με τη μάθηση και την εξέλιξη του φοιτητή» (Cerbini & Hutchings, 1993, p. 1). Τα δείγματα εργασίας έχουν χρησιμοποιηθεί στις επιχειρήσεις και στη βιομηχανία για τη μέτρηση της επίδοσης των εργαζομένων πάνω από 50

χρόνια. Η έρευνα για την αποτελεσματικότητά τους κατά την αξιολόγηση των επιδόσεων έχει διεξαχθεί στον τομέα της βιομηχανικής / οργανωτικής ψυχολογίας (Asher & Sciarrino, 1974; Siegel, 1986). Άλλοι ορισμοί περιέχουν αυτά τα βασικά στοιχεία (Berk, 1999, 2002; Cox, 1995; Edgerton, Hutchings, & Quinlan, 1991; Knapper & Wright, 2001; Murray, 1995; Seldin, Annis, & Zubizarreta, 1995).

Ο Knapper (1995) σκιαγράφησε την πιο πρόσφατη καταγωγή του χαρτοφυλακίου διδασκαλίας στην εργασία μίας επιτροπής της Canadian Association of University Teachers (CAUT). Ο πρόεδρος, Shore (1975), ισχυρίστηκε ότι τα μέλη ΔΕΠ πρέπει να προετοιμάζουν τα αποδεικτικά τους για την αποτελεσματικότητα της διδασκαλίας τους σε «ένα χαρτοφυλάκιο με αποδείξεις» (σελ 8). Αυτό που προέκυψε ήταν *The Teaching Dossier: A Guide to Its Preparation and Use* (Shore & Associates, 1980, 1986). Τη δεκαετία του '80, αυτός ο οδηγός έγινε χαρτοφυλάκιο βίβλος και η ιδέα εξαπλώθηκε σαν τη γρίπη στον Καναδά ως το “dossier”, στις ΗΠΑ ως το χαρτοφυλάκιο (portfolio) (Seldin, 1980, 2004), στην Αυστραλία (Roe, 1987) και στο Ηνωμένο Βασίλειο ως το προφίλ (profile) (Gibbs, 1988).

Ο *Guide* συνιστά 49 κατηγορίες ομαδοποιημένες κάτω από τρεις τίτλους: (α) προϊόντα καλής διδασκαλίας, (β) τα ίδια τα υλικά, και (γ) πληροφορίες από άλλους. Οι Knapper και Wright (2001) προσφέρουν ένα κατάλογο με τα 10 πιο συχνά χρησιμοποιημένα στοιχεία από μία έρευνα σε μέλη ΔΕΠ κολλεγίων και πανεπιστημίων στη Βόρεια Αμερική (O' Neil & Wright, 1995):

1. Μαθήματα φοιτητών και αξιολόγηση των δεδομένων της διδασκαλίας τα οποία προτείνουν βελτιώσεις ή παράγουν μία επί του συνόλου βαθμολογία της αποτελεσματικότητας ή της ικανοποίησης.
2. Κατάλογος τίτλων μαθημάτων και αριθμών, πιστώσεις, εγγραφές με σύντομη επεξεργασία.
3. Κατάλογος με τα υλικά των μαθημάτων που προετοιμάστηκαν για τους φοιτητές.
4. Συμμετοχή σε σεμινάρια, εργαστήρια, και επαγγελματικές συναντήσεις που στόχευαν στη βελτίωση της διδασκαλίας.
5. Δηλώσεις από συναδέλφους οι οποίοι έχουν παρακολουθήσει τη διδασκαλία είτε ως μέλη μίας διδασκαλικής ομάδας είτε ως ανεξάρτητοι παρατηρητές μίας σειράς μαθημάτων, ή διδάσκουν άλλα τμήματα της ίδιας σειράς μαθημάτων.
6. Προσπάθειες για καινοτομίες στη διδασκαλία και αξιολογήσεις της αποτελεσματικότητάς τους.

7. Μη δομημένες και (πιθανά ανεπιθύμητες) γραπτές αξιολογήσεις από τους φοιτητές που περιλαμβάνουν γραπτά σχόλια για τις εξετάσεις και γράμματα τα οποία έφτασαν μετά την ολοκλήρωση των μαθημάτων.

8. Συμμετοχή στην πορεία ή στην ανάπτυξη των προγραμμάτων σπουδών.

9. Απόδειξη για αποτελεσματική επίβλεψη τιμητικών διακρίσεων, μεταπτυχιακών και διδακτορικών διατριβών.

10. Εκθέσεις φοιτητών, δημιουργική εργασία και ερευνητική εργασία ή αναφορές για εργασίες πεδίου (field work) (σελ. 22-23).

Οι συγγραφείς προτείνουν τρεις κατηγορίες δηλώσεων: (α) των ευθυνών της διδασκαλίας, (β) της προσέγγισης της διδασκαλίας ή της φιλοσοφίας, και (γ) των δεδομένων από τους φοιτητές.

Ο Berk τονίζει ότι «οι επικεφαλίδες του χαρτοφυλακίου και ο μακρύς κατάλογος των πηγών απόδειξης της αποτελεσματικότητας της διδασκαλίας έχουν σχεδιαστεί για να εντυπωσιάσουν τον πιο κυνικό, μειωμένης αντίληψης, μεροληπτικό, και / ή ανίδεο μέλος ΔΕΠ σε μία επιτροπή αξιολόγησης, αναφορικά με το γεγονός ότι η διδασκαλία είναι μία ακαδημαϊκή δραστηριότητα η οποία είναι συγκρίσιμη με τον κατάλογο των εκδόσεων, των παρουσιάσεων, των επιχορηγήσεων, και των τιμητικών διακρίσεων που παρουσιάζονται ως απόδειξη για την υποτροφία έρευνας (research scholarship). Η πρακτική της διδασκαλίας δεν είναι απλά ένας κατάλογος μαθημάτων και αθροισμάτων των βαθμολογιών από τους φοιτητές» (σελ. 36).

Μία σύνθεση στοιχείων της βιβλιογραφίας τα οποία εμφανίζονται στα χαρτοφυλάκια διδασκαλίας και χρησιμοποιούνται σε διάφορα ιδρύματα, παρουσιάζει ένα πλήρη κατάλογο από στοιχεία, ταξινομημένα σε τρεις κατηγορίες:

1. Περιγραφή των ευθυνών διδασκαλίας

- α. Τα μαθήματα που διδάσκονται
- β. Παρουσιάσεις φιλοξενούμενων
- γ. Μία-προς-μία διδασκαλία (π. χ επιστημονικές ερευνητικές εργασίες, ανεξάρτητες μελέτες, επιτροπές διατριβών)
- δ. Ανάπτυξη καινούργιων προγραμμάτων ή σειράς μαθημάτων
- ε. Υπηρεσία σε επιτροπές προγραμμάτων σπουδών
- στ. Επιχορηγήσεις κατάρτισης.

2. Στοχαστική Ανάλυση (5 – 10 σελίδες)

- α. Φιλοσοφία της διδασκαλίας

- β. Καινοτόμες και δημιουργικές τεχνικές διδασκαλίας
- γ. Καθοδήγηση φοιτητών και διδασκόντων
- δ. Συμμετοχή σε δραστηριότητες για την ανάπτυξη της σχολής
- ε. Κατάρτιση διδασκαλίας
- στ. Αναγνώριση αποτελεσματικής διδασκαλίας

3. Αντικείμενα (Παραρτήματα – απόδειξη που υποστηρίζουν τα παραπάνω)

- α. Διδακτέα ύλη
- β. Φυλλάδια
- γ. Εξετάσεις
- δ. Δείγματα εργασίας φοιτητών
- ε. Χρήση τεχνολογίας
- στ. Βαθμολογίες από τους φοιτητές
- η. Βαθμολογίες από τους ομοβάθμιους
- θ. Βαθμολογίες από τους απόφοιτους
- ι. Μαγνητοσκοπήσεις διδασκαλίας
- κ. Κατάρτιση διδασκαλίας
- λ. Διαβουλεύσεις για τη διδασκαλία.

Δεδομένου ότι το χαρτοφυλάκιο απαιτεί χρόνο για την προετοιμασία του, η πρωταρχική του χρήση αφορά στη διαμόρφωση αποφάσεων καριέρας δηλαδή αποφάσεων σχετικών με την προαγωγή και την παραμονή του διδάσκοντα (Diamond, 2004; Seldin, 2004). Ωστόσο, μία παραλλαγή αυτού του θέματος ονομαζόμενη «Σωκρατικό χαρτοφυλάκιο» έχει προταθεί να παρέχει την κατεύθυνση και το έναυσμα στη διδασκαλική καριέρα ενός μεταπτυχιακού φοιτητή (Border, 2002, 2005). Αυτό το όνομα αναφέρεται στο διάλογο και στη συνεργασία μεταξύ του μαθητευόμενου φοιτητή και των πολλαπλών μεντόρων που εμφανίζονται στη διάρκεια του προγράμματος. Αυτός ο τύπος επικοινωνίας αποτελεί την γνωστή πλατωνική μέθοδο (Berk, 2006, σελ. 37).

Προκαταρκτικές εκτιμήσεις για την αξιοπιστία των αποφάσεων της επιτροπής προώθησης από «συναδέλφους» που βασίζονται σε χαρτοφυλάκια είναι εντυπωσιακές (Anderson, 1993; Centra, 1999). Αυτές οι αποφάσεις πρέπει να βασίζονται σε μία κλίμακα αξιολόγησης για να επαληθεύουν συστηματικά την αξιολόγηση της διδασκαλίας. Μία ποικιλία από φόρμες έχουν δημοσιευτεί στη βιβλιογραφία οι οποίες θα μπορούσαν να χρησιμεύσουν ως πρωτότυπα από τα οποία

κάποιο θα μπορούσε να προσαρμοστεί στον ορισμό του διδάσκοντα για την αποτελεσματικότητα της διδασκαλίας (Braskamp & Ory, 1994; Centra, 1973, 1993; Centra, Froth, Gray & Lambert, 1987; French-Lazovik, 1981).

Συμπέρασμα: Ως μία συλλογή πολλών από τις προγενέστερες πηγές, το χαρτοφυλάκιο διδασκαλίας πρέπει να προορίζεται πρωταρχικά για αποφάσεις συγκεντρωτικές να παρουσιάζει μία πλήρη εικόνα της αποτελεσματικότητας της διδασκαλίας και να συμπληρώνει τον κατάλογο των δημοσιεύσεων για την έρευνα.

5. Ζητήματα Μεθοδολογίας

Α. Ζητήματα Εγκυρότητας

5.1 Αποτελεσματικότητα διδασκαλίας και εγκυρότητα αξιολογήσεων

Αρκετό μέρος της συζήτησης σχετικά με την αξιολόγηση της διδασκαλίας από τους φοιτητές δεν έχει να κάνει με την ανάγκη για αξιολόγηση αλλά με το εάν τα ερωτηματολόγια εξασφαλίζουν μία έγκυρη εφαρμογή με αυτό το σκοπό. Η σχετική έρευνα μειονεκτεί εξαιτίας ενός θεμελιώδους προβλήματος: κανένας δεν έχει δώσει ένα ευρέως αποδεκτό ορισμό για το τι είναι «καλή» διδασκαλία ούτε έχει εδραιωθεί ένα καθολικά αποδεκτό κριτήριο για την αποτελεσματικότητα της διδασκαλίας (J. V. Adams, 1997; Kulik, 2001; Babin, Shaffer, & Morgan, 2002; Casey, Gentile, & Bigger, 1997; Murphy & MacLaren, 2007).

Ορίζοντας την αποτελεσματική διδασκαλία

Η Chism (1999) εντόπισε την έρευνα η οποία με συνέπεια έχει δείξει ότι υπάρχει ομοφωνία σε μεγάλο βαθμό για το τι χαρακτηρίζει την αποτελεσματική διδασκαλία. Ανάμεσα στους παράγοντες που προβάλλονται με συνέπεια είναι η αρμοδιότητα του διδάσκοντα στο αντικείμενο διδασκαλίας, η προετοιμασία και η οργάνωση, η διαύγεια, ο ενθουσιασμός, και η διαπροσωπική συμπάθεια. Ο Arreola (2007) αναγνώρισε πέντε ευρείες κατευθύνσεις δεξιοτήτων που απαιτείται να εμφανίζει η διδασκαλία: εξειδίκευση περιεχομένου, δεξιότητες σχεδιασμού της διδασκαλίας, δεξιότητες παράδοσης της διδασκαλίας, δεξιότητες αξιολόγησης της διδασκαλίας, και δεξιότητες διαχείρισης των μαθημάτων. Αυτές οι διαστάσεις βρίσκονται σε ευρεία συμφωνία με τα τέσσερα θεμελιώδη καθήκοντα διδασκαλίας που εισηγείται ο Fink (2008): σχεδιασμός μαθημάτων, αλληλεπίδραση καθηγητή –φοιτητή, ποιότητα της μάθησης των φοιτητών, και συνεχής βελτίωση.

Οι προσεγγίσεις των Arreola και Fink συγχρονίζονται με την παρουσίαση της διδασκαλίας από τον Hatch ως «μίας πολύπλοκης πνευματικής ενασχόλησης η οποία απαιτεί εξειδίκευση στο επιστημονικό πεδίο, μία βαθιά κατανόηση των φοιτητών, και εκλεπτυσμένες παιδαγωγικές δεξιότητες» (Hatch, 2006, σελ. 11).

Δημιουργώντας το ερωτηματολόγιο

Ο Ory (2001) και οι Theall και Franklin (2001) επισημαίνουν ότι, για να είναι οι μετρήσεις της αποτελεσματικότητας της διδασκαλίας έγκυρες, οι ερωτήσεις στο ερωτηματολόγιο αξιολόγησης πρέπει να αντανακλούν τόσο 1) τους τρόπους με τους

οποίους οι αξιολογήσεις χρησιμοποιούνται για διαμορφωτική και συγκεντρωτική αξιολόγηση της διδασκαλίας και 2) τους τρέχοντες παιδαγωγικούς και διδασκαλικούς στόχους του ιδρύματος.

Οι Ory και Ryan (2001) επίσης επισημαίνουν την σπουδαιότητα της επιβεβαίωσης ότι τα ερωτήματα αξιολόγησης ταιριάζουν όλα και μόνο με αυτό που η αξιολόγηση επιδιώκει να μετρήσει. Ο κεντρικός ρόλος των έγκυρων ερωτήσεων να εξασφαλίσουν τις έγκυρες αξιολογήσεις δεν μπορεί να υπερεκτιμηθεί.

Ο σχεδιασμός του ερωτηματολογίου επίσης παίζει ένα σημαντικό ρόλο στην επιβεβαίωση της εγκυρότητας. Ο Scriven (1995) υποστηρίζει ότι η εγκυρότητα των αξιολογήσεων μπορεί επίσης να επηρεάζεται από το μήκος της φόρμας ενώ ο Sedlmeier (2006) συζητά την επίδραση ειδικών κλιμάκων αξιολόγησης πάνω στα αποτελέσματα αξιολόγησης.

Οι Ory και Ryan (2001) γράφουν ότι «για να διατυπώσεις σωστά συμπεράσματα με βάση τις αξιολογήσεις των μαθημάτων από τους φοιτητές, οι δηλώσεις αξιολόγησης πρέπει να είναι σχετικές και αντιπροσωπευτικές των διαδικασιών, των στρατηγικών και του γνωστικού πεδίου που αφορούν στην ποιότητα της διδασκαλίας» (σελ. 32). Για να υπάρξουν στη φάση των αξιολογήσεων των μαθημάτων, έγκυρες μετρήσεις της αποτελεσματικότητας της διδασκαλίας, όχι μόνο πρέπει οι ερωτήσεις να αντανακλούν εκείνες τις πλευρές της διδασκαλίας που αναγνωρίστηκαν ως αποτελεσματικές, αλλά πρέπει να υφίσταται αναγνώριση και συμφωνία του ίδιου του ορισμού της αποτελεσματικής διδασκαλίας. Είναι δύσκολο να ορισθεί η συμπεριφορά της αποτελεσματικής διδασκαλίας, ως εκ τούτου οι Ory και Ryan υποστηρίζουν ότι «κανένα ενιαίο σύνολο χαρακτηριστικών των αποτελεσματικών καθηγητών και μαθημάτων που θα πρέπει να χρησιμοποιούνται ως στόχος ... (δεν) εμφανίζεται να υπάρχει» (σελ. 32).

Διάφορα εργαλεία αξιολόγησης έχουν προσπαθήσει να αναγνωρίσουν ειδικές συμπεριφορές διδασκαλίας ή μέσα δόμησης και παρουσίασης του υλικού των μαθημάτων τα οποία συμβάλλουν στην αποτελεσματική διδασκαλία. Για παράδειγμα, το SEEQ (Marsh, 1987), το Teaching Behaviours Inventory (Murray, 1987) και το IDEA (Cashin, 1992) (βλέπε και τα τρία στο Παράρτημα) κάθε ένα αξιολογούσε την

έρευνα για την εκμάθηση των φοιτητών και τους φοιτητές και τα μέλη ΔΕΠ που ρωτήθηκαν για να αναγνωρίσουν χαρακτηριστικά της αποτελεσματικής διδασκαλίας. Αυτές οι δηλώσεις μπορεί να αξιολογηθούν περαιτέρω διαμέσου συγκρίσεων με άλλες μετρήσεις της εκμάθησης των φοιτητών (για παράδειγμα, ακαδημαϊκή επίδοση).

Όμως υπάρχουν ελάχιστα μέσα πέρα από τη λογική ανάλυση (Marsh, 1987) που να μπορούν να επιβεβαιώσουν ότι αυτά τα χαρακτηριστικά της αποτελεσματικής διδασκαλίας είναι αντιπροσωπευτικά, περιεκτικά ή γενικεύσιμα. Ο Wachtel (1998) επισημαίνει ότι πολλοί είναι οι μελετητές που έχουν προειδοποιήσει ότι πολλά μέσα αποτίμησης της εγκυρότητας των αξιολογήσεων από τους φοιτητές «υποθέτουν μία ομοφωνία η οποία όμως δεν υπάρχει ακόμα. Αυτό σημαίνει, με ποιο τρόπο μπορούμε να αξιολογήσουμε την αποτελεσματικότητα της διδασκαλίας κατάλληλα εάν δεν μπορούμε ακόμα να συμφωνήσουμε πάνω στο τι συνιστά την αποτελεσματική διδασκαλία;» (σελ. 193).

Πράγματι, πολλοί μελετητές έχουν προειδοποιήσει για τις αρνητικές επιπτώσεις λόγω διαγραφής ειδικών χαρακτηριστικών της αποτελεσματικής διδασκαλίας. Ο McKeachie (1997) αντλεί πορίσματα από τον Scriven (1981) για να ισχυριστεί ότι «καμία αξιολόγηση του ύφους (στυλ) διδασκαλίας (π. χ ενθουσιασμός, οργάνωση, ζεστασιά) δεν πρέπει να χρησιμοποιείται, επειδή η αποτελεσματικότητα της διδασκαλίας μπορεί να επιτευχθεί με συνδυασμό ποικίλων τρόπων. Χρησιμοποιώντας χαρακτηριστικά τα οποία γενικά έχουν θετικές συσχετίσεις με την αποτελεσματικότητα «τιμωρείς» τον καθηγητή ο οποίος είναι αποτελεσματικός παρά τις μικρότερες από τα πάνω βαθμολογίες σε μία ή περισσότερες από τις διαστάσεις που συνήθως έχουν σχέση με την αποτελεσματικότητα» (σελ. 1218-1219). Ο McKeachie επίσης επισημαίνει ότι «τα μέλη ΔΕΠ και οι διαχειριστές έχουν στερεότυπα σχετικά με το τι περιλαμβάνει η καλή διδασκαλία» και ότι «οι καθηγητές που δεν ανταποκρίνονται σε αυτά φαίνεται ότι κρίνονται ως μη αποτελεσματικοί παρά την άλλη απόδειξη της αποτελεσματικότητας» (σελ. 1219).

Οι Ory και Ryan (2001) κατηγοριοποιούν τις θεσμικές επιπτώσεις των αξιολογήσεων σε εκείνες με «προβλεπόμενες» συνέπειες και στις άλλες με «ακούσιες». Ανάμεσα σε εκείνες με «ακούσιες» συνέπειες υφίσταται η πιθανότητα οι διδάσκοντες να χρησιμοποιούν δηλώσεις στις φόρμες αξιολόγησης για να διαμορφώνουν τις

διδασκαλίες ή τα μαθήματα τους. Εάν οι δηλώσεις σε μία φόρμα αξιολόγησης δεν αντανακλούν προτεραιότητες του ιδρύματος ή τρόπους αποτελεσματικής διδασκαλίας, αυτές οι αξιολογήσεις μπορεί να έχουν αρνητική επίδραση για τη διδασκαλία στο ίδρυμα.

Η μελέτη της Neumann (2001) πάνω στις διαφορετικές πρακτικές της διδασκαλίας στα διάφορα επιστημονικά πεδία έχει επίσης σημαντικές εφαρμογές στον ορισμό της αποτελεσματικής διδασκαλίας και κατά συνέπεια στις αξιολογήσεις των μαθημάτων. Η συγγραφέας συμφωνεί ότι οι στόχοι μάθησης και το ύφος (στυλ) διδασκαλίας διαφέρουν σημαντικά ανάμεσα στα επιστημονικά πεδία (σελ. 138). Αυτές οι διαφορές θα μπορούσαν να επηρεάσουν ισχυρά τους τρόπους με τους οποίους δομούνται οι αξιολογήσεις και αξιολογείται η εγκυρότητά τους καθώς και τον βαθμό στον οποίο αυτές (οι αξιολογήσεις) αντανακλούν με συνέπεια την αποτελεσματικότητα της διδασκαλίας σύμφωνα με τα στάνταρ κάθε συγκεκριμένου πεδίου. Ωστόσο, η Neumann επισημαίνει ότι «στις περισσότερες περιπτώσεις, τα ερωτηματολόγια αξιολόγησης που χρησιμοποιούνται είναι γενικά» (σελ. 143), υπονοώντας ότι η διδασκαλία σε όλα τα επιστημονικά πεδία αντιμετωπίζεται ως να είναι ίδια.

Σε αντίθεση με αυτές τις ανησυχίες, ο Wachtel (1998) υποστηρίζει ότι «οι φοιτητές και τα μέλη ΔΕΠ γενικά συμφωνούν πάνω στο ποια είναι τα συστατικά της αποτελεσματικής διδασκαλίας και της σχετικής τους σημαντικότητας» (σελ. 192) και ότι οι ερωτήσεις που αναπτύσσονται από αυτά τα συστατικά μπορούν με ακρίβεια να αντανακλούν τη μάθηση των φοιτητών και την αποτελεσματικότητα της διδασκαλίας.

5.2 Γενικές έναντι πολλαπλών διαστάσεων αξιολογήσεις διδασκαλίας

Οι περισσότερες προσπάθειες να αναγνωρισθούν τα ειδικότερα χαρακτηριστικά της αποτελεσματικής διδασκαλίας προέρχονται από την πεποίθηση ότι η διδασκαλία πρέπει να μετριέται περισσότερο σύμφωνα με τις πολλαπλές όψεις ή κατηγορίες της διδασκαλικής δραστηριότητας – για παράδειγμα την προσβασιμότητα, τον ενθουσιασμό, ή την οργάνωση – παρά από ερωτήσεις που ψάχνουν γενικά μία ευρεία, γενική (global) απάντηση για τα μαθήματα ή τον καθηγητή. Αυτό το επιχείρημα έχει προβληθεί περισσότερο από το Marsh (1987) και τους Marsh και Roche (1997). Αυτές οι πολλαπλές μετρήσεις βασίζονται στην προϋπόθεση ότι η διδασκαλία είναι πολυδιάστατη – που σημαίνει, ότι οι καθηγητές μπορεί να υπερέχουν στη μία πλευρά της διδασκαλίας αλλά όχι στην άλλη. Ο Ory (2001) τονίζει ότι ενώ δεν είναι εύκολο

να αναγνωριστεί ένα ειδικό σύνολο συμπεριφορών αποτελεσματικής διδασκαλίας, τα αποτελέσματα ερευνών πάνω στις αντιλήψεις των φοιτητών για την αποτελεσματική διδασκαλία «υποστηρίζουν την επισήμανση ότι οι φοιτητές βλέπουν την ποιότητα της διδασκαλίας ως πολυδιάστατη» (σελ. 10). Σύμφωνα με αυτούς τους ερευνητές, οι πολυδιάστατες αξιολογήσεις, σε συνδυασμό με μία πιο ακριβή απεικόνιση της αποτελεσματικής διδασκαλίας, μπορεί να αποτελέσουν πολύτιμα διαγνωστικά εργαλεία στη διαμορφωτική αξιολόγηση (Marsh & Roche, 1997) καθώς μπορούν να βοηθήσουν τους διδάσκοντες να αναγνωρίσουν περιοχές δύναμης και αδυναμίας. Επιπλέον, οι Marsh και Roche υποστηρίζουν ότι η πολυδιάστατη ανατροφοδότηση είναι σημαντική στη συνεχιζόμενη έρευνα με αντικείμενο τις αξιολογήσεις της διδασκαλίας. Αυτοί επισημαίνουν διάφορα παραδείγματα στα οποία μία μεταβλητή που μπορεί να έχει αναγνωριστεί ως σημαντική πηγή προκατάληψης να μπορεί εν προκειμένω, διαμέσου της αξιολόγησης των απαντήσεων των φοιτητών βασισμένων σε μία ειδική διάσταση της διδασκαλίας, να δείχτει ότι έχει θετική επίδραση στη μάθηση των φοιτητών. Με άλλα λόγια, «μία επίδραση η οποία έχει ερμηνευτεί ως προκατάληψη (bias) στις αξιολογήσεις της διδασκαλίας από τους φοιτητές μπορεί να ερμηνευτεί ορθότερα ως υποστήριξη για την εγκυρότητά τους σε σχέση με μία συνιστώσα της αποτελεσματικής διδασκαλίας» (σελ. 1193).

Ο McKeachie (1997) συμφωνεί με τον Marsh όσον αφορά στην ομαδοποίηση των συμπεριφορών διδασκαλίας σε διαστάσεις, επισημαίνοντας ότι αυτή βοηθά στην ανάγνωση και στην ερμηνεία δεδομένων και ως εκ τούτου είναι πολύ πιθανό να οδηγήσει σε βελτίωση. Ωστόσο, αυτός δηλώνει αβεβαιότητα για τον αριθμό των διαστάσεων που πρέπει να αναφέρονται στους σκοπούς της συγκεντρωτικής αξιολόγησης αναρωτώμενος: «Πρέπει ένα αποτέλεσμα που αντιπροσωπεύει μία σταθμισμένη συνοπτική παρουσίαση των παραγόντων να τους εκπροσωπεί (όπως υποστηρίζουν οι Marsh & Roche, 1997), ή απλά πρέπει να χρησιμοποιούνται τα αποτελέσματα από μία ή περισσότερες συνολικές αξιολογήσεις της αποτελεσματικότητας της διδασκαλίας (όπως ισχυρίζονται οι d' Apollonia & Abrami, 1997);».

Πολλοί ερευνητές, ωστόσο, υποστηρίζουν ότι οι γενικές ερωτήσεις αξιολόγησης είναι ισότιμες, αν όχι και περισσότερο αξιοπρόσεκτες για τη συγκεντρωτική αξιολόγηση της διδασκαλίας από τις πολλαπλών διαστάσεων μετρήσεις (Algozinne et al., 2004; Abrami, 2001; Cashin, 1995, 1992). Στη μελέτη τους πάνω στους τρόπους με τους

οποίους οι αξιολογήσεις χρησιμοποιούνται από τα μέλη ΔΕΠ και τα μέλη της διοίκησης οι Beran, Violato και Kline (2007) επισημαίνουν ότι οι γενικές μετρήσεις είναι συχνά οι μόνες αξιολογούμενες άτυπα διαδικασίες αξιολόγησης. Οι Abrami (2001) και Cashin (1995, 1992) επισημαίνουν ότι υπάρχει μία δυνατή συσχέτιση μεταξύ των γενικών ερωτήσεων και άλλων μετρήσεων της αποτελεσματικότητας της διδασκαλίας. Όταν έγινε επισκόπηση ενός περιορισμένου αριθμού αποτελεσμάτων, οι γενικές ερωτήσεις μπόρεσαν με ακρίβεια να λειτουργήσουν ως ένας πληρεξούσιος για περισσότερες πολύπλοκες μετρήσεις και ως εκ τούτου σε αυτές τις περιπτώσεις, μία αληθινή γενική ερώτηση είναι περισσότερο αντιπροσωπευτική της αποτελεσματικότητας της διδασκαλίας από ό, τι είναι μία περιορισμένη επιλογή με μετρήσεις πολλαπλών διαστάσεων. Ο McKeachie (1997) συμφωνεί ότι τέτοιες γενικές αξιολογήσεις είναι πολύτιμες για τη συγκεντρωτική αξιολόγηση της διδασκαλίας ακόμα και αν αυτές δεν είναι ιδιαίτερα πολύτιμες για να παρέχουν διαμορφωτική ανατροφοδότηση. Οι March και Roche (1997), ωστόσο, διαφωνούν με αυτή την προοπτική, υποστηρίζοντας ότι «οι γενικές ή συνολικές αξιολογήσεις δεν μπορούν να εκπροσωπούν ικανοποιητικά τον πολλαπλών διαστάσεων χαρακτήρα της διδασκαλίας. Αυτές επίσης μπορεί να είναι περισσότερο ευπαθείς στο πλαίσιο, στη διάθεση (mood) και σε άλλες πιθανές προκαταλήψεις από ό, τι είναι οι ειδικές δηλώσεις οι οποίες είναι περισσότερο στενά συνδεδεμένες με τις πραγματικές συμπεριφορές της διδασκαλίας» (σελ. 1188).

Στη σύγχρονη βιβλιογραφία υπάρχει περιορισμένη συζήτηση σχετικά με την διατύπωση των γενικών ερωτήσεων. Ο Scriven (1995) έχει επισημάνει ότι πολλές συγκεντρωτικές αξιολογήσεις ρωτούν τις λάθος γενικές ερωτήσεις. Αυτός αναφέρει το ακόλουθο σύνηθες λάθος: ερωτήσεις που απαιτούν συγκρίσεις μεταξύ καθηγητών, ερωτήσεις που ρωτούν τους φοιτητές εάν αυτοί θα πρότειναν το μάθημα σε άλλους, και ερωτήσεις που ρωτούν τους φοιτητές εάν μία σειρά μαθημάτων είναι η «καλύτερη» που είχαν ποτέ.

5.3 Το ζήτημα της μέτρησης της μάθησης φοιτητών

Υπερασπιστές και επικριτές των αξιολογήσεων της διδασκαλίας από τους φοιτητές (ΦΑΔ) συμφωνούν γενικά ότι οι φοιτητές μαθαίνουν περισσότερα από τους καλούς καθηγητές. Με άλλα λόγια, εάν η διαδικασία είναι έγκυρη, τότε θα πρέπει να υπάρχει μία σχέση μεταξύ της μάθησης των φοιτητών και των αξιολογήσεων που οι φοιτητές δίνουν στα μαθήματα και στους καθηγητές. Όμως όπως συμβαίνει με κάθε θέμα στο

χώρο της εκπαίδευσης, υπάρχουν και εκείνοι που διαφωνούν. Ο Scriven (1981) παρατήρησε ότι «η καλύτερη διδασκαλία δεν είναι εκείνη που παράγει την περισσότερη μάθηση» (σελ. 248), αλλά αυτή είναι η θέση της μειοψηφίας. Ο Cohen (1981) επιβεβαιώνει, «ακόμα κι αν υπάρχει μία έλλειψη ομοφωνίας πάνω σε ένα ορισμό της καλής διδασκαλίας, οι περισσότεροι ερευνητές σε αυτή την περιοχή συμφωνούν ότι η μάθηση των φοιτητών αποτελεί το πιο σημαντικό κριτήριο της αποτελεσματικότητας της διδασκαλίας» (σελ. 283).

Από την εξέταση της βιβλιογραφίας, υπάρχουν πολλά θέματα τα οποία εξελίσσονται σε σημαντικά όταν επιδιώκεται η αξιολόγηση της σχέσης μάθησης / ΦΑΔ. Εάν η μάθηση και οι ΦΑΔ σχετίζονται με την καλή διδασκαλία τότε θα πρέπει να βρεθεί ότι οι ΦΑΔ σχετίζονται με τη μάθηση. Διάφορες μεθοδολογικού τύπου δυσκολίες, η πιο θεμελιώδης από τις οποίες είναι **με ποιο τρόπο μπορεί να μετρηθεί η μάθηση** έχουν παρεμποδίσει τον έλεγχο αυτού του ισχυρισμού. Τα περισσότερα από τα ζητήματα αυτά μπορούν να συνοψιστούν σε τρεις περιοχές: 1) στη σχέση βαθμών / ΦΑΔ, 2) στις αντιλήψεις των φοιτητών για τη μάθηση και 3) στη σχέση των πραγματικών βαθμών με τη μάθηση.

6.6.3 Βαθμοί φοιτητών & αξιολογήσεις διδασκαλίας από φοιτητές

Οι βαθμοί είναι πολύτιμοι για τους φοιτητές. Αυτοί είναι πρόθυμοι να εργαστούν γι' αυτούς. Οι φοιτητές δεν είναι πάντοτε πρόθυμοι να εργαστούν για να μάθουν περισσότερα. Η έκταση στην οποία οι φοιτητές δίνουν υψηλότερες αξιολογήσεις στους διδάσκοντες οι οποίοι τους δίνουν μεγαλύτερους βαθμούς μπορεί να είναι ή και να μην είναι ένας δείκτης της σχέσης μάθησης / ΦΑΔ. Η έκταση στην οποία οι υψηλές αξιολογήσεις μπορούν να «αγοραστούν» με τους βαθμούς, οι οποίοι δεν είναι άμεσα σχετικοί με τη μάθηση, στην πραγματικότητα ακυρώνει τις ΦΑΔ. Αυτό το θέμα είναι τόσης ιστορικής σημασίας που είναι αναγκαίο να συζητηθεί σύντομα πριν η συζήτηση για τη σχέση μάθησης / ΦΑΔ μπορέσει να αντιμετωπιστεί επαρκώς.

- *Οι υποστηρικτές των ΦΑΔ* παραδέχονται ότι υπάρχει εμφανής σχέση βαθμών / αξιολόγησης (Braskamp & Ory, 1994; Cashin, 1995; Marsh & Dunkin, 1992), αλλά πολλοί συνεχίζουν να ισχυρίζονται ότι τα στάνταρ της βαθμολόγησης δεν αλλάζουν σημαντικά τις ΦΑΔ, ειδικότερα εφόσον άλλες μεταβλητές όπως είναι η αυστηρότητα και το προγενέστερο ενδιαφέρον των φοιτητών ληφθούν υπόψη και ελέγχονται επαρκώς (Cashin, 1995; Greenwald & Gillmore, 1997; Kaplan, Mets, & Cook, 2000; Marsh & Roche, 2000; Powell, 1977; Schwab, 1976; Seiver, 1983; Stumpf & Freedman, 1979). Μερικοί απλά ισχυρίζονται

ότι η συσχέτιση βαθμών / αξιολόγησης είναι τόσο μικρή για να είναι σημαντική (Marsh & Roche, 1999) ή ότι οι βαθμοί και η καλή διδασκαλία σχετίζονται με εγκυρότητα (Cohen, 1981; Marsh, Hau, Chung, & Siu, 1997; Marsh & Roche, 1997).

- *Πολλοί επικριτές* επιμένουν ότι οι βαθμοί επηρεάζουν τις ΦΑΔ με ένα τρόπο ο οποίος ακυρώνει τα ερωτηματολόγια (Gillmore & Greenwald, 1999; Greenwald & Gillmore, 1997b; βλέπε Greenwald, 1997 για χρήση στοιχείων). Δύο αρνητικές υποθέσεις έχουν προμοδοτηθεί. Η αρχαιότερη από τις δύο, της *επιείκειας*, ισχυρίζεται ότι οι φοιτητές δίνουν υψηλότερη βαθμολογία στους διδάσκοντες που βαθμολογούν με επιείκεια (lenient-grading). Ένας αριθμός ερευνητών έχει βρει αποδεικτικά στοιχεία γι' αυτή την επίδραση (Greenwald & Gillmore, 1997b; βλέπε Johnson, 2003, για μία εκτενή επισκόπηση). Μία δεύτερη υπόθεση, της *αμοιβαιότητας*, υποστηρίζει ότι οι φοιτητές οι οποίοι λαμβάνουν καλύτερους βαθμούς δίνουν καλύτερες αξιολογήσεις, ανεξάρτητα από την οποιαδήποτε τάση επιείκειας του διδάσκοντα (Clayson, 2004; Clayson, Frost, & Sheffet, 2005). Οι δύο υποθετικές και στατιστικές έννοιες είναι ευδιάκριτα διαφορετικές αλλά δεν είναι αναγκαίο να αλληλοαναιρούνται (Clayson, 2007; Stumpf & Freedman, 1979). Ο Johnson (2003) βρήκε ισχυρή απόδειξη τόσο για την επίδραση της επιείκειας όσο και για της αμοιβαιότητας. Μία άλλη εκτενής μελέτη βρήκε μία ισχυρή επίδραση της επιείκειας σε μαθήματα λογιστικής ακόμα κι κατά τον έλεγχο για τις άλλες μετρήσεις της μάθησης (Weinberg, Fleisher, & Hashimoto, 2007). Οι March και Roche (2000) και ο Centra (2003) διεξήγαγαν μεγάλες και προσεκτικές μελέτες αλλά δεν βρήκαν απόδειξη για την επίδραση της επιείκειας. Οι διαφορές στα ευρήματα μπορεί να είναι το αποτέλεσμα των προβλημάτων που δημιουργούνται από τη χρήση μόνο των μεταξύ-τάξης (between-class) δεδομένων, που είναι τόσο στατιστικά όσο και λογικά (Clayson, 2007). Εάν χρησιμοποιούνται οι μέσοι όροι της τάξης τότε οι επιδράσεις αμοιβαιότητας και επιείκειας συγχέονται. Μπορεί να βρεθεί ότι δεν υπάρχουν επιδράσεις επιείκειας σε επίπεδο φοιτητή, εκτός αν το μέγεθος της τάξης και ο μέσος όρος των βαθμών διατηρούνται σταθερά. Ανεξάρτητα από τις απαιτήσεις των ολιστικών μοντέλων, οι διδάσκοντες σε τελική ανάλυση αξιολογούνται από τους φοιτητές, όχι από τις τάξεις (classes). Ο Clayson και οι συνεργάτες του

(2005) επιμένουν ότι η μόνη υπόθεση που υποστηρίζεται από όλα τα δεδομένα είναι της αμοιβαιότητας.

Η ειρωνεία είναι ότι οι βαθμοί μπορεί να επηρεάζουν τις ΦΑΔ με αρνητικό τρόπο ανεξάρτητα από το ποια πλευρά του θέματος μπορεί να είναι η πλέον σωστή. Παρά το γεγονός ότι η πραγματική σχέση μπορεί να συζητηθεί, αυτό δεν αλλάζει τη συμπεριφορά τόσο των μελών ΔΕΠ όσο και των φοιτητών οι οποίοι συνεχίζουν να πιστεύουν ότι υπάρχει μία σχέση μεταξύ των βαθμών και της αξιολόγησης (Clayson, 2009, σελ. 18). Υφίσταται απόδειξη ότι οι πεποιθήσεις μόνες τους τροποποιούν τη συμπεριφορά των μελών ΔΕΠ και των φοιτητών (Birnbbaum, 2000; Goldman, 1985; Kolevzon, 1981; Marsh, 1987; Moore & Trahan, 1998; Redding, 1998; Ryan, Anderson, & Birchler, 1980; Simpson & Siguaw, 2000). Είναι προφανές ότι οι δύο πλευρές της συζήτησης συμφωνούν στο ότι οι βαθμοί, όπως συνήθως εμφανίζονται στις μελέτες, είτε σχετίζονται μόνο οριακά με τη μάθηση είτε δεν σχετίζονται καθόλου. Οι ερευνητές ισχυρίζονται ότι οι φοιτητές δεν πρέπει να πιστεύουν ότι οι βαθμοί τους είναι ισοδύναμοι των αντιλήψεών τους για τη μάθηση. Ο Cashin (1995) συνοψίζει τη σκέψη πολλών άλλων όταν επιμένει ότι οι βαθμοί δεν μπορούν να χρησιμοποιηθούν ως ένας δείκτης της μάθησης χωρίς κατάλληλο έλεγχο. Σε αυτό συμφωνεί με τους ερευνητές εκτός των εκπαιδευτικών επιστημονικών κλάδων (Clayson, 2004; Greenwald & Gillmore, 1997a; Johnson, 2003).

5.3.2 Αντιλήψεις φοιτητών για τη μάθηση

Το δεύτερο μεθοδολογικό πρόβλημα προκύπτει από τη χρήση των αντιλήψεων των φοιτητών ως μία μέτρηση της μάθησης. Μερικοί ερευνητές έχουν υποστηρίξει ότι οι φοιτητές είναι οι καλύτεροι κριτές για το τι μαθαίνουν (Cruse, 1987; Machina, 1987). Ο ισχυρισμός δεν έχει ευρέως υποστηριχτεί από πραγματικά αποτελέσματα. Οι βαθμοί που αντιλαμβάνονται οι φοιτητές δεν είναι ανάγκη να σχετίζονται δυνατά με τους πραγματικούς τους βαθμούς (Baird, 1987; Clayson, 2005a; Flowers, Osterlind, Pascarella, & Pierson, 2001; Sheehan & DuPrey, 1999; Williams & Ceci, 1997). Οι φοιτητές κάνουν συνεπή / σταθερά λάθη στην εκτίμηση των βαθμών τους τα οποία έχουν ερμηνευτεί ως μία «μετά-γνωστική» επίδραση. Με άλλα λόγια, οι πιο αδύνατοι φοιτητές δεν γνωρίζουν τι δεν γνωρίζουν και κατά συνέπεια υπερεκτιμούν τις γνώσεις τους στα υλικά ελέγχων, ενώ οι καλύτεροι φοιτητές γνωρίζουν τι δεν γνωρίζουν και υποτιμούν τις γνώσεις τους (Grimes, 2002; Kennedy, Lawton, & Plumlee, 2002; Moreland, Miller, & Laucka, 1981). Ένα πρόσφατο εύρημα δείχνει ότι αυτή η

ερμηνεία μπορεί να είναι υπεραπλουστευμένη. Οι φοιτητές εμφανίζονται να γνωρίζουν τι αυτοί δεν γνωρίζουν αλλά αυτοί χρησιμοποιούν αυτή την αναγνώριση μόνο κατά μέσο όρο ως θεμέλιο που συνδυάζεται με τους κανόνες της ομάδας για να εκτιμήσουν τις μαθησιακές τους επιδόσεις (Clayson, 2005a).

5.3.3 Πραγματικοί βαθμοί και μάθηση

Η χρήση των πραγματικών βαθμών του μαθήματος επίσης δημιουργεί προβλήματα. Υπάρχουν πολλές μεταβλητές που μπορούν να επηρεάσουν τους βαθμούς των μαθημάτων άσχετες με τις δεξιότητες και τις γνώσεις των φοιτητών οι οποίες χωρίς ισχυρούς στατιστικούς ή εμπειρικούς ελέγχους καθιστούν τους βαθμούς αμφισβητήσιμους ως μία μέτρηση της μάθησης. Οι Pollio και Beck (2000) επισημαίνουν ότι «Η παρατήρηση ότι οι βαθμοί παρέχουν ακριβείς αξιολογήσεις για το πόσο καλά ένας φοιτητής τα πάει σε ένα κολλέγιο και το πόσο καλά αυτός ή αυτή θα τα πάει σε ένα μελλοντικό επάγγελμα δεν υποστηρίζεται από την εμπειρική βιβλιογραφία» (σελ 100).

5.4 Μεθοδολογικού τύπου επισημάνσεις για τη μέτρηση της μάθησης

Ο Clayson (2009) στη μετά-ανάλυση και επισκόπηση της βιβλιογραφίας πάνω σε ζητήματα αξιολόγησης των μαθημάτων από τους φοιτητές και της μάθησής τους επισημαίνει τις πέντε λύσεις που έχουν προβληθεί εις απάντηση της έκκλησης της Cashin (1995) για αυστηρούς ελέγχους όταν γίνεται μέτρηση της μάθησης:

1. Όταν η μεταβλητή βαθμός χρησιμοποιείται για να μετρηθεί η μάθηση τότε θα πρέπει αυτή η μεταβλητή να είναι ο μέσος όρος των βαθμών της τάξης και όχι οι ξεχωριστοί βαθμοί των φοιτητών (Abrami, d'Appolonia, & Cohen, 1990; Cohen, 1981; Marsh & Roche, 2000). Με τον τρόπο αυτό αφαιρούνται, ή υπολογίζονται κατά μέσο όρο, οι ατομικές διαφορές και αντανakλάται περισσότερο η επίδραση της διδασκαλίας και / ή του διδάσκοντα. Όπως αναφέρθηκε νωρίτερα, αυτή η λύση έχει δεχτεί κριτική επειδή μπορεί να καταλήξει σε αλληλεπιδράσεις που δεν μπορούν να ξετυλιχτούν εύκολα στατιστικά. Η μεταβλητότητα (variability) μεταξύ των φοιτητών, ακόμα και του μέσου όρου, μπορεί να προκαλέσει σύγχυση στη μεταβλητότητα μεταξύ των ομάδων (Weinberg et al., 2007). Είναι πιθανό στα άτομα φοιτητές να μη προκύψει κάποια σχέση μεταξύ των αξιολογήσεων και της μάθησης, ακόμα κι αν οι μεταξύ-τάξεων μέσοι όροι θα μπορούσαν να δείξουν μία σημαντική σχέση (Clayson, 2007).

2. Κοινοί έλεγχοι θα μπορούσαν να χρησιμοποιηθούν σε πολλαπλά κεφάλαια (sections), ιδίως εάν μπορεί να ελεγχθεί η διακύμανση μεταξύ των διδασκόντων (Cohen, 1981; Williams & Ceci, 1997).
3. Η αλλαγή των βαθμών με βάση τους πριν και μετά ελέγχους θα μπορούσε να συνιστά τη μέτρηση της μάθησης (Hake, 2002).
4. Η μάθηση μπορεί να μετρηθεί από την επίδοση σε μελλοντικά μαθήματα ελεγχόμενη για τα χαρακτηριστικά των φοιτητών και την επίδοση σε προαπαιτούμενα μαθήματα (Johnson, 2003; Weinberg et al., 2007; Yunker & Yunker, 2003).
5. Σε έκπληξη, από μία μεθοδολογική σκοπιά, ελάχιστες πηγές βρέθηκαν να συστήνουν ή να χρησιμοποιούν έξωθεν σταθμισμένες μετρήσεις. Αυτό πιθανό να οφείλεται σε πρακτικούς λόγους. Σταθμισμένες μετρήσεις μάθησης είναι σχετικά σπάνιες και, για να είναι χρήσιμες, οι σταθμισμένοι έλεγχοι πρέπει να είναι κατάλληλοι τόσο για το αντικείμενο μελέτης όσο και για το ακαδημαϊκό επίπεδο οποιουδήποτε δοσμένου μαθήματος.

5.5 Το παράδοξο της αυστηρότητας /σχολαστικότητας (rigor)

Οι ερμηνείες που δίνονται σε μελέτες με αντικείμενο τη σχέση βαθμολογία / ΦΑΔ περιπλέκονται από ένα άλλο εύρημα. Στο μεγαλύτερο μέρος της βιβλιογραφίας έχει βρεθεί μία φαινομενική ασυνεπής σχέση μεταξύ αυστηρότητας και αξιολογήσεων (Clayson, 2009). Εάν η προσπάθεια και η πρόκληση (αυστηρότητα) σχετίζονται με τη μάθηση, και η μάθηση σχετίζεται με τη καλή διδασκαλία, τότε καθώς τα επίπεδα της προσπάθειας και της πρόκλησης αυξάνουν, το συνολικό επίπεδο της αξιολόγησης (ΦΑΔ) θα πρέπει να αυξάνει επίσης (Greenwald & Gillmore, 1997a). Εκείνοι που συνοψίζουν την έρευνα για τις ΦΑΔ, στο πλαίσιο των εκπαιδευτικών επιστημονικών τομέων, γενικά συμφωνούν για την ισχύ των προαναφερθέντων (Cashin, 1995; Βλέπε Sixbury & Cashin, 1995). Μετρήσεις της αυστηρότητας (φόρτος εργασίας, δυσκολία υλικών, χρόνου, κλπ) δεν συμπεριλαμβάνονται σε όλες τις μελέτες της σχέσης μάθηση / ΦΑΔ αλλά όταν προστίθενται τότε η σχέση μεταξύ αυστηρότητας και ΦΑΔ γίνεται γενικά αρνητική (Attiyeh & Lumsden, 1972; Clayson & Haley, 1990; Frey, Leonard, & Beatty, 1975; Steiner, Holley, Gerdes, & Cambell, 2006). Ο Centra (2003) βρήκε μία αρνητική σχέση μεταξύ της «προσπάθειας φοιτητή / βελτίωση» και των αξιολογήσεων. Στη μετά-ανάλυση του Cohen (1981), η «δυσκολία» βρέθηκε να μη σχετίζεται με τις μετρήσεις της μάθησης. Σε 24 μελέτες, ο Cohen βρήκε μόνο μία σημαντική σχέση μεταξύ δυσκολίας και μάθησης και αυτή ήταν αρνητική. Ο

Weinberg και οι συνεργάτες του (2007) τοποθέτησαν τη σχέση με οικονομικούς όρους όταν δήλωναν «η μεγαλύτερη παραγωγή ανθρώπινου κεφαλαίου σχετίζεται με τη λιγότερο ευχάριστη εμπειρία από τα μαθήματα επειδή απαιτεί περισσότερη εργασία από τους φοιτητές» (σελ. 7). Δύο επιφυλάξεις δικαιολογούνται σε αυτή τη διατύπωση. Σε μερικές μελέτες, ο λειτουργικός ορισμός της *δυσκολίας* θα μπορούσε να ειπωθεί ως αρνητικός, υποδεικνύοντας ότι ο διδάσκων παραβίαζε (in violation) κάποιο κανόνα. Δεύτερο, ένας φοιτητής θα ήταν λογικά αναμενόμενο να μάθει ελάχιστα εάν η διδασκαλία δεν ήταν αρκετά αυστηρή ή εφόσον αυτή κατέστη πολύ δύσκολη.

Ο Johnson (2003) ανέφερε ότι «η αυστηρή βαθμολόγηση σχετίζονταν με τα πιο υψηλά επίπεδα επίδοσης στη συνέχεια των μαθημάτων» (σελ 161), αλλά οι αυστηρές βαθμολογήσεις ήταν ισχυρά σχετιζόμενες με χαμηλά επίπεδα αξιολογήσεων (ΦΑΔ). Αυτός επίσης βρήκε ότι οι δηλώσεις οι οποίες είχαν μία υψηλή σχέση με τις αξιολογήσεις, όπως το (εκδηλωμένο) ενδιαφέρον του διδάσκοντα, οι ώρες που ξόδεψε στο μάθημα, η γνώση των μαθησιακών στόχων, και η αποτελεσματικότητα των εξετάσεων, δεν σχετίζονταν με τις μετρήσεις της μάθησης των φοιτητών. Ο Chacko (1983), χρησιμοποιώντας ένα πειραματικό σχεδιασμό, βρήκε ότι η πειραματική ομάδα η οποία βαθμολογήθηκε περισσότερο αυστηρά αξιολόγησε σημαντικά χαμηλότερα την προετοιμασία του καθηγητή, τη γνώση του αντικειμένου μάθησης, και τα πνευματικά κίνητρα από ό, τι έκανε η ομάδα ελέγχου. Ο διδάσκων βαθμολογήθηκε επίσης χαμηλότερα στα προσωπικά χαρακτηριστικά όπως είναι η αυτό-εκτίμηση, η αυτοπεποίθηση και η αίσθηση χιούμορ.

Συνοψίζοντας, και οι δύο, υποστηρικτές και επικριτές έχουν βρει με συνέπεια θετικές σχέσεις μεταξύ των αντιλήψεων των φοιτητών για τη μάθησή τους και των αξιολογήσεών τους για τη διδασκαλία, ενώ βρήκαν αρνητικές σχέσεις μεταξύ των ΦΑΔ και των αντιλήψεών τους (συμπεριλαμβανομένης της σχετικής συμπεριφοράς) για την αυστηρότητα.

5.5.1 Εξήγηση του παράδοξου της αυστηρότητας

Υπάρχουν τουλάχιστο πέντε εξηγήσεις γι' αυτό το φαινόμενο. Σύμφωνα με την πρώτη, τα αποτελέσματα μπορεί να είναι μεθοδολογικά τεχνήματα. Η έρευνα στις ΦΑΔ έχει υποστεί κριτική για τη μη αρκετή χρήση στατιστικού ελέγχου (Clayson, 1994; Gaski, 1987; Howard & Maxwell, 1980; Seiver, 1983). Η σχέση της αυστηρότητας με την αξιολόγηση (ΦΑΔ) μπορεί επίσης να εξαρτάται και από το

πότε γίνεται η μέτρηση (Clayson, 2009). Όταν η αυστηρότητα μετρήθηκε κατά τη διάρκεια των μαθημάτων του εξαμήνου, η σχέση της με την αξιολόγηση (ΦΑΔ) υπήρξε αρνητική. Όταν η μέτρηση διεξήχθη μετά την ολοκλήρωση των μαθημάτων, η προαναφερθείσα σχέση υπήρξε θετική.

Σύμφωνα με τη δεύτερη ερμηνεία, τα σχετιζόμενα με την αυστηρότητα δεδομένα έχουν ταυτόχρονα γραμμικές και καμπυλόγραμμες (curvilinear) συνιστώσες, υπονοώντας ότι τα δείγματα θα μπορούσαν να δείξουν διαφορετικές σχέσεις εξαρτώμενες από το επίπεδο δυσκολίας (Marsh & Roche, 2000). Οι φοιτητές εμφανίζονται να αξιολογούν πιο ευνοϊκά την αυστηρότητα που αυτοί αντιλαμβάνονται ως την κατάλληλη (Paswan & Young, 2002). Όπως διαπιστώνει ο Centra (2003) «Αυτό που αυτά τα ευρήματα δείχνουν είναι ότι οι καθηγητές λαμβάνουν καλύτερες αξιολογήσεις όταν οι φοιτητές μπορούν και διαχειρίζονται τα μαθήματά τους» (σελ. 515).

Σύμφωνα με την τρίτη, οι φοιτητές μπορούν επίσης να δείξουν την αντίδρασή τους στην αυστηρότητα αξιολογώντας με τρόπους που θα μπορούσαν να παραποιήσουν τα στατιστικά αποτελέσματα σε οποιοδήποτε δοσμένο δείγμα. Δια μέσου μίας μεγάλης κλίμακας επιλογών όσον αφορά στα μαθήματα οι φοιτητές μπορούν να κάνουν τις επιλογές τους με βάση τις προτιμήσεις τους για την αυστηρότητα. Η Wilhelm (2004) συνέκρινε τις αξιολογήσεις των μαθημάτων, την αξία των μαθημάτων, την επιείκεια βαθμολογίας, και το φόρτο μαθημάτων ως παράγοντες επιλογής μαθημάτων από τους φοιτητές του τμήματος διοίκησης επιχειρήσεων. Τα ευρήματα της έδειξαν ότι «οι φοιτητές φαίνονταν να διαλέγουν 10 φορές περισσότερο ένα μάθημα με ένα επιεική βαθμολογητή, με όλες τις άλλες μεταβλητές να παραμένουν αμετάβλητες» (σελ 24). Ο Johnson (2003) βρήκε ουσιαστική επίδραση των πολιτικών βαθμολόγησης στις αποφάσεις των φοιτητών σχετικά με την επιλογή των μαθημάτων ακόμα και για τόσο μικρές διαφορές στο μέσο βαθμό όσο μεταξύ των B+ και B.

Σύμφωνα με την τέταρτη, η αντίληψη της αυστηρότητας και της μάθησης μπορεί να είναι αναδρομική. Εάν οι φοιτητές πιστεύουν ότι η μάθησή τους είναι καλή, τότε μπορεί να αντιλαμβάνονται το επίπεδο της αυστηρότητας ως το κατάλληλο. Εάν αντιλαμβάνονται ότι η μάθησή τους είναι χαμηλή, τότε η αυστηρότητα, οποιουδήποτε επιπέδου, μπορεί να γίνεται αντιληπτή ως μη κατάλληλη. Οι αντιλήψεις των φοιτητών για τους βαθμούς σχετίζονται επίσης με τις αντιλήψεις τους για την προσωπικότητα του διδάσκοντα (Clayson & Sheffet, 2006). Συνεπώς, εάν αυτή η πρόταση είναι σωστή, αναμένεται να υπάρχει μία σχέση ανάμεσα σε βαθμούς /

αυστηρότητα / προσωπικότητα / αξιολόγηση. Αυτή ακριβώς βρέθηκε σε μία έρευνα στους φοιτητές του μάρκετινγκ. Οι Clayson και Haley (1990) βρήκαν την αυστηρότητα, έστω και οριζόμενη με μάλλον αρνητικούς όρους, να είναι σημαντικά και θετικά σχετιζόμενη με τις αντιλήψεις των φοιτητών για τη μάθηση, αλλά αρνητικά συνδεδεμένη με την εκπαιδευτική δικαιοσύνη που καθιστούσε αρνητική τη συνολική επίδρασή της στην αξιολόγηση (ΦΑΔ). Ο Marks (2000) επανέλαβε αυτή τη μελέτη σε φοιτητές που ήταν εγγεγραμμένοι σε μαθήματα διοίκησης επιχειρήσεων και βρήκε παρόμοια αποτελέσματα. Οι Bacon και Novothy (2002) τόνισαν ότι η αυστηρότητα αλληλεπιδρά όχι μόνο με τις αντιλήψεις των φοιτητών για τον διδάσκοντα, αλλά επίσης αλληλεπιδρά με τις προσωπικότητες των ίδιων των φοιτητών. Ένας επιεικής διδάσκων θα βελτιώνει τις αξιολογήσεις για τον ίδιο μέσω της προσέλευσης φοιτητών χαμηλών επιδόσεων, αλλά σε μικρότερο βαθμό φοιτητών με υψηλά κίνητρα.

Υπάρχει και μία πέμπτη εξήγηση η οποία είναι περισσότερο ολιστική ή φιλοσοφική από τις προαναφερθείσες. Πώς εκλαμβάνεται η εκπαίδευση; Δεν είναι απαραίτητο να είναι αλήθεια ότι οι προσδοκίες των φοιτητών, για τις σχέσεις που βρέθηκαν σημαντικές από τους ερευνητές, να είναι συμβατές με τις υποκείμενες υποθέσεις αυτών των ερευνητών. Σε ένα δείγμα φοιτητών της ψυχολογίας βρέθηκε ότι το 65% ήθελε «επιτυχία» ως αποτέλεσμα του μαθήματος, αλλά μόνο το 35% το όρισε ως μάθηση (Gaultney & Cann, 2001). Σε μία άλλη μελέτη, λιγότεροι του 2% των φοιτητών πίστευαν ότι ένα μέλος τάξης πρέπει να αποτυγχάνει σε μία liberal arts τάξη εφόσον αυτός ή αυτή αποτυγχάνει να λειτουργεί ικανοποιητικά και αποτυγχάνει να ανταποκρίνεται στις ελάχιστες απαιτήσεις της τάξης, όσο διάστημα καταβάλλει προσπάθεια σε αυτή τη τάξη (J. B. Adams, 2005). Μία έρευνα σε 750 πρωτοετείς φοιτητές της διοίκησης επιχειρήσεων έδειξε ότι σχεδόν το 86% δεν εξομοίωναν την εκπαιδευτική αριστεία με τη μάθηση. Οι περισσότεροι του 96% των φοιτητών δεν ανέφεραν «γνώστης» ως μία επιθυμητή ιδιότητα του καλού διδάσκοντα (Chonko, Tanner, & Davis, 2002). Οι φοιτητές γενικά δεν πιστεύουν ότι η απαίτηση για αυστηρότητα αποτελεί σημαντικό χαρακτηριστικό ενός καλού δασκάλου (Boex, 2000; Chonko et al., 2002; Clayson, 2005b). Επιπλέον, οι φοιτητές μοιάζουν να έχουν αποσυνδεδεμένες τις αντιλήψεις τους για τους βαθμούς από τις συνήθειες μελέτης. Στις τάξεις μάνατζμεντ, καμία σχέση δεν βρέθηκε μεταξύ της παραγωγής της μελέτης και της παραγωγής της μάθησης, «που σημαίνει ότι οι φοιτητές δεν σκέφτηκαν απαραίτητα ότι έμαθαν περισσότερα στα μαθήματα που διάβασαν περισσότερο

(Stapleton & Murkison, 2001, p. 281). Ο Johnson (2003) γράφει, «Σε κάθε περίπτωση, η έλλειψη μίας σχέσης μεταξύ βαθμών και συνηθειών μελέτης προτείνει ότι η κινητήριος επίδραση των βαθμών στην προσπάθεια των φοιτητών δεν είναι πλήρως κατανοήσιμη» (σελ. 79).

5.5.2 Συνοψίζοντας τα πορίσματα για τη σχέση μάθησης και αξιολογήσεων

Ο για 60 χρόνια διάλογος πάνω στη σχέση της μάθησης και των ΦΑΔ, ουσιαστικά είναι μία συζήτηση γύρω από την εγκυρότητα των εργαλείων. Ο Clayson (2009) ολοκληρώνοντας την επισκόπηση της βιβλιογραφίας και τη μετά-ανάλυσή της διαπιστώνει ότι, η διττή φύση του διαλόγου μπορεί να είναι άστοχη. Το ερώτημα δεν έχει να κάνει με το αν υφίσταται η σχέση, αλλά περισσότερο με το πότε υπάρχει η σχέση.

Οποιαδήποτε γενική εξήγηση που προσπαθεί να εξηγήσει τις αντιφάσεις και τα πρότυπα που επισημαίνονται στην ανασκόπηση της βιβλιογραφίας και στη μετά-ανάλυσή τους πρέπει να συνδυάζει πέντε πορίσματα (σελ. 26).

Πρώτο, οι αντιλήψεις των φοιτητών για του βαθμούς τους εμφανίζονται να σχετίζονται με τις αξιολογήσεις που κάνουν τόσο για το μάθημα όσο και για τον καθηγητή. Έχουν βρεθεί τόσο η επίδραση *επιείκειας* όσο και η επίδραση *αμοιβαιότητας*. Η έρευνα φαίνεται να υποδεικνύει ότι αυτές οι επιδράσεις τροποποιούνται από μία ποικιλία επιδράσεων ανεξάρτητων από τη «μάθηση». Είναι σημαντικό να θυμόμαστε ότι η σχέση των βαθμός / αξιολόγηση δεν είναι ισότιμη με τη σχέση των μάθηση / ΦΑΔ (Clayson, 2007).

Δεύτερο, η έρευνα σχεδόν παγκοσμίως βρίσκει μία αρνητική σχέση μεταξύ αυστηρότητας και μάθησης, στις ΦΑΔ. Οι φοιτητές φαίνεται να συσχετίζουν την αυστηρότητα με τα αρνητικά χαρακτηριστικά του διδάσκοντα που παρακάμπτουν τις θετικές σχέσεις μάθησης.

Τρίτο, όσο πιο αντικειμενικά γίνεται η διαδικασία μέτρησης, τόσο για τη μάθηση όσο και για τις ΦΑΔ, τόσο περισσότερο η σχέση μάθηση / ΦΑΔ μειώνεται. Οι πιο πρόσφατες και μεγαλύτερες μελέτες (Johnson, 2003; Weinberg et al., 2007) βρήκαν ότι αν και μία θετική σχέση υπάρχει μεταξύ των αντιλήψεων των φοιτητών για τη μάθηση και της αξιολόγησης, αυτή η σχέση δεν μπορεί να βρεθεί όταν γίνεται χρήση περισσότερο αντικειμενικών μετρήσεων της μάθησης. Όσο η στατιστική πολυπλοκότητα αυξάνεται με το χρόνο, η αναφερθείσα σχέση μάθησης / ΦΑΔ προκύπτει γενικά πιο αρνητική. Σχεδόν 40 χρόνια έχουν περάσει από τότε που ελήφθησαν τα θετικά στοιχεία από τη μελέτη των Sullivan και Skanes (1974). Καμία

μελέτη δεν βρέθηκε μετά το 1990 που να δείχνει μία θετική σημαντική σχέση μεταξύ της μάθησης και των ΦΑΔ. Ο τρόπος που η μάθηση γίνεται λειτουργική εμφανίζεται να έχει σημαντικές συνέπειες. Σε κάποιους επιστημονικούς κλάδους, οι βαθμοί τυπικά προκύπτουν από μία εργασία που έχει μία αντικειμενική «σωστή» απάντηση, σε αντίθεση με τις μετρήσεις σε ένα πιο υποκειμενικό πεδίο μελέτης. Μερικοί πιστεύουν ότι τα μαθήματα που χρησιμοποιούν ποσοτικές εφαρμογές τυπικά απαιτούν περισσότερες γνωστικές δεξιότητες κατά τη διάρκεια του ελέγχου από ό, τι οι περιοχές που μπορεί να βασίζονται περισσότερο σε μετρήσεις που δίνουν έμφαση στη μνημονική ανατροφοδότηση (McKeachie, 1987). Στη μελέτη των Rodin και Rodin (1972) στην οποία βρέθηκε η μεγαλύτερη αρνητική σχέση μεταξύ μάθησης και ΦΑΔ, όχι μόνο ελέγχθηκε η κατανόηση, αλλά ελέγχθηκε σε επίπεδο παραδείγματος (paradigm level). Η μεγαλύτερη θετική σχέση (0.91) βρέθηκε από το Frey και τους συνεργάτες του (1975) οι οποίοι χρησιμοποίησαν ένα σχετικά μικρό δείγμα μικρών εισαγωγικών μαθημάτων ψυχολογίας.

Τέταρτο, λόγω των διαφορών των επιστημονικών κλάδων η μετά-ανάλυση έδειξε ότι το ακαδημαϊκό επιστημονικό πεδίο αποτελεί μία σημαντική μεταβλητή. Αυτή έχει παραβλεφτεί σε μεγάλος μέρος της βιβλιογραφίας. Αν και έχει υπάρξει μία προσεκτική προσπάθεια να ελεγχθούν εξωτερικές μεταβλητές, δεν διευκρινίζονται πάντοτε οι ακαδημαϊκές περιοχές των μαθημάτων στις μελέτες των ΦΑΔ.

Πέμπτο, κάποιες διαφορές στα ευρήματα εμφανίζονται πραγματικές, και όχι εντελώς τεχνητά από διαφορετική μεθοδολογία. Υπάρχουν λίγα στοιχεία που να αποδεικνύουν ότι όλες οι διαφορές είναι μόνο καταστατικές ή οφείλονται στη μεθοδολογία. Για παράδειγμα, δεν υπάρχουν αρκετά μεγάλα λάθη στη μετά-ανάλυση του Cohen που να ακυρώνουν το πόρισμά του ότι η μάθηση σχετίζεται θετικά με τις ΦΑΔ. Από την άλλη, δεν βρέθηκαν λάθη στην εργασία του Johnson (2003) ή του Weinberg και των συνεργατών του (2007) τα οποία ακυρώνουν τα συμπεράσματά τους ότι η μάθηση και οι ΦΑΔ δεν σχετίζονται ή μπορεί ακόμη και να σχετίζονται αρνητικά.

5.6 Ανάπτυξη μέσων αξιολόγησης

Η συλλογή ερωτήσεων

Σύμφωνα με τους Marsh και Roche (1997) η επιλογή των ερωτήσεων αξιολόγησης συνιστά ένα σημαντικό παράγοντα για την επιβεβαίωση ότι οι αξιολογήσεις είναι έγκυρες μετρήσεις της αποτελεσματικότητας της διδασκαλίας και ότι η εγκυρότητα και η χρησιμότητα των αξιολογήσεων της διδασκαλίας από τους φοιτητές εξαρτάται

από το περιεχόμενο και το πεδίο κάλυψης των δηλώσεων (σελ. 187). Παρ' όλα αυτά οι προαναφερθέντες συγγραφείς γράφουν ότι «στην πράξη, τα περισσότερα εργαλεία βασίζονται σε ένα μείγμα από λογικές και πραγματικές θεωρήσεις, περιλαμβάνοντας περιστασιακά κάποια ψυχομετρική απόδειξη όπως είναι η αξιοπιστία ή η παραγοντική ανάλυση» (σελ. 1187). **Οι Ory και Ryan (2001) επισημαίνουν ότι «πολλές από τις φόρμες αξιολόγησης των μαθημάτων που χρησιμοποιούνται σήμερα έχουν αναπτυχθεί από άλλες υπάρχουσες φόρμες χωρίς πολλή σκέψη πάνω σε θέματα θεωρίας ή δομής» (σελ. 32).** Ωστόσο, η ανακριβής συλλογή ερωτήσεων και ανάπτυξη ερωτηματολογίου παραμένει ένα σημαντικό εμπόδιο στην αποτίμηση της εγκυρότητας (Gravestock, & Gregor-Greenleaf, 2008, σελ. 32).

Ψυχομετρικοί έλεγχοι

Η διατύπωση, η σειρά και η κλίμακα που χρησιμοποιούνται στις ερωτήσεις μπορούν από μόνες τους να ασκήσουν σημαντική επίδραση στις αξιολογήσεις. Συνεπώς ένα σημαντικό στοιχείο επιβεβαίωσης της εγκυρότητας των φορμών αξιολόγησης είναι ο ψυχομετρικός έλεγχος (Gravestock, & Gregor-Greenleaf, 2008, σελ. 33).

Επανεξέταση εργαλείου

Ο καθορισμός της βέλτιστης συχνότητας με την οποία αναθεωρούνται οι αξιολογήσεις είναι θέμα επίτευξης της ισορροπίας ανάμεσα στην εξασφάλιση ότι οι δηλώσεις αντιπροσωπεύουν τις τρέχουσες παιδαγωγικές και θεσμικές πρακτικές και προτεραιότητες και στη διασφάλιση ότι οι δηλώσεις αξιολόγησης συλλέγονται και αξιολογούνται αρκετά προσεκτικά ώστε να πληρούν τα κατασκευαστικά και ψυχομετρικά κριτήρια εγκυρότητας (Ory & Ryan, 2001; Theall & Franklin 2001).

5.7 Έκθεση και ερμηνεία των αποτελεσμάτων αξιολόγησης

Οι d' Apollonia και Abrami (1997) επισημαίνουν ότι «πολλοί ειδικοί στην αξιολόγηση των μελών ΔΕΠ εκτιμούν ότι η εγκυρότητα των συγκεντρωτικών αξιολογήσεων που βασίζεται στις βαθμολογίες από τους φοιτητές απειλείται από τη μη κατάλληλη: συλλογή στοιχείων, ανάλυση, έκθεση και ερμηνεία (σελ. 1203, εργασίες που αναφέρονται από Arreola, 1995; Theall, 1994; Franklin & Theall, 1990). Παρόμοια, αυτοί υποστηρίζουν ότι οι ειδικές ερωτήσεις που χρησιμοποιούνται, η χρήση γενικών ερωτήσεων σε σχέση με τις βαθμολογίες ενός παράγοντα, η πιθανή στρέβλωση (biasing) των μεταβλητών, και ούτω καθεξής, αποτελούν σχετικά μικρά προβλήματα. Το μεγαλύτερο πρόβλημα με την εγκυρότητα είναι «η χρήση των βαθμολογήσεων από το προσωπικό των επιτροπών και των μελών της διοίκησης» (σελ. 1222). Οι Franklin και Theall (1989) καταλήγουν σε παρόμοια

συμπεράσματα. Η κατάλληλη έκθεση των στοιχείων αξιολόγησης και η πρόσφορη και καταρτισμένη ερμηνεία αυτών των στοιχείων είναι σημαντική, ακόμα και κεντρική, για την εγκυρότητα των αξιολογήσεων.

5.7.1 Έκθεση αποτελεσμάτων Αξιολόγησης

Επιλέγοντας βαθμολογίες και σύνθεση βαθμολογιών

Ο Abrami (2001) επισημαίνει ότι οι βαθμολογίες των μεμονωμένων δηλώσεων στις φόρμες αξιολόγησης μπορούν και πρέπει να χρησιμοποιούνται για διαμορφωτικούς σκοπούς που βελτιώνουν την ποιότητα της διδασκαλίας. Ωστόσο, η παροχή αναφορών για κάθε δήλωση χωριστά στα μέλη ΔΕΠ θα μπορούσε να αποδειχτεί δύσκολη στο να αξιολογηθεί. Για την αντιμετώπιση, ο Algozzine και οι συνεργάτες του (2004) έχουν προτείνει οι δηλώσεις να ομαδοποιούνται ανά κατηγορία συμπεριφοράς διδασκαλίας ή στοιχείων της σειράς των μαθημάτων και τα μέλη ΔΕΠ να λαμβάνουν μόνο τις βαθμολογίες της κατηγορίας μάλλον παρά τις βαθμολογίες κάθε δήλωσης.

Η διδασκαλία είναι μία πολύπλευρη δραστηριότητα και γι' αυτό μερικοί μελετητές (βλέπε Marsh) επιμένουν στην ανάγκη για μία φόρμα αξιολόγησης πολλαπλών διαστάσεων για να αξιολογήσουν πλήρως την αποτελεσματικότητα της διδασκαλίας. Ωστόσο, ο Abrami (2001) έχει υποστηρίξει ότι ενώ η διδασκαλία είναι μία πολύπλευρη δραστηριότητα, οι συγκεντρωτικές αποφάσεις σχετικά με την αποτελεσματικότητα της διδασκαλίας δεν είναι, και ότι αυτό που είναι χρήσιμο για τα μέλη της διοίκησης είναι ένα αποτέλεσμα που εκπροσωπεί ένα ευρύ και περιεκτικό ορισμό της αποτελεσματικότητας της διδασκαλίας. Εν τούτοις οι μελετητές διαφωνούν στο εάν το προαναφερθέν αποτέλεσμα πρέπει να προέρχεται από α) ένα ευρύ ερώτημα που ζητά από τους φοιτητές να βαθμολογήσουν μία σειρά μαθημάτων ή ένα καθηγητή γενικά (μία θέση που την υπερασπίστηκε ο Abrami, 2001), ή β) αποτελεί ένα αποτέλεσμα το οποίο υπολογίζεται από το μέσο όρο των διαφόρων διαστάσεων της διδασκαλίας, σταθμισμένων σύμφωνα με τις προτεραιότητες του ιδρύματος (προτάθηκε από τους d' Apollonia & Abrami, 1997).

Οι Cashin και Downey (1992) συμφωνούν ότι οι βαθμολογίες των μεμονωμένων διαστάσεων χρησιμοποιούνται ελάχιστα από τους διοικούντες για τους σκοπούς της συγκεντρωτικής αξιολόγησης. Στη μελέτη τους σε στοιχεία από 17,183 σειρές μαθημάτων από 105 ιδρύματα βρήκαν ότι οι σύντομες, οικονομικές φόρμες (π. χ με γενικές ερωτήσεις) υπήρξαν ικανές να καλύψουν πολλές πληροφορίες που

χρειαζόντουσαν για τους συγκεντρωτικούς σκοπούς. Ωστόσο, πολλές λεπτομερειακές ερωτήσεις και αποτελέσματα μπορούν να βοηθηθούν με τις συγκεντρωτικές αξιολογήσεις.

5.7.2 Προκλήσεις στην ερμηνεία και χρήση για συγκεντρωτικούς σκοπούς

Στατιστική αξία ή αξιολόγηση στοιχείων

Ο Scriven (1995) υπογραμμίζει μία σειρά από συνήθη λάθη που σχετίζονται με τη χρήση των στοιχείων της αξιολόγησης των μαθημάτων, όπως τη χρήση των αποτελεσμάτων χωρίς αναφορά στη κατανομή (distribution), την αντιμετώπιση των μικρών διαφορών ως σημαντικών, και τη χρήση των στοιχείων αξιολόγησης ως πρωταρχικό εργαλείο στη συγκεντρωτική και διαμορφωτική αξιολόγηση. Ο Abrami (2001) διατυπώνει τις ανησυχίες του σχετικά με τις παρερμηνείες και την κακή χρήση των στοιχείων αξιολόγησης για αποφάσεις προσωπικού (personnel decisions). Αυτός προειδοποιεί τη διοίκηση να μην υπερεκτιμά τις μικρού μεγέθους διαφορές, ειδικότερα εάν αυτή δεν γνωρίζει στατιστική ανάλυση. Παρόμοια, οι McKeachie (1997), d' Arrolonia και Abrami (1997) και Wagenaar (1995) προειδοποιούν τα μέλη της διοίκησης και συνιστούν περισσότερο να υπερεκτιμούν την ακρίβεια των αποτελεσμάτων αξιολόγησης παρά να κάνουν χρήση ακατέργαστων αποτελεσμάτων που αναφέρονται με ένα ή δύο δεκαδικά ψηφία, για ερμηνευτικούς σκοπούς. Τα μέλη της διοίκησης πρέπει να ταξινομούν τα αποτελέσματα σε μία από τις τρεις κατηγορίες: εξαιρετικά, επαρκή, απαράδεκτα. Αυτές οι ευρείες κατηγορίες θα μπορούσαν να μετριάσουν κάθε διακύμανση ή προκατάληψη που εισάγεται με τα χαρακτηριστικά των μαθημάτων ή των επιστημονικών πεδίων με σκοπό τη διεξαγωγή θεμιτής σύγκρισης μεταξύ μαθημάτων ή διδασκόντων.

5.7.3 Χρησιμοποιώντας τις αξιολογήσεις για συγκριτικούς σκοπούς

Συγκρίνοντας τις σειρές των μαθημάτων

Πολλοί μελετητές συμφωνούν ότι τα αποτελέσματα της αξιολόγησης παρουσιάζουν καλύτερα την αποτελεσματικότητα της διδασκαλίας ενός διδάσκοντα εάν πιθανές διακυμάνσεις στις αξιολογήσεις που οφείλονται στα χαρακτηριστικά των μαθημάτων μετριάζονται με την ανάπτυξη μιας μέσης βαθμολογίας κατά μήκος πολλαπλών μαθημάτων. Ο Abrami (2001) προτείνει, κάθε καθηγητής να προσδιορίζει ένα σύνολο μαθημάτων έτσι ώστε να επέρχεται ισορροπία ανάμεσα στο χαμηλότερο και ανώτερο επίπεδο μαθημάτων, και ανάμεσα στα επιλεγόμενα και τα υποχρεωτικά μαθήματα.

Ο Franklin (2001) τονίζει ότι «ο αριθμός των μαθημάτων που απαιτούνται για να κατασκευάσουν ένα μέσο αποτελέσματα αυξάνει καθώς το μέγεθος της τάξης

μικραίνει. Γενικά, προτείνονται πέντε ή περισσότερα μαθήματα στις περισσότερες περιπτώσεις, αν και οι πολύ μικρές τάξεις βασικά χρειάζονται περισσότερα. Για παράδειγμα, τάξεις με τουλάχιστο πέντε φοιτητές μπορεί να χρειάζονται είκοσι μαθήματα (sections) για σύγκριση» (σελ. 92).

Ο Abramí προειδοποιεί ότι πρέπει να αναπτυχθούν καθαρές πολιτικές που να καθορίζουν ποια μαθήματα θα περιλαμβάνονται ή θα αποκλείονται από την συνολική αξιολόγηση οι οποίες πρέπει να εφαρμόζονται ισότιμα και με συνέπεια.

Συγκρίνοντας τους διδάσκοντες

Η έρευνα προτείνει ότι εξαιτίας των διαφορών μεταξύ των υφών (στυλ) διδασκαλίας και μεταξύ των στόχων των επιστημονικών πεδίων, οι αξιολογήσεις της διδασκαλίας – εφόσον συγκριθούν σε όλα –πρέπει να συγκρίνονται μόνο μεταξύ των διδασκόντων στα ίδια ή σε παρόμοια επιστημονικά πεδία. Οποιοσδήποτε συγκρίσεις μεταξύ των διδασκόντων πρέπει επίσης να παρέχουν άφθονες ευκαιρίες για συναφούς πλαισίου στοιχεία, και πρέπει να επιβεβαιώνουν ότι τα μαθήματα που συγκρίνονται μοιράζονται παρόμοια χαρακτηριστικά.

Η Neumann (2001) υπογραμμίζει τους διαφορετικούς ορισμούς της αποτελεσματικής διδασκαλίας σε διαφορετικά επιστημονικά πεδία. Ως εναλλακτική στη διεξαγωγή αξιολογήσεων γενικά κατά μήκος διαφορετικών τμημάτων, αυτή επισημαίνει την εργασία άλλων μελετητών οι οποίοι προτείνουν «την ανάπτυξη εργαλείων εξειδικευμένων στην αξιολόγηση της διδασκαλίας του κάθε επιστημονικού πεδίου» ή «την ανάπτυξη ενός αριθμού εργαλείων τα οποία να αντανακλούν την ποικιλία των φιλοσοφιών διδασκαλίας που ταιριάζουν στα επιστημονικά πεδία» (σελ. 143).

Ακόμα σε επίπεδο τμήματος, διακυμάνσεις ανά διδάσκοντα μπορεί να επηρεάσουν την εγκυρότητα των συγκρινόμενων στοιχείων (Addison, Best & Warrington, 2006).

Για παράδειγμα, οι Theall και Franklin (2001) συζητούν ότι ένα ειδικό παράδειγμα των προκαταλήψεων φύλου θα μπορούσε να εξηγηθεί από το γεγονός ότι τα περισσότερα από τα υποχρεωτικά, χαμηλότερου-επιπέδου μαθήματα στο τμήμα, τα δίδασκαν γυναίκες. Αυτό που παρουσιάζεται να είναι μία προκατάληψη φύλου στις αξιολογήσεις αντανακλά το γεγονός ότι συγκεκριμένοι διδάσκοντες δίδασκαν μαθήματα με συγκεκριμένα χαρακτηριστικά. Αυτό ενισχύει την ανάγκη να υπάρξει διασφάλιση για το ότι τα στοιχεία παρουσιάζονται από μία αντιπροσωπευτική ή ισοδύναμη επιλογή μαθημάτων.

5.8 Διαβούλευση ή συμβουλευτική ανατροφοδότηση

Οι Penny και Coe (2004) στη μετά ανάλυσή τους για την αποτελεσματικότητα της διαβούλευσης πάνω στις αξιολογήσεις από τους φοιτητές αναφέρουν ότι η χρήση της συμβουλευτικής ανατροφοδότησης (consultative feedback) ως μία αποτελεσματική στρατηγική βελτίωσης της διδασκαλίας έχει λάβει αρκετή εμπειρική υποστήριξη. Ωστόσο έχουν υπάρξει επίμονα ερωτήματα σχετικά με το τι λειτουργεί και τι δεν λειτουργεί στη διαβούλευση για τη διδασκαλία (instructional consultation). Αντικείμενο της μετά-αναλυτικής εξέτασης των Penny και Coe υπήρξε η αναγνώριση των στρατηγικών και των πρακτικών οι οποίες είναι δυνητικά οι πιο σημαντικές για την επιτυχία της συμβουλευτικής ανατροφοδότησης. Ένα από τα βασικά ευρήματά τους υπήρξε η υποστήριξη μιας θετικής μέτριας σχέσης μεταξύ της συμβουλευτικής ανατροφοδότησης και της αποτελεσματικότητας της διδασκαλίας. Τα αποτελέσματα ήταν συγκρίσιμα με εκείνα προηγούμενων μετά-αναλύσεων που διεξήγαγαν οι Cohen (1980) και οι Menges και Brinko (1986).

Το πρώτο και κυριότερο που παρουσίασε αυτή η μετά-ανάλυση υπήρξε η απόδειξη που προτείνει μία ποικιλία από φόρμες κατά τη διεξαγωγή της συμβουλευτικής ανατροφοδότησης. Συγκεκριμένα ότι, δεν υπάρχει μόνο μία προσέγγιση η οποία να μπορεί να εφαρμόζεται σε κάθε πλαίσιο όταν λαμβάνει χώρα διαβούλευση με τους καθηγητές πάνω στις αξιολογήσεις από τους φοιτητές. Ωστόσο οι προσεγγίσεις μπορεί να μην είναι το ίδιο αποτελεσματικές. Αυτό το εύρημα συμφωνεί με την επικρατούσα άποψη στη βιβλιογραφία ότι δεν υφίσταται «σωστός τρόπος» παροχής της διαβούλευσης. Από την άλλη πλευρά, έρχεται σε αντίθεση με την άποψη ότι δεν υπάρχει «ένα είδος συμβουλευτικής επί της διδασκαλίας που να είναι πιο αποτελεσματικός από τους άλλους» (Brinko, 1997, σελ. 3).

Οι Penny και Coe (2004, σελ. 244) διαπιστώνουν ότι η χρήση της συμβουλευτικής ανατροφοδότησης μπορεί να συμβάλει στη βελτίωση της διδασκαλίας στο πανεπιστήμιο μόνο με την αξιοποίηση ορισμένων πρακτικών διαβούλευσης και στρατηγικών. Αυτό σημαίνει ότι **τα σημαντικά στοιχεία της συμβουλευτικής ανατροφοδότησης ξεπερνούν την απλή υποστήριξη με όρους «συζήτησης» σχετικά με τις βαθμολογίες από τους φοιτητές και τη διατύπωση συμβουλών για το πώς να βελτιωθεί η διδασκαλία.** Η συμβουλευτική ανατροφοδότηση πρέπει να ειδωθεί ως κάτι περισσότερο από ερμηνεία αξιολογήσεων για ανατροφοδότηση και παροχή βοήθειας στους καθηγητές για να σχεδιάζουν και να εφαρμόζουν στρατηγικές βελτίωσης. Η συμβουλευτική ανατροφοδότηση μάλλον πρέπει να αντιμετωπίζεται ως

μία μορφή συνεργατικής έρευνας δράσης (action research) η οποία στοχεύει στο να βοηθήσει τους καθηγητές να κατανοήσουν καλύτερα τη φύση της εργασίας τους και το πώς μπορούν να βελτιώνουν την ποιότητα της πρακτικής τους με ένα τρόπο γεμάτο νόημα, αναστοχαζόμενοι τη διδασκαλία τους μέσα από τα μάτια των φοιτητών τους. Σύμφωνα με τους μελετητές, η συμμετοχή των καθηγητών στη συμβουλευτική ανατροφοδότηση πρέπει να τους οδηγεί σε ένα νέο τρόπο σκέψης για τη διδασκαλία τους, τους φοιτητές τους και την αλληλεπίδραση διδασκαλίας-μάθησης.

Οκτώ στρατηγικές διαβούλευσης βρέθηκαν ως οι πλέον σημαντικές για τη μεγαλύτερη αποτελεσματική συμβουλευτική ανατροφοδότηση. Σημειώνεται ότι, αν και η απόδειξη γι' αυτές μπορεί να μην είναι πολύ ισχυρή, παρ' όλα αυτά αυτή αντιπροσωπεύει την καλύτερη διαθέσιμη γνώση μέχρι σήμερα που μπορεί να αξιοποιηθεί στην καθοδήγηση των πρακτικών και της πολιτικής. Οι οκτώ πρακτικές είναι οι εξής (σελ. 245):

1. Ενεργός συμμετοχή των καθηγητών στη διαδικασία μάθησης.
2. Χρήση πολλαπλών πηγών πληροφόρησης.
3. Αλληλεπίδραση με ομοβάθμιους.
4. Επαρκής χρόνος για διάλογο και αλληλεπίδραση.
5. Χρήση των αυτό-αξιολογήσεων των καθηγητών.
6. Χρήση της υψηλής-ποιότητας πληροφορίας για ανατροφοδότηση.
7. Εξέταση του σχεδιασμού της διδασκαλίας.
8. Καθορισμός των στόχων βελτίωσης.

Οι μελετητές διευκρινίζουν ότι όσο σημαντικές κι αν φαίνονται να είναι αυτές οι πρακτικές και στρατηγικές, η επίδραση από τη χρήση τους εξαρτάται από την ποιότητα της αλληλεπίδρασης μεταξύ του συμβούλου και του καθηγητή. Ένα από τα πλέον σημαντικά ευρήματα αυτής της μετά-ανάλυσης, το οποίο συμπληρώνει τη βιβλιογραφία του θέματος, υπήρξε η παροχή στοιχείων που αποδεικνύουν τη μετατόπιση του ενδιαφέροντος από τον σύμβουλο ως παράγοντα ανατροφοδότησης συμβουλευτικού περιεχομένου, στις στρατηγικές και στις πρακτικές οι οποίες μπορούν να μεγιστοποιήσουν τα οφέλη της διαβούλευσης για τους εκπαιδευτικούς.

B: Ζητήματα Αξιοπιστίας

5.9 Χαρακτηριστικά υποβάθρου & αξιολόγηση

Αυτή η ενότητα εξετάζει την υπάρχουσα γνώση τη σχετική με τα χαρακτηριστικά **άλλων μεταβλητών**, πέρα από την αποτελεσματικότητα της διδασκαλίας, οι οποίες θα μπορούσαν θεωρητικά να ασκούν επίδραση στα αποτελέσματα της αξιολόγησης / βαθμολογίας από τον φοιτητή.

Κάποιοι συγγραφείς, μεταξύ των οποίων ο Brandenburg με τους συνεργάτες του (1977) και ο Tatro (1995), έχουν υποστηρίξει ότι οι μεταβλητές υποβάθρου μπορούν να παράγουν αξιοσημείωτες αλλαγές στις βαθμολογίες για τον διδάσκοντα. Σύμφωνα με τον Wachtel (1998, σελ. 194) η ύπαρξη και μόνο μίας συσχέτισης μεταξύ μίας μεταβλητής υποβάθρου και των αποτελεσμάτων της αξιολόγησης / βαθμολογίας δεν συνιστούν απαραίτητα μία προκατάληψη (bias) ή ένα κίνδυνο για την εγκυρότητα των βαθμολογιών από τους φοιτητές. Για παράδειγμα, εάν η αναμενόμενη από τον φοιτητή βαθμολογία του στο μάθημα βρέθηκε να συσχετίζεται με τη αξιολόγηση την οποία ο φοιτητής κάνει στον διδάσκοντα, αυτό δεν σημαίνει απαραίτητα ότι ένας διδάσκοντας μπορεί να εξασφαλίσει υψηλότερη βαθμολογία απλώς βάζοντας υψηλότερους βαθμούς. Εναλλακτικές εξηγήσεις περιλαμβάνουν την πιθανότητα ότι περισσότερο αποτελεσματική διδασκαλία θα εμπνεύσει τους φοιτητές να εργαστούν σκληρότερα και να κερδίσουν καλύτερους βαθμούς. Προσεκτική έρευνα είναι αναγκαία με σκοπό να καθορίσει τη φύση της σχέσης μεταξύ της μεταβλητής υποβάθρου και των βαθμολογιών των φοιτητών.

Μεγάλη προσοχή έχει δοθεί στην έρευνα για την ευρεία ακτίνα των μεταβλητών που μπορούν ή δεν μπορούν να επηρεάσουν την εγκυρότητα των στοιχείων αξιολόγησης από τους φοιτητές. (Zabaleta, 2007; Addison et al., 2006; Algozzine et al., 2004; Ory & Ryan, 2001; Theall & Franklin, 2001; Ali & Sell, 1998; Wachtel, 1998; Cashin, 1995, 1988). Η πληροφορία σε αυτό το τμήμα έχει αντληθεί κυρίως από αυτές τις εκθέσεις.

Οι μεταβλητές που συζητούνται στη βιβλιογραφία εμπίπτουν σε τέσσερεις κατηγορίες: συνθήκες διαχείρισης, χαρακτηριστικά μαθημάτων, χαρακτηριστικά καθηγητή και χαρακτηριστικά φοιτητών. Στη συνέχεια παρουσιάζονται οι ειδικοί παράγοντες που εμπίπτουν σε κάθε μία κατηγορία.

Πρέπει να τονισθεί ότι κάθε επίδραση στις συνολικές βαθμολογίες/αξιολογήσεις από κάθε μία από αυτές τις συγκεκριμένες μεταβλητές ακόμα και όταν είναι στατιστικά

σημαντική, είναι σχεδόν πάντα πολύ μικρή – συχνά αλλάζει τις βαθμολογίες/αξιολογήσεις λιγότερο από ένα δέκατο του 1%. Εξαιτίας αυτού, ακόμα και όταν αυτές οι μεταβλητές ασκούν κάποια επίδραση στα αποτελέσματα αξιολόγησης, η εγκυρότητα μπορεί σχεδόν πάντα να διατηρηθεί εκθέτοντας αποτελέσματα όχι με πάνω από ένα δεκαδικό ψηφίο ή ως μέρος μιας ευρείας κατηγορίας όπως περιγράφεται στο εδάφιο 5.9.1.

5.9.1 Συνθήκες διαχείρισης της έρευνας

Ο Seldin (1993) προειδοποιεί για την ακύρωση ενός καλού εντύπου αξιολόγησης λόγω ελλείψεων στις διαδικασίες διαχείρισης, στο χρονοδιάγραμμα ή στις οδηγίες. Τέτοιου τύπου χαρακτηριστικά διαχείρισης αναφέρονται στη συνέχεια.

Ο χρόνος διεξαγωγής της αξιολόγησης

Ο Feldman (1979) βρήκε ότι ο χρόνος διεξαγωγής των αξιολογήσεων από τους φοιτητές είτε είναι στο μέσο των μαθημάτων, είτε πριν, ή κατά τη διάρκεια, ή μετά την τελική αξιολόγηση, δεν έχει καμία επίδραση στα αποτελέσματα. Ο Frey (1976) βρήκε ότι οι βαθμολογίες που συλλέγονται κατά τη διάρκεια της τελευταίας εβδομάδας των μαθημάτων δεν υπήρξαν ιδιαίτερα διαφορετικές από εκείνες που συλλέχθηκαν στην πρώτη εβδομάδα του διαδεχόμενου όρου, αλλά ωστόσο ανησυχεί για την ενδεχόμενη επίδραση που μπορεί να έχει στις βαθμολογίες μια επικείμενη τελική εξέταση. Οι Marsh και Overall (1980) βρήκαν ότι οι ενδιάμεσου χρόνου (mid-term) βαθμολογίες ήταν υψηλά σχετιζόμενες με εκείνες προς το τέλος (end-of-term), αλλά ο L' Hommedieu και οι συνεργάτες του (1990) υποστηρίζουν τη μη ύπαρξη ισοδυναμίας μεταξύ των ενδιάμεσων και προς το τέλος βαθμολογιών και προτείνουν περισσότερη έρευνα πάνω στην επίδραση του χρόνου διεξαγωγής στις βαθμολογίες από τους φοιτητές. Η Aleamoni (1981) υποστηρίζει ότι εάν η διαχείριση γίνει πριν ή μετά από μία εξέταση τότε τα αποτελέσματα μπορεί και να διαφοροποιηθούν. Σύμφωνα με την ισχύουσα εμπειρία δεν είναι ασύνηθες ένα τμήμα να απαιτεί η διαχείριση των αξιολογήσεων να γίνεται σε μία συγκεκριμένη ημέρα ή εβδομάδα, η οποία θα μπορούσε να είναι μη θεμιτή από τους διδάσκοντες που έχουν μία εξέταση εκείνη την εβδομάδα. Ο Braskamp με τους συνεργάτες του (1984) και ο Ory (2001) υποστηρίζουν ότι οι βαθμολογίες μπορεί να είναι χαμηλότερες εάν η διαχείριση γίνει κατά τη διάρκεια της τελικής εξέτασης, και συνιστούν (το ίδιο και οι d'Apollonia & Abrami, 1997) η διαχείριση των αξιολογήσεων να γίνεται τις τελευταίες δύο εβδομάδες του ακαδημαϊκού χρόνου (αλλά όχι την τελευταία ημέρα του μαθήματος). Ο Seldin (1989) συμφωνεί με αυτή την πρόταση όταν οι βαθμολογίες πρόκειται να

χρησιμοποιηθούν συγκεντρωτικά, αλλά επιμένει ότι η αξιολόγηση πρέπει να γίνεται νωρίτερα, περίπου τέσσερεις ή έξι εβδομάδες μετά την έναρξη των μαθημάτων, όπου οι βαθμολογίες μπορούν να χρησιμοποιηθούν διαμορφωτικά.

Ανωνυμία των φοιτητών που βαθμολογούν

Οι Feldman (1979) και Blunt (1991) αναφέρουν ότι οι φοιτητές τείνουν να δίνουν κατά τι πιο υψηλές βαθμολογίες όταν είναι αναγνωρίσιμοι (επώνυμοι) συγκριτικά με εκείνους που παραμένουν στην ανωνυμία. Ωστόσο ο Feldman υποστηρίζει ότι, υπάρχουν άλλες συνθήκες οι οποίες μπορεί να αλληλεπιδρούν με την ανωνυμία, όπως όταν οι βαθμολογίες δίνονται πριν ή μετά που οι φοιτητές μαθαίνουν τους βαθμούς τους, όταν οι βαθμολογίες δίνονται σε «ειδικές πειραματικές περιόδους» στις οποίες λέγεται στους φοιτητές ότι οι βαθμοί θα χρησιμοποιηθούν για ερευνητικούς σκοπούς μόνο, και όταν οι φοιτητές πιστεύουν ότι υπάρχει η πιθανότητα μιας «αντιπαράθεσης» με τον διδάσκοντα (Abrami et al. 1976). Ο Blunt (1991) επίσης εκφράζει ανησυχία για το κατά πόσο ή όχι οι φοιτητές νιώθουν ότι αυτοί μπορούν να εμπιστευτούν τα μέλη ΔΕΠ και τη διοίκηση για τις διαβεβαιώσεις περί ανωνυμίας και εμπιστευτικότητας. Οι περισσότεροι συγγραφείς (για παράδειγμα Braskamp et al. 1984; Centra, 1993; McCallum, 1984; Cashin, 1995) συνιστούν οι φοιτητές να βαθμολογούν ανώνυμα.

Η παρουσία του διδάσκοντα στην τάξη

Ο Feldman (1979) επίσης αναφέρει ότι οι βαθμολογίες είναι κατά τι υψηλότερες όταν ο διδάσκων που αξιολογείται είναι παρόν στο χώρο που γίνεται η συμπλήρωση των φορμών. Οι περισσότεροι συγγραφείς, μεταξύ των οποίων οι Braskamp και Ory (1994), Centra (1993, p. 78), Eble (1970, p. 24) και Scriven (1981, p. 255), συνιστούν τη χρήση ενός τρίτου υπευθύνου στη συλλογή φορμών μαζί με διαβεβαιώσεις ότι η ανωνυμία των μαθητών προστατεύεται. Μία άλλη θεώρηση είναι η πιθανότητα να υπάρχει μία αφανής επίδραση στις βαθμολογίες ακόμα κι αν αυτός/-τη δεν είναι το πρόσωπο που συλλέγει τις φόρμες. Η Pulich (1984) υποστηρίζει ότι ακόμα κι αν ο διδάσκων φεύγει από το δωμάτιο ενώ οι φοιτητές συμπληρώνουν τις φόρμες, κάποιοι φοιτητές μπορεί ακόμη να αναστέλλονται από το γεγονός ότι ο διδάσκων ο ίδιος τα διένειμε. Αυτή ως εκ τούτου προτείνει όπως ένας ξεχωριστός αξιολογητής να μοιράζει και να συλλέγει τις φόρμες καθώς και να απαντά σε κάθε ερώτηση την οποία οι φοιτητές μπορεί να έχουν κατά τη διάρκεια της αξιολόγησης. Ο Algozzine και οι συνεργάτες του (2004) τονίζουν ότι η παρουσία του διδάσκοντα δεν επηρεάζει

σημαντικά την εγκυρότητα εκτός αν η πρακτική γίνεται με μη ανώνυμες βαθμολογήσεις.

Δηλωμένος στόχος της αξιολόγησης

Αν και κάποιες μελέτες έχουν βρει ότι οι βαθμολογίες των φοιτητών είναι κατά τι πιο υψηλές εάν ο δηλωμένος στόχος είναι για προαγωγή και παραμονή (Aleamoni & Hexner, 1980; Braskamp *et al.*, 1984; Centra, 1976; Feldman, 1979), ο Frankhouser (1984) κατέληξε στο συμπέρασμα ότι ο δηλωμένος στόχος της αξιολόγησης δεν είχε σημαντικές επιδράσεις στις βαθμολογίες. Ο Braskamp και οι συνεργάτες του (1984) συνιστούν, οι φοιτητές να ενημερώνονται για το αν οι βαθμολογίες πρόκειται να χρησιμοποιηθούν για αποφάσεις προσωπικού. Από την άλλη βρέθηκε ότι η έκθεση του σκοπού της αξιολόγησης, για παράδειγμα γίνεται για αποφάσεις προσωπικού, μπορεί να επηρεάσει θετικά τα αποτελέσματα (Algozzine *et al.*, 2004; Cashin, 1995). Ωστόσο, τα αποτελέσματα για αυτή τη μεταβλητή βρέθηκαν να είναι μικτά. Ο Cashin (1995) προτείνει ότι αυτό μπορεί να βρίσκεται υπό έλεγχο με τη χρήση τυποποιημένων οδηγιών. Ο Fox (2006) έχει τονίσει ότι οι βαθμολογίες/αξιολογήσεις μπορούν να βελτιωθούν όταν οι διδάσκοντες ζητούν από τους φοιτητές περισσότερη κριτική ανατροφοδότηση.

5.9.2 Χαρακτηριστικά των μαθημάτων

Προαιρετικό (electivity)

Οι ερευνητές έχουν βρει ότι οι καθηγητές των προαιρετικών ή μη-υποχρεωτικών μαθημάτων λαμβάνουν υψηλότερες βαθμολογίες από τους καθηγητές των υποχρεωτικών μαθημάτων (Algozzine *et al.*, 2004; Cashin, 1995, 1988). Ειδικότερα, η διάσταση «προαιρετικό» ενός μαθήματος μπορεί να οριστεί ως το ποσοστό των φοιτητών που παίρνουν αυτό το μάθημα ως μάθημα επιλογής (Feldman, 1978). Μία μικρή έως μέτρια θετική σχέση έχει βρεθεί μεταξύ της επιλογής ενός μαθήματος και των αξιολογήσεων (Brandenburg *et al.*, 1977; Feldman, 1978; McKeachie, 1979; Scherr & Scherr, 1990). Αυτό μπορεί να οφείλεται στο χαμηλότερο προγενέστερο υποκείμενο ενδιαφέρον για τα υποχρεωτικά έναντι των μη υποχρεωτικών μαθημάτων.

Η ώρα του μαθήματος

Από την περιορισμένη ποσότητα έρευνας σε αυτό το θέμα, η ομοφωνία βρίσκεται στο ότι καμία συνεπής σχέση δεν υπάρχει μεταξύ των αξιολογήσεων και της ώρας της ημέρας που διεξάγεται το μάθημα (Aleamoni, 1987; Centra, 1993; Feldman, 1978). Μία εξαίρεση σε αυτό είναι η μελέτη των Koushki και Kuhn (1982) οι οποίοι βρήκαν ότι τα πολύ νωρίς πρωινά μαθήματα, τα πολύ αργά απογευματινά, και τα μαθήματα

μετά το μεσημεριανό γεύμα λάμβαναν την χαμηλότερη βαθμολογία. Η επίδραση της ώρας της ημέρας βρέθηκε σε αυτή τη μελέτη ότι ήταν μεγαλύτερη από εκείνη άλλων μεταβλητών υποβάθρου όπως είναι το φύλο, ο χρόνος στη σχολή, το πεδίο μελέτης, και ο προσδοκώμενος βαθμός.

Επίπεδο σπουδών

Οι περισσότερες μελέτες έχουν βρει ότι τα υψηλότερου επιπέδου μαθήματα τείνουν να λαμβάνουν υψηλότερες βαθμολογίες (Aleamoni & Hexner, 1980; Algozzine et al., 2004; Cashin, 1988, 1995; Feldman, 1978; Marsh, 1987, 1997). Ωστόσο, καμία εξήγηση γι' αυτή τη σχέση δεν έχει δοθεί και ο Feldman επίσης αναφέρει ότι η σχέση μεταξύ του επιπέδου σπουδών και των βαθμολογιών ελαττώνεται όταν άλλες μεταβλητές υποβάθρου όπως το μέγεθος της τάξης, η προσδοκώμενη βαθμολογία και η επιλογή ελέγχονται για αυτό. Ως εκ τούτου, η επίδραση του επιπέδου σπουδών στις βαθμολογίες μπορεί να είναι άμεση, έμμεση, ή και τα δύο. Επίσης, η βιβλιογραφία έχει γενικά παραλείψει κάθε άλλο παράγοντα ο οποίος μπορεί να έχει σχέση εδώ, συγκεκριμένα, την ηλικία των φοιτητών. Εικάζεται ότι η διαφορά στο μέσο όρο της ηλικίας και /ή της ωριμότητας των φοιτητών κατά το χρόνο διαχείρισης των βαθμολογιών μπορεί να είναι μία μεγαλύτερη αιτία επίδρασης πάνω στις αξιολογήσεις από ότι οι ιδιότητες του μαθήματος καθ' εαυτές.

Το μέγεθος της τάξης

Αξιοσημείωτη προσοχή έχει δοθεί στη σχέση μεταξύ του μεγέθους της τάξης και των αξιολογήσεων / βαθμολογιών από τους φοιτητές. Οι περισσότεροι συγγραφείς αναφέρουν ότι οι μικρότερες τάξεις τείνουν να λαμβάνουν υψηλότερες βαθμολογίες (Algozzine et al., 2004; Feldman, 1978; Franklin et al., 1991; McKeachie, 1990). Ο Marsh (1987, σελ. 314; Marsh & Dunkin, 1992) αναφέρει ότι η επίδραση του μεγέθους της τάξης είναι συγκεκριμένη για ορισμένες διαστάσεις της αποτελεσματικής διδασκαλίας, συγκεκριμένα την αλληλεπίδραση της ομάδας και την επικοινωνία/ συμπάθεια. Αυτός επιπλέον υποστηρίζει ότι αυτή η ιδιαιτερότητα συνδυαζόμενη με παρεμφερή ευρήματα από τις αυτό-αξιολογήσεις των μελών ΔΕΠ έδειξε ότι το μέγεθος της τάξης δεν αποτελεί μία «προκατάληψη» στις βαθμολογίες από τους φοιτητές (βλέπε επίσης Cashin, 1992). Όμως ο Abrami (1989b) στην επισκόπησή του στη μονογραφία του Marsh (1987) υποστηρίζει ότι αυτό το επιχειρήμα δεν μπορεί να χρησιμοποιηθεί για να υποστηρίξει την εγκυρότητα των βαθμολογιών, και τουναντίον να επιδείξει ότι η αλληλεπίδραση και η επικοινωνία, όντας ευαίσθητες στο μέγεθος της τάξης, είναι διαστάσεις οι οποίες δεν θα πρέπει να

χρησιμοποιηθούν σε συγκεντρωτικές αποφάσεις. Ο Scott (1977) βρήκε ότι οι διδάσκοντες οι οποίοι αισθάνονταν ότι η τάξη τους ήταν τόσο μεγάλη για αυτούς για να εκθέτουν το υλικό των μαθημάτων κατάλληλα εξασφάλισαν χαμηλότερες βαθμολογίες από άλλους διδάσκοντες. Αυτό υπαινίσσεται ότι τα αισθήματα των διδασκόντων σχετικά με το μέγεθος της τάξης μπορεί να έχουν κάποια επίδραση στην επίδοσή τους και στις βαθμολογίες από τους φοιτητές (Feldman, 1978). Μία άλλη υπόθεση είναι ότι η σχέση μεταξύ μεγέθους τάξης και βαθμολογιών από τους φοιτητές δεν είναι γραμμική, αλλά μάλλον, U-shaped ή καμπυλόγραμμη (curvilinear) σχέση, με μικρές και μεγάλες τάξεις να λαμβάνουν υψηλότερες βαθμολογίες από ότι οι μεσαίου-μεγέθους τάξεις (Centra & Creech, 1976; Feldman, 1978, 1984; Koushki & Kuhn, 1982). (Ο Feldman επίσης υποστηρίζει ότι ακόμη περισσότεροι ερευνητές μπορεί να έχουν βρει μία καμπυλόγραμμη σχέση εφόσον γνώριζαν αυτή την πιθανότητα). Κάποιες εξηγήσεις οι οποίες έχουν διατυπωθεί για αυτή τη σχέση (αναπάντεχα υψηλότερες βαθμολογίες για πολύ μεγάλες τάξεις μαθημάτων) επικαλούνται ότι, τα τμήματα μπορεί να αναθέτουν σε γνωστούς ανώτερης βαθμίδας καθηγητές να κάνουν διάλεξη σε μεγάλες τάξεις. Οι καθηγητές ανώτερης βαθμίδας (superior) μπορεί να προσελκύουν περισσότερους φοιτητές στις τάξεις τους λόγω της φήμης τους, καθιστώντας έτσι το μέγεθος της τάξης μεγαλύτερο. Οι διδάσκοντες μπορεί να αισθάνονται μία αυξανόμενη πρόκληση διδασκαλίας για μία μεγάλη διάλεξη μαθημάτων και μπορεί να προετοιμάζουν τους εαυτούς τους καλά ή να προσαρμόζουν την παιδαγωγική τους σε αυτό τον τύπο μαθήματος, περιλαμβάνοντας και την πιο συχνή χρήση οπτικό-ακουστικών βοηθημάτων. Τέλος, πολλά μαθήματα με περισσότερους από εκατό φοιτητές συχνά έχουν μικρές συζητήσεις με τους αναπληρωτές / βοηθούς διδασκαλίας.

Γνωστικό αντικείμενο

Οι ερευνητές έχουν βρει ότι το γνωστικό αντικείμενο πράγματι ασκεί επίδραση στις βαθμολογίες από τους φοιτητές (Ramsden, 1991). Συγκεκριμένα, οι βαθμολογίες στις ανθρωπιστικές επιστήμες υπήρξαν υψηλότερες από ότι στις κοινωνικές και ακόμη χαμηλότερες στα μαθηματικά και στις φυσικές επιστήμες (Cashin, 1990, 1992; Cashin & Clegg, 1987; Centra & Creech, 1976; Feldman, 1978; Johnson, 2003; Neumann, 2001; Ory, 2001; Wachtel, 1998). Αυτό περισσότερο δείχνει τις διαφορές στο ύφος (στυλ) της διδασκαλίας και στους στόχους των επιστημονικών κλάδων παρά ότι είναι μία πηγή προκαταλήψεων. Γι' αυτό το λόγο οι Neumann (2001), Cashin (1995, 1988) και Ramsden (1991) ισχυρίζονται ότι οι συγκρίσεις στους

επιστημονικούς κλάδους μπορεί να μην είναι ακριβείς. Ο Centra (1993) διατυπώνει κάποιες πιθανές εξηγήσεις για τις χαμηλότερες βαθμολογίες στα μαθήματα μαθηματικών και φυσικών επιστημών, συγκεκριμένα ότι τα μαθήματα αυτά και οι διδάσκοντες είναι λιγότερο προσανατολισμένοι προς τους φοιτητές, τα μαθήματα είναι λιγότερο αποτελεσματικά στην παρουσίαση και είναι πιο γρήγορου ρυθμού, και υπάρχει η απαίτηση από τα μέλη ΔΕΠ να επενδύουν περισσότερο χρόνο στην έρευνα και στην αναζήτηση επιχορηγήσεων από ό, τι κάνουν οι συνάδελφοί τους σε άλλα επιστημονικά πεδία. Ο Centra επίσης παραθέτει μία προγενέστερη μελέτη (Centra, 1973) στην οποία βρήκε ότι φοιτητές και καθηγητές στις φυσικές επιστήμες έχουν πολύ διαφορετικές αντιλήψεις για την καταλληλότητα του ρυθμού και του φόρτου εργασίας, οι οποίες μπορούν να επηρεάζουν δυσμενώς τις βαθμολογίες. Η Cashin (1988) υποστηρίζει ότι εάν πράγματι η διδασκαλία είναι λιγότερο αποτελεσματική σε μαθήματα που βαθμολογούνται χαμηλότερα, τότε το γνωστικό αντικείμενο δεν θα έπρεπε να αποτελεί μία προκατάληψη (bias) για τις βαθμολογίες. Εάν ωστόσο τα αντικείμενα απαιτώντας δεξιότητες ποσοτικής αιτιολόγησης βαθμολογούνται χαμηλότερα επειδή οι φοιτητές είναι λιγότερο ικανοί σε αυτές τις δεξιότητες, τότε αυτό θα μπορούσε να αποτελέσει μία προκατάληψη στις βαθμολογίες. Ο Perry και οι συνεργάτες του (1974) επισημαίνουν ότι ένας «αδύνατος» καθηγητής που παρουσιάζει ενδιαφέρον υλικό βαθμολογείται σταθερά υψηλότερα σε κάποιες διαστάσεις της αποτελεσματικής διδασκαλίας από ένα «καλό» καθηγητή που παρουσιάζει βαρετό υλικό. Αυτό μπορεί να παρέχει επιπλέον λόγο αποφυγής της σύγκρισης των βαθμολογιών των καθηγητών με διαφορετικά γνωστικά αντικείμενα.

Φόρτος εργασίας μαθημάτων

Ο Ryan και οι συνεργάτες του (1980) ανέφεραν ότι η θέσπιση της υποχρεωτικής βαθμολόγησης από τους φοιτητές σε ένα πανεπιστήμιο των Κεντροδυτικών ΗΠΑ υποχρέωσε τα μέλη ΔΕΠ να ελαττώσουν το φόρτο εργασίας των μαθημάτων και να διεξάγουν πιο εύκολες εξετάσεις. Οι Dudley και Shawver (1991) βρήκαν ότι οι βαθμολογίες από τους φοιτητές σε ένα μάθημα για το marketing χωρίς εργασία στο σπίτι βελτιώθηκαν με τη καθιέρωση αναθέσεων σχετικών εργασιών. Ωστόσο αυτό το αποτέλεσμα μπορεί να μην είναι γενικό σε όλα τα γνωστικά αντικείμενα. (Επίσης, οι διδάσκοντες το marketing σε αυτή τη μελέτη αρχικά είχαν βαθμολογίες κάτω του μέσου όρου για τη σχολή επιχειρήσεων. Επισημαίνεται ότι η φόρμα βαθμολόγησης που χρησιμοποιήθηκε σε αυτή τη μελέτη δεν περιελάμβανε μία ερώτηση όπως «Ο φόρτος εργασίας για αυτό το μάθημα ήταν κατάλληλος;»). Ο Marsh (1987, σελ 316)

παρουσιάζει μελέτες οι οποίες βρήκαν μία θετική σχέση μεταξύ «φόρτου εργασίας / δυσκολίας» και βαθμολογιών από τους φοιτητές (περισσότερο δύσκολα μαθήματα βαθμολογήθηκαν περισσότερο ευνοϊκά), και σε αυτή τη βάση απέρριψαν αυτό τον παράγοντα ως μία πιθανή προκατάληψη στις βαθμολογίες. Ωστόσο, το επίπεδο του μαθήματος και η ηλικία των φοιτητών θα μπορούσε να αποτελούν παράγοντες σύγχυσης σε περισσότερο δύσκολα μαθήματα. Επίσης, η δυσκολία και ο φόρτος εργασίας δεν είναι εντελώς το ίδιο πράγμα. Για παράδειγμα, ο Franklin και οι συνεργάτες του (1991) αντιμετώπισαν το φόρτο εργασίας και τη δυσκολία ως ξεχωριστές μεταβλητές. Αυτοί βρήκαν ότι η δυσκολία, αλλά όχι ο φόρτος εργασίας, είχαν μία περιορισμένη σημασία θετική σχέση με τις βαθμολογίες για το διδάσκοντα. Όπως επισημάνθηκε σε προηγούμενο σημείο, οι αντιλήψεις των φοιτητών για το φόρτο εργασίας και το ρυθμό μπορεί να διαφέρουν αξιοσημείωτα από εκείνες του διδάσκοντα. Το συμπέρασμα ότι τα μαθήματα με μεγαλύτερο φόρτο εργασίας βαθμολογούνται υψηλότερα μπορεί να μη γενικευτεί στα μαθηματικά και στις φυσικές επιστήμες. Η έρευνα για να ελέγξει αυτή την υπόθεση σε αυτές τις περιοχές θα μπορούσε να ήταν πιο βοηθητική. Δεν αποκλείεται ο ρυθμός ενός μαθήματος και οι αντιλήψεις των φοιτητών γι' αυτό να έχουν δυσμενείς επιπτώσεις στις βαθμολογίες από τους φοιτητές.

5.9.3 Ένας τελικός αναστοχασμός στα χαρακτηριστικά των μαθημάτων

Μία μελέτη των Cranton και Smith (1986) έριξε περισσότερο φως στην επίδραση των χαρακτηριστικών των μαθημάτων όπως είναι το επίπεδο του μαθήματος και το μέγεθος της τάξης, στις βαθμολογίες από τους φοιτητές. Αν και η μελέτη επιβεβαίωσε τις συνολικές τάσεις που τεκμηρίωσαν οι προγενέστερες μελέτες, οι συγγραφείς βρήκαν ότι ανάμεσα στα επιμέρους τμήματα οι βαθμολογίες για τα χαρακτηριστικά των μαθημάτων διέφεραν δραστικά. Σε κάποια τμήματα τα χαρακτηριστικά των μαθημάτων δεν είχαν μία σημαντική επίδραση στις βαθμολογίες από τους φοιτητές, ενώ σε άλλες περιπτώσεις η επίδραση ήταν στην αντίθετη κατεύθυνση με την προβλεπόμενη. Οι συγγραφείς καταλήγουν στο συμπέρασμα ότι κάποιος δεν μπορεί να καθιερώνει νόρμες σε όλη την πανεπιστημιούπολη για την αποτελεσματικότητα της διδασκαλίας, και ακόμα για να τις επεκτείνει επί του συνόλου σε ένα τμήμα χρειάζεται μεγάλη προσοχή. Αφού η σχέση των χαρακτηριστικών του μαθήματος με τις βαθμολογίες από τους φοιτητές δεν ισχύουν το ίδιο για διαφορετικά τμήματα, θα ήταν χρήσιμο να διεξάγεται έρευνα για τις επιδράσεις των χαρακτηριστικών του μαθήματος πάνω στις βαθμολογίες σε ξεχωριστά αντικείμενα διδασκαλίας.

5.9.4 Χαρακτηριστικά του διδάσκοντα

Βαθμίδα διδάσκοντα και εμπειρία

Χωρίς έκπληξη, εκεί όπου οι βαθμολογίες των καθηγητών και των βοηθών διδασκαλίας έχουν υποστεί σύγκριση, οι καθηγητές βαθμολογούνται πολύ υψηλότερα (Brandenburg *et al.*, 1977; Centra & Creech, 1976; Marsh & Dunkin, 1992). Οι πρώτης χρονιάς καθηγητές λαμβάνουν χαμηλότερες βαθμολογίες από ό, τι λαμβάνουν στις επόμενες χρονιές (Centra, 1978). Εκτός από το θέμα των βοηθών διδασκαλίας, ο Feldman (1983) ανασκόπησε τη βιβλιογραφία που αφορούσε στις σχέσεις μεταξύ της αρχαιότητας (seniority) και των βαθμολογιών, και βρήκε ότι η πλειοψηφία των μελετών που αφορούσε στην ακαδημαϊκή βαθμίδα έδειξε μη σημαντική σχέση μεταξύ αυτής και των αξιολογήσεων διδασκαλίας. Στο ίδιο συμπέρασμα κατέληξαν και οι d'Arrollonia και Abrami (1997) και ο Arreola (2000). Ανάμεσα σε εκείνες τις μελέτες στις οποίες βρέθηκε μία σημαντική (αν και αδύνατη) σχέση, σχεδόν όλες βρήκαν ότι οι διδάσκοντες υψηλότερης βαθμίδας ελάμβαναν περισσότερο ευνοϊκές βαθμολογίες. Σε αναφορά με την ηλικία του διδάσκοντα και την εμπειρία, ο Feldman προειδοποιεί ότι τα χαρακτηριστικά αυτά δεν πρέπει να συγχέονται με την ακαδημαϊκή βαθμίδα. Οι μελέτες που ανασκόπησε ο Feldman και αφορούσαν στη σχέση μεταξύ ηλικίας / εμπειρίας και βαθμολογιών, ξανά η πλειοψηφία των μελετών δεν βρήκε σημαντική σχέση. Ωστόσο, μεταξύ εκείνων των μελετών στις οποίες βρέθηκε μία σημαντική σχέση, σχεδόν όλες βρήκαν μία ανάστροφη σχέση. Αυτό σημαίνει ότι, οι διδάσκοντες μεγαλύτερης ηλικίας και εμπειρίας στη διδασκαλία ελάμβαναν χαμηλότερες βαθμολογίες (Marsh, 2007). Ο Centra (1993, σελ 73) προειδοποιεί ότι μπορούμε να βγάλουμε μόνο περιορισμένα συμπεράσματα που να αφορούν στην επίδραση της βαθμίδας και της εμπειρίας εξαιτίας του ό, τι οι περισσότερες μελέτες είναι συγχρονικές (cross-sectional) παρά διαχρονικές (longitudinal).

Φήμη του διδάσκοντα

Υπάρχει σχετικά λίγη συζήτηση γύρω από αυτό το χαρακτηριστικό στη βιβλιογραφία. Ο Perry και οι συνεργάτες του (1974) βρήκαν ότι οι προσδοκίες προτέρων φοιτητών για την επίδοση της διδασκαλίας που βασιζόνταν στη φήμη του διδάσκοντα επηρέαζαν τις βαθμολογίες. Οι συγγραφείς υποστηρίζουν ότι αυτό το αποτέλεσμα έχει επιπτώσεις στη σκοπιμότητα της δημοσίευσης των πληροφοριών για τις βαθμολογίες από τους φοιτητές στο σύνολο της ακαδημαϊκής κοινότητας. Ο Leventhal και οι συνεργάτες του (1976) βρήκαν ότι οι φοιτητές οι οποίοι

χρησιμοποίησαν τη φήμη του καθηγητή για να επιλέξουν τα μαθήματα έδωσαν υψηλότερες βαθμολογίες στους καθηγητές τους από ό, τι έκαναν οι συμφοιτητές τους. Ο Perry και οι συνεργάτες του (1979) βρήκαν μία αλληλεπίδραση μεταξύ της φήμης του διδάσκοντα και της εκφραστικότητας, συγκεκριμένα ότι, η υψηλής εκφραστικότητας διδάσκοντες με μία αρνητική φήμη ελάμβαναν χαμηλότερες βαθμολογίες από άλλους υψηλής εκφραστικότητας διδάσκοντες με περισσότερο θετικές φήμες. Οι διδάσκοντες οι οποίοι παρουσίαζαν ένα χαμηλότερο βαθμό εκφραστικότητας δεν επηρεάζονταν σημαντικά από τον παράγοντα φήμη. Ο Cashin (1995) βρήκε ότι μερικές μελέτες έχουν τεκμηριώσει με σαφήνεια ότι η εκφραστικότητα τείνει να βελτιώνει τη μάθηση και ως εκ τούτου δεν μπορεί να θεωρηθεί παράγοντας στρέβλωσης.

Ερευνητική παραγωγικότητα (research productivity)

Συχνά πριμοδοτούνται τα επιχειρήματα ότι η ερευνητική παραγωγικότητα βοηθά να βελτιωθεί η αποτελεσματικότητα της διδασκαλίας επειδή αυτή συμβάλλει στην ενημέρωση, ή ότι εμποδίζει την αποτελεσματικότητα στη διδασκαλία επειδή τα μέλη ΔΕΠ δαπανώντας περισσότερο χρόνο για κατάρτιση (scholarship) έχουν λιγότερο χρόνο για τη διδασκαλία. Η βιβλιογραφία πάνω στις βαθμολογίες από τους φοιτητές δεν υποστηρίζει καμία από αυτές τις υποθέσεις. Αντιθέτως, οι περισσότερες μελέτες έχουν δείξει ότι δεν υπάρχει σχέση ή μία αδύνατη θετική σχέση μεταξύ της ερευνητικής παραγωγικότητας και της διδασκαλίας (Aleamoni, 1987; Centra, 1983; Feldman, 1987; Marsh, 1979, 1987; Marsh & Dunkin, 1992). Μία εξαίρεση αποτελεί μία έρευνα από τον Allen (1995) η οποία βρήκε μία μικρή αλλά σημαντική θετική σχέση μεταξύ της ερευνητικής παραγωγικότητας και των αξιολογήσεων της διδασκαλίας.

Η μελέτη από τον Centra (1983) επίσης βρήκε ότι, αν και μία μέτρια θετική σχέση υπάρχει μεταξύ έρευνας και αποτελεσματικότητας της διδασκαλίας στις κοινωνικές επιστήμες, αυτή ήταν σχεδόν μηδενική στις φυσικές επιστήμες. Σε αντίθεση με αυτό, υπάρχει απόδειξη ότι τα ιδρύματα συνεχίζουν να χρησιμοποιούν την έρευνα και τη δημοσίευση ως ένα παράγοντα αξιολόγησης της αποτελεσματικότητας της διδασκαλίας (Seldin 1984), μία πρακτική η οποία έχει ευρέως υποστεί κριτική (Aubrecht, 1984; Centra, 1993, p. 74).

Προσωπικότητα του διδάσκοντα

Αυτός ο παράγοντας έχει εξεταστεί εκπληκτικά σπάνια στη βιβλιογραφία. Ο Feldman (1986) στην επισκόπησή του βρήκε ότι όταν η προσωπικότητα του διδάσκοντα προέκυπτε ως συμπέρασμα από τις προσωπικές αναφορές των ίδιων των διδασκόντων, ελάχιστες από τις 14 συστάδες (clusters) των χαρακτηριστικών παρουσίασαν σημαντικές συσχετίσεις με τις συνολικές αξιολογήσεις από τους φοιτητές. Από την άλλη, όταν η προσωπικότητα του διδάσκοντα προέκυπτε ως συμπέρασμα από τις αντιλήψεις των φοιτητών ή των συναδέλφων, οι περισσότερες από τις συστάδες των χαρακτηριστικών έδειξαν σημαντικές συσχετίσεις με τις συνολικές αξιολογήσεις από τους φοιτητές. Ο Feldman ωστόσο δεν διέκρινε εάν (ή όχι) αυτή η σχέση ήταν μία έγκυρη επιρροή στις βαθμολογίες από τους φοιτητές ή μία πιθανή πηγή προκατάληψης (βλέπε επίσης Marsh 1987). Μία άλλη έρευνα (Murray et al., 1990) βρήκε ότι σε κάποια έκταση η αποτελεσματικότητα της διδασκαλίας θα μπορούσε να προβλεφτεί από τις αντιλήψεις των συναδέλφων για την προσωπικότητα. Οι συγγραφείς επίσης καθόρισαν ότι η αποτελεσματικότητα του διδάσκοντα ποικίλλει σημαντικά κατά μήκος των μαθημάτων, και τα χαρακτηριστικά της προσωπικότητας τα οποία μπορεί να φανούν βοηθητικά σε ένα συγκεκριμένο μάθημα θα μπορούσαν ενδεχομένως να αποτελούν επιβάρυνση σε ένα άλλο μάθημα. (Αυτό υποστηρίζει την άποψη ότι οι βαθμολογίες που χρησιμοποιούνται για αποφάσεις προσωπικού και αποφάσεις «θητείας» θα πρέπει να λαμβάνονται από διάφορες πηγές και όχι μόνο από μία) (Wachtel, 1998 σελ. 199). Αυτοί οι συγγραφείς συμφωνούν ότι η σχέση μεταξύ προσωπικότητας και βαθμολογιών από τους φοιτητές υποστηρίζει μάλλον παρά αντικρούει την εγκυρότητα των βαθμολογιών από τους φοιτητές.

Σαγηνευτικότητα: «Η επίδραση του Δρ Fox»

Στη συνέχεια εξετάζεται το θέμα εάν ένας διδάσκων ο οποίος είναι διασκεδαστικός ή εκφραστικός μπορεί να πάρει εξαιρετικά υψηλές βαθμολογίες. Στην αρχική «Δρ. Φοξ» μελέτη (Naftulin et al 1973), ένας επαγγελματίας ηθοποιός, αποκαλούμενος ως «Δρ. Φοξ» έδωσε μία υψηλής εκφραστικότητας και ενθουσιασμού διάλεξη η οποία ήταν άνευ περιεχομένου, και έλαβε υψηλές βαθμολογίες. Η σκέψη ήταν να δειχτεί ότι ένας πολύ εκφραστικός και χαρισματικός ομιλητής μπορεί να ξελογιάσει το κοινό να δώσει άδικα υψηλές βαθμολογίες. Ωστόσο, αυτή η μελέτη έχει δεχτεί πολλές κριτικές και το θέμα της αποπλάνησης αναλύθηκε ξανά από πολλές μετέπειτα μελέτες (Abrami, 2001; Kulik, 2001; Marsh, 1987; Marsh & Dunkin, 1992) και τώρα η αποκαλούμενη «Δρ. Φοξ» επίδραση έχει λίγη υποστήριξη στη βιβλιογραφία (Perry,

1990). Μία επισκόπηση των μελετών πάνω στην εκπαιδευτική αποπλάνηση από τον Abramí και τους συνεργάτες του (1982a) βρήκε ότι οι βαθμολογίες των φοιτητών ήταν πολύ περισσότερο ευαίσθητες στην εκφραστικότητα παρά στο περιεχόμενο της διάλεξης. (Σε ένα μετέπειτα άρθρο, ωστόσο, ο Abramí (1989a) έγραψε: « ... Η εκφραστικότητα επιδρά στις βαθμολογίες περισσότερο από τη μάθηση των φοιτητών και έτσι αντιπροσωπεύει μία προκατάληψη»). Οι Marsh και Ware (1982) επαναλαμβάνοντας την ανάλυσή τους βρήκαν ότι όταν στους φοιτητές δεν δίνονται κίνητρα να μάθουν, η εκφραστικότητα είχε μεγαλύτερη επίδραση στη βαθμολόγηση από τους φοιτητές από την κάλυψη της ύλης. Εν τούτοις, όταν τους δίνονταν κίνητρα, μία περίπτωση περισσότερο σαν μία τάξη κολλεγίου, η εκφραστικότητα ήταν λιγότερο σημαντική και δεν υπήρχε η επίδραση του «Δρ. Φοξ».

Το φύλο του διδάσκοντα

Συζήτηση για την επίδραση που ασκεί στις αξιολογήσεις της διδασκαλίας από τους φοιτητές το φύλο του καθηγητή εμφανίζεται να ποικίλει εντελώς. Πολλοί συγγραφείς αναφέρουν ότι οι βαθμολογίες από τους φοιτητές είναι προκατειλημμένες απέναντι στις γυναίκες διδάσκοντες (για παράδειγμα Basow, 1994; Basow & Silberg, 1987; Kaschak, 1978; Koblitz, 1990; Martin, 1984; Rutland, 1990). Λίγες μελέτες (Bennet, 1982; Kierstead, 1978 et al., 1988) έχουν βρει ότι οι γυναίκες διδάσκοντες χρειάζεται να συμπεριφέρονται σύμφωνα με στερεότυπες γυναικείες συμπεριφορές με σκοπό να αποφύγουν να λάβουν χαμηλότερες βαθμολογίες από τους άντρες συναδέλφους τους. Από αυτή την άποψη, ο Koblitz (1990) βλέπει μια δυσκολία για τις γυναίκες διδάσκοντες οι οποίες χρειάζονται να υιοθετήσουν την προσέγγιση «να σκληρύνουν τη στάση» (“get approach”). Από την άλλη, ο Tatro (1995) βρήκε ότι οι γυναίκες ελάμβαναν σημαντικά υψηλότερες βαθμολογίες από τους άντρες διδάσκοντες. Στο δεύτερο μέρος μιας μετά-ανάλυσης ο Feldman (1992, 1993) επισκόπησε υπάρχουσα έρευνα πάνω στις βαθμολογίες των φοιτητών για τις γυναίκες και τους άντρες διδάσκοντες στο περιβάλλον του εργαστήριου και της τάξης. Σε αυτή την επισκόπηση των μελετών σε περιβάλλον εργαστήριου, ο Feldman (1992) αναφέρει ότι η πλειοψηφία των μελετών που επισκόπησε δεν έδειξαν καμία διαφορά στις συνολικές αξιολογήσεις των αντρών και των γυναικών διδασκόντων. Στη μειοψηφία των μελετών στις οποίες βρέθηκαν διαφορές οι άντρες διδάσκοντες έλαβαν υψηλότερες συνολικά βαθμολογίες από τις γυναίκες. Πολύ λίγες μελέτες σε αυτή την επισκόπηση ανέφεραν επιδράσεις αλληλεπίδρασης μεταξύ του φύλου του φοιτητή και του καθηγητή, και αυτές οι επιδράσεις όταν αναφέρονταν ήταν ασυνεπείς. Συνεπώς, στην

επισκόπησή του σε μελέτες περιβαλλόντων, ο Feldman (1993) ξανά ανέφερε ότι η πλειοψηφία των μελετών έδειχναν μη σημαντικές διαφορές μεταξύ των φύλων. Ωστόσο, στις λίγες μελέτες που αναφέρθηκαν διαφορές, αυτή τη φορά οι γυναίκες διδάσκοντες έλαβαν ελαφρά υψηλότερες συνολικά βαθμολογίες από τους άντρες. Ένα αποτέλεσμα αλληλεπίδρασης βρέθηκε να είναι ότι οι φοιτητές έτειναν να βαθμολογούν τους καθηγητές ίδιου-φύλου ελαφρά υψηλότερα από τους καθηγητές του άλλου φύλου. Γενικά, οι μελέτες που συσχετίζουν το φύλο έχουν παράγει αμφίβολα αποτελέσματα, αλλά οι περισσότερες έχουν δείξει ότι αυτή η μεταβλητή έχει ελάχιστη ή και καθόλου επίδραση στις αξιολογήσεις (Algozzine et al., 2004; Theall & Franklin, 2001; Marsh & Roche, 1997; Cashin, 1995; Arreola, 2000; Aleamoni & Hexner, 1980).

5.9.5 Χαρακτηριστικά των φοιτητών

Χαρακτηριστικά προσωπικότητας

Δεν υφίσταται σαφής σχέση μεταξύ των χαρακτηριστικών της προσωπικότητας του φοιτητή και των αξιολογήσεων της διδασκαλίας (Abrami, et al., 1982b).

Προγενέστερο ενδιαφέρον για το αντικείμενο μελέτης

Υφιστάμενη απόδειξη υποστηρίζει ότι οι φοιτητές με το μεγαλύτερο ενδιαφέρον για το αντικείμενο μελέτης πριν από το μάθημα τείνουν να δίνουν πιο ευνοϊκές βαθμολογίες στον καθηγητή ((Feldman, 1977, p. 236; Marsh & Cooper, 1981; Prave & Baril, 1993). Οι Marsh και Dunkin (1992) υποστηρίζουν ότι η επίδραση του προγενέστερου ενδιαφέροντος στη βαθμολογία που δίνουν οι φοιτητές δεν συνιστά μία προκατάληψη, αν και αυτοί παραδέχονται ότι όταν οι βαθμολογίες χρησιμοποιούνται συγκεντρωτικά αυτή η επίδραση μπορεί να αποτελέσει πηγή αδικίας στο ό, τι αυτή είναι μία λειτουργία του μαθήματος και όχι του καθηγητή. Ο έλεγχος για το προγενέστερο ενδιαφέρον στο μάθημα μπορεί να είναι ιδιαίτερα σημαντικός σε μελέτες στις οποίες ένα μεγάλο ποσοστό φοιτητών εγγράφονται σε μαθήματα, όπως είναι τα μαθηματικά και τα αγγλικά, για παράδειγμα. Οι Prave και Baril (1993) συζητούν για μέσα μέτρησης και ελέγχου αυτού του παράγοντα.

Το φύλο του φοιτητή

Αντικρουόμενα ευρήματα έχουν προκύψει από μελέτες της σχέσης μεταξύ του φύλου του μαθητή και των βαθμολογιών (Aleamoni, 1987; Aleamoni & Hexner, 1980; Feldman, 1977). Πολλές μελέτες έχουν αναφέρει ότι δεν υπάρχει σημαντική διαφορά στις βαθμολογίες από τα αγόρια και τα κορίτσια φοιτητές, αλλά ελάχιστες έχουν καταλήξει επίσης σε διαφορετικά συμπεράσματα. Ο Feldman (1977) αναφέρει ότι

ανάμεσα σε εκείνες τις μελέτες στις οποίες πράγματι προκύπτουν σημαντικές σχέσεις μεταξύ του φύλου του φοιτητή και των βαθμολογιών, οι περισσότερες βρήκαν ότι τα κορίτσια φοιτήτριες έδωσαν υψηλότερες βαθμολογίες από τα αγόρια. Ο Tatro (1995) επίσης βρήκε ότι τα κορίτσια φοιτήτριες έδωσαν πιο υψηλές βαθμολογίες από τα αγόρια φοιτητές. Ωστόσο, οι Koushki και Kuhn (1982) βρήκαν στη μελέτη τους ότι τα αγόρια φοιτητές έδωσαν υψηλότερες βαθμολογίες από τα κορίτσια φοιτήτριες. Επιπρόσθετα, μερικές μελέτες έχουν αναφέρει μία τάση των φοιτητών να βαθμολογούν τους ίδιου φύλου διδάσκοντες ελαφρά υψηλότερα από ό, τι του αντίθετου φύλου (Feldman, 1993; Centra, 1993; Ory, 2001).

Παρακολούθηση μαθημάτων

Μία πρόσφατη μελέτη η οποία διεξήχθη σε 9.000 φοιτητές σε κολλέγια του Ισραήλ βρήκε ότι υπήρχε μία θετική σχέση μεταξύ των υψηλών ποσοστών παρακολούθησης (παρουσία) και των θετικών αξιολογήσεων των μαθημάτων. Γενικά, αυτό δεν μπορεί να θεωρηθεί ως μεταβλητή στρέβλωσης (biasing) επειδή η μεγαλύτερη συμμετοχή οδηγεί στη βελτιωμένη μάθηση (Davidovitch & Soen, 2006). Πρέπει να διευκρινιστεί ότι αυτή είναι η μόνη μέχρι τώρα πλήρους-κλίμακας μελέτη που εξέτασε αυτό το θέμα.

Αναμενόμενοι βαθμοί και η υπόθεση Επιείκειας

Η επίδραση της αναμονής των βαθμών από τους φοιτητές σε ένα μάθημα πάνω στις αξιολογήσεις τους για τον καθηγητή αυτού του μαθήματος έχει υπάρξει ένα από τα πιο αμφιλεγόμενα θέματα στη βιβλιογραφία για την αξιολόγηση της διδασκαλίας από τους φοιτητές. Μέχρι τώρα υφίσταται ομοφωνία ότι υπάρχει μία μέτρια θετική σχέση μεταξύ των αναμενόμενων βαθμών και των βαθμολογιών από τους φοιτητές (οι φοιτητές που περιμένουν υψηλότερους βαθμούς θα δώσουν πιο ευνοϊκές βαθμολογίες) (Braskamp & Ory, 1994; Centra, 1979, p. 32; DuCette & Kenney, 1982; Feldman, 1976a; Howard & Maxwell, 1980; Marsh, 1987; Marsh & Dunkin, 1992). Το αμφιλεγόμενο αφορά την ερμηνεία αυτής της σχέσης. Ο Marsh (1987; Marsh & Dunkin, 1992) υποστηρίζει τρεις πιθανές ερμηνείες:

- Την υπόθεση επιείκειας. Οι διδάσκοντες με πιο επιεική βαθμολογικά στάνταρ λαμβάνουν πιο ευνοϊκές βαθμολογίες, π. χ ο καθηγητής μπορεί να «αγοράσει» αξιολογήσεις αξιολογώντας ο ίδιος με μεγαλύτερους βαθμούς. Κάτι τέτοιο θα πρότεινε ότι η σχέση βαθμών-βαθμολογίας συνιστούν μία προκατάληψη και ένα ακίνδυνο στην εγκυρότητα των βαθμολογιών από τους φοιτητές.

- Η υπόθεση εγκυρότητας. Οι περισσότερο αποτελεσματικοί διδάσκοντες γίνονται αιτία να εργάζονται σκληρότερα οι φοιτητές, να μαθαίνουν περισσότερο, και να κερδίζουν καλύτερους βαθμούς. Αυτό συνεπάγεται ότι η σχέση βαθμών-βαθμολογιών υποστηρίζει την εγκυρότητα των βαθμολογιών από τους φοιτητές.
- Η υπόθεση των χαρακτηριστικών του φοιτητή. Τα προ-υπάρχοντα χαρακτηριστικά του φοιτητή όπως είναι το προγενέστερο προς το αντικείμενο μελέτης ενδιαφέρον επηρεάζει τόσο την αποτελεσματικότητα της διδασκαλίας όσο και τις βαθμολογίες από τους φοιτητές.

Πολλοί συγγραφείς έχουν ταχτεί υπέρ της υπόθεσης επιεικείας (Chacko, 1983; Koshland, 1991; Nimmer & Stone, 1991; Powell, 1978; Snyder & Clair, 1976; Vasta & Sarmiento, 1979; see also DuCette & Kenney, 1982) και αντίθετα σε αυτή (Abrami *et al*, 1980; Howard & Maxwell, 1980; Marsh, 1987; Theall & Franklin, 1991; Βλέπε επίσης Franklin *et al*, 1991). Ο Chacko (1983) για παράδειγμα, έδειξε ότι τα πιο αυστηρά στάνταρ βαθμολόγησης οδηγούσαν τους φοιτητές να βαθμολογούν τους διδάσκοντες χαμηλότερα ακόμα και για συνιστώσες της διδασκαλίας μη σχετικές με τη δίκαιη βαθμολογία, όπως το χιούμορ, την αυτάρκεια και τις στάσεις απέναντι στους φοιτητές. Οι Goldberg και Gallahan (1991) στη μελέτη τους βρήκαν ότι τα συμπληρωματικά μέλη ΔΕΠ έτειναν να δίνουν υψηλότερους βαθμούς και να λαμβάνουν υψηλότερες βαθμολογίες από τους πλήρους ωραρίου διδάσκοντες, ακόμα και όταν οι περισσότεροι των φοιτητών δεν γνώριζαν το καθεστώς του διδάσκοντά τους. Αυτοί πρότειναν την περαιτέρω μελέτη της σχέσης μεταξύ του καθεστώτος συμπλήρωσης και των βαθμολογιών από τους φοιτητές. Οι Nimmer και Stone (1991) βρήκαν ότι η προκατάληψη επιεικείας ήταν ένα σημαντικό πρόβλημα, και επιπλέον, ότι ο βαθμός προκατάληψης επιεικείας επηρεαζόταν από το χρόνο στον οποίο γινόταν η διαχείριση των βαθμολογιών, για παράδειγμα η προκατάληψη ήταν μεγαλύτερη όταν οι αξιολογήσεις γινόντουσαν μετά από εξετάσεις. Μέχρι τώρα η διαμάχη γύρω από την πιθανότητα της προκατάληψης περί επιεικείας δεν έχει βρει λύση.

Μία άλλη άποψη στο αμφιλεγόμενο θέμα για τον αναμενόμενο βαθμό δίνουν η Hewett και οι συνεργάτες της (1988), η οποία προτείνει ότι αυτό που πειράζει δεν είναι τόσο οι αναμενόμενοι από τους φοιτητές βαθμοί, αλλά οι εντυπώσεις

των φοιτητών για το μάθημα όπως υπαγορεύονται από την επίδοσή τους στις εξετάσεις. Αυτοί οι συγγραφείς βρήκαν ότι η σχέση μεταξύ των αποτελεσμάτων των φοιτητών σε μία εξέταση και των αξιολογήσεων στο τέλος του μαθήματος ήταν πιο δυνατές με την πρώτη εξέταση, πιο αδύνατες με κάθε επιτυχημένη εξέταση, και πιο αδύνατη με την τελευταία εξέταση που δόθηκε πριν τη διαχείριση της αξιολόγησης από τους φοιτητές. Η Hewett και οι συνεργάτες της αναφέρουν αυτό ως μία πιθανή εξήγηση για τα μικτά αποτελέσματα των μελετών για την επίδραση του αναμενόμενου βαθμού, σαν η επίδραση στις βαθμολογίες του συνολικού βαθμού να είναι πιο αδύνατη από το βαθμό μόνο της πρώτης εξέτασης. Αυτοί επίσης συμπεραίνουν ότι οι βαθμολογίες από τους φοιτητές μπορεί να μη μετρούν την όλη αποτελεσματικότητα της διδασκαλίας, αλλά μόνο τις πρώτες εντυπώσεις των φοιτητών για την αποτελεσματικότητα του διδάσκοντα. Αυτή η συζήτηση για την «επίδραση της πρώτης εντύπωσης» επίσης φέρνει στην επιφάνεια το ερώτημα εάν ένας διδάσκοντας μπορεί να παράγει υψηλότερες βαθμολογίες από τους φοιτητές όχι δίνοντας πιο υψηλούς βαθμούς, αλλά δίνοντας μία εύκολη πρώτη εξέταση.

Προσδοκίες φοιτητή

Ο McKeachie (1979) υποστηρίζει ότι το μόνο πιο σημαντικό χαρακτηριστικό του φοιτητή που επηρεάζει τις βαθμολογίες από τους φοιτητές είναι οι προσδοκίες των φοιτητών, π. χ οι φοιτητές που προσδοκούν ένα διδάσκοντα να είναι καλός συνήθως βρίσκουν να ισχύει αυτό. Η έρευνα πάνω στην προσδοκία τάξης είναι αξιοπρόσεκτη, αλλά πολύ λίγη έμφαση έχει δοθεί στις προσδοκίες των φοιτητών από ό, τι έχει δοθεί στις προσδοκίες των καθηγητών (Gigliotti, 1987; Koermer & Petelle, 1991; Perry *et al.*, 1979). Οι προσδοκίες των φοιτητών, ένας παράγοντας έξω από τον έλεγχο του διδάσκοντα, έχει δείχτει ότι επηρεάζει τις βαθμολογίες για το μάθημα. Ειδικότερα, οι φοιτητές με υψηλές προσδοκίες και υψηλές εμπειρίες δίνουν υψηλότερες βαθμολογίες από εκείνους με χαμηλές προσδοκίες και υψηλές εμπειρίες ή με χαμηλές εμπειρίες (Koermer & Petelle, 1991).

Συναισθηματική κατάσταση

Μία μόνο μελέτη είναι γνωστή η οποία εισάγει για συζήτηση αυτή τη μεταβλητή. Ο Small και οι συνεργάτες του (1982) βρήκαν ότι η συναισθηματική κατάσταση των φοιτητών στο τέλος του εξαμήνου, όταν οι βαθμολογίες από τους φοιτητές συνήθως είναι συγκεντρωμένες (αλλά όχι νωρίτερα) σχετίζονταν με τις βαθμολογίες τους για τη διδασκαλία. Ειδικότερα, όσο περισσότερο εχθρικοί, αγχωμένοι και πιεσμένοι

ένιωθαν οι φοιτητές στο τέλος του εξαμήνου, τόσο χαμηλότερες βαθμολογίες έδιναν στον διδάσκοντά τους. Αν και οι συγγραφείς πιστεύουν ότι αυτό συνιστά ένα σοβαρό κίνδυνο για την εγκυρότητα των τέλος-του-εξαμήνου βαθμολογιών από τους φοιτητές, δεν έχουμε δει τα αποτελέσματα τους να επιβεβαιώνονται πιο πρόσφατα.

Η ηλικία του φοιτητή

Για έκπληξη, στη γνώση μας δεν έχουν υπάρξει πρόσφατες μελέτες οι οποίες εξετάζουν συγκεκριμένα την επίδραση της ηλικίας των φοιτητών (σε αντίθεση με τις σχετικές μεταβλητές του επιπέδου του μαθήματος και του χρόνου φοίτησης) πάνω στις βαθμολογίες από τους φοιτητές για τη διδασκαλία. Δεν γνωρίζουμε, για παράδειγμα, εάν οι υψηλές βαθμολογίες για τα ανωτέρου- επιπέδου μαθήματα είναι ένα αποτέλεσμα του πιο προχωρημένου επιπέδου αντικειμένου μελέτης, ή ό, τι οι φοιτητές είναι μεγαλύτερης ηλικίας και πιο ώριμοι. Μία μελέτη των Klann και Hoff (1976) σε φοιτητές των μαθηματικών σε ένα κοινοτικό κολλέγιο δεν βρήκε διαφορά στις βαθμολογίες των φοιτητών κάτω των 20 χρόνων και των φοιτητών με ηλικία τουλάχιστον 20 χρόνων. Ενδείκνυται μελέτες σε δύο και τεσσάρων ετών φοίτησης κολλέγια με περισσότερες ηλικίες.

6. Κατευθύνσεις και μοντέλα αξιολόγησης

6.1 Κατευθύνσεις αξιολόγησης

Οι Gravestock και Gregor-Greenleaf (2008) σε μία αντίστοιχη Έκθεσή τους στην οποία παρουσιάζουν τις γενικές νέες κατευθύνσεις στην αξιολόγηση των μαθημάτων από τους φοιτητές, διαπιστώνουν ότι αυτές βασίζονται στις ισχύουσες πρακτικές αξιολόγησης που κατά τη γνώμη τους είναι σε θέση να οδηγήσουν τις αλλαγές στη διαχείριση και στη χρήση των αξιολογήσεων στο σχετικά εγγύς μέλλον. Συγκεκριμένα αυτές οι πρακτικές είναι α) τα ηλεκτρονικά εργαλεία αξιολόγησης β) η σύνδεση της αξιολόγησης με τις μετρήσεις λογοδοσίας (accountability) και με τα αποτελέσματα μάθησης που βασίζονται στην επάρκεια (competency-based), γ) οι αξιολογήσεις εκπαιδευτικού έργου για διαμορφωτικούς σκοπούς δ) οι αξιολογήσεις εκπαιδευτικού έργου για συγκεντρωτικούς σκοπούς. Στη συνέχεια γίνεται μία σύντομη και κατατοπιστική αναφορά σε κάθε μία από αυτές.

6.1.1 Τα ηλεκτρονικά εργαλεία αξιολόγησης

Η κίνηση της προσφοράς ηλεκτρονικών ή με βάση τον υπολογιστή αξιολογήσεων των μαθημάτων ξεκίνησε πριν από περίπου 20 χρόνια, με τη πιο διαδεδομένη θέσπισή τους να λαμβάνει χώρα κατά την τελευταία δεκαετία. Η έρευνα σε αυτό τον τομέα είναι σε ανοδική πορεία και ο διάλογος που αφορά ειδικά θέματα σχετικά με την υλοποίηση τους είναι ταχύτατα αυξανόμενος. Περισσότερες πληροφορίες από ιδρύματα που εφήρμοσαν τα ηλεκτρονικά συστήματα αξιολόγησης δίνονται στην ηλεκτρονική διεύθυνση: <http://onset.byu.edu>.

6.1.2 Η σύνδεση της αξιολόγησης με τις μετρήσεις λογοδοσίας και τα αποτελέσματα μάθησης με βάση την επάρκεια

Η αξιολόγηση φοιτητών και προγραμμάτων και η αποτίμηση στην ανώτατη εκπαίδευση για πολλά χρόνια έχουν κινηθεί προς ένα μοντέλο αξιολόγησης με βάση την επάρκεια (*competency-based*). Αυτό το μοντέλο επικεντρώνεται στη μέτρηση των προ-καθορισμένων αποτελεσμάτων ή στόχων και στον προσδιορισμό των πραγματικών μέσων όρων αξιολόγησης όπου και όποτε είναι δυνατό (δηλαδή, μετρήσεις αξιολόγησης που ζητούν οι φοιτητές να επιδείξουν τα αποτελέσματα άμεσα παρά να χρησιμοποιούν υποκατάστατα όπως βαθμούς μαθημάτων). Οι αξιολογήσεις των μαθημάτων και της μάθησης και της εμπειρίας των φοιτητών αρχίζουν να ενσωματώνονται στις θεσμικές και σε επίπεδο συστήματος διαδικασίες

αξιολόγησης, λογοδοσίας και προγραμματισμού. Όμως αυτές οι διαδικασίες παραμένουν ακόμα ανεπαρκείς επειδή οι αξιολογήσεις των μαθημάτων δεν είναι κατ' ανάγκη σχεδιασμένες να μετρούν αποτελέσματα από ειδικά προγράμματα και γενικά δεν προσφέρουν πραγματικούς μέσους όρους αξιολόγησης. Ως εκ τούτου, τα εργαλεία αξιολόγησης των μαθημάτων χρειάζεται να τροποποιηθούν για να αντανακλούν τα αποτελέσματα του προγράμματος περισσότερο άμεσα. Αυτό σχετίζεται με αυτά που ισχυρίζονται οι Ory και Ryan (2001) και οι Theall και Franklin (2001) ότι οι ερωτήσεις στις αξιολογήσεις πρέπει να αντανακλούν προτεραιότητες του ιδρύματος. Οι Hoyt και Pallet (1999) ισχυρίζονται παρόμοια ότι όταν οι αποτιμήσεις πρόκειται να χρησιμοποιηθούν κυρίως για συγκεντρωτικούς σκοπούς τότε τα εργαλεία πρέπει να επικεντρώνονται στη μέτρηση προσδιορισμένων αποτελεσμάτων (όπως «ποσό επιτυχημένοι υπήρξαν οι στόχοι που εισηγήθηκαν τα μαθήματα;»).

6.1.3 Αξιολογήσεις εκπαιδευτικού έργου για διαμορφωτικούς σκοπούς

Η αξιολόγηση της διδασκαλίας αποβαίνει συνεχώς όλο και περισσότερο πολλαπλών-όψεων και διαμορφωτική (Ory, 2001). Επιπλέον όπως ισχυρίζεται στη μελέτη του ο Marsh (2007) η απλή κοινοποίηση των αποτελεσμάτων των συγκεντρωτικών αξιολογήσεων στους διδάσκοντες προσφέρει πολύ λίγα στη βελτίωση της διδασκαλίας. Απεναντίας, αρκετά επίμονες διαδικασίες διαβούλευσης απαιτούνται προκειμένου να προκύψει ουσιαστική και διαρκής βελτίωση της διδασκαλίας. Ακόμα, μελετητές όπως οι Abrami (2001) και McKeachie (1997) έχουν επισημάνει ότι οι αξιολογήσεις των μαθημάτων δεν σχεδιαστήκαν εξ αρχής για να χρησιμοποιηθούν σε διαμορφωτικού τύπου αξιολογήσεις, αν και κάποια από τα εργαλεία μπορούν να έχουν κάποια διαγνωστική χρησιμότητα. Αντιθέτως, οι ενδιάμεσες των μαθημάτων αξιολογήσεις καθιστούν δυνατή τη μεμονωμένη και συχνά ποιοτικού τύπου ανατροφοδότηση και ως εκ τούτου μπορούν να ωφελήσουν τους διδάσκοντες που ευελπιστούν σε μία πληροφόρηση σχετικά με το πώς να βελτιώσουν τη διδασκαλία τους. Οι ενδιάμεσες των μαθημάτων αξιολογήσεις εξελίσσονται σε ένα πολύ δημοφιλές εργαλείο (Aultman, 2006), αν και εξαιτίας του σχετικού νεωτερισμού τους ελάχιστη έρευνα υπάρχει προς το παρόν σχετικά με τις καλύτερες πρακτικές στην ανάπτυξη και στη διαχείρισή τους ή σχετικά με τα μέσα για την καλύτερη ενσωμάτωση των αποτελεσμάτων τους στη βελτίωση της διδασκαλίας και στην ανάπτυξη δραστηριοτήτων.

6.1.4 Αξιολογήσεις εκπαιδευτικού έργου για συγκεντρωτικούς σκοπούς

Τα τελευταία 20 + χρόνια, η έρευνα έχει να επιδείξει ένα όλο και αυξανόμενο ενδιαφέρον σε θέματα που σχετίζονται με τη βελτίωση των πρακτικών αξιολόγησης. Από τη μία, αυτή έχει συμπεριλάβει την εισαγωγή και την εξερεύνηση των νέων εργαλείων αξιολόγησης όπως είναι το χαρτοφυλάκιο διδασκαλίας και η χρήση των μέσα-στην τάξη εκθέσεων από τους ομότιμους. Από την άλλη, περιλαμβάνει την αυξανόμενη προσοχή σε θέματα θεσμικών πρακτικών και πολιτικών.

Στο πεδίο της εξέλιξης των μελών ΔΕΠ έχει υπάρξει μία κίνηση που αντιμετωπίζει ορισμένες από τις ανησυχίες σχετικά με τη χρήση και την ερμηνεία των στοιχείων αξιολόγησης των μαθημάτων για αποφάσεις προσωπικού (personnel decisions). Ως μέρος των πρακτικών της, οι υπεύθυνοι σε θέματα εκπαιδευτικής ανάπτυξης εργάζονται ξεχωριστά με τα μέλη ΔΕΠ για να αποκωδικοποιήσουν τα αριθμητικά και τα ποιοτικά δεδομένα, τόσο για διαμορφωτικούς όσο και για συγκεντρωτικούς σκοπούς. Διαμέσου της διαβούλευσης επίσης, αυτοί βοηθούν τα μέλη ΔΕΠ στην υποστήριξη κατάλληλων αλλαγών στις στρατηγικές διδασκαλίας ή στο σχεδιασμό των μαθημάτων σε αντιστοιχία με τις απαντήσεις και τα σχόλια των φοιτητών. Και μερικοί όπως η Franklin (2001) συστήνουν όπως τα μέλη ΔΕΠ αναπτύσσουν αφηγήσεις (narratives) για να συμπληρώνουν και να πλαισιώνουν τα στοιχεία της αξιολόγησης για τους θεσμικού τύπου αξιολογητές, όπως είναι οι επιτροπές για τη θητεία και την προαγωγή. Αυτή διευκρινίζει ότι οι αριθμοί φανερώνουν μόνο ένα μέρος της ιστορίας. Μία αφήγηση μπορεί να μιλήσει για την πολλαπλότητα των διαστάσεων της διδασκαλίας και μπορεί να εισηγηθεί ειδικές πλευρές των μαθημάτων η του ύφους διδασκαλίας του διδάσκοντα που δεν μπορεί να κάνει μία περίληψη αποτελεσμάτων.

6.2 Η αναγκαιότητα των πολλαπλών πηγών αξιολόγησης

Σύμφωνα με τον Berk (2009) συνεχώς πληθαίνουν τα επιχειρήματα υπέρ της χρήσης των πολλαπλών πηγών απόδειξης στη διαμόρφωση αποφάσεων προσωπικού. Τρία από αυτά ειδικά, εμφανίζονται πιο συχνά στην αξιολόγηση των μελών ΔΕΠ: (Α) οι ψυχομετρικοί περιορισμοί του πάνω – κάτω (top-down) μοντέλου, (β) το ενδεχόμενο λάθους ή πλάνης των πηγών απόδειξης, και (γ) οι συστάσεις ειδικών.

Α. Οι ψυχομετρικοί περιορισμοί του μοντέλου πάνω –κάτω (top – down)

Παραδοσιακά, οι εργαζόμενοι στους περισσότερους οργανισμούς έχουν αξιολογηθεί με τη χρήση του «πάνω – κάτω» μοντέλου εποπτείας. Ο υπεύθυνος επόπτης βαθμολογεί την επίδοση του εργαζόμενου στο χρόνο της προγραμματισμένης

αξιολόγησης των επιδόσεων. Υπάρχουν τρία θέματα τα οποία ταλανίζουν αυτή την προσέγγιση:

- (1) η πιθανή αδικία και προκατάληψη του ενός βαθμολογητή,
- (2) η μη εγκυρότητα ενός μόνο συνόλου βαθμολογιών, η οποία μπορεί να βασίζεται σε μία μη επαρκή αξιολόγηση όλων των σχετικών εργασιακών δεξιοτήτων, στη μη άμεση παρατήρηση της επίδοσης της εργασίας, και / ή στην εξάρτηση από δευτερογενή πληροφόρηση, και
- (3) η έλλειψη αξιοπιστίας της μίας μόνο πηγής απόδειξης.

Αυτά τα ζητήματα θα μπορούσαν να προκύψουν στην περίπτωση που ο πρόεδρος του τμήματος, ο κοσμήτορας, ή ο αναπληρωτής κοσμήτορας μόνος του διεξάγει την αξιολόγηση του καθηγητή.

B. Το ενδεχόμενο λάθους ή πλάνης των πηγών απόδειξης

Τα θέματα προκατάληψης, μη εγκυρότητας, μη αξιοπιστίας της μία μόνο πηγής μπορούν επίσης να ισχύουν και για άλλες πηγές απόδειξης. Απλά, δεν υφίσταται τέλεια πηγή ή συνδυασμός πηγών απόδειξης. Στην πραγματικότητα, σχεδόν όλες οι πηγές βασίζονται σε ανθρώπινες αποφάσεις με τη μορφή είτε της ατομικής είτε της συλλογικής γνωμοδότησης από άλλους ανθρώπους. « Δεν υπάρχει καμία γνωστή αντικειμενική μέθοδος η οποία να μετρά την επίδοση στη διδασκαλία» (Berk, 2009, σελ. 1074).

Κάθε πηγή απόδειξης μπορεί να στηρίζει μοναδική πληροφορία, αλλά επίσης είναι σφαλερή κατά ένα ή περισσότερους τρόπους, συνήθως διαφορετικούς από των άλλων πηγών. Για παράδειγμα, η αναξιοπιστία και οι προκαταλήψεις των βαθμολογιών από τους ομότιμους δεν είναι ίδιες με εκείνες των βαθμολογιών από τους φοιτητές, οι οποίες έχουν τις δικές τους αδυναμίες. Με τον τριγωνισμό (triangulation) τριών ή περισσότερων διαφορετικών πηγών απόδειξης, «τα δυνατά σημεία κάθε πηγής μπορούν να αναχαιτίσουν τα αδύνατα σημεία των άλλων πηγών», συγκλίνοντας με αυτόν τον τρόπο σε μία απόφαση σχετικά με την αποτελεσματικότητα της διδασκαλίας που είναι περισσότερο ακριβής, αξιόπιστη, δίκαιη και ισορροπημένη από εκείνη που βασίζεται μόνο σε μία πηγή (Appling et al., 2001).

Δοθείσης της πολυπλοκότητας της μέτρησης, η πράξη της διδασκαλίας σε ένα πραγματικού-χρόνου περιβάλλον τάξης, η χρήση των πολλαπλών πηγών «πληροφορημένων αποφάσεων» είναι μία συνετή, υποστηρίξιμη στρατηγική. Αυτός που παίρνει αποφάσεις πρέπει να ενσωματώνει την πληροφορία μόνο από εκείνες τις

πηγές οι οποίες παρέχουν απόδειξη εγκυρότητας. Η ποιότητα των πηγών που επιλέγονται πρέπει να είναι άψογη.

Γ. Συστάσεις εμπειρογνομόνων

Ιστορικά, οι βαθμολογίες από τους φοιτητές έχουν καθιερωθεί ως η πρωταρχική μέτρηση της αποτελεσματικότητας της διδασκαλίας για τις προηγούμενες πέντε δεκαετίες (Seldin, 1999). Μόνο πρόσφατα έχει υπάρξει μία τάση απέναντι στην υπεράσπιση εκείνων των βαθμολογιών από άλλες πηγές δεδομένων για να βαθύνει και να διευρυνθεί η βάση δεδομένων (Centra 1993; Braskamp & Ory 1994; Knapper & Cranton 2001; Berk 2006; Seldin 2006; Arreola 2007).

6.3 Παραδείγματα Μοντέλων Αξιολόγησης

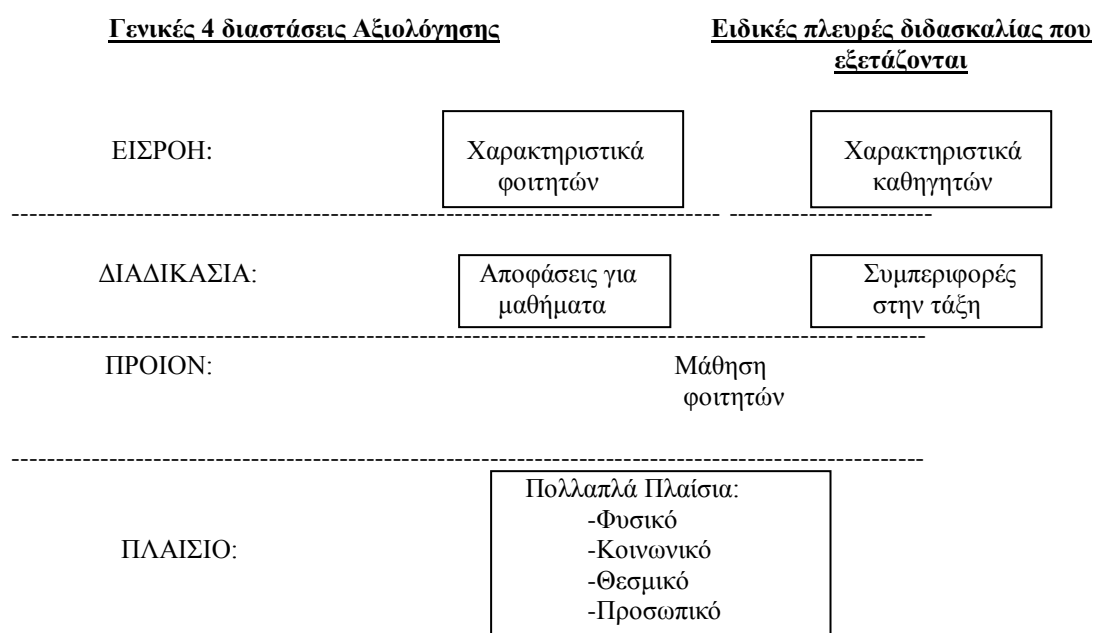
Όπως ήδη έχει αναφερθεί η κυρίαρχη τάση στη διαμόρφωση συγκεντρωτικού και διαμορφωτικού τύπου αποφάσεων σχετικών με τις συμπεριφορές διδασκαλίας των μελών ΔΕΠ περιλαμβάνει τη χρήση πολλαπλών πηγών απόδειξης. Στη συνέχεια επελέγησαν και παρουσιάζονται τα πιο εγκεκριμένα με βάση τη βιβλιογραφία μοντέλα αξιολόγησης του εκπαιδευτικού έργου. Αυτά είναι:

- το μοντέλο που προτείνει ο Fink για μία Πολλαπλών Διαστάσεων Αξιολόγηση,
- το 360^ο Πολλαπλών Πηγών Ανατροφοδότησης μοντέλο (ΠΠΑ)
- το μοντέλο της Περιβάλλουσας Αξιολόγησης Δεδομένων (ΠΑΔ),
- ένα μοντέλο έρευνας μικτών μεθόδων για την αξιολόγηση του περιβάλλοντος μάθησης με βάση την αλλαγή των πεποιθήσεων από αποτελεσματικότητας των φοιτητών

6.4 Μοντέλο Πολλαπλών Διαστάσεων του Fink

Πηγή: Fink, D. (2008). Improving the Evaluation of College Teaching. *To Improve the Academy* 26, an annual collection of articles produced by the POD Network in Higher Education and published by Jossey-Bass.

Σύμφωνα με τον Fink (2008) η αξιολόγηση που είναι κατάλληλη για χρήση στην τριτοβάθμια εκπαίδευση είναι τεσσάρων διαστάσεων. Αυτή καλείται να εξετάσει την *εισροή*, τη *διαδικασία*, το *προϊόν* και το *πλαίσιο* ενός γεγονότος ή μιας δράσης. Όταν αυτό το γενικό πλαίσιο εφαρμόζεται σε κάθε περίπτωση μαθημάτων, έχει ως αποτέλεσμα τα έξι θέματα που παρουσιάζονται στο Διάγραμμα 1.



Διάγραμμα 1: Πολλαπλές Διαστάσεις της Διδασκαλίας

Η πρώτη διάσταση της διδασκαλίας στην τριτοβάθμια εκπαίδευση είναι η «εισροή», και αυτή συνίσταται από δύο παράγοντες *τα χαρακτηριστικά των φοιτητών και τα χαρακτηριστικά του καθηγητή*. Ατομικά οι φοιτητές διαφέρουν αξιοσημείωτα στις γνώσεις, στις αξίες, και στις πεποιθήσεις που φέρνουν στη διαδικασία μάθησης. Επιπρόσθετα, οι προσωπικότητες των φοιτητών σε μία συγκεκριμένη τάξη μπορεί επίσης να αποτελούν ένα σημαντικό παράγοντα στην επιτυχία ενός δοσμένου μαθήματος. Παρόμοια, οι καθηγητές διαφέρουν στην ετοιμότητά τους να διδάξουν οποιοδήποτε δοσμένο μάθημα. Μερικές φορές το αντικείμενο μελέτης αποτελεί ένα θέμα το οποίο έχει υπάρξει ενδιαφέρον για ένα μέλος ΔΕΠ για πολλά χρόνια. Σε μία άλλη στιγμή ένα μέλος ΔΕΠ μπορεί να έχει να διδάξει ένα μάθημα για το οποίο αυτή ή αυτός διαθέτει περιορισμένο υπόβαθρο γνώσεων και περιορισμένο κίνητρο. Μία άλλη σημαντική μεταβλητή για τους καθηγητές είναι ο βαθμός στον οποίο αυτοί έχουν μάθει τον τρόπο να διδάσκουν σε διαφορετικές καταστάσεις: χαμηλότερη κατηγορία καθώς και μεταπτυχιακά μαθήματα, μεγάλες τάξεις καθώς και μικρές τάξεις, ή μαθήματα που απαιτούν διαδικασίες ενεργού μάθησης καθώς και μαθήματα με διαλέξεις. Ο καθηγητής που είναι ενήμερος στον τομέα του και έχει πραγματοποιήσει την απαιτούμενη έρευνα και προετοιμασία για ένα μάθημα παρέχει την αναγκαία *εισροή* για σημαντική μάθηση. Επομένως η ενδεχόμενη κατάρρευση οποιουδήποτε παράγοντα αυτής της συνιστώσας- *εισροής* περιορίζει τη μαθησιακή διαδικασία.

Η δεύτερη διάσταση της αξιολόγησης, η «*διαδικασία*», περιλαμβάνει δύο ξεχωριστές δραστηριότητες στη διδασκαλία: *τις αποφάσεις για τα μαθήματα και τις συμπεριφορές στην αίθουσα διδασκαλίας*. Όταν οι καθηγητές διδάσκουν μαθήματα, αυτοί παίρνουν αποφάσεις για το σκοπό του αντικειμένου μελέτης που καλύπτουν, τη στρατηγική διδασκαλίας που χρησιμοποιούν, το σύστημα βαθμολόγησης, και τις πολιτικές μαθήματος. Για να πάρουν τέτοιες αποφάσεις, οι καθηγητές χρειάζεται να αναλογιστούν κάποιους παράγοντες όπως είναι η φύση του προγράμματος σπουδών και τα χαρακτηριστικά των φοιτητών και μετά ανάλογα να σχεδιάσουν τη σειρά των μαθημάτων. Ακόμα, μέρος της «*διαδικασίας*» της διδασκαλίας, αλλά τελείως διαφορετικό στη φύση, είναι το τι κάνουν οι καθηγητές στην αίθουσα διδασκαλίας. Από τη στιγμή που έχουν ληφθεί οι αποφάσεις για τη σειρά μαθημάτων και οι καθηγητές προχωρούν προς την τάξη, αυτοί πρέπει να ενεργοποιήσουν όποιες δεξιότητες επικοινωνιακές και αλληλεπίδρασης διαθέτουν να παραδώσουν με

διαλέξεις, να καθοδηγήσουν συζητήσεις, να ρωτήσουν, να παρακινήσουν τους φοιτητές και να δημιουργήσουν ενδιαφέρον.

Η τρίτη διάσταση ασχολείται με το «προϊόν» της διδασκαλίας, που είναι η ποσότητα και ο τύπος της μάθησης η οποία συμβαίνει σε ένα δοσμένο μάθημα. Σχεδόν σε κάθε μάθημα οι φοιτητές πηγαίνουν για να μάθουν κάτι. Στα καλά μαθήματα ένα μεγάλο ποσοστό των φοιτητών μαθαίνει πολλά και αυτοί μαθαίνουν πράγματα τα οποία είναι σημαντικά παρά κοινότοπα. Ένα άλλο σημαντικό κομμάτι της μάθησης αποτελεί το ενδιαφέρον των φοιτητών για περισσότερη μάθηση. Προώθησε το μάθημα ένα τέτοιο ενδιαφέρον ή όχι;

Η τέταρτη διάσταση είναι το πλαίσιο. Στη διδασκαλία στη τριτοβάθμια εκπαίδευση υπάρχουν διάφορα πλαίσια τα οποία επηρεάζουν την ποιότητα μίας δοσμένης περίπτωσης διδασκαλίας: το φυσικό (π. χ τα χαρακτηριστικά της αίθουσας διδασκαλίας), το κοινωνικό (π. χ η φύση των φοιτητών), το θεσμικό (π. χ η υποστήριξη που δίνεται στη διδασκαλία), και το προσωπικό (π. χ άλλα συμβάντα στη ζωή του καθηγητή).

6.4.1 Τα πέντε γενικά ερωτήματα

Τα πέντε γενικά ερωτήματα και τα σχετικά επιμέρους σημεία που παρουσιάζονται στη συνέχεια είναι εφαρμόσιμα στη διδασκαλία σε όλα τα μαθήματα, σε ένα σκηνικό της τριτοβάθμιας εκπαίδευσης. Ο τρόπος με τον οποίο οι απαντήσεις διατυπώνονται σε αυτές τις ερωτήσεις διαφέρουν από τμήμα σε τμήμα αλλά οι ερωτήσεις από μόνες τους είναι εγγενείς της φύσης της διδασκαλίας και της φύσης της αξιολόγησης. Αυτές οι σημαντικές ερωτήσεις είναι οι επόμενες:

- Διαθέτει ο καθηγητής κατάλληλες και σύγχρονες γνώσεις στο αντικείμενο μελέτης του περιλαμβάνοντας την ακαδημαϊκή και / ή την πρακτική εμπειρία και τις προσπάθειες να τις βελτιώσει;
- Πόσο καλές ήταν οι αποφάσεις του καθηγητή σχετικά με: τη σειρά μαθημάτων, τους συμπεριλαμβανομένους στόχους, τις στρατηγικές διδασκαλίας, τις αναθέσεις μελέτης στο σπίτι, τον έλεγχο, και τη βαθμολόγηση;
- Πόσο καλά η συμπεριφορά του καθηγητή στην τάξη προώθησε την καλή μάθηση, συμπεριλαμβανομένης της οργάνωσης και της σαφήνειας, του ενθουσιασμού, της αλληλεπίδρασης με την τάξη ως ενός συνόλου, και των σχέσεων ατομικά με τους φοιτητές;

- Πόσο καλά ήταν τα εκπαιδευτικά αποτελέσματα του μαθήματος, συμπεριλαμβανομένης και της ποσότητας της μάθησης, της σημαντικότητας αυτού που έμαθαν οι φοιτητές, και των στάσεων απέναντι στην παραπάνω μάθηση σχετικά με το αντικείμενο;
- Πόσο η ποιότητα της διδασκαλίας και της μάθησης επηρεάστηκε από παράγοντες πλαισίου, συμπεριλαμβάνοντας το φυσικό πλαίσιο, το κοινωνικό πλαίσιο, το θεσμικό πλαίσιο και το προσωπικό πλαίσιο;

6.4.2 Αξιολογώντας την ποιότητα της διδασκαλίας

Η ποιότητα της διδασκαλίας, ωστόσο, μπορεί να εκληφθεί ως αποτελούμενη από έξι στοιχεία:

- τα χαρακτηριστικά των φοιτητών,
- τα χαρακτηριστικά του καθηγητή,
- τις αποφάσεις του καθηγητή για το μάθημα,
- τη συμπεριφορά του καθηγητή στην τάξη,
- την ποσότητα της σημαντικής μάθησης, και
- την επίδραση των παραγόντων πλαισίου.

Με σκοπό να αξιολογηθεί αποτελεσματικά κάθε ειδική περίπτωση διδασκαλίας, κάποιος πρέπει να ασχοληθεί με τη συλλογή και την ανάλυση της πληροφορίας κάθε ενός από αυτά τα συστατικά.

Καμία πηγή πληροφόρησης από μόνη της, ωστόσο, δεν είναι κατάλληλη για να αξιολογήσει και τις έξι συνιστώσες της διδασκαλίας. Αυτό σημαίνει ότι πολλαπλές πηγές πληροφορίας δεν είναι μόνο ενδεδειγμένες αλλά στην πραγματικότητα είναι και αναγκαίες. Συνεπώς οι διάφορες πηγές πληροφόρησης πρέπει να αξιολογούνται για να καθορίσουν τη σχετική τους αξία να απαντήσουν για κάθε μία από τις έξι συνιστώσες.

Ένας αριθμός από διαφορετικές **πηγές πληροφόρησης** είναι διαθέσιμος για αυτό το συγκεκριμένο καθήκον, και αυτές είναι:

- τα υλικά του μαθήματος,
- οι φοιτητές (οι τωρινοί, οι προηγούμενοι και οι απόφοιτοι),
- οι παρατηρήσεις του καθηγητή,
- οι συνάδελφοι (π. χ άλλα μέλη ΔΕΠ)

- οι διοικούντες, και
- οι παρατηρήσεις ενός συμβούλου εκπαίδευσης.

Μία επιπλέον διάκριση πρέπει να γίνει η οποία αφορά στους τρεις τύπους των περιπτώσεων αξιολόγησης, κοινούς στο ακαδημαϊκό σκηνικό:

- Ετήσιες αποφάσεις προσωπικού,
- Περιοδικές αποφάσεις προσωπικού (π. χ θητεία, προαγωγή, βραβείο διδασκαλίας,
- Διάγνωση αυτό-βελτίωσης.

Αυτές οι τρεις περιπτώσεις έχουν σε κάποιο βαθμό ομοιότητες, αλλά οι διαφορές είναι επαρκείς για να δικαιολογούν μια ιδιαίτερη θεώρηση των σκοπών της αξιολόγησης.

Όταν η διδασκαλία αξιολογείται ως μέρος μίας ετήσιας αξιολόγησης της επίδοσης, το ερώτημα είναι: **«Πόσο καλά αυτό το πρόσωπο δίδαξε αυτό τον χρόνο, συγκριτικά με τους άλλους σε αυτή την ακαδημαϊκή μονάδα;»**. Αυτό σημαίνει ότι η μόνη σχετική πληροφορία είναι η πληροφορία που αφορά στην επίδοση του καθηγητή αυτό τον χρόνο. Ως εκ τούτου, η πληροφορία από τους πρεσβύτερους φοιτητές είναι εγγενώς άσχετη.

Στη δεύτερη περίπτωση, περιοδικές αποφάσεις προσωπικού, διατυπώνεται μία διαφορετική ερώτηση **«Αυτό το άτομο γενικά διδάσκει αρκετά καλά για να αξίζει τη παραμονή, την προαγωγή ή το βραβείο διδασκαλίας;»**. Σε αυτή την περίπτωση, η πληροφορία από τους πρεσβύτερους φοιτητές δεν είναι μόνο σχετική αλλά και σημαντική. Τέλος, στην περίπτωση της διάγνωσης για αυτό-βελτίωση, το ερώτημα είναι **«Ποιες πλευρές της διδασκαλίας μου μπορούν να βελτιωθούν περισσότερο παραγωγικά;»**. Σε αυτή την περίπτωση, όλες οι πηγές πληροφόρησης είναι σχετικές σε κάποια έκταση. Η βασική διαφορά είναι ότι εδώ ένας σύμβουλος εκπαίδευσης είναι έτοιμος να γίνει μία σημαντική πηγή πληροφοριών.

Εκτιμάται ως αναγκαίο να γίνει αναφορά σε μία ειδική διαδικασία αξιολόγησης η οποία έχει κερδίσει ευρύ ενδιαφέρον σε εθνικό και διεθνές επίπεδο τα τελευταία χρόνια: **το χαρτοφυλάκιο διδασκαλίας**. Ο Seldin (1991, 1997) υπήρξε ο πρωτοπόρος δείχνοντας τις πολλές χρήσεις των χαρτοφυλακίων διδασκαλίας και περιγράφοντας τον τρόπο δημιουργίας τους. Βασικά ένα χαρτοφυλάκιο διδασκαλίας αποτελείται από μια σύντομη περιγραφή (π. χ 6-9 σελίδες) που περιγράφει και αξιολογεί τη διδασκαλία του, μαζί με ένα κατάλογο που περιέχει όλα τα αποδεικτικά

στοιχεία. Αν και τα χαρτοφυλάκια διδασκαλίας είναι ιδιαίτερης αξίας για την αυτό-αξιολόγηση, τα μέλη ΔΕΠ μπορούν να τα χρησιμοποιούν σε οποιαδήποτε από τις τρεις περιπτώσεις αξιολόγησης που αναφέρθηκαν προθύστερα.

6.4.3 Προτεινόμενες διαδικασίες

Εάν κάποιος αποδέχεται τις δύο αρχές που εξηγούνται ήδη από τις πολλαπλές διαστάσεις της διδασκαλίας και τις πολλαπλές πηγές πληροφόρησης και ακόμα, τη διαφορά στις τρεις περιπτώσεις αξιολόγησης, τότε οι δύο αρχές της ορθής αξιολόγησης γίνονται προφανείς:

- Η χρήση διαφορετικών πηγών πληροφόρησης για την αξιολόγηση διαφορετικών πλευρών της διδασκαλίας.
- Η χρήση διαφορετικών πηγών σε διαφορετικές περιπτώσεις αξιολόγησης.

Το Διάγραμμα 2 συγκεντρώνει τις πηγές πληροφόρησης που συνιστώνται για κάθε μία από τις έξι διαστάσεις της διδασκαλίας σε κάθε μία από τις τρεις περιπτώσεις αξιολόγησης.

ΠΑΡΑΓΟΝΤΕΣ ΠΟΙΟΤΗΤΑΣ	Υλικά μαθήματος	ΦΟΙΤΗΤΕΣ			Παρατηρήσεις καθηγητή	Συνάδελφοι	Διοίκηση	Σύμβουλος Εκπαίδευσης
		Τωρινού εξάμηνο	Πρεσβύτεροι	Απόφοιτοι				
1. Χαρακτηριστικά φοιτητών		E	Π		E Π Δ	ε π	ε δ	δ
2. Γνώσεις καθηγητή	ε				E Π Δ			
3. Αποφάσεις για μαθήματα	E Δ	ε δ	π	Π	ε Π Δ	ε δ	ε	Δ
4. Συμπεριφορά στην τάξη		E Δ	π	Π	ε π Δ	ε π δ	ε π	Δ
5. Μάθηση		E	π	Π	ε Π Δ	ε π δ	ε π	Δ
6. Πλαίσιο		E Δ			E Π Δ	ε δ	ε	Δ

Διάγραμμα 2: Προτεινόμενες Πηγές Πληροφόρησης για την Αξιολόγηση της Διδασκαλίας

Τρεις Περιπτώσεις Αξιολόγησης

E / **ε** = ετήσιες αποφάσεις προσωπικού

Π / **π** = περιοδικές αποφάσεις προσωπικού

Δ / **δ** = διάγνωση αυτό-αξιολόγησης

E / **Π** / **Δ** = πρωτεύουσα πηγή

ε / **π** / **δ** = δευτερεύουσα πηγή

6.5 Το μοντέλο 360^ο Πολλαπλών Πηγών Ανατροφοδότησης (ΠΠΑ)

Πηγή: Berk A. R. (2009). Using the 360o multisource feedback model to evaluate teaching and professionalism, *Medical Teacher*, 31, p: 1073-1080.

Κρίσιμες αναθεωρήσεις των στρατηγικών για την αξιολόγηση των συμπεριφορών διδασκαλίας στη βιβλιογραφία για την τριτοβάθμια εκπαίδευση προτείνουν μία ποικιλία από πιθανούς βαθμολογητές, περιλαμβάνοντας τους φοιτητές, τον εαυτό, τους ομοβάθμιους, τους εξωτερικούς εμπειρογνώμονες, τους καθοδηγητές (mentors), τους απόφοιτους, τους εργαζόμενους και τα μέλη της διοίκησης, και 14 διαφορετικές πιθανές πηγές απόδειξης (Berk 2005, 2006). Αυτές όλες είναι εφαρμόσιμες στην αξιολόγηση των μελών ΔΕΠ.

Το 360^ο Πολλαπλών Πηγών Ανατροφοδότησης μοντέλο (360^ο ΠΠΑ) εξαρτάται από τις ειδικές αποφάσεις που θα παρθούν. Στην αξιολόγηση των μελών ΔΕΠ, υπάρχουν τρεις εφαρμογές του μοντέλου που μπορούν να δημιουργηθούν για αποφάσεις διαμορφωτικές και συγκεντρωτικές σχετικά με τις συμπεριφορές διδασκαλίας και διαμορφωτικές αποφάσεις σχετικά με τις επαγγελματικές συμπεριφορές. Στη συνέχεια περιγράφεται ο τρόπος με τον οποίο οι βαθμολογητές και οι παραπάνω πηγές μπορούν να ταιριάξουν σε αυτά τα διαφορετικά μοντέλα (Berk, 2009, σελ. 1076-78). Ο «καθηγητής είναι το κομβικό σημείο» σε όλες τις εφαρμογές. Οι βαθμολογητές / πηγές ποικίλουν για κάθε απόφαση. Ακόμα, αυτές μπορεί να είναι διαφορετικές για κάθε τμήμα και ίδρυμα. Για να καταδειχτούν εδώ τα μοντέλα ένα συνδυασμός των «καλύτερων πρακτικών» των βαθμολογητών / πηγών, βασιζόμενος σε προηγούμενες επισκοπήσεις, συνιστάται για κάθε απόφαση. Αυτές οι επιλογές μπορεί να μην είναι οι καλύτερες για κάθε ίδρυμα.

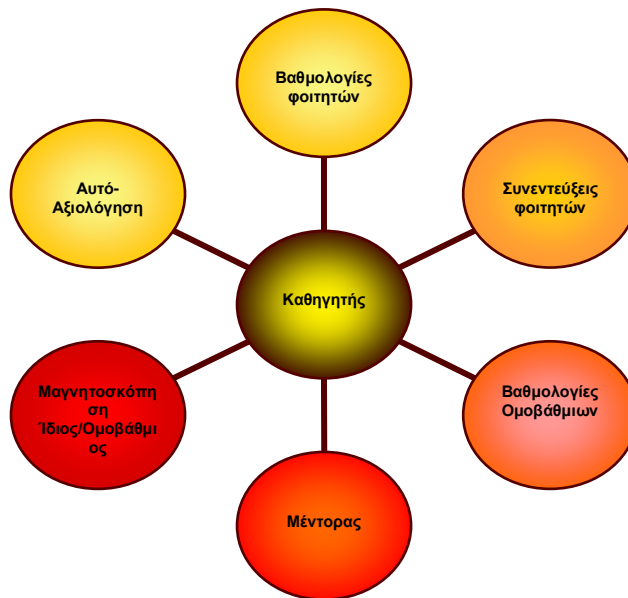
6.5.1 Διαμορφωτικές αποφάσεις σχετικές με τις συμπεριφορές διδασκαλίας

Θεωρώντας ότι οι κλίμακες αξιολόγησης είναι κατάλληλα δομημένες, με επαρκή απόδειξη για την αξιοπιστία και την εγκυρότητά τους, πέντε από τις καλύτερες πηγές που θα μπορούσαν να χρησιμοποιηθούν για να παρέχουν ακριβή πληροφόρηση για τα δυνατά σημεία και τις αδυναμίες του μέλους ΔΕΠ καθώς και προτάσεις για το πώς να βελτιωθούν είναι:

- ο καθηγητής μέντορας (από πάνω βαθμίδα),
- οι βαθμολογίες από τους ομοβάθμιους και
- η μαγνητοσκόπηση με τον ίδιο / ομοβάθμιο (ίδια βαθμίδα),
- οι βαθμολογίες από τους φοιτητές και

- οι συνεντεύξεις των φοιτητών (από κάτω επίπεδο),
- η αυτό-αξιολόγηση.

Διαφορετικές κλίμακες πρέπει να δοθούν στον μέντορα, στους ομοβάθμιους και στους φοιτητές. Ο καθηγητής θα μπορούσε επίσης να συμπληρώσει αυτές τις κλίμακες. Ανακολουθίες μεταξύ των δικών του/ή της βαθμολογιών και εκείνων από τους άλλους τρεις βαθμολογητές μπορούν να δώσουν ένα πορτραίτο των δυνατών και των αδύνατων σημείων και να τονίσουν ειδικές συμπεριφορές στο μάθημα που χρειάζονται προσοχή. Το 360^ο ΠΠΑ μοντέλο παρουσιάζεται στο Διάγραμμα 3.



Διάγραμμα 3: Το 360^ο ΠΠΑ μοντέλο αξιολόγησης ενός καθηγητή τριτοβάθμιας εκπαίδευσης (αποφάσεις διαμορφωτικού τύπου για τη διδασκαλία)

Τα χαρακτηριστικά του διαμορφωτικού τύπου μοντέλου είναι τα κάτωθι:

1. Ο καθηγητής επιλέγει τους βαθμολογητές και τις πηγές απόδειξης.
2. Οι βαθμολογητές θα πρέπει να είναι σε θέση να γνωρίζουν τις συμπεριφορές διδασκαλίας του καθηγητή.
3. Οι συμπεριφορές διδασκαλίας όπως ορίζονται για κάθε πηγή μπορεί να είναι διαφορετικές και συμπληρωματικές, αν και μπορεί να υπάρχει αλληλοεπικάλυψη μεταξύ των υποκλιμάκων (subscales) των κλιμάκων που προορίζονται για τους φοιτητές και τους ομότιμους.
4. Το δείγμα των βαθμολογητών μπορεί να είναι μεγάλο, περιλαμβάνοντας φοιτητές, ένα ή περισσότερους ομοβάθμιους, και ένα μέντορα.

5. Αναπτύσσονται διαφορετικού τύπου κλίμακες Likert για διαφορετικούς τύπους βαθμολογητών (βλέπε Παράρτημα) οι οποίες μετρούν διαφορετικές συμπεριφορές των μελών ΔΕΠ. Για παράδειγμα «οι κλίμακες βαθμολογίας από τους φοιτητές» επικεντρώνεται σε συμπεριφορές διδασκαλίας που σχετίζονται με το περιεχόμενο και την οργάνωση, με τα μαθησιακά αποτελέσματα, με τις μεθόδους διδασκαλίας, το κλίμα και τις μεθόδους αξιολόγησης. Επίσης, οι «κλίμακες παρατήρησης από τους ομοβάθμιους» επικεντρώνονται στο περιεχόμενο και στην οργάνωση, στο στυλ επικοινωνίας, στις δεξιότητες διατύπωσης ερωτήσεων (ή έρευνας), στις δεξιότητες κριτικής σκέψης, της αμοιβαίας συμπάθειας με τους μαθητές, το μαθησιακό περιβάλλον και τις μεθόδους διδασκαλίας.
6. Η ποιότητα πολλών εντόπιων κλιμάκων διαφέρει από πολύ καλή σε φτωχή, αλλά οι εμπορίου (commercial) κλίμακες βαθμολόγησης από τους φοιτητές είναι καλύτερες.
7. Ο αριθμός των σημείων-επιλογών ποικίλουν σε κάθε κλίμακα, για παράδειγμα, κλίμακες 4-7 σημείων από «συμφωνώ απόλυτα» μέχρι «διαφωνώ απόλυτα» και από «πάντα» έως «ποτέ» ή κλίμακες 4-σημείων από «εξαιρετικό» έως «χρειάζεται βελτίωση».
8. Η διαχείριση των κλιμάκων βαθμολόγησης από τους φοιτητές γίνεται ηλεκτρονικά και με χαρτί και μολύβι, με βαθμό συμπλήρωσης από το 30% μέχρι και το 90%, αλλά όλες οι άλλες κλίμακες είναι τυπικά διαχειριζόμενες με χαρτί και μολύβι. Τα δεδομένα από διαφορετικές πηγές μπορεί να συλλεχθούν σε διαφορετικούς χρόνους κατά τη διάρκεια του εξαμήνου.
9. Η ανατροφοδότηση του μέλους ΔΕΠ με βάση τις βαθμολογίες από τους φοιτητές μπορεί να διαρκέσει λιγότερο από 2 βδομάδες (online), ή περισσότερο από ένα μήνα (χαρτί-και-μολύβι), άμεση από τις συνεντεύξεις των φοιτητών, ή μέσα σε μερικές ημέρες και πρόσωπο-με-πρόσωπο με τους ομότιμους και τον μέντορα.
10. Ο καθηγητής παρακολουθεί την τροχιά των αλλαγών και την πρόοδο στη βελτίωση της διδασκαλίας από τις διαφορετικές πηγές απόδειξης και από τα έγγραφα βελτιώσεων.

6.5.2 Συγκεντρωτικές αποφάσεις σχετικές με τις συμπεριφορές διδασκαλίας

Σύμφωνα με την ίδια «λιμνούλα» των 13 στρατηγικών που παρουσιάστηκαν στο Τμήμα 3, το ζητούμενο είναι ποιες θα συνιστούσαν την καλύτερη επιλογή για τον πρόεδρο του τμήματος, τον κοσμήτορα, ή τον αναπληρωτή κοσμήτορα με τις οποίες αυτοί μπορούν να καθορίσουν τον μισθό ή την ανανέωση της σύμβασης (για μη-μόνιμο, πλήρες- και μερικής απασχόλησης) του μέλους ΔΕΠ. Συνεπώς το ζητούμενο αφορά στις πηγές εκείνες που μπορούν να παρέχουν ακριβή πληροφορία για την αποτελεσματικότητα της διδασκαλίας αλλά επίσης να μπορούν να συγκλίνουν από κοινού σε μία απόφαση η οποία να είναι δίκαιη και σωστή.

Ενδιαφέρον παρουσιάζει το γεγονός ότι, οι «καλύτερες» πηγές σε αυτή την περίπτωση είναι σχεδόν πανομοιότυπες με κάποιες από εκείνες που επελέγησαν για τις διαμορφωτικού τύπου αποφάσεις. Αυτές περιλαμβάνουν τις βαθμολογίες από τον πρόεδρο του τμήματος και τον μέντορα (από πάνω), τις βαθμολογίες από τους ομοβάθμιους (προαιρετικά) και τη μαγνητοσκόπηση (προαιρετικά) (ίδιο επίπεδο), τις βαθμολογίες από τους φοιτητές (από κάτω), συν την αυτό-αξιολόγηση. Η χρήση των βαθμολογιών από τους ομοβάθμιους, της επίδοσης και της ανάδρασης από τη μαγνητοσκόπηση και της βαθμολογίας από τον μέντορα πρέπει να επαφίενται στη διακριτική ευχέρεια του μέλους ΔΕΠ. Διαφορετικά, θα αποτελούσε παραβίαση του απορρήτου (Berk et al. 2004). Τα στοιχεία αυτά συλλέχθηκαν αρχικά για τη διαμορφωτική χρήση, με εμπιστοσύνη προς τον αντίστοιχο ομοβάθμιο ή μέντορα. Οι συγκεντρωτικές αποφάσεις είναι διαφορετικές. Είτε η πρόθεση της πληροφορίας για τους δύο τύπους αποφάσεων συμφωνεί στα περισσότερα είτε όχι, η τελική χρήση οποιασδήποτε από εκείνες τις πηγές πρέπει να επαφίεται στο μέλος ΔΕΠ.

Τα χαρακτηριστικά αυτού του συγκεντρωτικού μοντέλου είναι πολύ διαφορετικά από του διαμορφωτικό, αν και οι πηγές μπορεί να είναι οι ίδιες:

1. Το μέλος ΔΕΠ και οι διοικούντες ορίζουν τους βαθμολογητές και τις πηγές απόδειξης.
2. Πολλαπλής προέλευσης βαθμολογητές πρέπει να επιλεγούν για την εξειδίκευσή τους και για να ελαχιστοποιήσουν τους διάφορους τύπους των προκαταλήψεων γύρω από τη βαθμολόγηση (Berk, 2006).
3. Οι συμπεριφορές διδασκαλίας μπορεί να ορίζονται διαφορετικά ως προς κάθε πηγή.

4. Το μεγάλο δείγμα βαθμολογητών μπορεί να περιλαμβάνει όλους τους φοιτητές, ένα ή δύο ομοβάθμιους που συμφωνήθηκαν από τον καθηγητή και τον διευθυντή, τον μέντορα και τον ίδιο τον καθηγητή, και τον διευθυντή.
5. Ξεχωριστές εκθέσεις με τις παρατηρήσεις των ομοβάθμιων, διαφορετικού περιεχομένου και πεδίου εφαρμογής από ένα, μπορεί να προετοιμαστούν από τους βαθμολογητές και να υποβληθούν στον πρόεδρο του τμήματος (Berk et al. 2004).
6. Ανάδραση στον καθηγητή από τον πρόεδρο του τμήματος πάνω σε όλη τη πληροφορία που προέκυψε από τα δεδομένα όλων των πηγών συμβαίνει πρόσωπο-με πρόσωπο στο τέλος του ακαδημαϊκού χρόνου.
7. Ο πρόεδρος του τμήματος παρακολουθεί τις αλλαγές και την πρόοδο στη διδασκαλία διαχρονικά (longitudinally) καθώς και τα έγγραφα βελτιώσεων. Αυτός καθορίζει την αξία πληρωμών και την ανανέωση της σύμβασης. Απόδειξη μπορεί ακόμα να περιλαμβάνεται στο χαρτοφυλάκιο διδασκαλίας για την προαγωγή και την ανανέωση της θητείας.

6.5.3 Διαμορφωτικές αποφάσεις σχετικές με επαγγελματικές συμπεριφορές

Οι επαγγελματικές συμπεριφορές που μετρώνται με βάση το μοντέλο 360^ο ΠΠΑ (<http://professionalbehaviors.nbme.org/2008listofbehaviors.pdf>) μπορούν να επεκταθούν στις επαγγελματικές συμπεριφορές των μελών ΔΕΠ μίας σχολής. Σε αυτή την εφαρμογή για διαμορφωτικού τύπου αποφάσεις ο καθηγητής βρίσκεται πάλι στο κέντρο του τροχού. Ένα μεγάλο δείγμα από βαθμολογητές επιλέγεται από το μέλος ΔΕΠ και τον πρόεδρο του τμήματος, το οποίο, σε αυτή την περίπτωση, είναι διαφορετικό και περισσότερο ποικίλο από εκείνο στα προαναφερθέντα μοντέλα. Μαζί με τις προαναφερθείσες πηγές των φοιτητών (από κάτω), των ομοβάθμιων (ίδιο επίπεδο) και του προέδρου του τμήματος (από πάνω), αυτό το μοντέλο θα μπορούσε επίσης να περιλαμβάνει συναδέλφους και φοιτητές στους ρόλους των μεντόρων, επιστημονικούς συνεργάτες, βοηθούς στη διδασκαλία και βοηθούς εργαστηρίου, συν διοικητικούς βοηθούς, ΤΠ προσωπικό, προσωπικό υποδοχής, και άλλο προσωπικό. Διευκρινίζεται ότι στην ακαδημαϊκή «τροφική» αλυσίδα οι περισσότεροι βαθμολογητές είναι κάτω από τον καθηγητή. Αυτοί συχνά θεωρούν τους εαυτούς τους ως τους πλέον ευάλωτους και, στην πραγματικότητα, μπορεί να υποκινούν τα προβλήματα στις σχέσεις, για παράδειγμα με, φωνές, σκληρά λόγια, αγένεια, γελοιοποίηση, υπονοούμενα και άσχημα σχόλια, ή με ύπουλη, παθητική επιθετική ή βία συμπεριφορά κατά των μελών ΔΕΠ. Τα πανεπιστήμια που υιοθετούν κώδικες πολιτισμού (Forni, 2002) για τις συμπεριφορές των φοιτητών στην τάξη, φαίνεται

αναγκαίο να δίνουν οδηγίες στα μέλη ΔΕΠ τις οποίες να τηρούν έτσι ώστε οι φοιτητές να έχουν κατάλληλα μοντέλα ρόλων για να παραδειγματίζονται.

Οι επαγγελματικές συμπεριφορές των καθηγητών θα μπορούσαν να κατηγοριοποιηθούν περιλαμβάνοντας: τη συναισθηματική νοημοσύνη στις διαπροσωπικές και διαπροσωπικές δεξιότητες, την ομαδική εργασία, την επικοινωνία, την προσβασιμότητα, την υπευθυνότητα, τον αλτρουισμό, τη δικαιοσύνη, την αξιοπρέπεια, τη τιμιότητα, την εμπιστοσύνη, το σεβασμό, την φροντίδα, την υπομονή και τη συμπόνια.

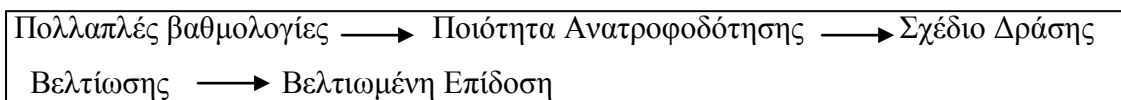
Οι διαστάσεις του χαρακτήρα ή οι ιδιότητές όπως, η ηγεσία, η αρίστευση, η δημιουργικότητα, τα κίνητρα, η ποικιλομορφία, οι αξίες, οι φιλοδοξίες, η αυτό-εμπιστοσύνη και η πρωτοβουλία θα μπορούσαν επίσης να υπόκεινται στο επαγγελματισμό. Καμία από αυτές τις συμπεριφορές δεν μετριέται συνήθως με τις βαθμολογίες από τους φοιτητές ή με άλλες πηγές. Όλα αυτά τα χαρακτηριστικά επίσης μπορούν να αξιολογούνται για τα μέλη της διοίκησης και για το άλλο ακαδημαϊκό προσωπικό.

Μέρος της διαδικασίας μέτρησης μπορεί να αποτελέσει η παρατήρηση από πολλαπλούς βαθμολογητές και η μέτρηση με τη βοήθεια κατάλληλης κλίμακας ενός κοινού πυρήνα συμπεριφορών. Ακόμα, μία ξεχωριστή υπό-κλίμακα στην κλίμακα βαθμολόγησης από τους φοιτητές μπορεί να εισάγει αυτές τις συμπεριφορές. Για να διατηρηθεί η ανωνυμία και η εμπιστοσύνη, αυτές οι κλίμακες πρέπει να διαμοιράζονται μέσω του διαδικτύου δύο φορές το χρόνο (ή μία το εξάμηνο) σε σταθερούς χρόνους. Ο καθηγητής πρέπει επίσης να συμπληρώνει την κλίμακα.

Όλες οι αναλύσεις που έχουν διεξαχθεί υποβάλλονται στον πρόεδρο του τμήματος, ο οποίος στη συνέχεια προγραμματίζει συναντήσεις ατομικά με κάθε μέλος ΔΕΠ. Στις συναντήσεις παρέχεται άμεση πρόσωπο-με-πρόσωπο διαμορφωτική ανατροφοδότηση από όλους τους βαθμολογητές σχετικά με τις θετικές συμπεριφορές και τις τάσεις που αναφέρονται στις αρνητικές συμπεριφορές. Αυτές οι βαθμολογίες μπορούν επίσης να αντιπαρατεθούν με τις αυτό-αξιολογήσεις και τις βαθμολογίες του προέδρου του τμήματος για να αναγνωρισθούν οι ανακολουθίες. Ένα σχέδιο δράσης πρέπει τότε να αναπτυχθεί για να εισάγει τις αρνητικές συμπεριφορές και να παρακολουθεί τις βελτιώσεις στους μήνες που μεσολαβούν μέχρι την επόμενη εξαμηνιαία διαχείριση.

6.5.4 10 κορυφαία προβλήματα

Αυτό που μένει αμετάβλητο στις εκάστοτε εφαρμογές του μοντέλου είναι το αρχικό του πνεύμα και η αρχική του λειτουργία η οποία περιγράφεται:



Επισημαίνεται ότι αν και οι βαθμολογίες προορίζονται για διαμορφωτικού τύπου αποφάσεις σε πολλές περιπτώσεις χρησιμοποιούνται για συγκεντρωτικές αποφάσεις.

Όλες αυτές οι εφαρμογές του 360^ο ΠΠΑ μοντέλου έχουν τα πλεονεκτήματα και τα μειονεκτήματά τους. Στην πραγματικότητα, είναι πιθανό να αποστάξουν διάφορα επίμονα και, ίσως, δυσεπίλυτα ψυχομετρικά θέματα κατά την υλοποίηση του. Στη συνέχεια παρουσιάζονται 10 κορυφαία προβλήματα που χρειάζονται προσοχή:

1. Δεν υφίσταται αντικειμενική μέτρηση της επίδοσης της διδασκαλίας. Όλες οι πηγές απόδειξης είναι επιδεκτικές σφάλματος ή πλάνης.
2. Σχεδόν όλη η απόδειξη για την επίδοση προέρχεται από «πληροφορημένες» αποφάσεις των φοιτητών και εκείνων των ατόμων στο τμήμα με τα οποία ένας καθηγητής συνεργάζεται.
3. Υφίσταται δυσκολία στον ορισμό των τομέων της συμπεριφοράς «εργασίας» και «επαγγελματικής» συμπεριφοράς και στην εξασφάλιση συναίνεσης για το τι πρέπει να μετρά κάθε κλίμακα.
4. Η ποιότητα πολλών κλιμάκων που «αναπτύσσονται στο σπίτι» (homegrown) είναι ακατάλληλη με όρους αξιοπιστίας και εγκυρότητας όσον αφορά στις αποφάσεις για τις οποίες χρησιμοποιούνται.
5. Οι κλίμακες για την παρατήρηση από τους ομότιμους και τα πρωτόκολλα είναι εν μέρει αδύνατα, αν και οι μεταξύ-εκτιμητών αξιοπιστία είναι από μέτρια έως υψηλή σε πολλές περιπτώσεις.
6. Εννέα πηγές προκατάληψης του βαθμολογητή αποτελούν πρόκληση για εξάλειψη ή ακόμα για ελαχιστοποίηση.
7. Η τυποποίηση της διαχείρισης της κλίμακας μέσω του διαδικτύου ή με τη μολύβι-και-χαρτί λειτουργία συνήθως είναι ανεπαρκής για να επιβεβαιώνει τις σωστές ερμηνείες των αποτελεσμάτων.
8. Τα ποσοστά ανταπόκρισης από τη διαχείριση κλιμάκων μέσω του διαδικτύου είναι μεταβλητά και απαράδεκτα χαμηλά σε κάποια ιδρύματα (π. χ 30-50%).

9. Σαφή, κατανοητά, με νόημα και κατάλληλα έντυπα έκθεσης απαιτούνται για τα αποτελέσματα της αξιολόγησης για τα μέλη ΔΕΠ και τους διοικούντες.
10. Απαραίτητες είναι οι οδηγίες και η εκπαίδευση για την επαρκή, δίκαια και ισότιμη ερμηνεία των αποτελεσμάτων από πολλαπλούς βαθμολογητές της επίδοσης, η οποία έτσι θα παρέχει ευαίσθητη και κατάλληλη ανατροφοδότηση.

Όλες αυτές οι περιοχές απαιτούν συνεχή έρευνα. Αν και πολλά έχουν γίνει γνωστά στη διάρκεια των 80-χρόνων ιστορίας της διαμόρφωσης κλιμάκων (scaling), των 60-χρόνων ιστορίας της αξιολόγησης των μελών ΔΕΠ, και των 50-χρόνων ιστορίας του 360° ΠΠΑ μοντέλου στο χώρο της διοίκησης (management)/ βιομηχανίας, ένα σωρό δουλειά είναι αναγκαία ακόμα να γίνει για να συλλάβει το αληθινό νόημα των «καλύτερων πρακτικών» στην αξιολόγηση της διδασκαλίας και του επαγγελματισμού.

6.6 Το μοντέλο της Περιβάλλουσας Αξιολόγησης Δεδομένων (ΠΑΔ)

Πηγή: Martin, E. (2006). Efficiency and Quality in the Current Higher Education Context in Europe: an application of the data envelopment analysis methodology to performance assessment of departments within the University of Zaragoza. *Quality in Higher Education*, 12, (1), p:57-79.

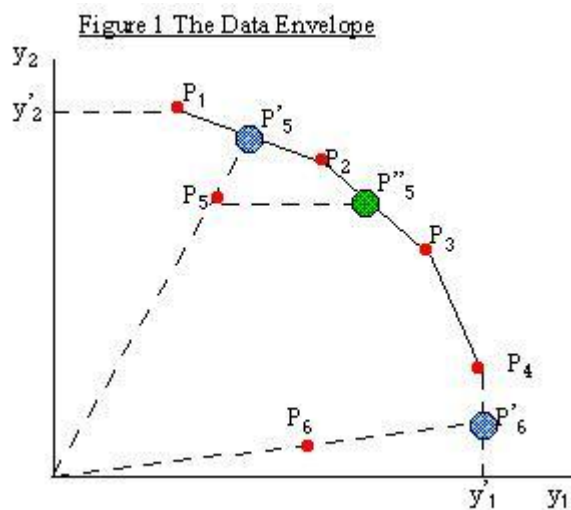
Τα ιδρύματα τριτοβάθμιας εκπαίδευσης ενδιαφέρονται για την καθιέρωση των δικών τους εσωτερικών μηχανισμών αξιολόγησης με σκοπό να καθοδηγούν την ολοκλήρωση των δικών τους αντικειμένων. Η αυτό-αξιολόγηση παίζει ένα σημαντικό ρόλο για τις στρατηγικές που εφαρμόζονται και στοχεύουν στη μέτρηση της επίδοσης και κατ' επέκταση στη βελτίωση της ποιότητας των ιδρυμάτων της τριτοβάθμιας εκπαίδευσης (Thune, 1998). Αυτή παρέχει κρίσιμη πληροφόρηση σχετικά με τις επιδόσεις των μονάδων που απαρτίζουν το ίδρυμα και έτσι επιτρέπει τη λήψη αποφάσεων που στοχεύουν στη βελτίωση της αποτελεσματικότητας και της ποιότητας των διαδικασιών. Εμπειρική ανάλυση των επιδόσεων των πανεπιστημίων λαμβάνει τυπικά με τη μορφή της αποτίμησης του λειτουργικού κόστους με επίκεντρο το μέγεθος και το σκοπό των οικονομικών ή με την ανάλυση της αποτελεσματικότητας χρησιμοποιώντας τη Περιβάλλουσα Ανάλυση Δεδομένων (ΠΑΔ) (Data Envelopment Analysis, DEA) ή τις λειτουργίες αιχμής (frontier functions) (Abbot & Doucouliagos, 2003).

Το μοντέλο της Περιβάλλουσας Ανάλυσης δεδομένων (ΠΑΔ) το εισήγαγε ο Charnes και οι συνεργάτες του (1978) και το επέκτεινε ο Banker με τους συνεργάτες του (1989). Το ΠΑΔ αποτελεί μία τεχνική με βάση το γραμμικό προγραμματισμό (programming – based) που χρησιμοποιείται για τη μέτρηση της επίδοσης των οργανωτικών μονάδων, σε περιπτώσεις που παρατηρούνται πολλαπλοί παράγοντες εισροών και εκροών και οι οποίοι δεν είναι δυνατό να μετατραπούν σε ένα σύνολο εισροών ή σε ένα παράγοντα εκροών. Από το 1978 χιλιάδες άρθρα έχουν δημοσιεύσει την εφαρμογή αυτής της μεθοδολογίας σε διαφορετικά πεδία.

Η μέθοδος είναι ιδιαίτερα κατάλληλη για την αξιολόγηση της αποτελεσματικότητας των μη κερδοσκοπικών φορέων όπως είναι τα δημόσια πανεπιστήμια, τα οποία λειτουργούν έξω από τις αγορές, δεδομένου ότι, γι' αυτά, οι μετρήσεις αποτελεσματικότητας όπως τα έσοδα και η κερδοφορία δεν είναι ικανοποιητικά. Οι δύο βασικοί λόγοι είναι: 1) αυτοί οι φορείς δεν επικεντρώνονται στην εξασφάλιση κερδών, και 2) η βασική πηγή χρηματοδότησής τους δεν προέρχεται από την πώληση αγαθών ή υπηρεσιών.

Το μοντέλο ΠΑΔ παρέχει ένα δείκτη συγκρίσιμης αποτελεσματικότητας των μονάδων που αξιολογούνται. Οι μονάδες που αναλύονται ορίζονται ως Μονάδες Λήψης Απόφασης (ΛΑΜς). Στο ΠΑΔ η *σχετική αποτελεσματικότητα* μίας ΛΑΜ ορίζεται ως ο λόγος των συνολικών σταθμισμένων εκροών προς τις συνολικές σταθμισμένες εισροές. Ως εκ τούτου, όσο υψηλότερος ο λόγος, τόσο μεγαλύτερη η αποτελεσματικότητα. Εάν διατηρείται η ομοιογένεια, οι εκροές και οι εισροές μπορούν να εκφράζονται σε οποιαδήποτε μονάδα μέτρησης.

Σε αντίθεση με τη λειτουργία της παραδοσιακής παραμετρικής παραγωγής, όπου μία συγκεκριμένη λειτουργική μορφή υποτίθεται ότι ισχύει για κάθε παρατήρηση, το ΠΑΔ δεν κάνει υποθέσεις σχετικά με τη μορφή της συνάρτησης παραγωγής. Οι πραγματικές εισροές και εκροές που παρατηρούνται χρησιμοποιούνται για την εκτίμηση ενός σημείου αναφοράς (ή κριτηρίου αξιολόγησης) στις λειτουργίες αιχμής (production frontier) (Διάγραμμα 4). Για το λόγο αυτό, ο δείκτης αποτελεσματικότητας που λαμβάνεται είναι σχετικός, δεδομένου ότι έχει αναπτυχτεί αναφορικά με τις μονάδες λήψης απόφασης (ΛΑΜς). Το ΠΑΔ επιτρέπει σε κάθε ΛΑΜ να επιλέγει τα διανύσματα βάρους των εισροών και των εκροών τα οποία μεγιστοποιούν το δικό της λόγο των σταθμισμένων εκροών προς τις σταθμισμένες εισροές, υπό τον περιορισμό ότι το διάνυσμα βάρους που επιλέγει για τη ΛΑΜ κατόπι ανάλυσης, δεν θα επιτρέψει στη ΛΑΜ να επιτύχαι ένα λόγο (σταθμισμένων εκροών προς τις σταθμισμένες εισροές) ο οποίος να υπερβαίνει τη μονάδα. Ως εκ τούτου, κάθε ΛΑΜ κρίνεται με βάση κάποια στάνταρ που καθορίζονται από την ίδια. Με άλλα λόγια, τα βάρη είναι τα πλέον ευνοϊκά για τη ΛΑΜ.



Διάγραμμα 4: Ο φάκελος δεδομένων (πηγή: Emrouznejad, 1995-2001)

Οι $P_1 \dots P_6$ είναι οι ΛΑΜς που αξιολογούνται και οι y_1, y_2 είναι οι εκροές. Όπως φαίνεται οι P_1, P_2, P_3 και P_4 είναι αποτελεσματικές.

Η στοιχειώδης μαθηματική έκφραση του μοντέλου για την j_0 th μονάδα είναι η ακόλουθη:

$$\max_{h_0} = (\sum_r u_r y_{rj_0}) / (\sum_i v_i x_{ij_0})$$

υπόκειται σε

$$\begin{aligned} (\sum_r u_r y_{rj}) / (\sum_i v_i x_{ij}) &\leq 1 \text{ για κάθε μονάδα } j \\ u_r, v_i &\geq \varepsilon \end{aligned}$$

όπου

u_r είναι το βάρος που δόθηκε στην εκροή r , y_{rj} είναι η εκροή ' r ' από τη μονάδα j_0
 v_i είναι το βάρος που δόθηκε στην εισροή i , x_{ij} είναι η εισροή ' i ' από τη μονάδα j_0
 r : 1, ..., s (αριθμός εκροών) και i : 1, ..., m (αριθμός εισροών)

Η παραπάνω εξίσωση παρουσιάζει την αξιολόγηση της επίδοσης για τη ΛΑΜ j_0 . Όπως μπορεί να ειπωθεί, ο σκοπός του μοντέλου είναι να αποφασιστούν τα u_r και v_i που μεγιστοποιούν το λόγο των σταθμισμένων εκροών προς τις σταθμισμένες εισροές. Ωστόσο, όπως ανατακλάται στους περιορισμούς, όταν εφαρμόζονται τα επιλεχθέντα βάρη σε άλλη ΛΑΜς η τιμή του λόγου δεν μπορεί να υπερβεί τη μονάδα. Κατά συνέπεια, η μέγιστη τιμή για την αντικειμενική συνάρτηση είναι 1. Εάν αυτό το αποτέλεσμα επιτυγχάνεται μπορεί να ειπωθεί ότι γι' αυτά τα ειδικά βάρη των εκροών και των εισροών καμία ΛΑΜ δεν δείχνει μία καλύτερη επίδοση από την j_0 . Σε αυτή την περίπτωση η j_0 θα αξιολογηθεί ως αποτελεσματική. Η ίδια διαδικασία θα τρέξει για κάθε ΛΑΜς που συνιστά το δείγμα.

Αυτό το μοντέλο παρέχει μία σχετική μέτρηση της γενικής (global) επίδοσης κάθε ΛΑΜς που μελετάται. Επιπλέον, ο Banker (1984) και ο Banker με τους συνεργάτες του (1989) εισήγαγαν την έννοια της «πιο παραγωγικής κλίμακας μεγέθους» (ΠΠΚΜ) για να ορίσουν το μέγεθος κλίμακας που μεγιστοποιεί την αποδοτικότητα των ΛΑΜς που μελετώνται. Την ίδια ώρα αυτοί απέδειξαν ότι είναι δυνατό να διαχωριστεί η προαναφερθείσα γενική αποδοτικότητα σε δύο συνιστώσες, συγκεκριμένα, στην *τεχνική* αποδοτικότητα, η οποία καθορίζεται από την κατάσταση της ΛΑΜ στα όρια της αποδοτικής παραγωγής, και στην *αποδοτικότητα της κλίμακας*, η οποία δίνεται από το μέγεθος της ΛΑΜ, γεγονός που την καθιστά

λειτουργική σε μία ζώνη σταθερών ή μεταβλητών αποδόσεων. Έτσι, εάν το μέγεθος της ΛΑΜ επηρεάζει τα αποτελέσματα της δραστηριότητας, αυτή η παραγοντοποίηση της επίδοσης μας καθιστά ικανούς να αναλύουμε, από τη μια πλευρά, εάν η διαχείριση των πηγών είναι αποδοτική και, από την άλλη, εάν οι ΛΑΜς λειτουργούν στην πλέον παραγωγική κλίμακα μεγέθους τους.

6.6.1 Βασικά προτερήματα της χρήση του ΠΑΔ

Στη συνέχεια παρουσιάζονται κάποια από τα βασικά προτερήματα της χρήσης του ΠΑΔ στην αξιολόγηση της αποδοτικότητας των πανεπιστημιακών τμημάτων:

- Μπορεί εύκολα να ενσωματώσει πολλαπλές εισροές και εκροές χωρίς να απαιτείται κάποια κρίση και με μόνη απαίτηση τις πληροφορίες για την παραγωγή (εκροές) και τις ποσότητες των εισροών (όχι τις τιμές). Αυτό το καθιστά ιδιαίτερα κατάλληλο για την ανάλυση της αποδοτικότητας των ιδρυμάτων που παρέχουν εκπαιδευτικού τύπου υπηρεσίες όπου είναι δύσκολη ή έως αδύνατη η εκχώρηση τιμών στα περισσότερα από τα αποτελέσματα (εκροές) (Johnes & Johnes, 1993).
- Αυτό καθορίζει τις πιθανές πηγές της μη αποδοτικότητας καθώς και τα επίπεδα αποδοτικότητας. Έτσι, παρέχει ένα μέσο «αποικοδόμησης» της οικονομικής μη αποδοτικότητας σε τεχνική και μη αποδοτική κατανομή των πόρων (Tomkins & Green, 1988).
- Αυτό καθορίζει τους στόχους για τις μη αποδοτικές μονάδες και υποδεικνύει τι βελτιώσεις μπορούν να γίνουν για να βελτιωθεί η αποδοτικότητα. Αυτό καθιστά το ΠΑΔ ένα δυνητικά χρήσιμο εργαλείο για τη συγκριτική αξιολόγηση και την αλλαγή της πορείας υλοποίησης των προγραμμάτων (McMillan & Datta, 1998).
- Αυτό επιτρέπει στους διαχειριστές του χώρου της εκπαίδευσης, ή στους υποψήφιους ερευνητές, να διαλέξουν εισροές και εκροές για να παρουσιάσουν μία ειδική προσέγγιση (Ankiran, 2001).

6.6.2 Επικρίσεις για το μοντέλο ΠΑΔ

Από την άλλη πλευρά, κάποιες επικρίσεις για το μοντέλο έχουν ως εξής:

- Είναι δυνατό να περιλαμβάνει μεγάλο αριθμό οργανισμών σε ένα αποδοτικό σύνολο, μειώνοντας έτσι την διακριτική ισχύ του μοντέλου (Emrouznejad, 1995–2001). Μία μονάδα η οποία έχει τον υψηλότερο λόγο σε μία από τις εκροές προς μία από τις εισροές, θα είναι αποτελεσματική ή θα έχει μία

απόδοση πολύ κοντά στη μονάδα εάν τοποθετείται στο λόγο τόσο πολύ βάρος όσο είναι δυνατό, και ελάχιστο βάρος σε άλλες εισροές και εκροές. Η χρησιμότητα των αποτελεσμάτων είναι μεγαλύτερη για ΛΑΜς που αξιολογήθηκαν ως αναποτελεσματικές, δεδομένου ότι σε αυτές τις περιπτώσεις το μοντέλο παρέχει πληροφορίες σχετικά με τους τομείς στους οποίους αυτές πρέπει να βελτιώσουν την επίδοση.

- Μερικές φορές οι ΛΑΜς πετυχαίνουν 100% απόδοση απλά επειδή καμία άλλη δεν είναι ανταγωνιστική σε εκείνη τη θέση (Johnes & Johnes, 1993). Μία μονάδα μπορεί να εμφανίζεται αποτελεσματική μόνο λόγω των δικών της τύπων εισροών και εκροών και όχι λόγω κάποιας εγγενούς αποδοτικότητας. Έτσι, ένα τμήμα εξέχον σε ένα ιδιαίτερο τύπο εκροών ή χρησιμοποιώντας λιγότερο μία σημαντική πηγή μπορεί να πετύχει ένα υψηλό βαθμό αποδοτικότητας με το να είναι απλώς διαφορετικό (στο δικό του μίγμα εισροών και εκροών) από τις άλλες μονάδες.
- Το ΠΑΔ βοηθά στην τοποθέτηση στόχων αλλά δεν καθοδηγεί τον χρήστη με ποιο τρόπο να τους φτάσει (Ankiran, 2001).
- Η αποδοτικότητα που μετριέται με το ΠΑΔ είναι σχετική. Υπό αυτή την έννοια, τα τμήματα μπορεί να πετύχουν ένα διαφορετικό επίπεδο αποδοτικότητας εξαρτώμενο από το δείγμα με το οποίο αυτά συγκρίνονται, και είναι δυνατό να βελτιωθεί η επίδοση ακόμα και των αποδοτικών μονάδων. Ωστόσο, το μοντέλο δεν παρέχει κανένα κλειδί για τον τρόπο με τον οποίο μπορεί να βελτιωθεί η επίδοση των εμπειρικά αποδοτικών μονάδων (Sowlati & Paradi, 2004).

6.7 Ένα μοντέλο έρευνας μικτών μεθόδων για την αξιολόγηση του περιβάλλοντος μάθησης με βάση την αλλαγή των πεποιθήσεων αυτό-αποτελεσματικότητας των φοιτητών

Όπως ήδη επισημάνθηκε, πολλά από τα εργαλεία με τα οποία οι φοιτητές αξιολογούν τη διδασκαλία αποτυγχάνουν να παρέχουν μία ολοκληρωμένη αξιολόγηση της θεωρητικά ορθής, πολλαπλών διαστάσεων ποιότητας της διδασκαλίας οπότε υπονομεύουν τη χρησιμότητά τους, ειδικότερα τη διαγνωστική ανατροφοδότηση (διαμορφωτική αξιολόγηση) η οποία συμβάλει στη βελτίωση της αποτελεσματικότητας της διδασκαλίας. Οι έρευνες με «τοπικού» τύπου εργαλεία αξιολόγησης της διδασκαλίας τα οποία κατασκευάζουν διδάσκοντες ή επιτροπές σπάνια αξιολογούνται με βάση τις αυστηρές ψυχομετρικές εκτιμήσεις (θέματα εγκυρότητας με βάση ένα θεωρητικό πλαίσιο και αξιοπιστίας) και ως εκ τούτου δεν αναθεωρούνται αναλόγως. Αυτή η διακύμανση στην ποιότητα περιπλέκει επίσης την ικανότητα ερμηνείας και τη δυνατότητα γενίκευσης των αποτελεσμάτων μίας έρευνας που διεξάγεται σε αυτές τις συνθήκες (Marsh & Rosche, 1997, σελ. 1188).

Αναλυτικότερα επισημαίνεται ότι επειδή δεν υπάρχει ένας επαρκής ενιαίος δείκτης για την αποτελεσματική διδασκαλία η εγκυρότητα των εργαλείων που χρησιμοποιούνται στην αξιολόγησή της από τους φοιτητές (και από άλλες πηγές) πρέπει μεταξύ άλλων να αποδεικνύεται διαμέσου μίας αξιολόγησης με την οποία να επικυρώνεται η δομή τους (construct-validation approach). Συγκεκριμένα, είναι σημαντικό τα εργαλεία για την αξιολόγηση της διδασκαλίας από τους φοιτητές να σχετίζονται με μία ευρεία ποικιλία κριτηρίων της αποτελεσματικής διδασκαλίας. Για παράδειγμα, η έρευνα να αφορά σημαντικά αποτελέσματα μάθησης όπως είναι τα κίνητρα, την έννοια-εαυτού, το συναίσθημα, τη μετά-γνώση, τις στρατηγικές μελέτης, την επιλογή μαθήματος-εργασίας, τις επαγγελματικές φιλοδοξίες, και ούτω καθεξής. Ένα άλλο σημαντικό στοιχείο της επικύρωσης της δομής των εργαλείων που χρησιμοποιούνται στη διαδικασία της αξιολόγησης της διδασκαλίας από τους φοιτητές είναι η ανάγκη τα εργαλεία να αφορούν στις πραγματικές διαδικασίες που συμβαίνουν στην τάξη (Marsh & Rosche, 1997, σελ. 1190).

Ένα από τα θέματα που απασχολούν τους υπεύθυνους στην εκπαίδευση των φοιτητών οι οποίοι θα διδάξουν τα μαθήματα των ΦΕ στο δημοτικό και στο νηπιαγωγείο, μελλοντικά, είναι η ανάγκη να εξερευνηθούν οι παράγοντες του

πλαίσιου μάθησης που επηρεάζουν τις πεποιθήσεις τους σχετικά με τα μαθήματα των ΦΕ και τη διδασκαλία τους. Μία γόνιμη προσέγγιση θα μπορούσε να ακολουθήσει το επιχείρημα του Bandura (1997) ότι η εκπλήρωση των επιδόσεων (performance) αποτελεί την καλύτερη πρόβλεψη όσον αφορά στην αυτό-αποτελεσματικότητα, το οποίο υπαινίσσεται ότι οι φοιτητές που αποκτούν την εμπειρία της επιτυχούς μάθησης θα έχουν θετικές πεποιθήσεις αυτό-αποτελεσματικότητας. Σύμφωνα με την επιστημολογία του εποικοδομητισμού η επιτυχής μάθηση λαμβάνει χώρα σε ένα *κοινωνικό* πλαίσιο το οποίο εγείρει *συναισθήματα* (emotional) και στο οποίο η γνώση δομείται *συνεργατικά* από αυτούς που μαθαίνουν (Printrich, Marx & Boyle, 1993). Ως εκ τούτου τα ερωτήματα που απασχολούν τους υπεύθυνους του χώρου έχουν ως επίκεντρο το βαθμό που τέτοια πλαίσια παρέχονται στους φοιτητές. Συγκεκριμένα, αφενός στην έκταση που οι φοιτητές αποκτούν την εμπειρία τέτοιων πλαισίων μάθησης και αφετέρου στη βελτίωση των επιδόσεών τους σε αυτά, για παράδειγμα στη μάθηση και στην προοπτική της εκ μέρους τους διδασκαλίας με τις στρατηγικές που διδάχτηκαν.

Οι Watters και Ginns (1995) μελέτησαν τους παράγοντες που συνέβαλαν στο άγχος απέναντι στις φυσικές επιστήμες (Φ.Ε) και στην αυτό-αποτελεσματικότητα στη διδασκαλία των φυσικών επιστημών ανάμεσα στους υποψήφιους δασκάλους και νηπιαγωγούς σε ένα πανεπιστήμιο της Αυστραλίας. Οι μελετητές με βάση τη βιβλιογραφία ισχυρίζονται ότι ο τρόπος με τον οποίο ένα άτομο αποδίδει ή συμπεριφέρεται σε μία δοσμένη περίπτωση εξαρτάται από τις στάσεις του, οι οποίες συνιστούν εκδηλώσεις τόσο των γνωστικών όσο και των συναισθηματικών ιδιοτήτων του ατόμου (Bandura, 1977; Ajzen, 1985; Prawat, 1985; Shrigley, Koballa, & Simpson, 1988; Hewson & Hewson, 1989). Ακόμα επισημαίνουν ότι η έκταση στην οποία οι δάσκαλοι και οι νηπιαγωγοί θα διδάξουν τις Φ.Ε επηρεάζεται από τη γνώση (cognition) τους στις Φ.Ε και σε θέματα της διδασκαλίας των Φ.Ε καθώς και από τα αισθήματα ή τις στάσεις τους απέναντι σε αυτή τη γνώση (Morrissey, 1981). Αυτές οι στάσεις μπορεί να αναπτύσσονται κατά τη διάρκεια της σχολικής τους ζωής αλλά μπορεί επίσης να επηρεάζονται από τις εμπειρίες τους κατά την πανεπιστημιακή τους εκπαίδευση και προετοιμασία.

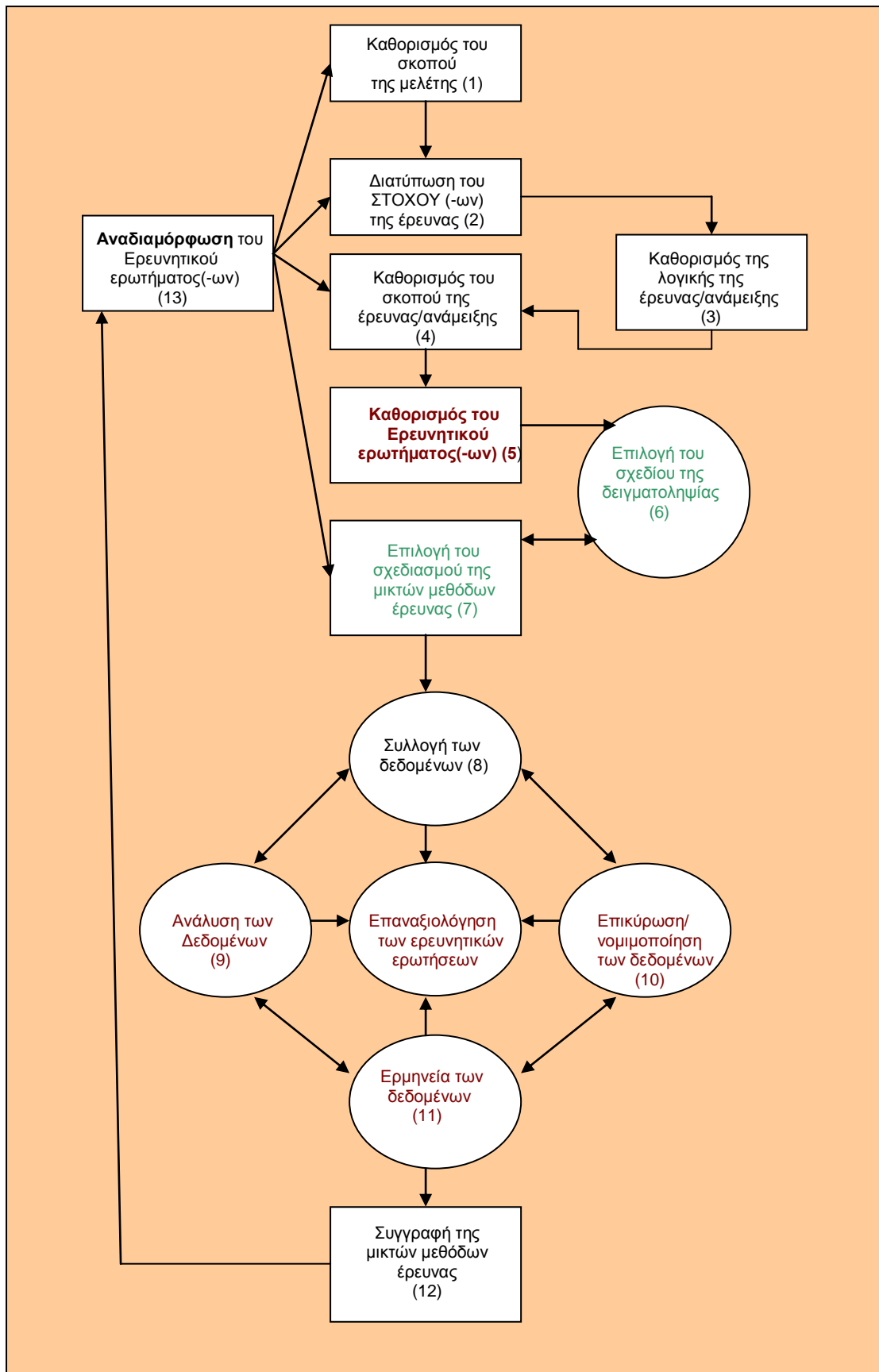
Ο σχεδιασμός της μελέτης των Watters και Ginns (1995) περιελάμβανε τη χρήση ποσοτικών και ποιοτικών προσεγγίσεων κατάλληλα διαμορφωμένων. Τα ποσοτικά στοιχεία προέκυψαν από ένα προ-και-μετά-έλεγχο σχεδιασμό, με τη χρήση έγκυρων

και σταθμισμένων ερωτηματολογίων STEBI-B (Bleicher, 2004), CLES (Johnson & McClure, 2002) ενώ πλούσιες περιγραφές επιλεγμένων συμμετεχόντων περιελήφθησαν διαμέσου συνεντεύξεων, σημειώσεων πεδίου και παρατήρησης.

Το θεωρητικό μοντέλο του Germann (1988) είναι σε θέση να δικαιολογήσει την ανάπτυξη των αδύνατων στάσεων των φοιτητών απέναντι στις φυσικές επιστήμες. Η μοιρολατρία των φοιτητών (fatalism), οι αντιλήψεις τους για την αξία των Φ.Ε, η ποιότητα του καθηγητή, το κοινωνικό περιβάλλον της τάξης και η οργάνωση εμφανίζονται να είναι σημαντικοί παράγοντες οι οποίοι συνιστούν αυτό το μοντέλο. Το μοντέλο, δίνει έμφαση ειδικότερα, στη συμβολή του καθηγητή στην κοινωνική αλληλεπίδραση και στο περιβάλλον της τάξης και έτσι το σημείο εστίασης γίνεται η αλληλεπίδραση καθηγητή-φοιτητή-προγράμματος σπουδών. Η συνεισφορά του καθηγητή στην αλληλεπίδραση καθοδηγείται από τη δική του/ ή της θέαση του κόσμου, τις στάσεις, τις ανάγκες και τις προτεραιότητες (Clark & Peterson, 1986; Shulman, 1987). Προσπάθειες κατανόησης της ικανότητας του καθηγητή να αποδίδει σε τέτοιες περίπλοκες αλληλεπιδράσεις, ή της αυτό-αποτελεσματικότητας, έχουν υπάρξει το αντικείμενο έρευνας η οποία προέρχεται από μία κοινωνικό-γνωστική προσέγγιση (Watters & Ginns, 1994).

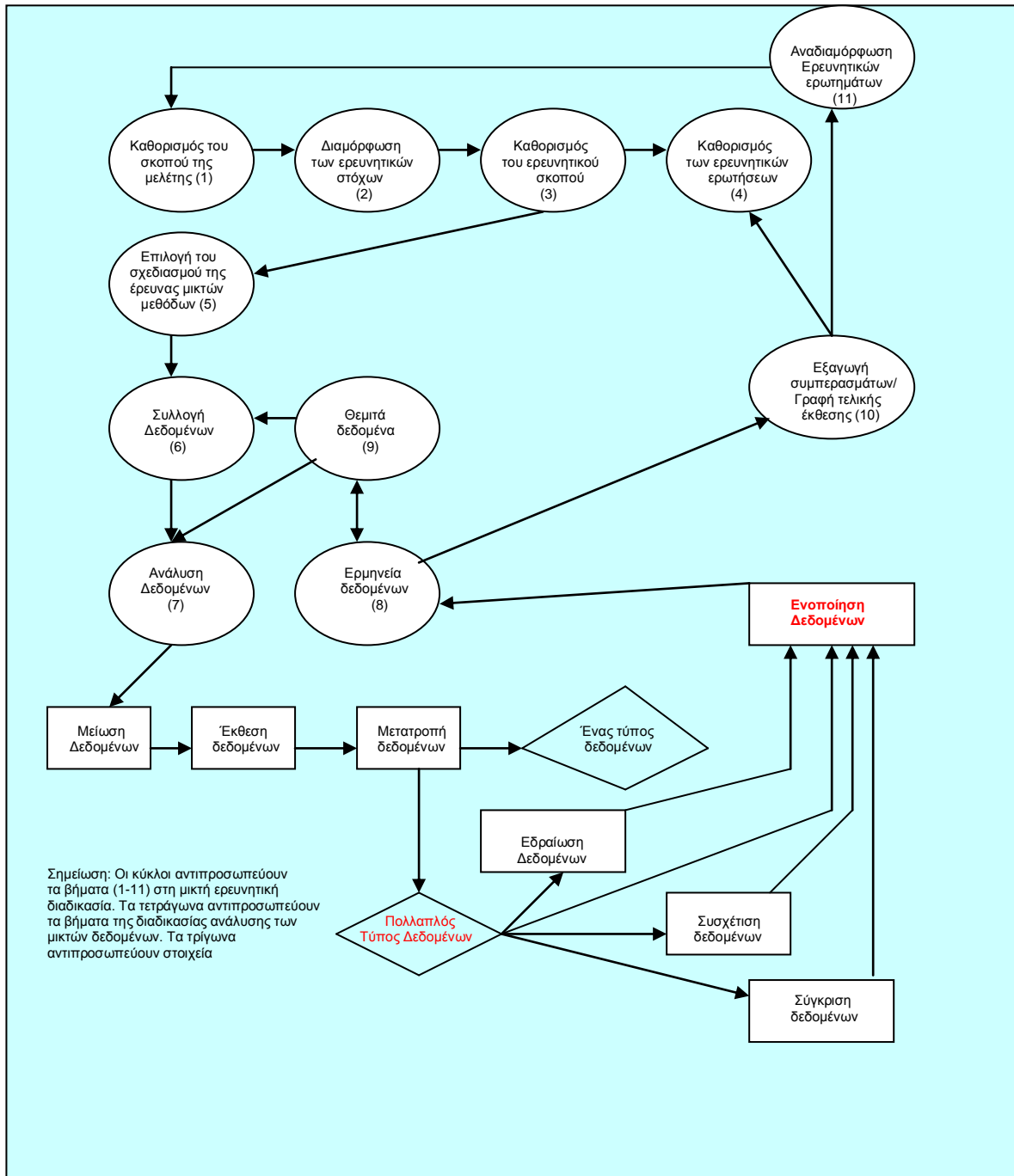
Στ' αχνάρια της μελέτης των Watters και Ginns διεξήχθη μία Έρευνα Μικτών Μεθόδων με σκοπό την αξιολόγηση του περιβάλλοντος μάθησης στο μάθημα επιλογής Α' εξαμήνου «Οι έννοιες των ΦΕ» στο ΤΕΑΠΗ / ΕΚΠΑ, το 2010-2011. Η έρευνα διεξήχθη στα πλαίσια της εκπόνησης της μετα-διδακτορικής μελέτης της συγγραφέως της παρούσας Έκθεσης. Το κεντρικό ερώτημα της έρευνας μικτών μεθόδων υπήρξε: *«Ποιες είναι οι συνέπειες του μαθήματος 'Οι έννοιες των ΦΕ' στις πεποιθήσεις των φοιτητών σχετικά με την αποτελεσματικότητά τους στη διδασκαλία των ΦΕ, μελλοντικά στο νηπιαγωγείο;»* Το διαδραστικό πλαίσιο το οποίο αξιοποίησαν οι ερευνητές υπήρξε εκείνο που συγκρότησαν οι Collins, Onwuegbuzie, και Sutton (2006). Στη συνέχεια με τη βοήθεια των Διαγραμμάτων 1, 2 και 3 παρουσιάζονται οι διαδικασίες σχεδιασμού της έρευνας, ανάλυσης δεδομένων και της σύνδεσης των τύπων των ερευνητικών ερωτημάτων με τα παραπάνω στάδια. Τα ερωτηματολόγια που χρησιμοποιήθηκαν στην ποσοτική φάση βρίσκονται στο Παράρτημα.

Διάγραμμα 1. Βήματα της Έρευνας Μικτών Μεθόδων



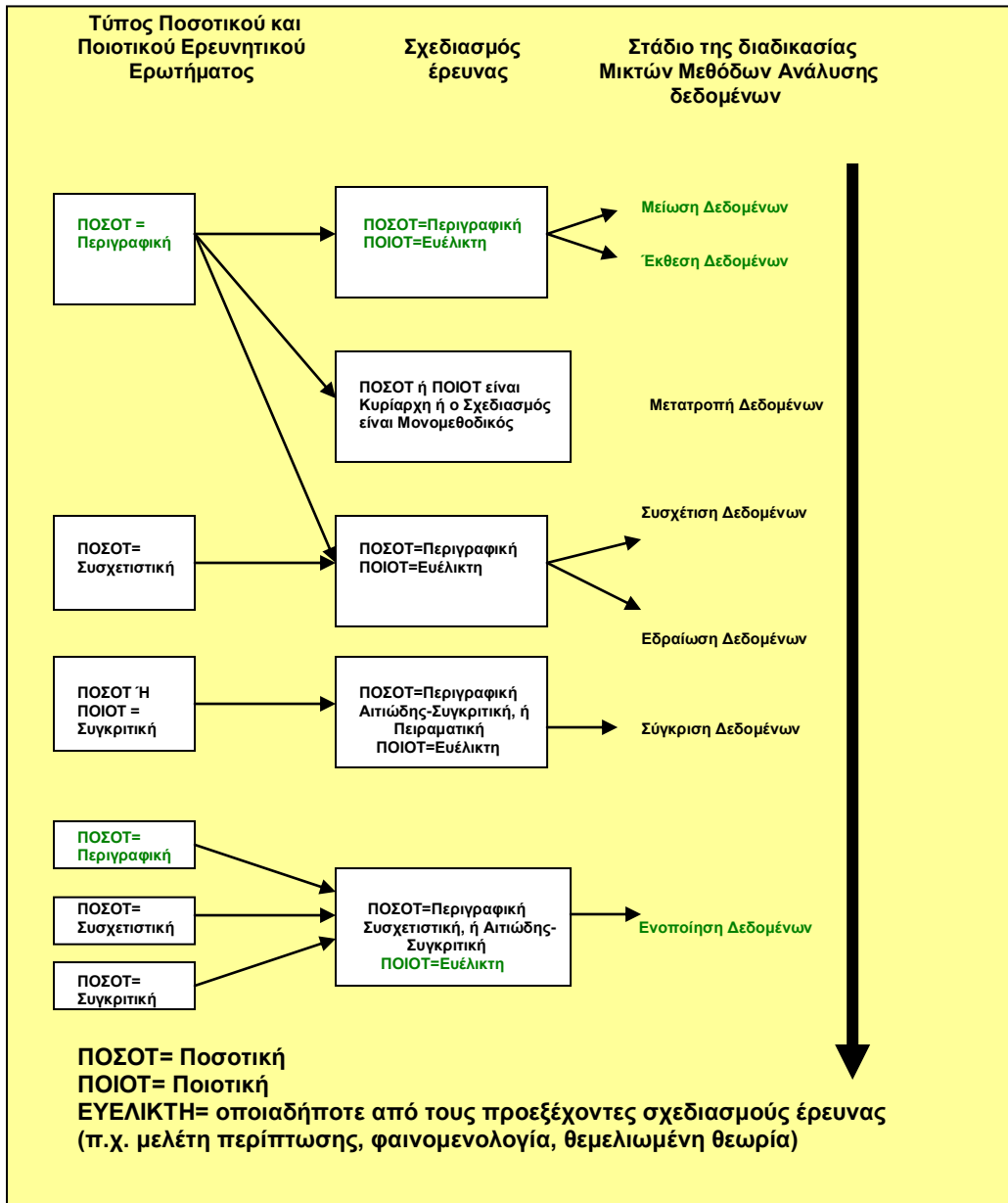
Πηγή: Onwuegbuzie, J. A., & Leech, L. N. (2006). Linking Research Questions to Mixed Methods Data Analysis Procedures. *The Qualitative Report*, 11 (3), p: 474-498.

Διάγραμμα 2: Βήματα της ΜΙΚΤΩΝ ΜΕΘΟΔΩΝ ΑΝΑΛΥΣΗΣ δεδομένων



Πηγή: Onwuegbuzie, J. A., & Leech, L. N. (2006). Linking Research Questions to Mixed Methods Data Analysis Procedures. *The Qualitative Report*, 11 (3), p: 474-498.

Διάγραμμα 3: Συνδυασμοί σχεδιασμού έρευνας, τύπου ερευνητικών ερωτήματα, και Μικτών Μεθόδων Αναλύσεων δεδομένων



Πηγή: Onwuegbuzie, J. A., & Leech, L. N. (2006). Linking Research Questions to Mixed Methods Data Analysis Procedures. *The Qualitative Report*, 11 (3), p: 474-498.

7. Τελικές Επισημάνσεις

Η κυρίαρχη επισήμανση των σχετικών μελετών αναφορικά με την αξιολόγηση του εκπαιδευτικού έργου στην Ανώτατη Εκπαίδευση αφορά στην αναγνώριση της βασικής προϋπόθεσης για την ομαλή εξέλιξη της διαδικασίας η οποία είναι η ορθή αντιμετώπιση θεμάτων που άπτονται της εγκυρότητάς της. Ειδικότερα:

- Οι βαθμολογίες από τους φοιτητές πρέπει να αντιμετωπίζονται ως ένα μέρος μιας συνολικής εικόνας της αποτελεσματικής διδασκαλίας ενός διδάσκοντα.
- Δεν πρέπει να υφίσταται υπερβολή όσον αφορά στην αποδεικτική αξία των βαθμολογιών από τους φοιτητές. Οι φόρμες βαθμολόγησης δεν είναι ικανοποιητικά έγκυρες και ως εκ τούτου κοντινές διακρίσεις μεταξύ των αποτελεσμάτων δεν είναι αρκετά σημαντικές για να υποστηρίξουν λεπτές διακρίσεις στις αριθμητικές βαθμολογίες.
- Αυτός που χρησιμοποιεί τις βαθμολογίες, συμπεριλαμβανομένου και των μελών της διοίκησης των πανεπιστημίων, πρέπει να είναι εκπαιδευμένος πάνω στη σωστή εφαρμογή τους, με σκοπό να αποφευχθεί η κατάχρηση των δεδομένων τους. Ανάμεσα σε άλλα πράγματα, οι χρήστες πρέπει να θυμούνται ότι τα ευρήματα της έρευνας γενικεύονται από ένα δείγμα σε ένα πληθυσμό και δεν βεβαιώνουν ότι κάθε περίπτωση θα εξηγηθεί. Ακόμα, πρέπει να γίνει αντιληπτό ότι είναι άδικο να συγκρίνονται βαθμολογίες ανάμεσα σε μαθήματα.
- Η απλή παροχή των αξιολογήσεων των δεδομένων στους καθηγητές δεν βελτιώνει σημαντικά τη διδασκαλία. Είναι επιτακτική ανάγκη να τους παρέχονται πόροι για τη βελτίωσή τους. Οι ερευνητές προειδοποιούν ότι η χρήση των βαθμολογιών χωρίς την παροχή τέτοιων πόρων για βελτίωση λειτουργεί ως «τιμωρία».

Για τα ιδρύματα- χρήστες των αξιολογήσεων

Ο βαθμός στον οποίο οι αξιολογήσεις συνιστούν ένα αποτελεσματικό εργαλείο μέτρησης της ποιότητας του εκπαιδευτικού έργου στην ανώτατη εκπαίδευση καθορίζεται πάνω απ' όλα από τις πολιτικές, τις διαδικασίες και τις πρακτικές που εφαρμόζονται στα ιδρύματα. Τα ιδρύματα πρέπει να υιοθετούν δικές τους απόψεις και χρήσεις για τις αξιολογήσεις. Αυτό σημαίνει ότι:

- Οι αξιολογήσεις πρέπει να σχεδιάζονται προσεκτικά έτσι ώστε να ταιριάζουν με το θεσμικό πλαίσιο των ιδρυμάτων και να συνοδεύονται από ουσιαστική θεσμική υποστήριξη.
- Οι πολιτικές σχεδιασμού πρέπει να είναι συνεκτικές και ισότιμες.
- Η εκπαίδευση των χρηστών – φοιτητών, μελών ΔΕΠ και διοίκησης- πάνω στην αξιολόγηση επιβάλλεται για να διαλύει τους μύθους, τις παρανοήσεις και να βελτιώνει τις δεξιότητες και τη διαφάνεια.
- Οι αξιολογήσεις πρέπει να συνοδεύονται από συνεχή διάλογο και μηχανισμούς υποστήριξης, περιλαμβάνοντας τη διαβούλευση, για να επιβεβαιώνεται ότι αυτοί συμβάλλουν στην υποστήριξη και στη βελτίωση της διδασκαλίας στο ίδρυμα.

Συμπληρωματικά, αυτοί που χαράσσουν πολιτική πρέπει να έχουν υπόψη τους τα προβλήματα που χαρακτηρίζουν τις αξιολογήσεις του εκπαιδευτικού έργου από τους φοιτητές. Αυτά συνοψίζονται (Moore, 2009) στα εξής: Οι αξιολογήσεις 1) είναι εύκολες στο χειρισμό, 2) έχουν αμφισβητούμενο κύρος ως μέτρο της μάθησης των φοιτητών, 3) μπορεί να γίνει "κατάχρηση" από τα μέλη της διοίκησης, 4) μπορούν να δημιουργήσουν "στρεβλά" κίνητρα στα μέλη ΔΕΠ, 5) οι αναρμόδιοι φοιτητές αποκτούν υπερβολικά πάρα πολλή δύναμη πάνω στα μέλη ΔΕΠ, και 6) δημιουργούν άγχος στο μέλος ΔΕΠ που αξιολογείται.

Για τη συνέχιση της έρευνας

Οι σύγχρονες τάσεις της έρευνας εστιάζουν τα ενδιαφέροντα τους προς τα γενικότερα θέματα σχεδιασμού της έρευνας, εφαρμογής και ερμηνείας των αποτελεσμάτων. Ως εκ τούτου, η έρευνα εξελίσσεται σε μία ολιστικού τύπου και αφορά στα εξής θέματα:

- Στη βελτίωση της εκπαίδευσης και της πληροφορίας για τους χρήστες της αξιολόγησης και των ελεγχόμενων αποτελεσμάτων.
- Στην ανάπτυξη και στον έλεγχο αποτελεσματικών μέσων για τη δημοσιοποίηση των αποτελεσμάτων και των εργαλείων για την ερμηνεία (σε σχέση με τις ανάγκες των χρηστών).
- Στη συνειδητή εμπλοκή των μελών ΔΕΠ και των φοιτητών στη διαδικασία αξιολόγησης.
- Στην τακτική επανεξέταση των μέσων αξιολόγησης με βάση τις θεσμικές ανάγκες και τους στόχους και σε σχέση με τα τρέχοντα ερευνητικά ευρήματα.

Τέλος, κατά την κρίση της συγγραφέα εκτιμάται ως απαραίτητο να αναφερθεί η διαπίστωση του Clayson (2011) ότι, μετά από σχεδόν 30 χρόνια έρευνας, δεν υφίσταται κάποια πειστική δημοσίευση τα πορίσματα της οποίας να δείχνουν ότι η υιοθέτηση του συστήματος αξιολόγησης συνέβαλε ώστε να βελτιωθεί η διδασκαλία, οι φοιτητές να μαθαίνουν περισσότερα, οι εργοδότες να βρίσκουν καλύτερα προετοιμασμένους εργαζόμενους, η κοινωνία να έχει βελτιωθεί, ή ότι οι καθηγητές ή οι φοιτητές να είναι πιο ευτυχισμένοι λόγω της αξιολόγησης της διδασκαλίας από τους φοιτητές. Αυτό μπορεί να θεωρηθεί ως άδικος δείκτης-πρότυπο όσον αφορά στην εφαρμογή της διαδικασίας, αλλά όταν οι αξιολογήσεις καταναλώνουν σημαντικούς πόρους, και αποτελούν ένα σημαντικό (και σε ορισμένες περιπτώσεις, το μόνο) μέτρο των διδακτικών ικανοτήτων των μελών ΔΕΠ, τότε δικαιολογείται η απαίτηση για κάτι περισσότερο από τη θεωρία και την καλή πίστη.

Βιβλιογραφία

A

Abbot, M. & Dougouliagos, C., (2003). The efficiency of Australian universities: a data envelopment analysis. *Economics of Education Review*, 22, pp. 89–97.

Abrami, P. C. (1989a). How should we use student ratings to evaluate teaching?, *Research in Higher Education*, 30, pp. 221-227.

Abrami, P. C. (1989b). SEEQing the truth about student evaluations of instruction, *Educational Researcher*, 18, 43-45.

Abrami, P.C. (2001). Improving judgments about teaching effectiveness using teacher rating forms. In M. Theall, P.C. Abrami, and L.A. Mets (Eds.). *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research* 109, 59-87.

Abrami, P. C., d'Appollonia, S., & Cohen, P. A. (1990). Validity of student ratings of instruction: What we know and what we do not. *Journal of Educational Psychology*, 82, 219-231.

Abrami, P. C, Dickens, W. J., Perry, R. P. & Leventhal, L. (1980). Do teacher standards for assigning grades affect student evaluations of instruction?, *Journal of Educational Psychology*, 72, pp. 107-118.

Abrami, P. C, Leventhal, L., Perry, R. P. & Breen, L. J., L. J. (1976) Course evaluation: How?, *Journal of Educational Psychology*, 68, pp. 300-304

Abrami, P. C, Leventhal, L. & Perry, R. P. (1982a). Educational seduction, *Review of Educational Research*, 52, pp. 446-464.

Adams, J. B. (2005). What makes the grade? Faculty and student perceptions. *Teaching of Psychology*, 32 (1), 21-24.

Addison, W.E., Best, J., & Warrington, J. D. (2006). Students' perceptions of course difficulty and their ratings of the instructor. *College Student Journal*, 40(2), 409-416.

Aleamoni, L. M. (1981) Student ratings of instruction, in: J. MILLMAN (Ed.) *Handbook of Teacher Evaluation*, pp. 110-145 (Beverly Hills, Sage).

Aleamoni, L. M. (1982). Components of the instructional setting. *Instructional Evaluation*, 7, 11–16.

Aleamoni, L. M. (1987) Student rating myths versus research facts, *Journal of Personnel Evaluation in Education*, 1, pp. 111-119

Aleamoni, L.M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, 13(2), 153-166.

Aleamoni, L. M. & Hexner, P. Z. (1980) A review of the research on student evaluation and a report on the effect of different sets of instructions on student course and instructor evaluation, *Instructional Science*, 9, pp. 67-84.

Allen, M. (1995). *Research productivity and positive teaching evaluations: Examining the relationship using meta-analysis*. Paper presented at the Annual Meeting of the Western States Communication Association, Portland, Oregon.

Algozzine, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L., Mohanty, G., & Spooner, F. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52(4), 134-141.

Anderson, E. (Ed.). (1993). *Campus use of the teaching portfolio: Twenty-five profiles*. Washington, DC: American Association for Higher Education.

Appling, S. E., Naumann, P. L., & Berk, R. A. (2001). Using a faculty evaluation triad to achieve evidence-based teaching. *Nursing and Health Care Perspectives*, 22, 247–251.

Arreola, R. A. (1983). Establishing successful faculty evaluation and development programs. In A. Smith (Ed.). *Evaluating faculty and staff* [Special issue]. *New Directions for Community Colleges*, 41, 83-93.

Arreola, R.A. (1995). *Developing a comprehensive faculty evaluation system*. Bolton, MA: Anker.

Arreola, R. A. (2000). *Developing a comprehensive faculty evaluation system: A handbook for college faculty and administrators on designing and operating a comprehensive faculty evaluation system (2nd ed.)*. Bolton, MA: Anker.

Arreola, R. A. (2007). *Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems (3rd ed.)*. Bolton, MA: Anker.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Springer Series in Social Psychology. Action Control: From Cognition to Behavior* (pp. 11-39). Berlin: Springer-Verlag.

Ajzen, I. (1988). *Attitudes, Personality, and Behavior*. Milton-Keynes, England: Open University Press.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice. Hall.

Asher, J. J., & Sciarrino, J. A. (1974). Realistic work sample tests: A review. *Personnel Psychology*, 27, 519–533

Ashworth, K.W. (1994) Performance based funding in higher education: the Texas case study, *Change*, November/December, pp. 8-15.

Attiyeh, R., & Lumsden, K. G. (1972). Some modern myths in teaching economics: The U. K. experience. *American Economic Review*, 62, 429-433.

Aubrecht, J. D. (1984). Better faculty evaluation systems, in: P. Seldin (Ed.) *Changing Practices in Faculty Evaluation* (San Francisco, Jossey-Bass).

Aultman, L. P. (2006). An Unexpected Benefit of Formative Student Evaluations. *College Teaching*, 54(3), 251.

Avkiran, N., (2001). Investigating technical and scale efficiencies of Australian universities through data envelopment analysis', *Socio-economic Planning Sciences*, 35, pp. 57-80.

B

Babin, L. A., Shaffer, T. R., & Morgan, T. A. (2002). Teaching portfolios: Uses and development. *Journal of Marketing Education*, 24(1), 35.

Bacon, D. R., & Novotny, J. (2002). Exploring achievement striving as a moderator of the grading leniency effect. *Journal of Marketing Education*, 24, 4-14.

Baird, J. S. (1987). Perceived learning in relation to student evaluation of university instruction. *Journal of Educational Psychology*, 79, 9091.

Baldwin, T. & Blattner, N. (2003). Guarding against potential bias in student evaluations: What every faculty member needs to know. *College Teaching*, 51(1), 27-32.

Bandura, A. (1977). Self efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.

Banker, R., 1984, 'Estimating most productive scale size using data envelopment analysis', *European Journal of Operational Research*, 17, pp. 35-44.

Banker, R., Charnes, A., Cooper, W., Swarts, J. & Thomas, D., 1989, 'An introduction to data envelopment analysis with some of its models and their uses', *Research in Governmental and Non Profit Accounting*, 5, pp. 125-63.

Barnett, R. (1992). *Improving higher education: Total quality care*. London: The Society for Research into Higher Education and The Open University.

Basow, S. A. (1994). *Student ratings of professors are not gender blind*. Paper presented at the meeting of the Society of Teaching and Learning in Higher Education, Vancouver, British Columbia.

Basow, S. A. & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently?, *Journal of Educational Psychology*, 79, pp. 308-314.

Bennett, S. K. (1982) Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation, *Journal of Educational Psychology*, 74, pp. 170-179.

Beran, T., Violato, C., Kline, D., & Frideres, J. (2005). The utility of student ratings of instruction for students, faculty, and administrators: A “consequential validity” study. *Canadian Journal of Higher Education*, 35(2), 49-70.

Beran, T., Violato, C., & Kline, D. (2007). What’s the ‘use’ of student ratings of instruction for administrators? One university’s experience. *Canadian Journal of Higher Education*, 17(1), 27-43.

Berk, R. A. (1999). Assessment for measuring professional performance. In D. P. Ely., L. E. Odenthal, & T. J. Plomp (Eds.), *Educational science and technology: perspectives for the future* (pp. 29–48). Enschede, The Netherlands: Twente University Press.

Berk, R. A. (2002). Teaching portfolios used for high-stakes decisions: You have technical issues! In National Evaluation Systems, *How to find and support tomorrow’s teachers* (pp. 45–56). Amherst, MA: Author

Berk A. R. (2006). Survey of 12 Strategies to Measure Teaching Effectiveness. *International Journal of Teaching and Learning in Higher Education* 17, 1, 48-62.

Berk, A. R. (2006). “*Thirteen Strategies to Measure College Teaching*”, Sterling, Virginia: Stylus.

Berk A. R. (2009). Using the 360o multisource feedback model to evaluate teaching and professionalism, *Medical Teacher*, 31, p: 1073-1080.

Berk, R. A., Naumann, P. L., & Appling. S. E. (2004). Beyond student ratings: Peer observation of classroom and clinical teaching. *International Journal of Nursing Education Scholarship*, 1(1), 1–26.

Birnbaum, M. H. (2000). *A survey of faculty opinions concerning student evaluation of teaching*. Retrieved June 21, 2008, from <http://psych.fullerton.edu/mbirnbaum/faculty3.htm>

Bleicher, R. E. (2004). Revisiting the STEBI-B: Measuring Self-Efficacy in Preservice Elementary Teachers. *School Science and Mathematics Journal*, 104, 383–391.

Blunt, A. (1991). The effects of anonymity and manipulated grades on student ratings of instructors, *Community College Review*, 18, pp. 48-54

Boex, L. F. J. (2000). Attributes of effective economics instructors: An analysis of student evaluations. *Research in Economic Education*, 31, 211-227.

Bo-Linn, C., Gentry, J., Lowman, J., Pratt, R. W., & Zhu, R. (2004, November). *Learning from exemplary teachers*. Paper presented at the Lilly Conference on College Teaching, Miami University, Oxford, OH.

Border, Laura L. B. (2002). The socratic portfolio: A guide for future faculty. *PS: Political Science & Politics* 35 (4), 739-734 www.apsanet.org/section_223.cfm.

Border, Laura L. B. (2005). Teaching portfolios for graduate students: Process, content, product, and benefits. *Teaching Excellence*, 17(4).

Bologna Declaration, The (1999). *The European Higher Education Area*. Joint declaration of the European Ministers of Education, Bologna, 19 June, <http://www.unizar.es/ees/doc/Declaration%20Bolonia%201999.pdf>

Boyer, E. (1990). *Scholarship reconsidered: New priorities for the professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

Branderburg, D. C, Slinde, J. A. & Batista, E. E. (1977). Student ratings of instruction: Validity and normative interpretations, *Research in Higher Education*, 7, pp. 67-78.

Braskamp, L. A., Brandenburg, D. C. & Ory, J. C. (1984). *Evaluating Teaching Effectiveness* (Newbury Park, CA, Sage).

Braskamp, L. A., Caulley, D. N., & Costin, F. (1979). Student ratings and instructor self-ratings and their relationship to student achievement. *American Educational Research Journal*, 16, 295–306

Braskamp, L.A. & Ory, J.C. (1994). *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco: Jossey-Bass.

Braskamp, L. A., Ory, J. C., & Pieper, D. M. (1981). Student written comments: Dimensions of instructional quality. *Journal of Educational Psychology*, 73, 65–70.

Brinko, K. T. (1993). The practice of giving feedback to improve teaching: What is effective? *Journal of Higher Education*, 64(5), 54–68.

Brinko, K. T. (1997). The interactions of teaching improvement. In K. T. Brinko & R. J. Menges (Eds.), *Practically speaking: A sourcebook for instructional consultants in higher education* (pp. 3–8). Stillwater OK: New Forum. (Reprinted from *New Directions for Teaching and Learning Series*, No. 48, 1991, San Francisco: Jossey-Bass.)

Burke, J.C. & Serban, A.M. (1997) *Performance Funding and Budgeting for Public Higher Education: Current status and future prospects* (Albany, NY, Rockefeller Institute of Government).

C

Campbell, J.P. & Bozeman, W.C. (2008). The value of student ratings: Perceptions of students, teachers and administrators. *Community College Journal of Research and Practice*, 32(1), 13-24.

Carnegie Foundation for the Advancement of Teaching. (1994). *National survey on the reexamination of faculty roles and rewards*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

Carter, K. (1994). The performance budget revisited: A report on state budget reform. Legislative Finance Paper No. 91. Denver, CO: National Conference of State Legislatures.

Carusetta, E. (2001). Evaluating teaching through teaching awards. In C. Knapper & P. Cranton (Eds.), *Fresh approaches to the evaluation of teaching* (New Directions for Teaching and Learning, No. 88) (pp. 31–46). San Francisco: Jossey-Bass.

Casey, J. R., Gentile P., & Bigger, S. W. (1997). Teaching appraisal in higher education: An Australian perspective. *Higher Education*, 34(4), 459–482.

Cashin, W. E. (1988). *Student Ratings of Teaching: A Summary of the Research* (Manhattan, KS, Center for Faculty Evaluation and Development, Kansas State University).

Cashin, W. E. (1989). Defining and evaluating college teaching (IDEA Paper No. 21). Manhattan, KS: Center for Faculty Evaluation and Development, Kansas State University.

Cashin, W. E. (1990). Student ratings of teaching: Recommendations for use (IDEA Paper No. 22). Manhattan, KS: Center for Faculty Evaluation and Development, Kansas State University.

Cashin, W.E. (1992). Student ratings: The need for comparative data. *Instructional Evaluation and Faculty Development*, 12(2), 1-6.

Cashin, W.E. (1995). Student ratings of teaching: The research revisited (IDEA Paper No. 32). Manhattan, KS: Kansas State University Center for Faculty Evaluation and Development.

Cashin, W. E. & Clegg, V. L. (1987). *Are student ratings of different academic fields different?* Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

Cashin, W.E., & Downey, R.G. (1992). Using global student rating items for summative evaluation. *Journal of Educational Psychology*, 84(4), 563-572.

Cave, M., Hanney, S., Henkel, M., & Kogan M. (1997). *The use of performance indicators in higher education: The challenge of the quality movement*. London: Jessica Kingsley.

Cederblom, D., & Lounsbury, J. W. (1980). An investigation of user acceptance of peer evaluations. *Personnel Psychology*, 33, 567–580.

Centra, J. A. (1973). Self-ratings of college teachers: A comparison with student ratings. *Journal of Educational Measurement*, 10, 287–295.

Centra, J. A. (1976). The influence of different directions on student ratings of instruction, *Journal of Educational Measurement*, 13, pp. 277-282.

Centra, J. A. (1978). Using student assessments to improve performance and vitality, in: W. R. Kirschling (Ed.) *Evaluating Faculty Performance and Vitality*, pp. 31-49 (San Francisco, Jossey-Bass).

Centra, J.A. (1979). *Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions and improvement*. San Francisco: Jossey-Bass.

Centra, J. A. (1983). Research productivity and teaching effectiveness, *Research in Higher Education*, 18, pp. 379-389.

Centra, J.A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass.

Centra, J. A. (1999). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass.

Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education*, 44, 495-518

Centra, J., Froth, R.C., Gray, P.J. and Lambert, L.M. 1987. *A guide to evaluating teaching for promotion and tenure*. R.M. Drummond, Editor. Littleton, MA: Copley.

Centra, J. A. & Creech, F. R. (1976). *The relationship between student teachers and course characteristics and student ratings of teacher effectiveness*. Project Report 76-1 (Princeton, NJ, Educational Testing Service).

Cerbin, W., & Hutchings, P. (1993, June). *The teaching portfolio*. Paper presented at the Bush Summer Institute, Minneapolis, MN.

Chacko, T. I. (1983). Student ratings of instruction: A function of grading standards. *Educational Research Quarterly*, 8(2), 19-25.

Charnes, A., Cooper, W. & Rhodes E. (1978). Measuring the efficiency of efficiency of decision-making units, *European Journal of Operational Research*, 2, pp. 429-44.

Chism, N. (1999). *Peer review of teaching: A sourcebook*. Bolton, MA: Anker Publishing.

Chacko, T. I. (1983). Student ratings of instruction: A function of grading standards, *Educational Research Quarterly*, 8, pp. 19-25.

Chonko, L. B., Tanner, J. F., & Davis, R. (2002). What are they thinking? Students' expectations and self-assessments. *Journal of Education for Business*, 77, 271- 281.

Clark, C.M., & Peterson, P.L. (1986). Teachers' Thought Processes, in Wittrock, M.C. (Ed.), *Hand book of research on teaching*, NY: MacMillan Publishing Company, 255-296.

Clayson, D. E. (1994). Contrasting results of three methodological approaches on the interpretation of a student evaluation of instruction. In E. W. Chandler (Ed.), *Proceedings of the Midwest Marketing Association* (pp. 209-214). Chicago: Midwest Marketing Association

Clayson, D. E. (2004). A test of the reciprocity effect in the student evaluation of instructors in marketing classes. *Marketing Education Review*, 14(2), 11-21.

Clayson, D. E. (2005a). Performance overconfidence: Meta-cognitive effects or misplaced student experience. *Journal of Marketing Education*, 27, 122-129.

Clayson, D. E. (2005b). Within-class variability in student-teacher evaluations: Example and problems. *Decision Sciences Journal of Innovative Education*, 3(1), 109-124.

Clayson, D. E. (2007). Conceptual and statistical problems of using between-class data in educational research. *Journal of Marketing Education*, 29, 34-38.

Clayson, D. E. (2009). Student Evaluations of Teaching: Are They Related to What Students Learn?: A Meta-Analysis and Review of the Literature. *Journal of Marketing Education* 31(1), 16-30.

Clayson, D. E., & Haley, D. A. (1990). Student evaluations in marketing: What is actually being measured? *Journal of Marketing Education*, 12, 9-17.

Clayson, D. E., & Sheffet, M. J. (2006). Personality and the student evaluation of teaching. *Journal of Marketing Education*, 28, 149-160.

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multi-section validity studies. *Review of Educational Research*, 51, 281-309.

Cohen P. A., & McKeachie, W. J. (1980). The role of colleagues in the evaluation of teaching. *Improving College and University Teaching*, 28, 147-154.

Colbert, A., Levary, R. & Shaner, M. (2000). Determining the relative efficiency of MBA barograms using DEA. *European Journal of Operational Research*, 125, pp. 656-69.

Collins, K. M. T., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed methods research in special education and beyond. *Learning Disabilities: A Contemporary Journal*, 4, 67-100.

Commission of the European Communities (2003). *The role of the universities in the Europe of knowledge*. European Commission, Brussels, 5 February.

Cox, M. D. (1995). A department-based approach to developing teaching portfolios: Perspectives for faculty and development chairs. *Journal on Excellence in College Teaching*, 6(1), 117–143.

Cranton, P. (2001). Interpretive and critical evaluation. In C. Knapper & P. Cranton (Eds.), *Fresh approaches to the evaluation of teaching* (New Directions for Teaching and Learning, No. 88) (pp. 11–18). San Francisco: Jossey-Bass.

Cranton, P. A. & Smith, R. A. (1986). A new look at the effect of course characteristics on student ratings of instruction, *American Educational Research Journal*, 23, pp. 117-128.

Crosson, A.C., Boston, M., Levison, A., Matsumura, L.C., Matsumura, L.C., Resnick, L.B., et al. (2006). *Beyond summative evaluation: The instructional quality assessment as a professional development tool* (CSE Technical Report 691). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Cruse, D. B. (1987). Student evaluations of the university professor: Caveat professor. *Higher Education*, 16, 723-737.

D

d'Apollonia, S., & Abrami, P.C. (1997). Navigating student ratings of instruction. *American Psychologist*, 52(11), 1198-1208.

Davidovitch, N., & Soen, D. (2006). Class attendance and students' evaluation of their college instructors. *College Student Journal*, 40(3), 691-703.

Diamond, M. R. (2004). The usefulness of structured mid-term feedback as a catalyst for change in higher education classes. *Active Learning in Higher Education* 5(3), 217-231.

Doyle, K. O., & Crichton, L. I. (1978). Student, peer, and self-evaluation of college instruction. *Journal of Educational Psychology*, 70, 815–826.

DuCette, J., & Kenney, J. (1982) Do grading standards affect student evaluations of teaching? Some evidence on an old question, *Journal of Educational Psychology*, 74, pp. 308-314.

Dudley, S. & Shawver, D. L. (1991). The effect of homework on students' perceptions of teaching effectiveness, *Journal of Education for Business*, 67, pp. 21-25.

E

Eble, K. E. (1970). *The Recognition and Evaluation of Teaching* (Salt Lake City, Project to Improve College Teaching).

Edgerton, R., Hutchings, P., & Quinlan, K. (1991). *The teaching portfolio: Capturing the scholarship in teaching*. Washington, DC: American Association for Higher Education.

Eicher J. C. (1998). The costs and financing of higher education in Europe. *European Journal of Education*, 33, 31-33.

Eiszler, C.F. (2002). College students' evaluations of teaching and grade inflation. *Research in Higher Education*, 43(4), 483-501.

El-Khawas, E., & Massey, W. F. (1996). Britain's "performance-based" system. In W. F. Massey (Ed.), *Resource allocation in higher education* (pp. 223-240). Ann Arbor, MI: The University of Michigan Press.

Emery, C. R., Kramer, T. R., & Tian, R. G. (2003). Return to academic standards: A critique of students' evaluations of teaching effectiveness. *Quality Assurance in Education: An International Perspective*, 11(1), 37-47.

Emrouznejad, A. (1995-2001). Ali Emrouznejad's DEA homepage, Warwick Business School, <http://www.deazone.com/>

Ewell, P.T. & Jones, D. (1994). Pointing the way: indicators as policy tools in higher education, in: S. Ruppert (Ed.) *Charting Higher Education Accountability: A Sourcebook on State Level Performance Indicators* (Denver, Education Commission of the States).

F

Feldman, K. A. (1976a). Grades and college students' evaluations of their courses and teachers, *Research in Higher Education*, 4, pp. 69-111.

Feldman, K. A. (1976b). The superior college teacher from the students' view, *Research in Higher Education*, 5, pp. 243-288.

Feldman, K.A. (1976). Grades and college students' evaluations of their courses and teachers. *Research in Higher Education*, 4(1), 69-111.

Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers and courses: A review and analysis, *Research in Higher Education*, 6, pp. 223-274.

Feldman, K. A. (1978). Course characteristics and college students' ratings of their teachers: What we know and what we don't, *Research in Higher Education*, 9, pp. 199-242.

Feldman, K. A. (1979). The significance of circumstances for college students' ratings of their teachers and courses, *Research in Higher Education*, 10, pp. 149-172.

Feldman, K. A. (1983). Seniority and experience of college teachers as related to evaluations they receive, *Research in Higher Education*, 18, pp. 3-124.

Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: A review and synthesis, *Research in Higher Education*, 24, pp. 139-213.

Feldman, K. A. (1987). Research productivity and scholarly accomplishments: A review and exploration, *Research in Higher Education*, 26, pp. 227-298.

Feldman, K.A. (1989). Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral) observers. *Research in Higher Education*, 30(2), 137-194.

Feldman, K. A. (1992). College students' views of male and female college teachers: Part I—Evidence from the social laboratory and experiments, *Research in Higher Education*, 33, pp. 317-375.

Feldman, K. A. (1993). College students' views of male and female college teachers: Part II—Evidence from students' evaluations of their classroom teachers, *Research in Higher Education*, 34, pp. 151-211.

Fink, D. (2008). Improving the Evaluation of College Teaching. *To Improve the Academy* 26, an annual collection of articles produced by the POD Network in Higher Education and published by Jossey-Bass

Flowers, L., Osterlind, S. J., Pascarella, E. T., & Pierson, C. T. (2001). How much do students learn in college? *Journal of Higher Education*, 72, 565-583.

Fox, C.R. (2006). The availability heuristic in the classroom: How soliciting more criticism can boost your course ratings. *Judgment and Decision Making*, 1(1), 86-90.

Frankhouser, W. M. (1984). The effects of different oral directions as to disposition of results on student ratings of college instruction, *Research in Higher Education*, 20, pp. 367-374.

Franklin, J. (2001). Interpreting the numbers: Using a narrative to help others read student evaluations of your teaching accurately. In K.G. Lewis (Ed.), Techniques and strategies for interpreting student evaluations [Special issue]. *New Directions for Teaching and Learning*, 87, 85-100.

Franklin, J., & Theall, M. (1989). *Who reads ratings: Knowledge, attitude and practice of users of student ratings of instruction?* Paper presented at the Annual meeting of the American Educational Research Association, San Francisco.

Franklin, J., & Theall, M. (1990). Communicating student ratings to decision makers: Design for good practice. In M. Theall & J. Franklin (Eds.), Student ratings of instruction: Issues for improving practice [Special issue]. *New Directions for Teaching and Learning*, 43, 75-93.

Franklin, J., Theall, M. & Ludlow, L. (1991). *Grade inflation and student ratings: A closer look*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

French-Lazovik, G. (1981), 'Peer review: documentary evidence in the evaluation of teaching' in Jason Millman (ed), *Handbook of Teacher Evaluation*, London, Sage, pp. 73-89.

Frey, P. W., Leonard, D. W., & Beatty, W. W. (1975). Student ratings of instruction: Validation research. *American Educational Research Journal*, 12, 435-447.

Frey, P. W. (1976) Validity of student instructional ratings: Does timing matter? *Journal of Higher Education*, 47, pp. 327-336.

G

Gaither, G.H. (1997). Performance indicator systems as instruments for accountability and assessment, *Assessment Update*, 9, pp. 1-2, 14-15.

Gaski, J. F. (1987). On “Construct validity of measures of college teaching effectiveness.” *Journal of Educational Psychology*, 79, 326-330.

Gaultney, J. F., & Cann, A. (2001). Grade expectations. *Teaching of Psychology*, 28 (2), 84-87.

Germann, P. J. (1988). Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal of Research in Science Teaching*, 25(8), 689-703.

Gibbs, G. (1988). *Creating a teaching profile*. Bristol, England: Teachin

Gigliotti, R. J. (1987). Are they getting what they expect? *Teaching Sociology*, 15, pp. 365-375.

Gillmore, G. M., & Greenwald, A. G. (1999). Using statistical adjustment to reduce biases in student ratings. *American Psychologist*, 54, 518-519.

Goldberk, G. & Callahan, J. (1991). Objectivity of student evaluations of instructors, *Journal of Education for Business*, 66, pp. 377-378.

Goldman, L. (1985). The betrayal of the gatekeepers: Grade inflation. *Journal of General Education*, 37(2), 97-121.

Goldschmid, M.L. (1978). The evaluation and improvement of teaching in higher education. *Higher Education*, 7(2), 221-245.

Gravestock, P., & Greco-Greenleaf, E. (2008). *Student Course Evaluations: Research, Model and Trends*, University of Toronto, Higher Education Quality Council of Ontario. Toronto, ON Canada.

Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, 52, 1182-1186.

Greenwald, A. G., & Gillmore, G. M. (1997a). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52, 1209-1217.

Greimel-Fuhrmann, B., & Geyer, A. (2003). Students' evaluation of teachers and instructional quality: Analysis of relevant factors based on empirical evaluation research. *Assessment & Evaluation in Higher Education*, 28(3), 229–239.

Grimes, P. W. (2002). The overconfident principles of economics students: An examination of meta-cognitive skill. *Journal of Economic Education*, 33(1), 15-30.

H

Hake, R. R. (2002). *Problems with student evaluations: Is assessment the remedy?* Retrieved June 19, 2007, and February 13, 2008, from <http://physics.indiana.edu/~hake/assesstherem1.pdf>

Hamalainen, K., & Moitus, S. (1999). High-quality education as the criteria for university funding in Finland. *Quality in Higher Education*, 5(1), 51-60.

Hamilton, J. B. III, Smith, M., Heady, R. B., & Carson, P. P. (1997). Using open-ended questions on senior exit surveys to evaluate and improve faculty performance: Results from a school of business administration. *Journal on Excellence in College Teaching*, 8(1), 23–48.

Harper, S.R., & Kuh, G. (2007). Myths and misconceptions about using qualitative methods in assessment. In S.R. Harper & S.D. Museus (Eds.), Using qualitative methods in institutional assessment [Special issue]. *New Directions for Institutional Research*, 136, 5-14.

Harvey, L. & Knight, P.T. (1996). *Transforming Higher Education* (Buckingham, UK, Society for Research into Higher Education/Open University Press).

Haskell, R.E. (1997). Academic freedom, tenure, and the student evaluation of faculty: Galloping polls in the 21st century. *Education Policy Analysis*, 5(6).

Hatch, T. (2006). *Into the classroom: Developing the scholarship of teaching and learning*. San Francisco: Jossey-Bass.

Havelka, D., Neal, C. S., & Beasley, F. (2003, November). Student evaluation of teaching effectiveness: What criteria are most important? Paper presented at the annual Lilly Conference College Teaching, Miami University, Oxford, OH.

Hewett, L., Chastain, G. & Thurberg, S. (1988). *Course evaluations: Are students' ratings dictated by first impressions?* Paper presented at the annual meeting of the Rocky Mountain Psychological Association, Snowbird, UT.

Hewson, P., W., & Hewson, M. G. A'B. (1989). Analysis and use of a task for identifying conceptions of teaching science. *Journal of Education for Teaching*, 15, 191-209.

Hodges, L.C., & Stanton, K. (2007). Translating comments on student evaluations into the language of learning. *Innovative Higher Education*, 31, 279-286

Holtta, S. (1998). The funding of universities in Finland: Towards goal-oriented government steering. *European Journal of Education*, 33, 55-63.

Howard, G. S., & Maxwell, S. E. (1980). Correlation between student satisfaction and grades: A case of mistaken causation? *Journal of Educational Psychology*, 72, 810-820.

Hoyt, D.P., & Pallett, W.H. (1999). *Appraising teaching effectiveness: Beyond student ratings* (IDEA Paper No.36). Manhattan, KS: Kansas State University Center for Faculty Evaluation and Development.

Hutchings, P., & Shulman, L. S. (1999). The scholarship of teaching: New elaborations, new developments. *Change*, 31(5), 11–15.

J

Jenrette, M., & Hayes, K. (1996). Honoring exemplary teaching: The two-year college setting. In M. D. Svinicki & R. J. Menges (Eds.), *Honoring exemplary teaching* (New Directions for Teaching and Learning, No. 65). San Francisco: Jossey-Bass.

Jones, D. (1997). *Perspectives on Performing Funding: Concepts, practices, and principles* (Unpublished draft).

Johnes, G. & Johnes J. (1993). Measuring the research performance of UK economics departments: an application of data envelopment analysis. *Oxford Economic Papers*, 4(2), pp. 332–47.

Johnson, V. E. (2003). *Grade inflation: A crisis in college education*. New York: Springer.

Johnson, B., & McClure, R. (2002). Validity and Reliability of a Shortened, Revised version of the Constructivist Learning Environment Survey (CLES). *Learning Environments Research*, 7, 65-80.

K

Kaplan, M., Mets, L. A., & Cook, C. E. (2000). Questions frequently asked about student ratings forms: Summary of research findings. Retrieved March 31, 2006, from <http://www.crlt.umich.edu/tstrategies/studentratingfaq.html>

Kaschak, E. (1978). Sex bias in student evaluations of college professors, *Psychology of Women Quarterly*, 2, pp. 235-242.

Keig, L. W., & Waggoner, M. D. (1994). *Collaborative peer review: The role of faculty in improving college teaching* (ASHE/ERIC Higher Education Report, No. 2). Washington, DC: Association for the Study of Higher Education.

Keig, L. W., & Waggoner, M. D. (1995). Peer review of teaching: Improving college instruction through formative assessment. *Journal on Excellence in College Teaching*, 6(3), 51–83.

Kennedy, E. J., Lawton, L., & Plumlee, E. L. (2002). Bliss ignorance: The problem of unrecognized incompetence and academic performance. *Journal of Marketing Education*, 24, 243-252.

Kierstead, D., D'Agostino, P. & Dill, H. (1988). Sex role stereotyping of college professors: Bias in students' ratings of instructors, *Journal of Educational Psychology*, 80, pp. 342-344.

King, F. A. (2000). The Changing Face of Accountability: Monitoring and Assessing Institutional Performance in Higher Education. *The Journal of Higher Education*, 71 (4), 411-431

Klann, W. E. & Hoff, E. (1976). The use of judgment analysis in analyzing student evaluation of teachers, *Mathematical Association of Two-Year College Journal*, 10, pp. 137-139.

Knapper, C. (1995). The origins of teaching portfolios. *Journal on Excellence in College Teaching*, 6(1), 45–56.

Knapper, C., & Cranton, P. (Eds.). (2001). Fresh approaches to the evaluation of teaching (New Directions for Teaching and Learning, No. 88). San Francisco: Jossey-Bass.

Knapper, C., & Wright, W. A. (2001). Using portfolios to document good teaching: Premises, purposes, practices. In C. Knapper & P. Cranton (Eds.), *Fresh approaches to the evaluation of teaching* (New Directions for Teaching and Learning, No. 88) (pp. 19–29). San Francisco.

Koblitz, N. (1990). Are student ratings unfair to women?, *Newsletter of the Association for Women in Mathematics*, 20, pp. 17-19.

Koermer, C. D. & Petelle, J. L. (1991). Expectancy violation and student rating of instruction, *Communication Quarterly*, 39, pp. 341-350.

Kolevzon, M. S. (1981). Grade inflation in higher education: A comparative study. *Research in Higher Education*, 15, 195-212.

Koshland, D. E. (1991). Teaching and research, *Science*, 251, p. 249.

Koushki, P. A. & Kuhn, H. A. J. (1982) How reliable are student evaluations of teachers?, *Engineering Education*, 72, pp. 362-367.

Kulik, J. A. (2001). Student ratings: Validity, utility, and controversy. In M. Theall, P. C. Abrami, & L. A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* (New Directions for Institutional Research, No. 109) (pp. 9–25). San Francisco: Jossey-Bass.

L

Lattuca, L.R., & Domagal-Goldman, J.M. (2007). Using qualitative methods to assess teaching effectiveness. In S.R. Harper & S.D. Museus (Eds.), Using qualitative methods in institutional assessment [Special issue]. *New Directions for Institutional Research*, 136, 81-93.

Layzell T. D. (1998). Linking Performance to Funding Outcomes for Public Institutions of Higher Education: The US Experience. *European Journal of Education*, 33(1), Financing Higher Education: Innovation and Changes. 103-111

Leventhal, L., Abrami, P. C. & Perry, R. P. (1976) Do teacher rating forms reveal as much about students as about teachers?, *Journal of Educational Psychology*, 68, pp. 441-445.

L' Hommedieu, R., Menges, R. J. & Brinko, K. T. (1990) Methodological explanations for the modest effects of feedback from student ratings, *Journal of Educational Psychology*, 82, pp. 232-241.

Lewis, K.G. (2001). Using mid-semester student feedback and responding to it. In K.G. Lewis (Ed.), Techniques and strategies for interpreting student evaluations [Special issue]. *New Directions for Teaching and Learning*, 87, 33-44.

Love, K. G. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias, and user reaction. *Journal of Applied Psychology*, 66, 451-457.

M

Machina, K. (1987). Evaluating student evaluations. *Academe*, 73(3), 19-22.

Marks, R. B. (2000). Determinants of student evaluations of global measures of instructor and course value. *Journal of Marketing Education*, 22, 108-119.

Marsh, H. W. (1979) Annotated Bibliography of Research on the Relationship Between Quality of Teaching and Quality of Research in Higher Education (Los Angeles, Office of Institutional Studies, University of Southern California).

Marsh, H.W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 253-388.

Marsh, H.W. (2007). Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years. *Journal of Educational Psychology*, 99(4), 775-790.

Marsh, H. W. & Cooper, T. L. (1981). Prior subject interest, students' evaluations, and instructor effectiveness, *Multivariate Behavioral Research*, 16, pp. 82-104.

Marsh, H.W., & Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional approach. In J.C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 8) (pp. 143-223). New York: Agathon Press.

Marsh, H.W., & Roche, L. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias and utility. *American Psychologist*, 52(11), 1187-1197.

Marsh, H. W., Hau, K., Chung, C., & Siu, T. L. (1997). Students' evaluations of university teaching: Chinese version of the Students' Evaluations of Educational Quality instrument. *Journal of Educational Psychology*, 89, 568-572.

Marsh, H. W. & Overall, J. U. (1980) Validity of students' evaluations of teaching effectiveness: Cognitive and affective criteria, *Journal of Educational Psychology*, 72, pp. 468-475.

Marsh, H.W., Overall, J. U., & Kesler, S. P. (1979). The validity of students' evaluations of instructional effectiveness: A comparison of faculty self-evaluations and evaluations by their students. *Journal of Educational Psychology*, 71, 149-160.

Marsh, H.W., & Roche, L. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias and utility. *American Psychologist*, 52(11), 1187-1197.

Marsh, H. W., & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92, 202-228.

Marsh, H. W. & Ware, J. E., JR. (1982). Effects of expressiveness, content coverage, and incentive on multidimensional student rating scales: New interpretations of the 'Dr. Fox' effect, *Journal of Educational Psychology*, 74, pp. 126-13

Martin, E. (1984). Power and authority in the classroom: Sexist stereotypes in teaching evaluations, *Journal of Women in Culture and Society*, 9, pp. 482-492.

Martin, E. (2006). Efficiency and Quality in the Current Higher Education Context in Europe: an application of the data envelopment analysis methodology to performance assessment of departments within the University of Zaragoza. *Quality in Higher Education*, 12, (1), p:57-79.

McCallum, L. W. (1984) A meta-analysis of course evaluation data and its use in the tenure decision, *Research in Higher Education*, 21, pp. 150-158.

McKeachie, W. J. (1979) Student ratings of faculty: A reprise, *Academe*, 65, pp. 384-397.

McKeachie, W. J. (1990) Research on college teaching: The historical background, *Journal of Educational Psychology*, 82, pp. 189-200.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist*, 52(11), 1218-1225.

McKeachie, W. J. & Kaplan, M. (1996). Persistent problems in evaluating college teaching. *AAHE Bulletin*, 48(6), 5–8.

McKeachie, W.J., & Lin, W.G. (1975). Multiple discriminant analysis of student ratings of college teachers. *Journal of Educational Research*, 68(8), 300-305.

McMillan, M. & Datta, D. (1998). The relative efficiencies of Canadian universities: a DEA perspective', *Canadian Public Policy*, 24(4), pp. 485–511.

McNaught, C., & Anwyl, J. (1993). *Awards for teaching excellence at Australian Universities* (University of Melbourne Centre for the Study of Higher Education Research Working Paper No. 93.1). (ED 368-291)

Menges, R. J. (1996). Awards to individuals. In M. D. Svinicki & R. J. Menges (Eds.), *Honoring exemplary teaching* (New Directions for Teaching and Learning, No. 65). San Francisco: Jossey-Bass.

Menges, R. J., & Brinko, K. T. (1986). *Effects of student evaluation feedback: A meta-analysis of higher education research*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Millea, M., & Grimes, P. W. (2002). Grade expectations and student evaluation of teaching. *College Student Journal*, 36(4), 582–591.

Millis, B. J., & Kaplan, B. B. (1995). Enhance teaching through peer classroom observations. In P. Seldin & Associates (Eds.), *Improving college teaching* (pp. 137–152). Bolton, MA: Anker.

Moore, S., & Kuol, N. (2005). A punitive tool or a valuable resource? Using student evaluations to enhance your teaching. In G. O'Neill, S. Moore, & B. McMulline (Eds.), *Emerging issues in the practice of university learning and teaching* (pp. 141-148). Dublin: All Ireland Society for Higher Education.

Moore, M., & Trahan, R. (1998). Tenure status and grading practices. *Sociological Perspectives*, 41, 775-781.

Mora, J.G., 2003, 'La mejora de la eficacia de la enseñanza superior en el nuevo contexto europeo', *Papeles de Economía Española*, 95, pp. 188–94.

Moreland, R., Miller, J., & Laucka, F. (1981). Academic achievement and self-evaluations of academic performances. *Journal of Educational Psychology*, 73, 335-344.

Morrissey, J. T. (1981). An analysis of studies on changing the attitude of elementary student teachers toward science and science teaching. *Science Education*, 65(2), 157-177.

Muchinsky, P. M. (1995). Peer review of teaching: Lessons learned from military and industrial research on peer assessment. *Journal on Excellence in College Teaching*, 6(3), 17–30.

Murnane, R. J., & Cohen, D. K. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and a few survive. *Harvard Educational Review*, 56, 1-17.

Murphy, T., & MacLaren, I. (2007). Theory and practice in the development of teaching portfolios at NUI Galway. In C. O'Farrell (Ed.), *Teaching portfolio practice in Ireland: A Handbook* (p. 86-100). Dublin, Ireland: AISHE.

Murray, H.G. (1987). Acquiring student feedback that improves instruction. In Weimer, M.G. (Ed.), *Teaching large classes well* [Special issue]. *New Directions for Teaching and Learning*, 32, 85-96.

Murray, J. P. (1995). The teaching portfolio: A tool for department chairperson to create a climate for teaching excellence. *Innovative Higher Education*, 19(3), 163–175.

Murray, H. G., Rushton, P. J. & Paunonen, S. V. (1990). Teacher personality traits and student instructional ratings in six types of university courses, *Journal of Educational Psychology*, 82, pp. 250-261.

N

Naftulin, D. H., Ware, J. E. & Donnelly, F. A. (1973) The Doctor Fox lecture: A paradigm of educational seduction, *Journal of Medical Education*, 48, pp. 630-635.

Nasser, F. & Fresko, B. (2002). Faculty view of student evaluation of college teaching. *Assessment & Evaluation in Higher Education*, 27 (2), 187-198.

Neave, G. (1998). The evaluative state reconsidered. *European Journal of Education*, 33(3), pp. 265–84.

Neumann, R. (2001). Disciplinary differences and university teaching. *Studies in Higher Education*, 26(2), 135-146.

Newport, J.F. (1996). Rating teaching in the USA: Probing the qualifications of student raters and novice teachers. *Assessment & Evaluation in Higher Education*, 21(1), 17-21.

Nlimmer, J. G. & Stone, E. F. (1991). Effects of grading practices and time of rating on student ratings of faculty performance and student learning, *Research in Higher Education*, 32, pp. 195-215.

O

O'Neil, C., & Wright, W. A. (1995). *Recording teaching accomplishment: A Dalhousie guide to the teaching dossier* (5th ed.). Halifax, Canada: Dalhousie University Office of Instructional Development and Technology.

Onwuegbuzie, A. J., & Leech, N. L. (2006). Linking research questions to mixed methods data analysis procedures. *The Qualitative Report*, 11, 474-498. Retrieved June 8, 2007, from <http://www.nova.edu/ssss/QR/QR11-3/onwuegbuzie.pdf>

Organization For Economic Cooperation And Development, (OECD). (2005) 'Education at a glance', OECD indicators–Edition 2005.

Ory, J.C. (2001). Faculty thoughts and concerns about student ratings. In K.G. Lewis (Ed.), *Techniques and strategies for interpreting student evaluations* [Special issue]. *New Directions for Teaching and Learning*, 87, 3-15.

Ory, J. C., & Braskamp, L. A. (1981). Faculty perceptions of the quality and usefulness of three types of evaluative information. *Research in Higher Education*, 15, 271–282.

Ory, J.C., & Ryan, K. (2001). How do student ratings measure up to a new validity framework? In M. Theall, P.C Abrami, & L.A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research*, 109, 27-44.

Overall, J. U., & Marsh, H. W. (1980). Students' evaluations of instruction: A longitudinal study of their stability. *Journal of Educational Psychology*, 72, 321–325.

P

Paswan, A. K., & Young, J. A. (2002). Student evaluation of instructors: A nomological investigation using structural equation modeling. *Journal of Marketing Education*, 24, 193-202

Pencavel, J. H. (1997). Work effort, on-the-job screening, and alternative methods of remuneration. In R. Ehrenberg (Ed.), *Research in labor economics, 1*, (pp. 225–258). Greenwich, CT: JAI Press.

Penny, A.R., & Coe. R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, 74(2), 215-253.

Perlberg, A. E. (1983). When professors confront themselves: Towards a theoretical conceptualization of video self-confrontation in higher education. *Higher Education*, 12, 633–663.

Perry, R. P. (1990). Introduction to the special section: Instruction in higher education, *Journal of Educational Psychology*, 82, pp. 183-188.

Perry, R. P., Abrami, P. C., Leventhal, L. & Check, J. (1979). Instructor reputation: An Expectancy relationship involving student ratings and achievement, *Journal of Educational Psychology*, 71, pp. 776-787.

Pollio, H. R., & Beck, H. P. (2000). When the tail wags the dog. *Journal of Higher Education*, 71, 84-102.

- Powell, R. W. (1977). Grades, learning, and student evaluation of instructors. *Research in Higher Education*, 7, 193-205.
- Powell, R. (1978). Faculty rating scale validity: The selling of a myth, *College English*, 39, pp. 616-629.
- Prave, R. S. & Baril, G. L. (1993). Instructor ratings: Controlling for bias from initial student interest, *Journal of Education for Business*, 68, pp. 362-366.
- Prawat, R., S. (1985). Affective versus cognitive goal orientations in elementary teachers. *American Educational Research Journal*, 22, 587-604.
- Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63(2), 167-199.
- Pulich, M. A. (1984). Better use of student evaluations for teaching effectiveness, *Improving College and University Teaching*, 32, pp. 91-94.

R

- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire, *Studies in Higher Education*, 16, pp. 129-150.
- Read, W. J., Rama, D. V., & Raghunandan, K. (2001). The relationship between student evaluations of teaching and faculty evaluations. *Journal of Business Education*, 76(4), 189-193.
- Redding, R. E. (1998). Students' evaluation of teaching fuel grade inflation. *American Psychologist*, 53, 1227-1228.
- Rice, R. E. (1991). The new American scholar: Scholarship and the purposes of the university. *Metropolitan Universities*, 1(4) 7-18.
- Rich, H.E (1976). Attitudes of college and university faculty toward the use of student evaluations. *Educational Research Quarterly*, 1, 17-28
- Roe, E. (1987). *How to compile a teaching portfolio*. Kensington, Australia: Federation of Australian University Staff Associations.
- Romberg, E. (1985). Description of peer evaluation within a comprehensive evaluation program in a dental school. *Instructional Evaluation*, 8 (1), 10-16.
- Ryan, J.J., Anderson, J.A., & Birchler, A.B. (1980). Student evaluation: The faculty responds. *Research in Higher Education*, 12 (4), 317-333.
- Ruedrich, S. L., Cavey, C., Katz, K., & Grush, L. (1992). Recognition of teaching excellence through the use of teaching awards: A faculty perspective. *Academic Psychiatry*, 16(1), 10-13.

Rutland, P. (1990). Some considerations regarding teaching evaluations, *Political Science Teacher*, 3, pp. 1-2.

S

Scherr, F. C. & Scherr, S. S. (1990) Bias in student evaluations of teacher effectiveness, *Journal of Education for Business*, 65, pp. 356-358.

Schmelkin, L.P., Spencer, K.J. & Gellman, E.S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education*, 38(5), 575-592.

Schwab, D. P. (1976). *Manual for the Course Evaluation Instrument*. Madison: University of Wisconsin, School of Business.

Scott, C. S. (1977). Student ratings and instructor-defined extenuating circumstances, *Journal of Educational Psychology*, 69, pp. 744-747.

Scriven, M. (1981). *Evaluation thesaurus* (3rd ed.). Pt. Reyes, CA: Edgepress.

Scriven, M. (1995). Student ratings offer useful input to teacher evaluations. *Practical Assessment, Research & Evaluation*, 4(7).

Sedlmeier, P. (2006). The role of scales in student ratings. *Learning and Instruction*, 16(5), 401-415.

Seiver, D. A. (1983). Evaluations and grades: A simultaneous framework. *Journal of Economic Education*, 14(3), 32-38.

Seldin, P. (1980). *Successful faculty evaluation programs: A practical guide to improved faculty performance and promotion/tenure decisions*. Crugers, NY: Coventry Press.

Seldin, P. (1984). *Changing Practices in Faculty Evaluation* (San Francisco, Jossey-Bass).

Seldin, P. (1989) Using student feedback to improve teaching, in: A. F. LUCAS (Ed.) The Department Chairperson's Role in Enhancing College Teaching. *New Directions for Teaching and Learning*, Vol. 37, pp. 89-97 (San Francisco, Jossey-Bass).

Seldin, P. (1993). *Successful use of teaching portfolios*. Bolton, MA: Anker.

Seldin, P. (1991, 1997) *The Teaching Portfolio: A Practical Guide to Improved Performance and Promotion/Tenure Decisions*. 1st edition: 1991; 2nd edition, 1997. Bolton, MA: Anker.

Seldin, P. (1998, February). *The teaching portfolio*. Paper presented for the American Council on Education, Department Chairs Seminar, San Diego, CA.

Seldin, P. (1999a). Current practices – good and bad –nationally. In P. Seldin & Associates (Eds.), *Changing practices in evaluating teaching: A practical guide to*

improved faculty performance and promotion/tenure decisions (pp. 1–24). Bolton, MA: Anker.

Seldin, P. (1999b). Self-evaluation: What works? What doesn't? In P. Seldin & Associates (Eds.), *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions* (pp. 97–115). Bolton, MA: Anker.

Seldin, P. (1999). *Changing practices in evaluating teaching. A practical guide to improved faculty performance for promotion/tenure decisions*. Bolton, MA: Anker.

Seldin, P. (2004). *The teaching portfolio* (3rd ed.). Bolton, MA: Anker.

Seldin P, editor (2006). *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. Bolton, MA: Anker.

Seldin, P., Annis, L., & Zubizarreta, J. (1995). Answers to common questions about the teaching portfolio. *Journal on Excellence in College Teaching*, 6(1), 57–64.

Seldin, P., & Associates (Eds.). (1999). *Changing practices in evaluating teaching: A practical guide to improve faculty performance and promotion / tenure decisions*. Bolton, MA: Anker.

Seppanen, L. J. (1995). Linkages to the world of employment. In P. T. Ewell (Ed.), *Student tracking: New techniques, new demands*. San Francisco: Jossey-Bass.

Sewell, William F. 1992. "A Theory of Structure: Duality, Agency, and Transformation". *The American Journal of Sociology*, Volume 98, Number 1 (Jul., 1992), 1-29. Meredith F. Burnett

Shariff, S. H. (1999). Students' quality control circle: A case study on students' participation in the quality of control circles at the Faculty of Business and Management. *Assessment & Evaluation in Higher Education*, 24, 141–146.

Sheehan, E. P., & DuPrey, T. (1999). Student evaluations of university teaching. *Journal of Instructional Psychology*, 26, 188–193.

Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: Love me, love my lectures? *Assessment & Evaluation in Higher Education*, 25(4), 397–405.

Shore, B. M., & Associates. (1986). *The teaching dossier: A guide to its preparation and use* (rev.ed.). Ottawa: Canadian Association of University Teachers.

Shore, B. M., & Associates. (1986). *The teaching dossier: A guide to its preparation and use* (rev.ed.). Ottawa: Canadian Association of University Teachers.

Shrigley, R. L., Koballa, T. R., & Simpson, R. D. (1988). Defining attitude for science educators. *Journal of Research in Science Teaching*, 25, 659–678.

Shulman, L. (1987). Knowledge of teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–21.

Shulman, L. S. (1998). Course anatomy: The dissection and analysis of knowledge through teaching. In P. Hutchings (Ed.), *The course portfolio: How faculty can*

examine their teaching to advance practice and improve student learning. Washington, DC: American Association for Higher Education.

Siegel, A. I. (1986). Performance tests. In R. A. Berk (Ed.), *Performance assessment: Methods and applications* (pp. 121–142). Baltimore, MD: Johns Hopkins University Press.

Simpson, P. M., & Siguaw, J. A. (2000). Student evaluations of teaching: An exploratory study of the faculty response. *Journal of Marketing Education*, 22, 199-213.

Sixbury, G. R., & Cashin, W. E. (1995). *Description of database for the IDEA Diagnostic Form* (IDEA Technical Report No. 9). Manhattan: Center for Faculty Evaluation & Development, Division of Continuing Education, Kansas State University.

Small, A. C, Hollenbeck, A. R. & Haley, R. L. (1982). The effect of emotional state on student ratings of instructors, *Teaching of Psychology*, 9, pp. 205-208.

Snyder, C. R. & Clair, M. (1976). Effects of expected and obtained grades on teacher evaluation and attribution of performance, *Journal of Educational Psychology*, 68, pp. 75-82.

Soderberg, L.O. (1986). A credible model: Evaluating classroom teaching in higher education. *Instructional Evaluation*, 8, 13–27.

Sojka, J., Gupta, A. K., & Deeter-Schmelz, D. R. (2002). Student and faculty perceptions of student evaluations of teaching. *College Teaching*, 50(2), 44–49.

Sorey, K. E. (1968). A study of the distinguishing characteristics of college faculty who are superior in regard to the teaching function. *Dissertation Abstracts*, 28(12-A), 4916.

Sowlati, T. & Paradi, J. (2004). Establishing the “practical frontier” in data envelopment analysis, *Omega- International Journal of Management Science*, 32, pp. 261–272.

Sproule, R. (2000). Student evaluation of teaching: A methodological critique of evaluation practices. *Education Policy Analysis*, 8(50).

Sproule, R. (2002). The under-determination of instructor performance by data from the student evaluation of teaching. *Economics of Education Review*, 21(3), 287–295.

Stapleton, R. J., & Murkison, G. (2001). Optimizing the fairness of student evaluations: A study of correlations between instructor excellence, study production, learning production, and expected grades. *Journal of Management Education*, 25, 269-291.

Steiner, S., Holley, L. C., Gerdes, K., & Campbell, H. E. (2006). Evaluating teaching: Listening to students while acknowledging bias. *Journal of Social Work Education*, 42, 355-376.

Stumpf, S. A., & Freedman, R. D. (1979). Expected grade covariation with student ratings of instruction: Individual versus class effects. *Journal of Educational Psychology*, 71, 293-302.

Sullivan, A. M., & Skanes, G. R. (1974). Validity of student evaluation of teaching and the characteristics of successful instructors. *Journal of Educational Psychology*, 66, 584-590.

T

Tatro, C. N. (1995) Gender effects on student evaluations of faculty, *Journal of Research and Development in Education*, 28, pp. 169-173.

Theall, M. (1994). What's wrong with faculty evaluation: A debate on the state of practice. *Instructional Evaluation and Faculty Development*, 14, 27-34

Theall, M., Abrami, P. C., & Mets, L. A. (Eds.). (2001). *The student ratings debate: Are they valid? How can we best use them?* (New Directions for Institutional Research, No. 109). San Francisco: Jossey-Bass.

Theall, M., & Franklin, J. L. (1990). Student ratings in the context of complex evaluation systems. In M. Theall & J. L. Franklin (Eds.), *Student ratings of instruction: Issues for improving practice* (New Directions for Teaching and Learning, No. 43). San Francisco: Jossey-Bass.

Theall, M. & Franklin, J. (1991). Using student ratings for teaching improvement, in: M. Theall & J. Franklin (Eds.), *Effective Practices for Improving Teaching*, *New Directions for Teaching and Learning*, No. 48, pp. 83-96 (San Francisco, Jossey-Bass).

Theall, M., & Franklin, J. (2000). Creating responsive student ratings systems to improve evaluation practice. In K.E. Ryan (Ed.), *Evaluating teaching in higher education: A vision for the future* [Special issue]. *New Directions for Teaching and Learning*, 83, 95-107.

Theall, M., & Franklin, J. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction? In M. Theall, P.C Abrami, & L.A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research*, 109, 45-56.

Thune, C. (1998). The European systems of quality assurance - dimensions of harmonization and differentiation. *Higher Education Management*, 10(3), pp. 9-27.

Tomkins, C. & Green, R. (1988). An experiment in the use of data envelopment analysis of evaluating the efficiency of UK university departments of accounting, *Financial Accountability and Management*, (Summer) 4, pp. 147-64.

Trinka, J., (2002). Students' course and faculty evaluations: An informal look. *Psychological Reports*, 91, 988.

Τσελφές, Β., & Παρούση, Αντ. (2009). Θεατρική αφήγηση επιστημονικών ιδεών: αφορμή για μια διαλεκτική προσέγγιση της μάθησης, *Κριτική* 9 <http://www.hpdst.gr/el/publications/kritiki/issues/9/33>

V

Vasta, R. & Sarmiento, R. F. (1979). Liberal grading improves evaluations but not performance, *Journal of Educational Psychology*, 71, pp. 207-211.

W

Wachtel, H.K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, 29(2), 191-121.

Wagenaar, T.C. (1995). Student evaluation of teaching: Some cautions and suggestions. *Teaching Sociology*, 23 (1), 64-68.

Watters, J. J. & Ginns, I. S. (1994). Science anxiety and self efficacy among preservice primary teachers: origins and remedies. *Research in Science Education*, 24, 348-357.

Watters, J. J. & Ginns, I. S. (1995). *Origins of, and changes in pre-service teachers' science teaching self efficacy*, Paper presented at the Annual Meeting of National Association for Research in Science Teaching, April 22-25, 1995, San Francisco CA.

Webb, J., & McEnerney, K. (1995). The view from the back of the classroom: A faculty-based peer observation program. *Journal on Excellence in College Teaching*, 6(3), 145-160.

Weimer, M. E. (1990). *Improving college teaching: Strategies for developing instructional effectiveness*. San Francisco: Jossey-Bass.

Weimer, M. E. (1993). The disciplinary journal of pedagogy. *Change*, 25, 44-51.

Weinberg, B. A., Fleisher, B. M., & Hashimoto, M. (2007). *Evaluating methods for evaluating instruction: The case of higher education* (NBER Working Paper No. 12844). Retrieved June 21, 2008, from <http://www.nber.org/papers/w12844>.

Wergin, J. E. (1992, September). *Developing and using performance criteria*. Paper presented at the Virginia Commonwealth University Conference on Faculty Rewards, Richmond.

Wilhelm, W. B. (2004). The relative influence of published teaching evaluations and other instructor attributes on course choice. *Journal of Marketing Education*, 26, 17-30.

Williams, W. M., & Ceci, S. J. (1997). "How'm I Doing?": Problems with student ratings of instructors and courses. *Change*, 29(5), 13-23.

Wright, A. W., & Associates (1995). Teaching improvement practices: Successful strategies for higher education. Bolton, MA: Anker

Wright, R.E. (2006). Student Evaluations of Faculty: Concerns Raised in the Literature, and Possible Solutions, *College Student Journal*, 40:2, 417-422.

Wright, A. W., & Associates (1995). Teaching improvement practices: Successful strategies for higher education. Bolton, MA: Anker.

Z

Zabaleta, F. (2007). The use and misuse of student evaluations of teaching. *Teaching in Higher Education*, 12 (1), 55-76.

Zahorski, K. J. (1996). Honoring exemplary teaching in the liberal arts institution. In M. D. R. J. Menges (Eds.), Honoring exemplary teaching. (New Directions for Teaching and Learning, No. 65). San Francisco: Jossey-Bass.

Y

Yunker, P. J., & Yunker, J. (2003). Are student evaluations of teaching valid? Evidence from an analytical business core course. *Journal of Education for Business*, 78, 313-317.

ΠΑΡΑΡΤΗΜΑΤΑ

ΠΑΡΑΡΤΗΜΑ Α

Μετρήσεις της Αποτελεσματικότητας της Διδασκαλίας (Πολλαπλών Διαστάσεων)

A 1: Student Evaluation of Educational Quality (SEEQ) Instrument

A 2: Individual Development & Educational Assessment (IDEA) Student Ratings of Instruction

A 3: Teaching Behaviors Inventory (TBI)

A 1: Student Evaluation of Educational Quality (SEEQ) Instrument

Το ερωτηματολόγιο SEEQ αναπτύχθηκε από τον Δρ. Herbert Marsh το 1987. Η φόρμα ζητά από τους φοιτητές να βαθμολογήσουν τους διδάσκοντες με βάση διάφορες πολλαπλών διαστάσεων όψεις της αποτελεσματικότητας της διδασκαλίας που αναγνωρίστηκαν διαμέσου τη έρευνας πάνω στη μάθηση των φοιτητών και την ανάλυση παραγόντων άλλων εργαλείων αξιολόγησης. Αυτή εισηγείται μόνο δηλώσεις που οι φοιτητές μπορούν με ακρίβεια να βαθμολογήσουν με βάση την εμπειρία τους από τη σειρά μαθημάτων. Το SEEQ ή μία παραλλαγή του, χρησιμοποιείται από πολλά ιδρύματα. Η παρακάτω φόρμα αντιγράφηκε από Mt. Allison University's Puddy Crawford Teaching Centre at http://www.mta.ca/pctc/TONI_SEEQ/StudentEvaluation of Educational Quality SHORT.pdf.

Student Evaluation of Educational Quality (SEEQ) - SHORT VERSION
 Developed by Dr. Herbert W. Marsh
 Paper Version

Please read each question very carefully. Make sure you understand what is being asked. Use this scale and circle the number that is closest to your rating for that item:

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable

1) I found the course intellectually challenging and stimulating.

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable

2) I learned and understood the subject materials in this course.

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable

3) Instructor had a genuine interest in individual students

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable

4) Instructor was adequately accessible to students during office hours or after class.

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable

5) Instructor is enthusiastic about teaching the course.

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable

6) Feedback on examinations/graded materials was valuable.

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable

7) Methods of evaluating student work were fair and appropriate.

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable

8) Instructor's materials were well prepared and carefully explained.

Very Poor	Poor	Moderate	Good	Very Good	N/A
1	2	3	4	5	Not applicable


9) Students were encouraged to express their own ideas and/or question the instructor.

Very Poor	Poor	Moderate	Good	Very Good	N/A
-----------	------	----------	------	-----------	-----

- Τα κίνητρα των φοιτητών να πάρουν το μάθημα αναφορικά με το ποιος το διδάσκει
- Τις συνήθειες εργασίας των φοιτητών
- Το μέγεθος της τάξης
- Την προσπάθεια των φοιτητών που δεν αποδίδεται στον διδάσκοντα
- Τη δυσκολία των μαθημάτων που δεν αποδίδεται στον διδάσκοντα

Η φόρμα που ακολουθεί αντιγράφηκε από το


IDEA Center website at http://www.idea.ksu.edu/forms/IDEA_copyrightSR.html



IDEA CENTER


SURVEY FORM - STUDENT REACTIONS TO INSTRUCTION AND COURSES

IMPORTANT!




USE NO. 1 PENCIL ONLY

Proper Marks



Improper Marks



Institution: _____ Instructor: _____

Course Number: _____ Time and Days Class Meets: _____

Your thoughtful answers to these questions will provide helpful information to your instructor.

Describe the frequency of your instructor's teaching procedures, using the following code:

1=Hardly Ever 2=Occasionally 3=Sometimes 4=Frequently 5=Almost Always

The Instructor:

1. (1) (2) (3) (4) (5) Displayed a personal interest in students and their learning
2. (1) (2) (3) (4) (5) Found ways to help students answer their own questions
3. (1) (2) (3) (4) (5) Scheduled course work (class activities, tests, projects) in ways which encouraged students to stay up-to-date in their work
4. (1) (2) (3) (4) (5) Demonstrated the importance and significance of the subject matter
5. (1) (2) (3) (4) (5) Formed "teams" or "discussion groups" to facilitate learning
6. (1) (2) (3) (4) (5) Made it clear how each topic fit into the course
7. (1) (2) (3) (4) (5) Explained the reasons for criticisms of students' academic performance
8. (1) (2) (3) (4) (5) Stimulated students to intellectual effort beyond that required by most courses
9. (1) (2) (3) (4) (5) Encouraged students to use multiple resources (e.g. data banks, library holdings, outside experts) to improve understanding
10. (1) (2) (3) (4) (5) Explained course material clearly and concisely
11. (1) (2) (3) (4) (5) Related course material to real life situations
12. (1) (2) (3) (4) (5) Gave tests, projects, etc. that covered the most important points of the course
13. (1) (2) (3) (4) (5) Introduced stimulating ideas about the subject
14. (1) (2) (3) (4) (5) Involved students in "hands on" projects such as research, case studies, or "real life" activities
15. (1) (2) (3) (4) (5) Inspired students to set and achieve goals which really challenged them
16. (1) (2) (3) (4) (5) Asked students to share ideas and experiences with others whose backgrounds and viewpoints differ from their own
17. (1) (2) (3) (4) (5) Provided timely and frequent feedback on tests, reports, projects, etc. to help students improve
18. (1) (2) (3) (4) (5) Asked students to help each other understand ideas or concepts
19. (1) (2) (3) (4) (5) Gave projects, tests, or assignments that required original or creative thinking
20. (1) (2) (3) (4) (5) Encouraged student-faculty interaction outside of class (office visits, phone calls, e-mail, etc.)

Twelve possible learning objectives are listed below, not all of which will be relevant in this class. Describe the amount of progress you made on each (even those not pursued in this class) by using the following scale:

1-No apparent progress
 2-Slight progress; I made small gains on this objective.
 3-Moderate progress; I made some gains on this objective.
 4-Substantial progress; I made large gains on this objective.
 5-Exceptional progress; I made outstanding gains on this objective.

Progress on:

21. (1) (2) (3) (4) (5) Gaining factual knowledge (terminology, classifications, methods, trends)
22. (1) (2) (3) (4) (5) Learning fundamental principles, generalizations, or theories
23. (1) (2) (3) (4) (5) Learning to *apply* course material (to improve thinking, problem solving, and decisions)
24. (1) (2) (3) (4) (5) Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course
25. (1) (2) (3) (4) (5) Acquiring skills in working with others as a member of a team
26. (1) (2) (3) (4) (5) Developing creative capacities (writing, inventing, designing, performing in art, music, drama, etc.)
27. (1) (2) (3) (4) (5) Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)
28. (1) (2) (3) (4) (5) Developing skill in expressing myself orally or in writing
29. (1) (2) (3) (4) (5) Learning how to find and use resources for answering questions or solving problems
30. (1) (2) (3) (4) (5) Developing a clearer understanding of, and commitment to, personal values
31. (1) (2) (3) (4) (5) Learning to *analyze* and *critically evaluate* ideas, arguments, and points of view
32. (1) (2) (3) (4) (5) Acquiring an interest in learning more by asking my own questions and seeking answers

Mark Reflex® forms by Pearson NCS MM248694-2 654321 Printed in U.S.A.

Continued on back page

On the next three items, compare this course with others you have taken at this institution, using the following code:
 1=Much Less than Most Courses 2=Less than Most Courses 3=About Average 4=More than Most Courses 5=Much More than Most Courses

The Course:

33. ① ② ③ ④ ⑤ Amount of reading
 34. ① ② ③ ④ ⑤ Amount of work in other (non-reading) assignments
 35. ① ② ③ ④ ⑤ Difficulty of subject matter

Describe your attitudes and behavior in this course, using the following code:
 1=Definitely False 2=More False Than True 3=In Between 4=More True Than False 5=Definitely True

36. ① ② ③ ④ ⑤ I had a strong desire to take this course.
 37. ① ② ③ ④ ⑤ I worked harder on this course than on most courses I have taken.
 38. ① ② ③ ④ ⑤ I really wanted to take a course from this instructor.
 39. ① ② ③ ④ ⑤ I really wanted to take this course regardless of who taught it.
 40. ① ② ③ ④ ⑤ As a result of taking this course, I have more positive feelings toward this field of study.
 41. ① ② ③ ④ ⑤ Overall, I rate this instructor an excellent teacher.
 42. ① ② ③ ④ ⑤ Overall, I rate this course as excellent.

For the following items, blacken the space which best corresponds to your judgment:
 1=Definitely False 2=More False Than True 3=In Between 4=More True Than False 5=Definitely True

43. ① ② ③ ④ ⑤ As a rule, I put forth more effort than other students on academic work.
 44. ① ② ③ ④ ⑤ The instructor used a variety of methods—not only tests—to evaluate student progress on course objectives.
 45. ① ② ③ ④ ⑤ The instructor expected students to take their share of responsibility for learning.
 46. ① ② ③ ④ ⑤ The instructor had high achievement standards in this class.
 47. ① ② ③ ④ ⑤ The instructor used educational technology (e.g., Internet, e-mail, computer exercises, multi-media presentations, etc.) to promote learning.

EXTRA QUESTIONS

If your instructor has extra questions, answer them in the space designated below (questions 48-66):

- | | |
|---------------|---------------|
| 48. ① ② ③ ④ ⑤ | 58. ① ② ③ ④ ⑤ |
| 49. ① ② ③ ④ ⑤ | 59. ① ② ③ ④ ⑤ |
| 50. ① ② ③ ④ ⑤ | 60. ① ② ③ ④ ⑤ |
| 51. ① ② ③ ④ ⑤ | 61. ① ② ③ ④ ⑤ |
| 52. ① ② ③ ④ ⑤ | 62. ① ② ③ ④ ⑤ |
| 53. ① ② ③ ④ ⑤ | 63. ① ② ③ ④ ⑤ |
| 54. ① ② ③ ④ ⑤ | 64. ① ② ③ ④ ⑤ |
| 55. ① ② ③ ④ ⑤ | 65. ① ② ③ ④ ⑤ |
| 56. ① ② ③ ④ ⑤ | 66. ① ② ③ ④ ⑤ |
| 57. ① ② ③ ④ ⑤ | |

Use the space below for comments (unless otherwise directed).
 Note: Your written comments may be returned to the instructor. You may want to PRINT to protect your anonymity.

Comments: _____

A 3: Teaching Behaviors Inventory (TBI)

Το TBI σχεδιάστηκε από τον Δρ. Harry Murray, στο Πανεπιστήμιο του Δυτικού Οντάριο. Όπως και το SEEQ, αυτό επιτρέπει στους φοιτητές να βαθμολογήσουν τη διδασκαλία με βάση εννέα κατηγορίες της διδασκαλικής συμπεριφοράς για την οποία οι φοιτητές μπορούν να διαθέτουν επαρκή ανάδραση. Το εργαλείο είναι πολλαπλών διαστάσεων και έχει επιλεγεί να χρησιμοποιηθεί από πολλά ιδρύματα. Η πληροφορία που ακολουθεί αντιγράφηκε από το:

Wilfred Laurier University's Teaching Support Services Website at http://cubic.wlu.ca/page.php?grp_id=333&p=3335

Teacher Behaviours Inventory

(Adapted from the original Teacher Behaviours Inventory authored by: Professor H. G. Murray, Department of Psychology, University of Western Ontario)

Instructions to Student

PLEASE NOTE: At the request of your instructor, you are asked in this inventory to assess his/her specific classroom behaviours. Please try to be both thoughtful and candid in your responses, so as to maximize the value of feedback to your instructor. Your judgements should reflect the type of teaching you think is best for this particular course and your particular learning style. Try to assess each behaviour independently, rather than letting your overall impression of the instructor determine each individual rating.

Each section of the inventory begins with a definition of the category of teaching to be assessed in that section. You are asked to evaluate your instructor on 60 specific teaching behaviours. Using a pencil on the provided computer card, please indicate your judgement as to whether you agree or disagree with the statements below; do **not** indicate your identity on the computer card.

Please use the following rating scale in making your judgements for each of the 60 items listed below.

- 1 (Should Significantly Decrease)
- 2 (Should Slightly Decrease)
- 3 (No Change Needed)
- 4 (Should Slightly Increase)
- 5 (Should Significantly Increase)
- 10 (Not Applicable)

CLARITY: methods used to explain or clarify concepts and principles

1. Gives several examples of each concept.
2. Defines new or unfamiliar terms.
3. Repeats difficult ideas several times.
4. Stresses most important points.
5. Uses graphs or diagrams to facilitate explanation.
6. Points out practical applications of concepts.
7. Answers students' questions thoroughly.
8. Highlights key terms.
9. Explains subject matter in familiar conversational language.

EXPRESSION: use of non-verbal behaviour to solicit student attention and interest

10. Speaks in a dramatic or expressive way.
11. Moves about while lecturing.

12. Gestures with hands or arms.
13. Makes eye contact with students.
14. Gestures with head or body.
15. Tells jokes or humorous anecdotes.
16. Reads lecture verbatim from prepared notes or text.
17. Smiles or laughs while teaching.
18. Shows distracting mannerisms.

INTERACTION: techniques used to foster students' participation in class

19. Encourages students to ask questions or make comments during lectures.
20. Offers constructive criticism.
21. Praises students for good ideas.
22. Asks questions of individual students.
23. Asks questions of class as a whole.
24. Incorporates students' ideas into lecture.
25. Presents challenging, thought-provoking ideas.
26. Uses a variety of activities in class (e.g., group work, guest lecturers, etc.).
27. Asks rhetorical questions.

ORGANIZATION: ways of organizing or structuring the course's subject matter

28. Uses headings and subheadings to organize lectures.
29. Puts outline of lecture on blackboard or overhead screen.
30. Clearly indicates transition from one topic to the next.
31. Gives preliminary overview of lecture at beginning of class.
32. Explains how each topic fits into the course as a whole.
33. Reviews topics covered in previous lecture at beginning of each class.
34. Periodically summarizes points previously made.

PACING: rate of presentation of information; efficient use of class time

35. Clarifies fundamental points.
36. Digresses from major theme of lecture.
37. Covers very little material in class sessions.
38. Asks if students understand before proceeding to next topic.
39. Sticks to the point in answering students' questions.

DISCLOSURE: explicitness concerning course requirements and grading criteria

40. Advises students as to how to prepare for tests or exams.
41. Provides sample exam questions.
42. Provides clear expectations for all assessed work.
43. States objectives of each lecture.
44. Reminds students of test dates or assignment deadlines.
45. States objectives of course as a whole.

SPEECH: characteristics of voice relevant to classroom teaching

46. Speaks at an appropriate volume.
47. Speaks clearly.
48. Speaks at an appropriate pace.
49. Says "um" or "ah".
50. Speaks with voice modulation (fluctuates).

RAPPORT: quality of interpersonal relations between teacher and students

51. Addresses individual students by name.
52. Announces availability for consultation outside of class.
53. Offers to help students with course-related problems.
54. Shows tolerance of other points of view.
55. Talks with students before or after class, when possible.

TEACHING AIDS: use of media and materials to enhance learning

56. Uses visual teaching aids.
57. Makes effort to ensure readability of visual aids.
58. Uses audio, video, and computer equipment.
59. Uses presentation software.
60. Uses video programs.

ΠΑΡΑΡΤΗΜΑ Β

Έγγραφα Ιδρυμάτων σχετικά με την Αξιολόγηση Μαθημάτων

- B 1:** Δεξαμενή Ερωτήσεων για Αξιολόγηση Μαθημάτων, Πανεπιστήμιο του Μίσιγκαν
- B 2:** Κλίμακα Αξιολόγησης από Ομοβάθμιο (peer), Πανεπιστήμιο Τζων Χόπκινς, Νοσηλευτική Σχολή
- B 3:** Κλίμακα Αυτό-αξιολόγησης (self-evaluation), Πανεπιστήμιο Τζων Χόπκινς, Νοσηλευτική Σχολή
- B 4:** Κλίμακα Αποτελεσματικότητας της σχέσης με τον Μέντορα

- B 1:** Δεξαμενή Ερωτήσεων για Αξιολόγηση Μαθημάτων, Πανεπιστήμιο του Μίσιγκαν

Student Course-Guide Questions

These questions help students obtain information about University courses. If you include these eight questions on your teaching questionnaire, results from these questions and the University-wide questions will be released to the Michigan Student Assembly (MSA) for publication in the on-line and printed course evaluation guide *Advice*.

891. The workload for this *course* was (5=LIGHT ... 1=HEAVY)
892. Students felt comfortable asking questions.
893. Graded assignments reflected the material covered.
894. The grades in this *course* were fairly determined.
895. Students' difficulty with the material was recognized.
896. My expected grade in this *course* is (5=A ... 1=E).
897. The *course* requirements were clearly defined.
898. The *instructor* presented material clearly in lectures/discussions.

Please check the appropriate square below:

- I want to include the *Advice* questions on my questionnaire.
 I do not want to include the *Advice* questions on my questionnaire.

Teaching Improvement Questions

Student responses to these questions can help teachers find strengths and weaknesses in their teaching. Note that you may choose up to 26 questions from those printed below if you do not include the *Advice* questions on your questionnaire. If you include the *Advice* questions, you may choose only 18 items from the questions listed below. Indicate your choices by circling the numbers that appear before the questions.

STUDENT DEVELOPMENT

Knowledge

120. I learned a good deal of factual material in this *course*.
121. I gained a good understanding of concepts/principles in this field.
122. I learned to apply principles from this course to new situations.
123. I learned to identify main points and central issues in this field.
124. I learned to identify formal characteristics of works of art.
125. I developed the ability to solve real problems in this field.
126. I developed creative ability in this field.
127. I developed the ability to communicate clearly about this subject.
128. I developed ability to carry out original research in this area.
129. I developed an ability to evaluate new work in this field.
130. I learned to recognize the quality of works of art in this field.
131. I became more aware of multiple perspectives on issues of diversity.
132. I learned to think critically about difficult issues of diversity.

Interests and Values

140. I deepened my interest in the subject matter of this course.
141. I developed enthusiasm about the course material.
142. I was stimulated to do outside reading about the course material.
143. I was stimulated to discuss related topics outside of class.
144. I developed plans to take additional related courses.
145. I developed a set of overall values in this field.

Participation

160. I participated actively in class discussion.
161. I developed leadership skills in this class.
162. I developed new friendships in this class.

Social Awareness

163. I developed greater awareness of societal problems.
164. I became interested in community projects related to the course.
165. I learned to value new viewpoints.
166. I reconsidered many of my former attitudes.
167. I increased my appreciation of other students in this class.

Self-concept

170. I gained a better understanding of myself through this course.
171. I gained an understanding of some of my personal problems.
172. I developed a greater sense of personal responsibility.
173. I increased my awareness of my own interests and talents.
174. I developed more confidence in myself

Vocational Skills and Attitudes

180. I developed skills needed by professionals in this field.
181. I learned about career opportunities.
182. I developed a clearer sense of professional identity.

INSTRUCTOR EFFECTIVENESS

Instructor Skill

198. I was very satisfied with the educational experience this *instructor* provided.
199. The *instructor* explained material clearly and understandably.
200. The *instructor* handled questions well.
201. The *instructor* gave clear explanations.
202. The *instructor* made good use of examples and illustrations.
203. The *instructor* stressed important points in lectures/discussions.
204. The *instructor* was enthusiastic.
205. The *instructor* put material across in an interesting way.
206. The *instructor* seemed to enjoy teaching.
207. The *instructor* appeared to have a thorough knowledge of the subject.
208. The *instructor* seemed knowledgeable in many areas.
209. The *instructor* was not confused by unexpected questions.
210. The *instructor* was skillful in observing student reactions.
211. The *instructor* was sensitive to student difficulty with course work.
212. The *instructor* taught near the class level.
213. The *instructor* used examples that had relevance for me.
214. The *instructor* taught in a manner that served my needs as a student.
215. The *instructor* was sensitive to multicultural issues in the classroom.
216. The *instructor* was effective in handling multicultural issues and content.
217. The *instructor* promoted meaningful discussions of issues of diversity.
218. The *instructor* handled controversy in the classroom productively.
219. The *instructor* challenged stereotypic assumptions in discussions.
220. The *instructor* accommodated students with various learning needs.
221. The *instructor* accommodated the needs of students with disabilities.
222. The *instructor* tried to accommodate individual rates of learning.
223. The *instructor* tried to accommodate individual styles of learning.
224. The *instructor* responded to the different language needs of students.

Instructional Climate

213. The *instructor* was friendly.
214. The *instructor* was permissive and flexible.
215. The *instructor* maintained an atmosphere of good feeling in class.
216. The *instructor* acknowledged all questions insofar as possible.
217. The *instructor* treated students with respect.
218. The *instructor* encouraged constructive criticism.
219. The *instructor* was willing to meet and help students outside class.
220. The *instructor* gave individual attention to students in the class.
221. The *instructor* treated all students fairly.
222. The *instructor* encouraged student participation in an equitable way.
223. The *instructor* valued the diversity of life experiences among students.
224. The *instructor* tried to learn the names of all students.
225. The *instructor* made me feel known as an individual in this class.
226. The classroom's physical environment was conducive to learning.
227. The *instructor* appeared open to viewpoints besides her or his own.
228. The *instructor* was open to contributions from all class members.

251. The *instructor* saw cultural and personal differences as assets.

Interaction

221. Students frequently volunteered their own opinions.
 222. One real strength of this *course* was the classroom discussion.
 223. Students in this *course* were free to disagree and ask questions.
 252. The *instructor* made me feel valued in this class.
 253. I felt included and valued when working with other students.
 254. Group activities in this class contributed to my learning.
 255. Collaborative group activities helped me learn the material.
 256. Working with other students helped me learn more effectively.

Feedback

224. The *instructor* suggested specific ways students could improve.
 225. The *instructor* told students when they had done especially well.
 226. The *instructor* kept students informed of their progress.

Organization

227. The *instructor* had everything going according to schedule.
 228. The *instructor* followed an outline closely.
 229. The *instructor* used class time well.
 230. The *instructor* seemed well prepared for each class.
 231. The objectives of the *course* were clearly explained.
 232. Work requirements and grading system were clear from the beginning.

Difficulty

239. The amount of work required was appropriate for the credit received.
 240. The amount of material covered in the *course* was reasonable.
 241. The *instructor* set high standards for students.
 242. The *instructor* made the course difficult enough to be stimulating.

WRITING ASSIGNMENTS

318. Writing assignments seemed carefully chosen.
 319. Writing assignments were interesting and stimulating.
 320. Writing assignments made students think.
 321. Directions for writing assignments were clear and specific.
 322. Writing assignments required a reasonable amount of time and effort.
 323. Writing assignments were relevant to what was presented in class.
 324. Writing assignments were graded fairly.
 325. Writing assignments were returned promptly.
 257. Writing assignments encouraged the inclusion of diverse perspectives.

READING ASSIGNMENTS

326. Reading assignments seemed carefully chosen.
 327. Reading assignments were interesting and stimulating.
 328. Reading assignments made students think.
 329. Reading assignments required a reasonable amount of time and effort.
 330. Reading assignments were relevant to what was presented in class.
 258. Reading assignments covered material from diverse perspectives.
 259. The course pack covered material from diverse perspectives.

LABORATORY ASSIGNMENTS

331. The laboratory was a valuable part of this course.
 332. Laboratory assignments seemed carefully chosen.
 333. Laboratory assignments were interesting and stimulating.
 334. Laboratory assignments made students think.
 335. Directions for laboratory assignments were clear and specific.
 336. Laboratory assignments required a reasonable amount of time and effort.
 337. Laboratory assignments were relevant to what was presented in class.
 338. Laboratory reports were graded fairly.
 339. Laboratory reports were returned promptly.

OTHER ASSIGNMENTS

260. Group assignments helped me to learn the material.
 261. The term project was very useful in learning the material.

TEXTBOOK

340. The textbook made a valuable contribution to the course.
 341. The textbook was easy to read and understand.
 342. The textbook presented various sides of issues.
 343. A textbook would be a useful addition to this course.

AUDIOVISUAL MATERIALS

344. Films were a valuable part of this *course*.
 345. Audio materials were a valuable part of this *course*.
 346. Films used in this *course* were a great help to learning.
 347. Multimedia materials were a valuable part of this *course*.
 348. Audiovisual materials were a valuable part of this *course*.
 349. Videotapes used in this *course* were a great help to learning.
 350. Slides/overheads were a valuable part of this *course*.

INSTRUCTIONAL COMPUTING

351. Electronic presentations were a valuable part of this *course*.
 352. E-mail discussions were a valuable part of this *course*.
 353. Use of the World Wide Web was a valuable part of this *course*.
 354. Computer labs were a valuable part of this *course*.
 355. Computer tutorials were a valuable part of this *course*.

EXAMS

356. Examinations covered the important aspects of the course.
 357. The exams covered the reading assignments well.
 358. The exams covered the lecture material well.
 359. Exams were creative and required original thought.
 360. Exams were reasonable in length and difficulty.
 361. Examination items were clearly worded.
 362. The exams were returned in a reasonable amount of time.
 363. The examinations were graded very carefully and fairly.
 364. The test items were adequately explained after a test was given.

GRADING

365. Grades were assigned fairly and impartially.
 366. The grading system was clearly explained.
 367. The *instructor* had a realistic definition of good performance.

STUDENT RESPONSIBILITY

369. I tried to relate what I learned in this *course* to my own experience.
 370. I attended class regularly.
 371. I utilized all the learning opportunities provided in this course.
 372. I created my own learning experiences in connection with the course.
 373. I helped classmates learn.

Ερωτήσεις Ανοικτού – Κλειστού τύπου

Open-ended Questions

These questions ask students to write short answers. You may select up to five questions from the group below for inclusion on your questionnaire. Indicate your choices by circling the numbers that appear before the questions.

900. Comment on the quality of instruction in this *course*.
901. How can the *instructor* improve the teaching of this *course*?
902. Which aspects of this *course* did you like best?
903. Which aspects of this *course* did you like least?
904. What changes would you make in the lectures?
905. What changes would you make in the readings?
906. What changes would you make in the examinations?
907. How would you change this *course*?
908. Which aspects of this *course* were most valuable?
909. Which aspects of this *course* were least valuable?
910. How might the class climate be made more inclusive of diverse students?
911. How might the course content be more inclusive of diverse groups?
912. How might the course materials be more inclusive of diverse groups?
913. How might the teaching methods used be more sensitive to diverse needs?
914. How might working in groups be made more inclusive for diverse students?

B 2: Κλίμακα Αξιολόγησης από Ομοβάθμιο (peer), Πανεπιστήμιο Τζον Χόπκινς, Νοσηλευτική Σχολή

**JOHNS HOPKINS UNIVERSITY SCHOOL OF NURSING
PEER EVALUATION SCALE**

Instructor: _____
 Observer: _____
 Course/Room No.: _____
 Class Topic: _____
 Date: _____

PURPOSE: This scale is designed as an observation tool to rate an individual instructor's teaching performance. It is intended to provide a diagnostic profile for teaching improvement.

DIRECTIONS: Using the anchors below, check (✓) your rating for each teaching behavior that's applicable for the specific class observed. Check "NA" for items that do not apply.

- E = Excellent
- VG = Very Good
- G = Good
- NI = Needs Improvement
- NA = Not Applicable

	E	VG	G	NI	NA	Comments:
Content and Organization						
Started and ended class on time						
Presented overview of class content/objectives						
Presented rationale for topics covered						
Presented key concepts						
Presented current material						
Presented information in an organized manner						
Demonstrated accurate knowledge of content						
Used relevant examples to explain major ideas						
Used alternative explanations when necessary						
Made efficient use of class time						
Covered class content/objectives						
Communication Style						
Varied pace appropriately						
Enunciated clearly						
Varied modulation						
Varied tone						
Spoke with adequate volume						
Demonstrated confidence						
Demonstrated enthusiasm						
Moved easily about room during presentation						
Used speech fillers (um, ok, ah) rarely						
Established and maintained eye contact						
Maintained students' attention						
Questioning Skills						
Encouraged students' questions						
Listened carefully to students' questions						
Answered questions appropriately						
Restated students' questions or comments as necessary						

	E	VG	G	NI	NA	Comments:
Critical Thinking Skills						
Asked probing questions						
Used case studies or scenarios						
Used small-group discussion						
Encouraged students to answer difficult questions by providing cues or rephrasing						
Rapport with Students						
Greeted students at the beginning of class						
Responded appropriately to students' puzzlement or Boredom						
Asked students to clarify questions, when necessary						
Requested very difficult, time-consuming, or irrelevant questions be addressed at a later time						
Used humor and/or anecdotes appropriately						
Demonstrated respect for students and their thoughts/Concerns						
Learning Environment						
Physical characteristics (temperature, lighting, crowding, seating)						
Class Affect						
Conducive to learning						
Relaxed						
Controlled						
Teaching Methods						
Lecture						
Engagement techniques						
Q&A						
Discussion						
Small-group activities						
Student individual/panel presentations						
Active learning (e.g., think-pair-share)						
One-minute paper						
Other						
Role playing						
Demonstrations/skits						
Simulations						
Games						
Use or integration of technology						
Overheads						
PowerPoint						
Slides						
PC						
CD-ROM						
Course Web site						
Internet						
Videos						
Audiotapes						
Other						
Experimental/Innovative techniques						

B 3: Κλίμακα Αυτό-αξιολόγησης (self-evaluation), Πανεπιστήμιο Τζον Χόπκινς, Νοσηλευτική Σχολή

**JOHNS HOPKINS UNIVERSITY SCHOOL OF NURSING
CLINICAL COURSE FACULTY PERFORMANCE APPRAISAL FORM**

Course Coordinator: _____

Clinical Faculty: _____

Course Title (No.): _____

Date: _____

PURPOSE: This scale is designed as an evaluation tool to rate each clinical faculty member. It is intended to provide a diagnostic profile for teaching improvement.

PROCEDURES:

Course coordinator will visit new clinical faculty at least once during the semester to evaluate clinical instruction methods. Additional visits may occur as necessary. Following the initial evaluation, course coordinators will evaluate faculty at least bi-annually.

Clinical faculty member will complete a self-evaluation using this tool, including the two open-ended responses.

A *meeting* will be scheduled between the course coordinator and the faculty member to discuss the results following the clinical site visit. The meeting should occur as soon as possible after the site visit. The self-evaluation form will be used in conjunction with the course coordinator's appraisal to analyze faculty performance and suggest areas for improvement.

DIRECTIONS: Check (✓) your rating of AGREE (A) or DISAGREE (DA) for each of the teaching behaviors listed below. For each DA response, please explain in the comments section to the right or on the back of this form if more space is needed. For any behaviors NOT APPLICABLE (NA) or NOT OBSERVED (NOB), check as appropriate.

	A	DA	NA	NOB	Comments:
TEACHING METHODS					
Demonstrates:					
Professionalism					
Enthusiasm					
Respect for students, staff, and patients					
Current clinical knowledge					
Thorough preparation for clinical experience					
Sensitivity to gender and cultural differences					
Effective communication skills (e.g., constructive, nonthreatening feedback, appropriate verbal and nonverbal responses)					
Assures safe application of clinical care					
Assists students to apply theory to practice					
Assures clinical experiences are appropriate for course level					
Facilitates critical thinking skills					
Provides timely verbal feedback	✓				

	A	DA	NA	NOB	Comments:
Provides constructive verbal feedback					
Organizes clinical experiences to maximize learning					
Varies teaching strategies according to student characteristics and abilities					
Provides insightful written feedback to students					
Provides timely written feedback to students (e.g., nursing care plan comments, e-mail, memos)					
Grading clearly discriminates among different levels of performance					
COURSE MANAGEMENT					
Follows clinical course guidelines					
Adheres to established student clinical hours					
Attends clinical course-related meetings					
Develops a written, systematic plan for student improvement, as appropriate					
Provides course coordinator with appropriate feedback on selected students					
Provides course coordinator with timely feedback on selected students					
Seeks assistance from course coordinator as appropriate					
Assists with additional course activities as necessary (e.g., developing care plan guidelines, revising student evaluation tool)					
Submits students' grades promptly					

Strengths:

Areas for Improvement:

Signature _____

B 4: Κλίμακα Αποτελεσματικότητας της σχέσης με τον Μέντορα

MENTORSHIP EFFECTIVENESS SCALE

Your name: _____

Directions: The purpose of this scale is to evaluate the mentoring characteristics of _____, who has identified you as an individual with whom he/she has had a professional, mentor/mentee relationship. Indicate the extent to which you agree or disagree with each statement listed below. Circle the letters that correspond to your response. Your responses will be kept confidential.

SD = Strongly Disagree
D = Disagree
SID = Slightly Disagree
SIA = Slightly Agree
A = Agree
SA = Strongly Agree
NA = Not Applicable

	SD	D	SID	SIA	A	SA	NA
SAMPLE: My mentor was hilarious.						SA	
1. My mentor was accessible.	SD	D	SID	SIA	A	SA	NA
2. My mentor demonstrated professional integrity.	SD	D	SID	SIA	A	SA	NA
3. My mentor demonstrated content expertise in my area of need.	SD	D	SID	SIA	A	SA	NA
4. My mentor was approachable.	SD	D	SID	SIA	A	SA	NA
5. My mentor was supportive and encouraging.	SD	D	SID	SIA	A	SA	NA
6. My mentor provided constructive and useful critiques of my work.	SD	D	SID	SIA	A	SA	NA
7. My mentor motivated me to improve my work product.	SD	D	SID	SIA	A	SA	NA
8. My mentor was helpful in providing direction and guidance on professional issues. (e.g., networking).	SD	D	SID	SIA	A	SA	NA
9. My mentor answered my questions satisfactorily (e.g., timely response, clear, comprehensive).	SD	D	SID	SIA	A	SA	NA
10. My mentor acknowledged my contributions appropriately (e.g., committee contributions, awards).	SD	D	SID	SIA	A	SA	NA
11. My mentor suggested appropriate resources (e.g., experts, electronic contacts, source materials).	SD	D	SID	SIA	A	SA	NA
12. My mentor challenged me to extend my abilities (e.g., risk taking, try a new professional activity, draft a section of an article).	SD	D	SID	SIA	A	SA	NA

Please make additional comments on the back of this sheet. Copyright© 2002 The Johns Hopkins University School of Nursing

ΠΑΡΑΡΤΗΜΑ Γ

Δείγματα Κλιμάκων Αξιολόγησης του μοντέλου 360°

- Γ 1: Πολλαπλών-Πηγών Ανάδραση: 360° Αξιολόγηση της Συμπεριφοράς της Ομάδας (Team Assessment of Behavior, TAB)**
- Γ 2: Αξιολόγηση της Συμπεριφοράς της Ομάδας: Συνοπτική παρουσίαση για τη Σύσκεψη Αξιολόγησης**
- Γ 3: Πολλαπλών-Πηγών Ανάδραση: 360° Αξιολόγηση των Επαγγελματικών Δεξιοτήτων Διευθυντή Κλινικής**
- Γ 4: Μικρή φόρμα εργαλείου Αξιολόγησης από Ομοβάθμιο (mini-Peer Assessment Tool, mini-PAT)**
- Γ 5: Μικρή φόρμα εργαλείου Αυτό-αξιολόγησης (Self mini-Pat)**

Γ 1: Πολλαπλών-Πηγών Ανάδραση: 360° Αξιολόγηση της Συμπεριφοράς της Ομάδας (Team Assessment of Behavior, TAB)

MULTI-SOURCE FEEDBACK: 360° Team Assessment of Behaviour (TAB)*

Trainee's Name: _____ GMC No: _____ Current post: _____ Date started present post: _____

Please use the comments boxes to commend good behaviour and to describe any behaviour causing you concern. Give specific examples. This form will be sent to the trainee's educational supervisor, who may ask you privately to enlarge on any concern behaviour you report. At least 9 other forms will also be considered. The trainee will receive private feedback, but you will not be identified in person without advance discussion with you.

ATTITUDE AND/OR BEHAVIOUR	No concern	You have some concern	You have a major concern	COMMENTS: Anything especially good? If you cannot give an opinion due to lack of knowledge of the trainee say so here. You must specifically comment on any concern behaviour, and this should reflect the trainee's behaviour over time – not usually just a single incident.
Maintaining trust / Professional relationship with patients Listens. Is polite and caring. Shows respect for patients' opinions, privacy, dignity and confidentiality. Is unprejudiced.				
Verbal communication skills Gives understandable information. Speaks good English, at the appropriate level for the patient.				
Team-working / Working with colleagues Respects others' roles, and works constructively in the team. Hands over effectively, and communicates well. Is unprejudiced, supportive and fair.				
Accessibility Accessible. Takes proper responsibility. Only delegates appropriately. Does not shirk duty. Responds when called. Arranges cover for absence.				
Name of assessor:	Post/designation:			Signature:
				Date:

Please send the completed form, straight away, in a sealed envelope, to the postgraduate centre manager
 *Reprinted by permission of Dr. David Wall (9/05). Whitehouse, A., Hassell, A., Wood, L., Wall, D., Walzman, M., & Campbell, I. (2005). Development and reliability testing of TAB—A form for 360° assessment of senior house officers' professional behaviour, as specified by the General Medical Council. *Medical Teacher*, 27, 252–258. (Also see www.wmdeanery.org/.)

Γ 2: Αξιολόγηση της Συμπεριφοράς της Ομάδας: Συνοπτική παρουσίαση για τη Σύσκεψη Αξιολόγησης

Team Assessment of Behaviour (TAB) : SUMMARY FOR APPRAISAL MEETING

Name of Trainee _____ GMC No _____ Present post No: _____ Present speciality /attachment: _____

ATTITUDE AND/OR BEHAVIOUR	How many 'No concern'	How many 'Some concern'	How many 'Major concern'	How many 'No response'	SUMMARY OF COMMENTS
Maintaining trust / Professional relationship with patients Listens. Is polite and caring. Shows respect for patients opinions, privacy, dignity and confidentiality. Is unprejudiced.					
Verbal communication skills Gives understandable information. Speaks good English, at the appropriate level for the patient.					
Team-working / Working with colleagues Respects others' roles, and works constructively in the team. Hands over effectively, and communicates well. Is unprejudiced, supportive and fair.					
Accessibility Accessible. Takes proper responsibility. Only delegates appropriately. Does not shirk duty. Responds when called. Arranges cover for absence.					
Totals					
Name of ES: Date of appraisal meeting:					Signature: NB: This summary, and any other evidence of performance, must be made available to the trainee at the appraisal meeting.

Γ 3: Πολλαπλών-Πηγών Ανάδραση: 360° Αξιολόγηση των Επαγγελματιών Δεξιοτήτων Διευθυντή Κλινικής

Multi-Source Feedback: 360° Assessment of Professional Skills of Clinical Director*

Clinical Director's name: _____

Please use this form to commend good skills or to describe issues that cause you concern. Give specific examples. The clinical director has nominated you as one of the 10 individuals who have been sent this form within the directorate. Results will be collated by the medical director and no individual will be identified.

SKILLS	UNACCEPTABLE	BORDERLINE	ACCEPTABLE	GOOD	EXCELLENT	COMMENTS Please state anything exceptionally good. You must make specific comment if there is any area of concern.
1. Operational Management Writing and presenting reports, coping with pressure and deadlines, reasoned thinking						
2. Interpersonal Skills Conflict resolution, sensitivity, communication skills, credibility and adaptability						
3. Creative and Strategic Thinking Developing the learning environment, innovative thinking, leadership, flexibility and challenging traditional practice to introduce new initiatives						

Name of assessor (caps)Signature.....Date.....

*Reprinted by permission of Dr. David Wall (9/05). Wall, D., Rayner, H., & Palmer, R. (2005, September). *Multi-source feedback: 360° Assessment of professional skills of clinical directors*. Short communication presented at the annual conference of the Association for Medical Education in Europe, Amsterdam, The Netherlands.

**Γ 4: Μικρή φόρμα εργαλείου Αξιολόγησης από Ομοβάθμιο
(mini-Peer Assessment Tool, mini-PAT)**

Please refer to curriculum at www.mmc.nhs.uk for details of expected competencies for F1 and F2

mini-PAT (Peer Assessment Tool) - F2 Version

Please complete the questions using a cross: ☒ Please use black ink and CAPITAL LETTERS

Doctor's Surname:

Forename:

GMC Number:

How do you rate this Doctor in their:	Below expectations for F2 completion		Borderline for F2 completion	Meets expectations for F2 completion	Above expectations for F2 completion		U/C*
	1	2	3	4	5	6	
Good Clinical Care							
1 Ability to diagnose patient problems	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 Ability to formulate appropriate management plans	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 Awareness of their own limitations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 Ability to respond to psychosocial aspects of illness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 Appropriate utilisation of resources e.g. ordering investigations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Maintaining good medical practice							
6 Ability to manage time effectively / prioritise	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7 Technical skills (appropriate to current practice)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Teaching and Training, Appraising and Assessing							
8 Willingness and effectiveness when teaching/training colleagues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Relationship with Patients							
9 Communication with patients	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10 Communication with carers and/or family	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11 Respect for patients and their right to confidentiality	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Working with colleagues							
12 Verbal communication with colleagues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13 Written communication with colleagues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14 Ability to recognise and value the contribution of others	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15 Accessibility/Reliability	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16 Overall, how do you rate this doctor compared to a doctor ready to complete F2 training?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Do you have any concerns about this doctor's probity or health? Yes No

If yes please state your concerns:

*U/C Please mark this if you have not observed the behaviour and therefore feel unable to comment. 2500288218

Anything especially good?

Please describe any behaviour that has raised concerns or should be a particular focus for development:

Please continue your comments on a separate sheet if required

Your Gender: Male Female

Your ethnic group:

<input type="checkbox"/> British	<input type="checkbox"/> Bangladeshi
<input type="checkbox"/> Irish	<input type="checkbox"/> Other Asian Background
<input type="checkbox"/> Other White Background	<input type="checkbox"/> White and Black Caribbean
<input type="checkbox"/> Caribbean	<input type="checkbox"/> White and Black African
<input type="checkbox"/> African	<input type="checkbox"/> White and Asian
<input type="checkbox"/> Any other Black background	<input type="checkbox"/> Any other mixed background
<input type="checkbox"/> Indian	<input type="checkbox"/> Chinese
<input type="checkbox"/> Pakistani	<input type="checkbox"/> Any other ethnic group

Which environment have you primarily observed the doctor in?
(Please choose one answer only)

<input type="checkbox"/> Inpatients	<input type="checkbox"/> Intensive Care
<input type="checkbox"/> Outpatients	<input type="checkbox"/> Theatre
<input type="checkbox"/> Both In and Out-patients	<input type="checkbox"/> General Practice
<input type="checkbox"/> A&E/Admissions	<input type="checkbox"/> Other (Please specify)
<input type="checkbox"/> Community Speciality	<input style="width: 150px;" type="text"/>
<input type="checkbox"/> Laboratory/Research	

Your position:

<input type="checkbox"/> Consultant	<input type="checkbox"/> SASG	<input type="checkbox"/> SpR	<input type="checkbox"/> Foundation/PRHO
<input type="checkbox"/> Nurse	<input type="checkbox"/> SHO	<input type="checkbox"/> Allied Health Professional	
<input type="checkbox"/> GP			
<input type="checkbox"/> Other (Please specify)	<input style="width: 150px;" type="text"/>		

If you are a Nurse or AHP how long have you been qualified?: years Length of working relationship: months

What training have you had in the use of this assessment tool?: Face-to-Face Have Read Guidelines Web/CD rom

How long has it taken you to complete this form (in minutes)?:

Your Signature: Date: / /

Your Surname:

Your GMC Number: (Doctors only)

Acknowledgements: mini-PAT is derived from SPRAT (Sheffield Peer Review Assessment Tool)

9787288212

Reprinted with permission of the United Kingdom Department of Health, Policy Research Programme (10/05).

Γ 5: Αυτό-αξιολόγησης μικρή φόρμα εργαλείου (Self mini-Pat)

Self mini-PAT (Peer Assessment Tool)

Please complete the questions using a cross: ☒ Please use black ink and CAPITAL LETTERS

Your Surname:

GMC Number: **YOUR GMC NUMBER MUST BE COMPLETED**

How do you rate yourself in your:	Below expectations for F1 completion		Borderline for F1 completion	Meets expectations for F1 completion	Above expectations for F1 completion		U/C*
	1	2	3	4	5	6	
Good Clinical Care							
1 Ability to diagnose patient problems	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 Ability to formulate appropriate management plans	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 Awareness of their own limitations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 Ability to respond to psychosocial aspects of illness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 Appropriate utilisation of resources e.g. ordering investigations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Maintaining good medical practice							
6 Ability to manage time effectively / prioritise	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7 Technical skills (appropriate to current practice)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Teaching and Training, Appraising and Assessing							
8 Willingness and effectiveness when teaching/training colleagues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Relationship with Patients							
9 Communication with patients	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10 Communication with carers and/or family	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11 Respect for patients and their right to confidentiality	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Working with colleagues							
12 Verbal communication with colleagues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13 Written communication with colleagues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14 Ability to recognise and value the contribution of others	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15 Accessibility/Reliability	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16 Overall, how do compare yourself to a doctor ready to complete F2 training?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

*U/C Please mark this if you feel unable to comment.

PTO:

Acknowledgements: mini-PAT is derived from SPRAT (Sheffield Peer Review Assessment Tool)

6830612691

Anything going especially well?	Please describe any areas that you think you should particularly focus on for development:
Your Signature: Date: <input type="text"/> / <input type="text"/> / <input type="text"/>	
8727612698	

Reprinted with permission of the United Kingdom Department of Health, Policy Research Programme (10/05).

ΠΑΡΑΡΤΗΜΑ Δ

- Δ 1:** Ερωτηματολόγια Αξιολόγησης Διδακτικών-Μαθησιακών Διαδικασιών στο Πρόγραμμα Σπουδών του Τμήματος Εκπαίδευσης και Αγωγής στην Προσχολική Ηλικία (ΤΕΑΠΗ-ΕΚΠΑ)
- Δ 2:** Εργαλεία Αξιολόγησης Πεποιθήσεων Αποτελεσματικότητας των Φοιτητών και του Περιβάλλοντος Μάθησης στο μάθημα επιλογής «Οι έννοιες των ΦΕ» του Τμήματος Εκπαίδευσης και Αγωγής στην Προσχολική Ηλικία (ΤΕΑΠΗ-ΕΚΠΑ)

Δ 1: Ερωτηματολόγια Αξιολόγησης Διδακτικών-Μαθησιακών Διαδικασιών στο Πρόγραμμα Σπουδών του Τμήματος Εκπαίδευσης και Αγωγής στην Προσχολική Ηλικία (ΤΕΑΠΗ-ΕΚΠΑ)

Τα ερωτηματολόγια που ακολουθούν έχουν χρησιμοποιηθεί στο ΤΕΑΠΗ σε διάφορα μαθήματά του για να αξιολογήσουν διαμορφωτικά μια σειρά από συνιστώσες της διδακτικής-μαθησιακής διαδικασίας, ανάλογα με τα ενδιαφέροντα των διδασκόντων, τη δομή των μαθημάτων τους και τις ανάγκες της αξιολόγησης των προπτυχιακών ή και των μεταπτυχιακών Προγραμμάτων Σπουδών του Τμήματος.

Η γενική υπόθεση που στηρίζει τα ερωτηματολόγια αυτά συγκροτείται από τις παρακάτω προτάσεις:

1. Διδάσκοντες και εκπαιδευτικοί αποτελούν ενεργά και δρώντα υποκείμενα των εκπαιδευτικών δομών του Τμήματος (Sewell 1992, Τσελφές & Παρούση 2010).
2. Η απόκτηση γνώσεων, δεξιοτήτων και ικανοτήτων αποτελεί αποτέλεσμα ενέργειας-δράσης των υποκειμένων (των φοιτητών και φοιτητριών) και με κανένα τρόπο αποτέλεσμα «μεταφοράς» των γνώσεων, δεξιοτήτων ή ικανοτήτων από τον διδάσκοντα στους διδασκόμενους (όλες οι γνωστικές θεωρίες από τον Piaget και μετά).
3. Η απόφαση των υποκειμένων της γνώσης να εμπλακούν με συγκεκριμένο τρόπο με κάθε μάθημα που επιλέγουν ή υποχρεώνονται να παρακολουθήσουν εξαρτάται τουλάχιστον από (π.χ. Ajzen 1988, Ajzen & Fishbein 1980):
 - a. τη στάση (πεποιθήσεις περί προσωπικού κέρδους και ζημιάς) που διαμορφώνουν απέναντι στο γεγονός της συγκεκριμένης εμπλοκής,
 - b. κανονιστικές πεποιθήσεις που διαμορφώνουν σημαντικοί για τα υποκείμενα «τρίτοι»,
 - c. πεποιθήσεις εκτίμησης ελέγχου της εμπλοκής,
 - d. εκτιμήσεις για τις «αντικειμενικές» συνθήκες της εμπλοκής

Με βάση τα παραπάνω τα ερωτηματολόγια μπορούν να παρέχουν πληροφορίες στο διδάσκοντα, με βάση τις οποίες μπορεί να διαμορφώσει διάφορες συνιστώσες του μαθήματός του με τρόπο τέτοιο που να ενισχύσει την ενεργό εμπλοκή των φοιτητών του.

Τα ερωτηματολόγια περιλαμβάνουν κλειστές και ανοιχτές ερωτήσεις. Η ανάλυση των απαντήσεων στις κλειστές (με γραμμική παλινδρόμηση από την αρχή) υποδεικνύει τις ισχυρές και τις ασθενείς συνιστώσες του μαθήματος και η κατηγοριοποίηση των απαντήσεων στις ανοιχτές δικαιολογεί την ισχύ των συνιστωσών.

Τα ερωτηματολόγια πρέπει να προσαρμόζονται κάθε φορά στα ιδιαίτερα χαρακτηριστικά του μαθήματος και τα ερωτήματα που ενδιαφέρουν τον διδάσκοντα. Τα παραδείγματα που επισυνάπτω έχουν «τρέξει» σε διάφορα μαθήματα με τη βοήθεια συναδέλφων που ενδιαφέρθηκαν. Τα δεδομένα τους έχουν αναλυθεί με τη βοήθειά μου και κάποια αποτελέσματα έχουν συγκροτήσει το περιεχόμενο δημοσιεύσεων.

Βασίλης Τσελφές

ΠΑΡΑΔΕΙΓΜΑΤΑ ΕΡΩΤΗΜΑΤΟΛΟΓΙΩΝ

ΠΑΡΑΔΕΙΓΜΑ 1.

Το ερωτηματολόγιο αφορά εργαστηριακό μάθημα που έχει στόχους σχετικούς με την ανάπτυξη ικανοτήτων. Συμπληρώνεται κατά τη διάρκεια του τελευταίου μαθήματος, μετά την παρουσίαση/ κατάθεση της τελικής υποχρεωτικής εργασίας και πριν τις τελικές εξετάσεις.

ΔΕΛΤΙΟ ΑΞΙΟΛΟΓΗΣΗΣ

Η συμπλήρωση των δελτίων αξιολόγησης είναι ανώνυμη. **Στοχεύει στην αποτίμηση των αποτελεσμάτων του μαθήματος Οι απαντήσεις σας μας ενδιαφέρουν για να βελτιώσουμε το μάθημα και δεν έχουν καμία σχέση με το βαθμό σας.** Η συνεισφορά σας, μας είναι πολύτιμη!!!

Έχοντας στο μυαλό σας την παιδαγωγική δραστηριότητα (ή δραστηριότητες) που σχεδιάσατε και παρουσιάζετε στην εργασία σας, σας παρακαλούμε να απαντήσετε στις παρακάτω ερωτήσεις.

Υποθέστε ότι την πρώτη χρονιά που θα δουλέψετε σε νηπιαγωγείο σκοπεύετε να πραγματοποιήσετε την παιδαγωγική δραστηριότητα (ή δραστηριότητες) που σχεδιάσατε και παρουσιάζετε στην εργασία σας.

1. Τι νομίζετε ότι θα κερδίσετε, εσείς προσωπικά, από την πραγματοποίηση αυτής της δραστηριότητας;

.....
.....

2. Τι νομίζετε ότι θα χάσετε, εσείς προσωπικά, από την πραγματοποίηση αυτής της δραστηριότητας;

.....
.....

3. Τι νομίζετε ότι θα κερδίσουν τα παιδιά από την πραγματοποίηση αυτής της δραστηριότητας;

.....
.....

4. Ποιοι σημαντικοί για σας άνθρωποι νομίζετε ότι θα σας επαινέσουν αν πραγματοποιήσετε αυτή τη δραστηριότητα;

.....
.....

5. Ποιοι σημαντικοί για σας άνθρωποι νομίζετε ότι θα σας επικρίνουν / κοροϊδέψουν αν πραγματοποιήσετε αυτή τη δραστηριότητα;

.....
.....

6. Ποιες προσωπικές σας ικανότητες νομίζετε ότι θα σας βοηθήσουν να πραγματοποιήσετε με επιτυχία αυτή τη δραστηριότητα;

.....

 7. Ποιες προσωπικές σας αδυναμίες νομίζετε ότι θα σας δυσκολέψουν να πραγματοποιήσετε με επιτυχία αυτή τη δραστηριότητα;

Έχοντας στο μυαλό σας την παιδαγωγική δραστηριότητα (ή δραστηριότητες) που σχεδιάσατε και παρουσιάζετε στην εργασία σας, σας παρακαλούμε να απαντήσετε στις παρακάτω ερωτήσεις, κυκλώνοντας τον κατάλληλο αριθμό από 5 – 1, με το 5 να αντιστοιχεί στις εξαιρετικά θετικές ή πιθανές απαντήσεις και το 1 στις ελάχιστα θετικές ή πιθανές.

8. Πόσο πιθανό θεωρείτε να πραγματοποιήσετε την παιδαγωγική δραστηριότητα που σχεδιάσατε στην εργασία σας με τα παιδιά, την πρώτη χρονιά που θα δουλέψετε σε νηπιαγωγείο;

5 4 3 2 1

9. Πόσο σημαντικό νομίζετε ότι θα είναι το προσωπικό σας κέρδος από την εμπλοκή σας με την πραγματοποίηση της δραστηριότητας αυτής;

5 4 3 2 1

10. Πόσο σημαντικές νομίζετε ότι θα είναι οι προσωπικές σας απώλειες / ζημιές από την εμπλοκή σας με την πραγματοποίηση της δραστηριότητας αυτής;

5 4 3 2 1

11. Πόσο σημαντικό παιδαγωγικό αποτέλεσμα νομίζετε ότι θα έχει η πραγματοποίηση αυτής της δραστηριότητας;

5 4 3 2 1

12. Πόσο πιθανό θεωρείτε οι σημαντικοί για σας άνθρωποι να σας επαινέσουν αν πραγματοποιήσετε αυτή τη δραστηριότητα;

5 4 3 2 1

13. Πόσο πιθανό θεωρείτε οι σημαντικοί για σας άνθρωποι να σας επικρίνουν / κοροϊδέψουν αν πραγματοποιήσετε αυτή τη δραστηριότητα;

5 4 3 2 1

14. Πόσο ικανές νομίζετε ότι είστε για να πετύχετε την πραγματοποίηση αυτής της δραστηριότητας;

5 4 3 2 1

Ευχαριστούμε πολύ!

ΠΑΡΑΔΕΙΓΜΑ 2.

Το ερωτηματολόγιο αφορά επιλεγόμενο εργαστηριακό, σεμιναριακό ή θεωρητικό μάθημα που έχει στόχους σχετικούς, είτε με τη λειτουργική κατανόηση κάποιων θεωριών είτε με την ανάπτυξη ικανοτήτων. Συμπληρώνεται πριν από την έναρξη του εξαμήνου στο οποίο θα διδαχθεί από ένα κατά το δυνατόν τυχαίο δείγμα φοιτητών του κατάλληλου εξαμήνου σπουδών (συνήθως στο πλαίσιο ενός άλλου υποχρεωτικού μαθήματος του προηγούμενου εξαμήνου).

ΔΕΛΤΙΟ ΑΞΙΟΛΟΓΗΣΗΣ

Η συμπλήρωση των δελτίων αξιολόγησης είναι ανώνυμη. **Στοχεύει στην εκτίμηση συμμετοχής στο επιλεγόμενο μάθημα**, το οποίο θα διδαχθεί στο επόμενο εξάμηνο. **Οι απαντήσεις σας μας ενδιαφέρουν για να προσαρμόσουμε τη δομή του μαθήματος στις τρέχουσες ανάγκες και δεν έχουν καμία σχέση με το βαθμό σας.** Η συνεισφορά σας, μας είναι πολύτιμη!!!

Στο επόμενο εξάμηνο θα προσφερθεί στον κύκλο σπουδών..... το μάθημα..... με διδάσκοντα τον..... Η περιγραφή του μαθήματος όπως αυτή παρουσιάζεται στον οδηγό σπουδών είναι η ακόλουθη:

1. Τι νομίζετε ότι θα κερδίσετε, εσείς προσωπικά, αν επιλέξετε να παρακολουθήσετε αυτό το μάθημα;

2. Τι νομίζετε ότι θα χάσετε, εσείς προσωπικά, αν επιλέξετε να παρακολουθήσετε αυτό το μάθημα;

3. Ποιοι σημαντικοί για σας άνθρωποι νομίζετε ότι θα σας επαινέσουν αν επιλέξετε να παρακολουθήσετε αυτό το μάθημα;

4. Ποιοι σημαντικοί για σας άνθρωποι νομίζετε ότι θα σας επικρίνουν / κοροϊδέψουν αν επιλέξετε να παρακολουθήσετε αυτό το μάθημα;

5. Ποιες προσωπικές σας ικανότητες νομίζετε ότι θα σας βοηθήσουν να παρακολουθήσετε με επιτυχία αυτό το μάθημα;

6. Ποιες προσωπικές σας αδυναμίες νομίζετε ότι θα σας δυσκολέψουν να παρακολουθήσετε με επιτυχία αυτό το μάθημα;

.....

 7. Ποιες αντικειμενικές δυσκολίες εκτιμάτε ότι θα αντιμετωπίσετε αν επιλέξετε να παρακολουθήσετε αυτό το μάθημα;

.....

 Έχοντας στο μυαλό σας το ενδεχόμενο να επιλέξετε να παρακολουθήσετε το μάθημα που ήδη αναφέραμε, σας παρακαλούμε να απαντήσετε στις παρακάτω ερωτήσεις, κυκλώνοντας τον κατάλληλο αριθμό από 5 – 1, με το 5 να αντιστοιχεί στις εξαιρετικά θετικές ή πιθανές απαντήσεις και το 1 στις ελάχιστα θετικές ή πιθανές.

8. Πόσο πιθανό θεωρείτε να επιλέξετε να παρακολουθήσετε αυτό το μάθημα;
 5 4 3 2 1

9. Πόσο σημαντικό νομίζετε ότι θα είναι το προσωπικό σας κέρδος από την εμπλοκή σας με την ενεργή παρακολούθηση αυτού του μαθήματος;
 5 4 3 2 1

10. Πόσο σημαντικές νομίζετε ότι θα είναι οι προσωπικές σας απώλειες / ζημιές από την εμπλοκή σας με την ενεργή παρακολούθηση αυτού του μαθήματος;
 5 4 3 2 1

11. Πόσο πιθανό θεωρείτε οι σημαντικοί για σας άνθρωποι να σας επαινέσουν αν επιλέξετε να παρακολουθήσετε αυτό το μάθημα;
 5 4 3 2 1

13. Πόσο πιθανό θεωρείτε οι σημαντικοί για σας άνθρωποι να σας επικρίνουν / κοροϊδέψουν αν επιλέξετε να παρακολουθήσετε αυτό το μάθημα;
 5 4 3 2 1

14. Πόσο ικανές νομίζετε ότι είστε για να παρακολουθήσετε με επιτυχία αυτό το μάθημα;
 5 4 3 2 1

Ευχαριστούμε πολύ!

ΠΑΡΑΔΕΙΓΜΑ 3.

Το ερωτηματολόγιο αφορά υποχρεωτικό ή επιλεγόμενο εργαστηριακό, σεμιναριακό ή θεωρητικό μάθημα που έχει στόχους σχετικούς, είτε με τη λειτουργική κατανόηση κάποιων θεωριών είτε με την ανάπτυξη ικανοτήτων. Συμπληρώνεται όσο γίνεται πιο αργά (μετά την τρίτη ή τέταρτη βδομάδα μαθημάτων) και πάντως πριν από την υποβολή των δηλώσεων. Η συμπλήρωση γίνεται από τις φοιτήτριες στο πλαίσιο της διδασκαλίας του μαθήματος.

ΔΕΛΤΙΟ ΑΞΙΟΛΟΓΗΣΗΣ

Η συμπλήρωση των δελτίων αξιολόγησης είναι ανώνυμη. **Στοχεύει στη βελτίωση της πορείας του μαθήματος** **Οι απαντήσεις σας μας ενδιαφέρουν για να βελτιώσουμε το μάθημα και δεν έχουν καμία σχέση με το βαθμό σας.** Η συνεισφορά σας, μας είναι πολύτιμη!!!

Ένα οποιοδήποτε μάθημα, υποχρεωτικό ή επιλεγόμενο από αυτά που προσφέρονται από το Τμήμα κατά τη διάρκεια των σπουδών σας είναι δυνατόν να το παρακολουθήσετε συμμετέχοντας ενεργά σε όλη τη διάρκεια του εξαμήνου είτε να το «ξεπετάξετε» συμμετέχοντας κυρίως στις τελικές εξετάσεις.

Το μάθημα που αυτή τη στιγμή παρακολουθείτε είναι το με διδάσκοντα τον και η περιγραφή του, όπως αυτή παρουσιάζεται στον οδηγό σπουδών, είναι η ακόλουθη:

.....

1. Τι νομίζετε ότι θα κερδίσετε, εσείς προσωπικά, αν επιλέξετε να συμμετέχετε ενεργά στην παρακολούθηση αυτού του μαθήματος σε όλη τη διάρκεια του εξαμήνου;

.....

2. Τι νομίζετε ότι θα χάσετε, εσείς προσωπικά, αν επιλέξετε να συμμετέχετε ενεργά στην παρακολούθηση αυτού του μαθήματος σε όλη τη διάρκεια του εξαμήνου;

.....

3. Ποιοι σημαντικοί για σας άνθρωποι νομίζετε ότι θα σας επαινέσουν, αν επιλέξετε να συμμετέχετε ενεργά στην παρακολούθηση αυτού του μαθήματος σε όλη τη διάρκεια του εξαμήνου;

.....

4. Ποιοι σημαντικοί για σας άνθρωποι νομίζετε ότι θα σας επικρίνουν / κοροϊδέψουν, αν επιλέξετε να συμμετέχετε ενεργά στην παρακολούθηση αυτού του μαθήματος σε όλη τη διάρκεια του εξαμήνου;

.....

5. Ποιες προσωπικές σας ικανότητες νομίζετε ότι θα σας βοηθήσουν να συμμετέχετε ενεργά στην παρακολούθηση αυτού του μαθήματος σε όλη τη διάρκεια του εξαμήνου;

.....

6. Ποιες προσωπικές σας αδυναμίες νομίζετε ότι θα σας δυσκολέψουν να συμμετέχετε ενεργά στην παρακολούθηση αυτού του μαθήματος σε όλη τη διάρκεια του εξαμήνου;

.....

7. Ποιες αντικειμενικές δυσκολίες εκτιμάτε ότι θα αντιμετωπίσετε αν επιλέξετε να συμμετέχετε ενεργά στην παρακολούθηση αυτού του μαθήματος σε όλη τη διάρκεια του εξαμήνου;

.....

Έχοντας στο μυαλό σας το ενδεχόμενο να επιλέξετε να παρακολουθήσετε το μάθημα που πραγματοποιείται αυτή τη στιγμή, συμμετέχοντας ενεργά σε όλη τη διάρκεια του εξαμήνου, σας παρακαλούμε να απαντήσετε στις παρακάτω ερωτήσεις, κυκλώνοντας τον κατάλληλο αριθμό από 5 – 1, με το 5 να αντιστοιχεί στις εξαιρετικά θετικές ή πιθανές απαντήσεις και το 1 στις ελάχιστες θετικές ή πιθανές.

8. Πόσο πιθανό θεωρείτε να παρακολουθήσετε αυτό το μάθημα συμμετέχοντας ενεργά σε όλη τη διάρκεια του εξαμήνου;

5 4 3 2 1

9. Πόσο σημαντικό νομίζετε ότι θα είναι το προσωπικό σας κέρδος αν παρακολουθήσετε αυτό το μάθημα συμμετέχοντας ενεργά σε όλη τη διάρκεια του εξαμήνου;

5 4 3 2 1

10. Πόσο σημαντικές νομίζετε ότι θα είναι οι προσωπικές σας απώλειες / ζημιές αν παρακολουθήσετε αυτό το μάθημα συμμετέχοντας ενεργά σε όλη τη διάρκεια του εξαμήνου;

5 4 3 2 1

11. Πόσο πιθανό θεωρείτε οι σημαντικοί για σας άνθρωποι να σας επαινέσουν αν παρακολουθήσετε αυτό το μάθημα συμμετέχοντας ενεργά σε όλη τη διάρκεια του εξαμήνου;

5 4 3 2 1

13. Πόσο πιθανό θεωρείτε οι σημαντικοί για σας άνθρωποι να σας επικρίνουν / κοροϊδέψουν αν παρακολουθήσετε αυτό το μάθημα συμμετέχοντας ενεργά σε όλη τη διάρκεια του εξαμήνου;

5 4 3 2 1

14. Πόσο ικανές νομίζετε ότι είστε για να καταφέρετε να παρακολουθήσετε αυτό το μάθημα συμμετέχοντας ενεργά σε όλη τη διάρκεια του εξαμήνου;

5 4 3 2 1

Ευχαριστούμε πολύ!

Δ 2: Εργαλεία Αξιολόγησης Πεποιθήσεων Αποτελεσματικότητας των φοιτητών και του Μαθησιακού Περιβάλλοντος στο μάθημα επιλογής «Οι έννοιες των ΦΕ» του Τμήματος Εκπαίδευσης και Αγωγής στην Προσχολική Ηλικία (ΤΕΑΠΗ-ΕΚΠΑ)

ΗΜΕΡΟΜΗΝΙΑ:

ΚΩΔΙΚΟΣ
ΤΗΛ/ΝΟ:.....
e-mail:

**«ΠΕΠΟΙΘΗΣΕΙΣ ΩΣ ΠΡΟΣ ΤΗΝ ΑΠΟΤΕΛΕΣΜΑΤΙΚΟΤΗΤΑ ΣΟΥ
 ΣΤΗ ΔΙΕΞΑΓΩΓΗ ΔΡΑΣΤΗΡΙΟΤΗΤΩΝ ΣΧΕΤΙΚΑ ΜΕ ΤΗ ΔΙΔΑΣΚΑΛΙΑ ΤΩΝ
 ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ (ΦΕ) ΣΤΟ ΝΗΠΙΑΓΩΓΕΙΟ ΜΕΛΛΟΝΤΙΚΑ»**

ΦΥΛΟ	Άνδρας <input type="checkbox"/>	Γυναίκα <input type="checkbox"/>	
ΗΛΙΚΙΑ	18-19 <input type="checkbox"/>	20-21 <input type="checkbox"/>	22- <input type="checkbox"/>
Πόσα μαθήματα ΦΕ παρακολούθησα	Υποχρεωτικά <input type="checkbox"/>	Κυμαινόμενα <input type="checkbox"/>	Επιλογής <input type="checkbox"/>
Πόσα εργαστήρια ΦΕ παρακολούθησα	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Εμπειρία στη διδασκαλία ΦΕ	Καμία <input type="checkbox"/>	Εθελοντική <input type="checkbox"/>	Διδασκαλίες 1-2 <input type="checkbox"/>
Προγενέστερη εμπειρία από τις ΦΕ	ΘΕΤΙΚΗ <input type="checkbox"/>	ΑΡΝΗΤΙΚΗ <input type="checkbox"/>	

Οδηγίες: Παρακαλώ, δήλωσε το βαθμό που συμφωνείς ή διαφωνείς με κάθε μία από τις δηλώσεις που ακολουθούν κυκλώνοντας τα κατάλληλα γράμματα που βρίσκονται δεξιά σε κάθε δήλωση.

ΣΑ = συμφωνώ απόλυτα, ΣΣ = σχεδόν συμφωνώ, ΑΝΠ = αναποφάσιτος,					
ΣΔ = σχεδόν διαφωνώ, ΔΑ = διαφωνώ απόλυτα					
1	Όταν το παιδί τα πηγαίνει καλύτερα απ' ότι συνήθως με τις δραστηριότητες ΦΕ, αυτό οφείλεται στο ό, τι ο νηπιαγωγός κατέβαλε παραπάνω προσπάθεια.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ ΔΑ
2	Συνεχώς θα βρίσκω καλύτερους τρόπους για να διεξάγω τις δραστηριότητες ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ ΔΑ
3 R	Ακόμα κι αν θα προσπαθήσω σκληρά, δεν θα διεξάγω τις δραστηριότητες ΦΕ το ίδιο καλά όπως θα διεξάγω εκείνες στο μάθημα της γλώσσας.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ ΔΑ
4	Όταν ο βαθμός κατανόησης των παιδιών στις ΦΕ βελτιώνεται, αυτό τις περισσότερες φορές οφείλεται στο ότι ο νηπιαγωγός τους έχει βρει μία πιο αποτελεσματική προσέγγιση για να διεξάγει τις δραστηριότητες.	ΣΑ	Σ Σ	ΑΝΠ	ΣΔ ΔΑ
5	Γνωρίζω τα βήματα που πρέπει να ακολουθήσω για να διεξάγω αποτελεσματικά τις δραστηριότητες που σχετίζονται με τις έννοιες των ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ ΔΑ
6 R	Δεν θα είμαι πολύ αποτελεσματικός στο να χειρίζομαι τα πειράματα των ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ ΔΑ

7	Εάν οι κατανοήσεις των παιδιών είναι κατώτερες από τις προσδοκώμενες, το πιο πιθανό είναι αυτές να οφείλονται στη μη αποτελεσματική διεξαγωγή των δραστηριοτήτων από τον νηπιαγωγό.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
8 R	Γενικά θα διεξάγω μη αποτελεσματικά τις δραστηριότητες ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
9	Η ανεπάρκεια γνώσεων ενός παιδιού στις ΦΕ μπορεί να ξεπεραστεί με την καλή διεξαγωγή των δραστηριοτήτων από τον νηπιαγωγό.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
10 R	Ο χαμηλός βαθμός κατανόησης των παιδιών δεν μπορεί γενικά να αποδοθεί στο νηπιαγωγό τους.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
11	Όταν ένα παιδί με χαμηλό βαθμό κατανόησης στις ΦΕ προοδεύει, αυτό συνήθως οφείλεται στην επιπλέον προσοχή που δείχνει ο νηπιαγωγός.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
12	Κατανοώ τις έννοιες των ΦΕ αρκετά καλά για να είμαι αποτελεσματικός στη διεξαγωγή των σχετικών δραστηριοτήτων.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
13 R	Αυξανόμενη προσπάθεια από τον νηπιαγωγό κατά τη διεξαγωγή των δραστηριοτήτων ΦΕ επιφέρει μικρή αλλαγή στο βαθμό κατανόησης των παιδιών στο μάθημα.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
14	Ο νηπιαγωγός είναι γενικά υπεύθυνος για το βαθμό που τα παιδιά κατανοούν τις ΦΕ μέσα από τις δραστηριότητες.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
15	Ο βαθμός που τα παιδιά κατανοούν τις ΦΕ είναι άμεσα εξαρτώμενος από την αποτελεσματικότητα του νηπιαγωγού στη διεξαγωγή των δραστηριοτήτων.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
16	Εάν οι γονείς παρατηρούν ότι το παιδί τους δείχνει μεγαλύτερο ενδιαφέρον για τις ΦΕ, αυτό πιθανό να οφείλεται στο νηπιαγωγό τους.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
17 R	Θα το βρίσκω δύσκολο να εξηγήσω στα παιδιά αυτό που συμβαίνει κατά τη διάρκεια ενός πειράματος.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
18	Θα είμαι σε θέση να δίνω απαντήσεις στις ερωτήσεις των παιδιών σχετικά με τις δραστηριότητες ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
19 R	Αναρωτιέμαι αν θα έχω τις αναγκαίες δεξιότητες για να διεξάγω τις δραστηριότητες ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
20 R	Αν υπήρχε η δυνατότητα επιλογής, δεν θα καλούσα τον Σύμβουλο να εκτιμήσει τη διδασκαλία μου στις ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
21 R	Όταν ένα παιδί αντιμετωπίζει δυσκολία να κατανοήσει μια έννοια των ΦΕ, τότε μάλλον δεν θα ξέρω πώς να το βοηθήσω.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
22	Οι ερωτήσεις των παιδιών θα είναι ευπρόσδεκτες όταν διεξάγω τις δραστηριότητες ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ
23 R	Δεν θα ξέρω πώς να ενεργοποιήσω το ενδιαφέρον των παιδιών για τις ΦΕ.	ΣΑ	ΣΣ	ΑΝΠ	ΣΔ	ΔΑ

ΗΜΕΡΟΜΗΝΙΑ:**Τι συμβαίνει στη Τάξη του εργαστηρίου/ μαθήματος**

Παρακαλώ, σημείωσε ένα Χ στο πλαίσιο που αντιστοιχεί σε μία από τις πέντε επιλογές που βρίσκονται δεξιά σε κάθε δήλωση

Σε αυτό το εργαστήριο/μάθημα		ΣΧΕΔΟΝ ΠΟΤΕ	ΣΠΑΝΙΑ	ΜΕΡΙΚΕΣ ΦΟΡΕΣ	ΣΥΧΝΑ	ΣΧΕΔΟΝ ΠΑΝΤΑ
1	μαθαίνω σχετικά με τον κόσμο μέσα κι έξω από τη σχολή.					
2	μαθαίνω ότι οι ΦΕ δεν μπορούν πάντοτε να παρέχουν απαντήσεις σε προβλήματα.					
3	νιώθω ασφαλής να διατυπώνω απορίες σχετικές με αυτά που διδάσκομαι ή με τον τρόπο που τα διδάσκομαι.					
4	βοηθώ τον καθηγητή να σχεδιάσει αυτό που πρόκειται να μάθω.					
5	συζητώ με άλλους φοιτητές σχετικά με το πώς επιλύουν τα προβλήματα.					
6	μαθαίνω ότι η διαδικασία μάθησης σχετίζεται με τις εμπειρίες ή τις ερωτήσεις σχετικά με τον κόσμο μέσα κι έξω από τη σχολή.					
7	μαθαίνω ότι οι επιστημονικές εξηγήσεις αλλάζουν με την πάροδο του χρόνου.					
8	μπορώ να ρωτήσω για το <i>τι</i> ή και το <i>πώς</i> διδάσκομαι.					
9	βοηθώ τον καθηγητή να αποφανθεί με ποιο τρόπο ενδέχεται να μαθαίνω					
10	εξηγώ τις ιδέες μου σε άλλους φοιτητές.					
11	μαθαίνω πώς οι ΦΕ είναι μέρος της ζωής μου.					
12	μαθαίνω ότι οι ΦΕ επηρεάζονται από τις πολιτισμικές αξίες και τις γνώμες των ανθρώπων.					
13	ζητώ διευκρινίσεις για τις δραστηριότητες που είναι μπερδεμένες.					
14	βοηθώ τον καθηγητή να αποφασίσει ποιες δραστηριότητες είναι οι καλύτερες για μένα.					
15	ζητώ από τους άλλους φοιτητές να εξηγούν τις ιδέες τους.					
16	μαθαίνω ενδιαφέροντα πράγματα σχετικά με τον κόσμο μέσα κι έξω από τη σχολή.					
17	μαθαίνω ότι οι ΦΕ συνιστούν ένα τρόπο για να τεθούν ερωτήματα και να αναζητηθούν απαντήσεις.					
18	εκδηλώνω την έγνοια μου για κάθε τι που αφορά στην εκπαίδευσή μου					
19	βοηθώ τον καθηγητή να αποφασίσει τον χρόνο που χρειάζομαι για να ολοκληρώσω μία δραστηριότητα.					
20	οι άλλοι φοιτητές ζητούν να εξηγώ τις ιδέες μου.					