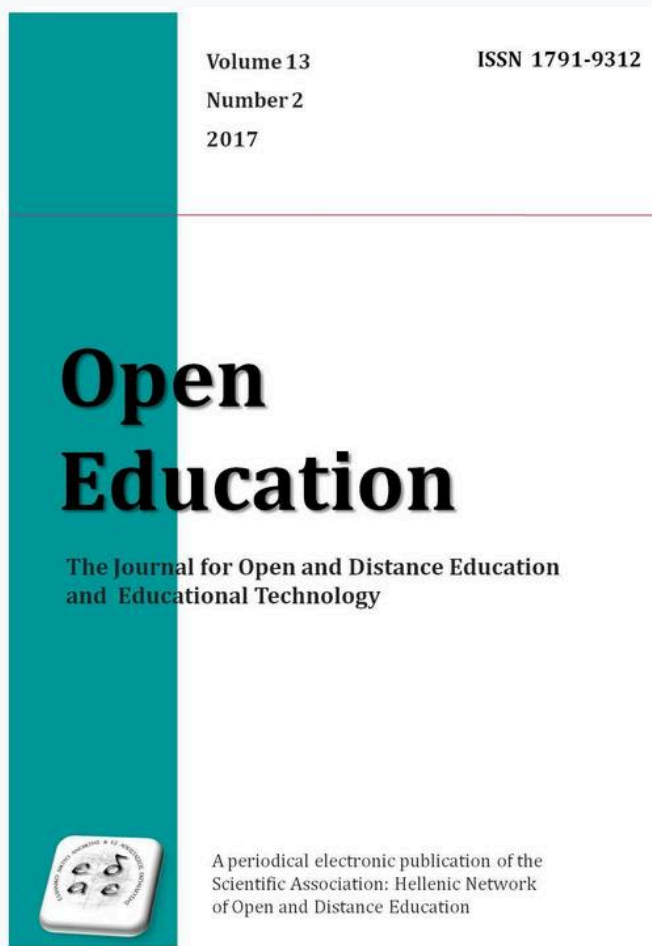


## Ανοικτή Εκπαίδευση: το περιοδικό για την Ανοικτή και εξ Αποστάσεως Εκπαίδευση και την Εκπαιδευτική Τεχνολογία

Τόμ. 13, Αρ. 2 (2017)



### Η ΘΕΩΡΙΑ ΤΗΣ ΑΠΟΚΡΙΣΗΣ ΕΡΩΤΗΜΑΤΟΣ ΓΙΑ ΤΗΝ ΑΞΙΟΛΟΓΗΣΗ ΕΠΙΔΟΣΗΣ ΧΡΗΣΤΩΝ ΕΙΚΟΝΙΚΟΥ ΕΡΓΑΣΤΗΡΙΟΥ ΒΙΟΛΟΓΙΑΣ

*Ευγενία Παξινού, Αργυρώ Σγουρού, Χρήστος Παναγιωτακόπουλος, Βασίλειος Βερούκιος*

doi: [10.12681/jode.14618](https://doi.org/10.12681/jode.14618)

Βιβλιογραφική αναφορά:

## Η θεωρία της απόκρισης ερωτήματος για την αξιολόγηση επίδοσης χρηστών εικονικού εργαστηρίου βιολογίας

**Ευγενία Παξινού**

Σχολή Θετικών Επιστημών και Τεχνολογίας  
Ελληνικό Ανοικτό Πανεπιστήμιο  
[paxinou.evgenia@ac.eap.gr](mailto:paxinou.evgenia@ac.eap.gr)

**Αργυρώ Σγουρού**

Σχολή Θετικών Επιστημών και Τεχνολογίας  
Ελληνικό Ανοικτό Πανεπιστήμιο  
[sgourou@eap.gr](mailto:sgourou@eap.gr)

**Χρήστος Παναγιωτακόπουλος**

Παιδαγωγικό Τμήμα Δημοτικής Εκπαίδευσης  
Πανεπιστήμιο Πατρών  
[cranag@upatras.gr](mailto:cranag@upatras.gr)

**Βασίλειος Βερύκιος**

Σχολή Θετικών Επιστημών και Τεχνολογίας  
Ελληνικό Ανοικτό Πανεπιστήμιο  
[verykios@eap.gr](mailto:verykios@eap.gr)

### Περίληψη

Τα στατιστικά μοντέλα που βασίζονται στη Θεωρία Απόκρισης Ερωτήματος (Item Response Theory-IRT), χρησιμοποιούνται επιτυχώς για τη μέτρηση λανθανουσών μεταβλητών σε ψυχομετρικά τεστ. Τα μοντέλα αυτά διαχωρίζουν τον βαθμό δυσκολίας των ερωτημάτων ενός τεστ από την ικανότητα του υποκειμένου του τεστ. Στην παρούσα μελέτη χρησιμοποιήσαμε ένα IRT μοντέλο, το μοντέλο Rasch, με σκοπό να λάβουμε μία έγκυρη και χρήσιμη αξιολόγηση φοιτητών οι οποίοι έκαναν χρήση ενός τρισδιάστατου εικονικού εργαστηρίου, σαν μέσο εξ αποστάσεως προετοιμασίας τους για το πραγματικό εργαστήριο Βιολογίας. Από την επεξεργασία των απαντήσεων τους στα ερωτήματα των τεστ, μέσω του μοντέλου Rasch, αξιολογήσαμε την ικανότητα του κάθε φοιτητή, η οποία είναι ανεξάρτητη από τον βαθμό δυσκολίας των ερωτημάτων.

### Λέξεις/έννοιες κλειδιά

*Θεωρία απόκρισης ερωτήματος (IRT); Στατιστικό μοντέλο Rasch; Αξιολόγηση επίδοσης φοιτητών; Εικονικό εργαστήριο Βιολογίας; Εξ αποστάσεως εκπαίδευση*

### 1. Εισαγωγή

Στην εξ αποστάσεως εκπαίδευση, έχει μεγάλη σημασία το προσφερόμενο στους φοιτητές εκπαιδευτικό υλικό, αφού το τελευταίο καθορίζει σε μεγάλο βαθμό την ποιότητα της εκπαίδευσής τους (Λιοναράκης, 2001). Στα εξ αποστάσεως

προγράμματα σπουδών Φυσικών Επιστημών η φυσική παρουσία των φοιτητών στα πραγματικά εργαστήρια είναι, λόγω της φύσης των σπουδών τους, χρονικά περιορισμένη, επομένως η ανάπτυξη ενός εξ αποστάσεως εκπαιδευτικού εργαλείου εξάσκησης, προετοιμασίας αλλά και μετάδοσης των εργαστηριακών δεξιοτήτων, είναι αναγκαία.

Αναφορικά με τα εξ αποστάσεως προγράμματα σπουδών στις Φυσικές Επιστήμες, τα διαδικτυακά εικονικά εργαστήρια αποτελούν ό,τι πιο σύγχρονο, αφού δίνουν τη δυνατότητα στους εκπαιδευόμενους να εξοικειωθούν, σε ένα πρώτο επίπεδο και χωρίς χρονικούς ή οικονομικούς περιορισμούς, με τους εργαστηριακούς χώρους, καθώς και με τη χρήση ακριβών και ευαίσθητων εργαστηριακών μηχανημάτων (Makransky, Thisgaard & Gadegaar, 2016). Επιπλέον, όπως όλες οι σύγχρονες εκπαιδευτικές εφαρμογές των ΤΠΕ, τα εικονικά εργαστήρια διαθέτουν γενικά χαρακτηριστικά που μπορούν να υποστηρίξουν την εποικοδομητική μάθηση (Mayer, 1999), ενώ παράλληλα είναι και πολύ αποτελεσματικά στο να εμπλέξουν δυναμικά τους φοιτητές στη μαθησιακή διαδικασία, (Makransky et al., 2016; Shute et al., 2009).

Το Ελληνικό Ανοικτό Πανεπιστήμιο (ΕΑΠ), ένα εκπαιδευτικό ίδρυμα το οποίο εξειδικεύεται στην εξ αποστάσεως εκπαίδευση, παρέχει στους φοιτητές που συμμετέχουν στο πρόγραμμα σπουδών «Σπουδές στις Φυσικές Επιστήμες» μία εφαρμογή τρισδιάστατου εικονικού εργαστηρίου Βιολογίας, την ονομαζόμενη OnLabs, (<https://sites.google.com/site/onlabseap/>) (Zafeiropoulos et al., 2014), Εικόνα 1.



Εικόνα 1: Στιγμιότυπο από το τρισδιάστατο εικονικό εργαστήριο Βιολογίας, OnLabs

Έναν τρόπον, για να διερευνηθεί εάν ένα εικονικό εργαστήριο Βιολογίας σαν το OnLabs, προετοιμάζει πιο αποτελεσματικά τους φοιτητές για τις εργαστηριακές ασκήσεις, από ότι η κλασική εκπαιδευτική μεθοδολογία, αποτελεί η αξιολόγηση μέσω τεστ εννοιολογικής κατανόησης, των επιδόσεων των φοιτητών εκείνων που συμπεριλαμβάνουν στην εκπαίδευσή τους την εμπειρία του εικονικού εργαστηρίου, καθώς και των επιδόσεων των φοιτητών οι οποίοι επιλέγουν μόνο την κλασική εκπαιδευτική μεθοδολογία.

Δύο από τις πιο βασικές μέριμνες των εκπαιδευτικών, όταν πρόκειται να αξιολογήσουν τις επιδόσεις των φοιτητών σε ένα γραπτό τεστ, είναι αρχικά, η σωστή επιλογή των ερωτημάτων και στη συνέχεια η χρήση ενός σωστού μοντέλου

βαθμολόγησης έτσι ώστε οι βαθμολογίες που θα προκύψουν να είναι «ακριβείς» και «χρήσιμες» (Wu, Tam & Jen, 2016).

Με τον όρο «ακριβείς» εννοούμε βαθμολογίες τις οποίες οι εκπαιδευτικοί θα μπορούν να εμπιστευτούν. Έτσι, αν ένας φοιτητής βαθμολογηθεί, και η ικανότητά του σε ένα τεστ αποτιμηθεί ως  $\theta$ , θα πρέπει αυτή η βαθμολογία να είναι μία αξιόπιστη αποτύπωση των γνώσεων του και άρα είναι αναμενόμενο, σε παρόμοιο τεστ ο συγκεκριμένος φοιτητής, να επιδείξει και πάλι τη ίδια ικανότητα  $\theta$ .

Με τον όρο «χρήσιμες» εννοούμε βαθμολογίες οι οποίες έχουν αξία και μπορούν να βοηθήσουν τον εκπαιδευτικό σε κάποιον σκοπό, π.χ. στην ομαδοποίηση των αδύναμων φοιτητών ενός τμήματος προκειμένου να τους προσφερθούν επιπλέον διδακτικές ώρες.

### 1.1. Κλασική Θεωρία Μέτρησης των Τεστ (Classical Test Theory-CTT)

Το πιο συνηθισμένο μέχρι και σήμερα μοντέλο για τη βαθμολόγηση της επίδοσης φοιτητών, είναι αυτό που βασίζεται στην Κλασική Θεωρία Μέτρησης των τεστ (Classical Test Theory-CTT) (Gulliksen, 1950). Η CTT, η οποία λέγεται και πραγματική βαθμολογία, έχει ένα χαρακτηριστικό το οποίο αποτελεί συγχρόνως και μία αδυναμία της. Θεωρεί ότι όλα τα ερωτήματα σε ένα τεστ συνεισφέρουν ισοδύναμα στη συνολική επίδοση του εξεταζόμενου. Έτσι, εφαρμόζοντας την CTT, σε ένα εύκολο τεστ ένας φοιτητής μετρίων δυνατοτήτων πλησιάζει σε επίδοση έναν πολύ καλό φοιτητή ενώ σε ένα δύσκολο τεστ ένας μέτριος φοιτητής πλησιάζει σε επίδοση έναν πολύ αδύναμο φοιτητή. Άρα, με αυτό το μοντέλο βαθμολόγησης τα αποτελέσματα που λαμβάνουμε για την επίδοση των εκπαιδευομένων, δεν είναι χρήσιμα αφού εάν βασιστούμε σε αυτά θα καταλήξουμε σε εσφαλμένα συμπεράσματα, π.χ. για την συνεισφορά ενός εκπαιδευτικού εργαλείου στη διαδικασία της μάθησης. Χρησιμοποιώντας την CTT μπορούμε μεν να κατατάξουμε τους φοιτητές σε 1<sup>ov</sup>, 2<sup>ov</sup>, 3<sup>ov</sup>,... κλπ, αλλά δεν μπορούμε να γνωρίζουμε τη σταθερή μεταξύ τους απόσταση σε μία κλίμακα ικανοτήτων (Wu et al., 2016).

### 1.2. Θεωρία Απόκρισης Ερωτήματος (Item Response Theory-IRT)

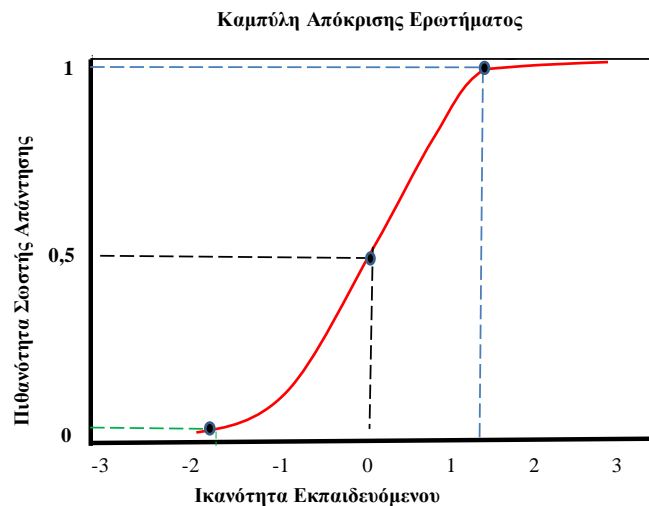
Οι παραπάνω απαιτήσεις, αναφορικά με την αμεροληψία στην τιμή που αντιστοιχεί στην επίδοση ενός φοιτητή, μπορούν να ικανοποιηθούν αν χρησιμοποιηθεί ένα μοντέλο βαθμολόγησης που βασίζεται στη Θεωρία Απόκρισης Ερωτήματος (Item Response Theory-IRT) (Αλεξόπουλος, 2004; Ritchardson, 1936; Tucker, 1946). Σύμφωνα με την IRT, η βαθμολογία ενός εξεταζόμενου δεν εξαρτάται από τις συγκεκριμένες ερωτήσεις που περιέχει ένα τεστ (de Ayala, 2008). Πιο αναλυτικά, παρότι ο εξεταζόμενος μπορεί να αποκτήσει υψηλή βαθμολογία σε ένα εύκολο τεστ και χαμηλή βαθμολογία σε ένα δύσκολο, στον εξεταζόμενο αντιστοιχεί μία ικανότητα, η οποία είναι ανεξάρτητη από το σεν των ερωτημάτων που καλείται να απαντήσει κάθε φορά (Almond et al., 2015). Η ικανότητα αυτή του εξεταζόμενου, στη σύγχρονη ψυχομετρική θεωρία αναφέρεται ως λανθάνουσα μεταβλητή (Thissen & Wainer, 2001) και βασικός στόχος στο χώρο της εκπαίδευσης είναι ο καθορισμός του βαθμού ύπαρξης αυτής της μεταβλητής (Τσιγγίλης, <http://research.edu.uoi.gr/sites/default/files/files/docs/irt.pdf>).

Το πιο απλό IRT μοντέλο είναι το μοντέλο Rasch το οποίο αναπτύχθηκε από τον Δανό Στατιστικό Georg Rasch (Rasch, 1960). Το μοντέλο Rasch, σε αντίθεση με τα άλλα IRT μοντέλα, από τις παραμέτρους ενός ερωτήματος (δυσκολία, διάκριση, παράγοντας «τυχαίας πρόβλεψης»), λαμβάνει υπόψη του μόνο την παράμετρο της δυσκολίας, προκειμένου να υπολογίσει την ικανότητα ενός φοιτητή. Το μοντέλο Rasch είναι δηλαδή μονοπαραμετρικό, γι' αυτό και λέγεται συχνά και 1PL μοντέλο,

(1 Parameter Logistic Model) (Verhelst & Glas, 1995). Επιπλέον είναι ένα διχοτομικό μοντέλο αφού όταν εφαρμόζεται, κάθε λανθασμένη απάντηση βαθμολογείται με 0 ενώ κάθε σωστή με 1. Το μοντέλο Rasch χρησιμοποιεί την ακόλουθη πιθανοτική συνάρτηση:

$$P(X_j|\theta, \beta_j) = \frac{1}{1 + e^{-(\theta - \beta_j)}}$$

όπου  $\beta_j$  είναι η παράμετρος του ερωτήματος που αντιστοιχεί στο βαθμό δυσκολίας του, και  $\theta$  είναι η ικανότητα του φοιτητή να απαντήσει σωστά στο ερώτημα δυσκολίας  $\beta_j$  (Almond et al., 2015). Εξ' ορισμού η δυσκολία  $\beta_j$  ενός ερωτήματος και η ικανότητα  $\theta$  ενός φοιτητή παίρνουν τιμές από την ίδια κλίμακα, από  $-\infty$  έως  $+\infty$ . Από την παραπάνω συνάρτηση, εύκολα προκύπτει ότι ένας φοιτητής που έχει ικανότητα  $\theta$  ίση με την δυσκολία  $\beta_j$  ενός ερωτήματος, έχει 50% πιθανότητα να απαντήσει σωστά στο ερώτημα αυτό. Τι σημαίνει όμως πρακτικά ότι ένας φοιτητής έχει 50% πιθανότητα να απαντήσει σωστά σε μία ερώτηση; Σημαίνει ότι αν στον συγκεκριμένο φοιτητή δώσουμε 100 ερωτήσεις παρόμοιας δυσκολίας  $\beta_j$  θα απαντήσει κατά μέσο όρο σωστά στις 50 από αυτές. Η σχέση μεταξύ της πιθανότητας να δώσει ένας φοιτητής μια σωστή απάντηση σε ένα ερώτημα, και της ικανότητας του φοιτητή, παρουσιάζεται οπτικά σε μία καμπύλη, την Καμπύλη Απόκρισης Ερωτήματος (Item Characteristic Curve-ICC) (Hambleton, Swaminathan & Rogers, 1991), στο Σχήμα 1.



Σχήμα 1: Καμπύλη Απόκρισης Ερωτήματος, ICC

Η καμπύλη του Σχήμα 1 είναι μία καμπύλη η οποία αντιστοιχεί σε ένα ερώτημα ενός τεστ το οποίο θα μπορούσε εύστοχα να χρησιμοποιηθεί προκειμένου να διαχωριστούν για εκπαιδευτικούς σκοπούς, οι φοιτητές ενός τμήματος σε αδύναμους, μέτριους και πολύ ικανούς. Η καμπύλη μας αποκαλύπτει ότι φοιτητές με πολύ μικρή ικανότητα έχουν και πολύ μικρή πιθανότητα να απαντήσουν σωστά στο συγκεκριμένο ερώτημα (πράσινη διακεκομμένη γραμμή), φοιτητές με μέτρια ικανότητα έχουν περίπου 50% πιθανότητα να δώσουν σωστή απάντηση, (μαύρη διακεκομμένη γραμμή), ενώ φοιτητές με μεγάλη ικανότητα έχουν και μεγάλη πιθανότητα να απαντήσουν σωστά (μπλε διακεκομμένη γραμμή). Χρησιμοποιώντας το μοντέλο Rasch για την επεξεργασία των απαντήσεων των εκπαιδευομένων σε ένα τεστ λαμβάνουμε τόσες καμπύλες απόκρισης ερωτήματος όσες και οι ερωτήσεις του



τεστ. Από τις καμπύλες αυτές, μπορούμε να υπολογίσουμε την πιθανότητα ένας φοιτητής να απαντήσει σωστά σε ένα συγκεκριμένο ερώτημα.

## 2. Εκπαιδευτικά σενάρια

Το δείγμα της έρευνας αποτέλεσαν 43 προπτυχιακοί φοιτητές οι οποίοι παρακολούθησαν τον Α΄ κύκλο της Εργαστηριακής Ενότητας της Βιολογίας, στο 3ο έτος του προγράμματος σπουδών «Σπουδές στις Φυσικές Επιστήμες», στη Σχολή Θετικών Επιστημών και Τεχνολογίας του Ελληνικού Ανοικτού Πανεπιστημίου.

### 2.1. Εκπαιδευτικό σενάριο χωρίς τη χρήση εικονικού εργαστηρίου

Ενημερωτικά, οι φοιτητές που εγγράφονται στο συγκεκριμένο πρόγραμμα σπουδών του ΕΑΠ, παρακολουθούν στο 3<sup>ο</sup> ακαδημαϊκό τους έτος, που διαρκεί 10 μήνες, από τον Οκτώβριο μέχρι και τον Ιούλιο, τη Θεματική Ενότητα ΦΥΕ31 με τίτλο «Δομή και Λειτουργία του Κυττάρου». Ταυτόχρονα, εγγράφονται και στην Εργαστηριακή Ενότητα της Βιολογίας και συγκεκριμένα στον Α΄ κύκλο την οποία παρακολουθούν τους μήνες Αύγουστο ή Σεπτέμβριο στα εργαστήρια του ΕΑΠ. Για τη πιο αποτελεσματική διαχείριση του χρόνου αλλά και του χώρου των εργαστηρίων, οι φοιτητές χωρίζονται σε ισομερή τμήματα. Το κάθε τμήμα πραγματοποιεί τις εργαστηριακές του ασκήσεις σε διαφορετικές και προκαθορισμένες χρονικές περιόδους. Ο διαχωρισμός των φοιτητών στα τμήματα αυτά είναι τυχαίος αφού οι φοιτητές είναι αυτοί που θα επιλέξουν το τμήμα που θα παρακολουθήσουν ανάλογα με το ποια ημερομηνία ταιριάζει στο πρόγραμμά τους. Σε όλους τους φοιτητές αποστέλλεται το εργαστηριακό βιβλίο «Εργαστηριακές Ασκήσεις Βιολογίας» ενώ στην εκπαιδευτική πλατφόρμα του ΕΑΠ (study.gr), αναρτώνται διαφάνειες PowerPoint, συμπληρωματικές της ύλης που καλύπτει το παραπάνω εργαστηριακό βιβλίο.

Οι φοιτητές όταν προσέλθουν στα εργαστήρια Βιολογίας, ασκούνται επί πέντε συνεχείς ημέρες εκτελώντας δύο ασκήσεις την ημέρα, μία το πρωί και μία απόγευμα. Η καθημερινή εκπαίδευση των φοιτητών για κάθε άσκηση περιλαμβάνει αρχικά μία «πρόσωπο με πρόσωπο» παράδοση μισής ώρας από τον καθηγητή-σύμβουλο. Μετά την περιληπτική παρουσίαση της θεωρίας, ο καθηγητής-σύμβουλος πραγματοποιεί, με τη μορφή επίδειξης, την εργαστηριακή άσκηση και στη συνέχεια ζητά από τους φοιτητές να την επαναλάβουν προκειμένου να εξασκηθούν στις διάφορες εργαστηριακές τεχνικές. Μετά το τέλος κάθε εργαστηριακής άσκησης, τόσο το πρωί όσο και το απόγευμα, ο φοιτητής εξετάζεται σε γραπτό τεστ. Η τυπική συνολική διάρκεια κάθε εργαστηριακής άσκησης, συμπεριλαμβανομένης και της γραπτής εξέτασης, είναι τέσσερις ώρες το πρωί και τέσσερις το απόγευμα. Στο τέλος του εργαστηριακού κύκλου οι φοιτητές συμμετέχουν στις τελικές εξετάσεις (ΕΑΠ, [https://www.eap.gr/images/stories/pdf/kanonismos\\_erg\\_fye\\_F21415.pdf](https://www.eap.gr/images/stories/pdf/kanonismos_erg_fye_F21415.pdf)).

### 2.2. Εκπαιδευτικό σενάριο με χρήση εικονικού εργαστηρίου

Η έρευνά μας πραγματοποιήθηκε την ακαδημαϊκή χρονιά 2016-2017 όπου δημιουργήθηκαν τρία τμήματα φοιτητών. Στο 1ο τμήμα γράφτηκαν 16 φοιτητές, στο 2ο τμήμα 13 φοιτητές και στο 3ο τμήμα γράφτηκαν 14 φοιτητές. Στοιχεία σχετικά με το φύλο και το μορφωτικό επίπεδο των φοιτητών που συμμετείχαν, παρουσιάζονται στον Πίνακα 1.

Τρεις μήνες πριν από την προσέλευση των φοιτητών στα εργαστήρια Βιολογίας, ανέβηκε στην εκπαιδευτική πλατφόρμα η εφαρμογή του τρισδιάστατου εικονικού εργαστηρίου OnLabs.

**Πίνακας 1:** Στοιχεία για το φύλο και το μορφωτικό επίπεδο των 43 φοιτητών που συμμετείχαν στην έρευνα

Φύλο	1 <sup>ο</sup> Τμήμα	2 <sup>ο</sup> Τμήμα	3 <sup>ο</sup> Τμήμα	Σύνολο	%
Ανδρες	12	10	8	30	70
Γυναίκες	4	3	6	13	30
<b>Μορφωτικό Επίπεδο</b>					
Απόφοιτος Λυκείου	3	2	5	10	23
Απόφοιτος ΙΕΚ	2	2	2	6	14
Απόφοιτος ΤΕΙ	2	4	5	11	25
Απόφοιτος ΑΕΙ	4	2	2	8	19
Κάτοχος Μεταπτυχιακού Διπλώματος Ειδίκευσης	4	2	0	6	14
Άλλο	1	1	0	2	5

Δύο εβδομάδες πριν από την προσέλευση της κάθε ομάδας στα εργαστήρια του ΕΑΠ, στάλθηκε στους φοιτητές της συγκεκριμένης κάθε φορά ομάδας, μέσω email στην πλατφόρμα επικοινωνίας τους, μία πρόσκληση ενδιαφέροντος η οποία τους καλούσε να λάβουν μέρος σε μία ομαδική Skype συνεδρία. Η συνεδρία θα γινόταν μέσω της πλατφόρμα σύγχρονης επικοινωνίας Skype for Business και θα είχε σαν στόχο την προετοιμασία των φοιτητών για την εργαστηριακή άσκηση: «Εκμάθηση Μικροσκοπίου», μιας άσκησης η οποία πραγματοποιείται σύμφωνα με το πρόγραμμα, τη δεύτερη εργαστηριακή μέρα και συγκεκριμένα το απόγευμα της ημέρας αυτής. Η Skype συνεδρία θα περιελάμβανε αρχικά παράδοση των βασικών εννοιών της μικροσκόπησης, μέσω διαφανειών PowerPoint, από τον υπεύθυνο της συνεδρίας. Η παράδοση αυτή θα διαρκούσε μισή ώρα. Στη συνέχεια, και πάλι μέσω της πλατφόρμας επικοινωνίας, θα γινόταν μία επίδειξη από τον υπεύθυνο μίας ολοκληρωμένης μικροσκόπησης με φωτονικό μικροσκόπιο χρησιμοποιώντας την εφαρμογή του εικονικού εργαστηρίου. Και αυτό το στάδιο της συνεδρίας θα διαρκούσε μισή ώρα.

Σύμφωνα με την πρόσκληση αυτή, όσοι φοιτητές θα δήλωναν συμμετοχή στην Skype συνεδρία θα πριμοδοτούνταν με το βαθμολογικό μπόνους της βάσης. Για τη συνεδρία αυτή ενδιαφέρθηκαν από το 1<sup>ο</sup> τμήμα οι 7 από τους 16 φοιτητές, από το 2<sup>ο</sup> τμήμα οι 9 από τους 13 φοιτητές και από το 3<sup>ο</sup> τμήμα οι 6 από τους 14 φοιτητές. Οι συνεδρίες προγραμματίστηκαν να γίνουν μία ή δύο μέρες πριν από την προσέλευση των φοιτητών στα εργαστήρια.

Τις παραμονές της κάθε συνεδρίας οι ενδιαφερόμενοι φοιτητές έλαβαν λεπτομερείς οδηγίες για τη διαδικασία σύνδεσής τους στην εικονική αίθουσα του Skype for Business καθώς και σαφή βήματα που θα έπρεπε να ακολουθήσουν για να «κατεβάσουν» στο PC ή το Laptop τους την εφαρμογή του εικονικού εργαστηρίου.

Η 1<sup>η</sup> συνεδρία ήταν μία πιλοτική συνεδρία αφού παρουσιάστηκαν διάφορα προβλήματα που αφορούσαν στην αδυναμία επικοινωνίας των φοιτητών με την πλατφόρμα. Τα προβλήματα αυτά οφείλονταν στην ασυμβατότητα του λογισμικού των φοιτητών με το λογισμικό που υποστηρίζει το Skype for Business, στην κακή ποιότητα του δικτύου των φοιτητών, ή τέλος στην απειρία των τελευταίων ως προς τη χρήση γενικά τέτοιων λογισμικών. Στους φοιτητές αυτούς στάλθηκε βιντεοσκοπημένη τόσο η παράδοση των διαφανειών, όσο και η επίδειξη της μικροσκόπησης μέσω του εικονικού εργαστηρίου.

Έχοντας υπόψη μας τα διάφορα προβλήματα σύνδεσης που αντιμετωπίσαμε στην 1<sup>η</sup> συνεδρία, τις παραμονές της 2<sup>ης</sup> και της 3<sup>ης</sup> συνεδρίας ζητήθηκε από τους

ενδιαφερόμενους φοιτητές να συνδεθούν δοκιμαστικά στην πλατφόρμα επικοινωνίας είτε ανά ζευγάρια είτε με τον υπεύθυνο της συνεδρίας προκειμένου να λυθούν, όσο ήταν δυνατόν, εκ των προτέρων τα διάφορα τεχνικά προβλήματα. Έτσι, τόσο η 2<sup>η</sup> όσο και η 3<sup>η</sup> συνεδρία πραγματοποιήθηκε με επιτυχία.

Μετά το τέλος της κάθε συνεδρίας ζητήθηκε από τους φοιτητές να χρησιμοποιήσουν μόνοι τους την εφαρμογή του εικονικού εργαστηρίου και μέσω οδηγιών που παρέχει η εφαρμογή αυτή, να πραγματοποιήσουν μία ολοκληρωμένη μικροσκόπηση χρησιμοποιώντας το εικονικό μικροσκόπιο. Σαν τελικό στάδιο της συμμετοχής τους στη συνεδρία αυτή, οι φοιτητές συμπλήρωσαν ένα ερωτηματολόγιο που τους στάλθηκε μέσω email με τη μορφή word αρχείου. Με το ερωτηματολόγιο αυτό αξιολογούσαν τόσο την εφαρμογή του εικονικού εργαστηρίου όσο και τη Skype συνεδρία στην οποία συμμετείχαν.

Όταν προσήλθαν όλοι οι φοιτητές στα πραγματικά εργαστήρια, και αυτοί που παρακολούθησαν τη συνεδρία και οι υπόλοιποι, συμμετείχαν σε μία γραπτή εξέταση (pre-test), προκειμένου να καθοριστεί το γνωστικό επίπεδό τους. Στη συνέχεια ακολούθησε η δια ζώσης παράδοση της μισής ώρας, από τον καθηγητή-σύμβουλο, καθώς και η επίδειξη της εργαστηριακής άσκησης. Αμέσως μετά την επίδειξη όλοι οι φοιτητές συμμετείχαν σε δεύτερη γραπτή εξέταση (post-test) προκειμένου να μετρηθεί η συνεισφορά της δια ζώσης παράδοσης και της επίδειξης του πειράματος στην εννοιολογική κατανόηση της μικροσκόπησης. Επίσης συμπλήρωσαν και ένα ερωτηματολόγιο με το οποίο αξιολόγησαν την παρουσίαση και την επίδειξη του καθηγητή-σύμβουλου. Είναι πολύ σημαντικό να αναφερθεί ότι τα pre-test και post-test περιείχαν τις ίδιες ερωτήσεις.

Σε επόμενο στάδιο, στον κάθε φοιτητή δόθηκε ένα φωτονικό μικροσκόπιο και του ζητήθηκε να πραγματοποιήσει μία ολοκληρωμένη μικροσκόπηση. Ταυτόχρονα με τη μικροσκόπηση ο φοιτητής συμπλήρωνε και ένα φύλλο «εργαστηριακών δεξιοτήτων». Το φύλλο αυτό περιελάμβανε τη διεξαγωγή 22 ενεργειών, όσα και τα βήματα στα οποία χωρίσαμε τη διαδικασία μιας ολοκληρωμένης μικροσκόπησης. Μετά την ολοκλήρωση της κάθε ενέργειας ο φοιτητής έπρεπε να διαλέξει ανάμεσα σε τρία εναλλακτικά αποτελέσματα. Για καλύτερη κατανόηση, στον Πίνακα 2 δίνεται τυχαία, η 11<sup>η</sup> από τις 22 ενέργειες που έπρεπε να πραγματοποιήσει ο φοιτητής.

**Πίνακας 2:** Η 11<sup>η</sup> από τα 22 ενέργειες του φύλλο των «εργαστηριακών δεξιοτήτων»

<b>Βήμα 11</b>
«Περιστρέψτε την κεφαλή των αντικειμενικών φακών έτσι ώστε να γίνει ενεργός ο φακός με τη μεγέθυνση 10X».
α) Έκανα το βήμα επιτυχώς και με ευκολία β) Έκανα το βήμα αυτό επιτυχώς αλλά με δυσκολία γ) Δεν μπόρεσα να κάνω μόνος μου το βήμα αυτό, ζήτησα τη βοήθεια του υπευθύνου ή κάποιου συμμαθητή μου

Τα pre-tests και τα post-tests, τα ερωτηματολόγια καθώς και τα φύλλα «εργαστηριακών δεξιοτήτων», συλλέχθηκαν προς ανάλυση.

Είναι αναγκαίο να αναφερθεί σε αυτό το σημείο, ότι πριν από την πραγματοποίηση της έρευνας οι φοιτητές δεν είχαν διδαχθεί τη διαδικασία της μικροσκόπησης σε κάποια άλλη Εργαστηριακή ή Θεματική Ενότητα, κατά τη διάρκεια των σπουδών τους στο ΕΑΠ. Παρόλα αυτά, από τα συγκεκριμένα τρία τμήματα, 5 συνολικά φοιτητές είχαν ήδη μία εμπειρία στην μικροσκόπηση είτε λόγω του ότι ήδη



εργάζονταν σε κάποιο μικροβιολογικό-αιματολογικό εργαστήριο (2 φοιτητές), είτε λόγω του ότι ήταν απόφοιτοι σχολών ΑΕΙ ή ΤΕΙ με προγράμματα σπουδών στα οποία διδάσκεται η εργαστηριακή άσκηση της μικροσκοπησης (3 φοιτητές).

### 3. Ανάλυση Αποτελεσμάτων

Στην παρούσα μελέτη, προκειμένου να αξιολογηθούν οι επιδόσεις των φοιτητών, χρησιμοποιήθηκαν μόνο οι βαθμολογίες που προέκυψαν από τα post-tests. Οι ερωτήσεις των post-tests επιλέχθηκαν από μία τράπεζα 64 ερωτήσεων, μικρής, μεσαίας και μεγάλης δυσκολίας. Όλες οι ερωτήσεις σχετίζονταν με τις βασικές έννοιες της μικροσκοπησης καθώς και με θέματα εξοικείωσης με τα διάφορα τμήματα ενός φωτονικού μικροσκοπίου. Στο post-test του 1<sup>ου</sup> τμήματος χρησιμοποιήθηκαν 36 ερωτήσεις, στο post-test του 2<sup>ου</sup> τμήματος 27 ερωτήσεις και στο post-test του 3<sup>ου</sup> τμήματος επίσης 27 ερωτήσεις. Στον Πίνακα 3 που ακολουθεί, παρουσιάζονται οι βαθμολογίες στα post-tests των φοιτητών και των τριών τμημάτων, σύμφωνα με την Κλασική Θεωρία Μέτρησης των τεστ, CTT.

**Πίνακας 3:** Οι βαθμολογίες των post-tests, με τη μέθοδο CTT, και ο μέσος όρος τους, για τους φοιτητές των τριών τμημάτων

<b>Βαθμολογία</b>			
<b>A/A Φοιτητών</b>	<b>Βαθμολογία 1<sup>ου</sup> Τμήματος</b>	<b>Βαθμολογία 2<sup>ου</sup> Τμήματος</b>	<b>Βαθμολογία 3<sup>ου</sup> Τμήματος</b>
<b>1</b>	8,89	8,80	6,67
<b>2</b>	9,16	10,00	9,26
<b>3</b>	9,17	9,26	7,78
<b>4</b>	9,17	8,51	6,67
<b>5</b>	9,72	8,15	7,78
<b>6</b>	9,17	9,63	6,30
<b>7</b>	9,44	9,63	6,30
<b>8</b>	9,44	7,78	7,04
<b>9</b>	9,44	9,26	7,41
<b>10</b>	8,89	9,63	6,67
<b>11</b>	9,44	8,52	7,03
<b>12</b>	9,17	10,00	6,67
<b>13</b>	9,44	5,92	4,80
<b>14</b>	9,17	-	4,01
<b>15</b>	8,61	-	-
<b>16</b>	9,16	-	-
<b>Μέσος Όρος</b>	<b>9,22</b>	<b>8,85</b>	<b>6,74</b>
<b>Τυπική Απόκλιση (SD)</b>	<b>0,27</b>	<b>1,08</b>	<b>1,26</b>

Σύμφωνα με τον Πίνακα 3, ο μέσος όρος της βαθμολογίας των φοιτητών στο 1<sup>ο</sup> τμήμα είναι 9,22, στο 2<sup>ο</sup> τμήμα 8,85 ενώ στο 3<sup>ο</sup> τμήμα είναι 6,74. Αν λάβουμε υπόψη μας ότι ο διαχωρισμός των φοιτητών στα τρία τμήματα έγινε με τυχαίο τρόπο η παραπάνω διαφορά στους μέσους όρους των βαθμολογιών θεωρούμε ότι οφείλεται στο διαφορετικό βαθμό δυσκολίας των ερωτημάτων στα τρία post-tests. Πράγματι, το post-test του 1<sup>ου</sup> τμήματος περιείχε αρκετά εύκολες ερωτήσεις, του 2<sup>ου</sup> τμήματος περιείχε ερωτήματα μεσαίας δυσκολίας, ενώ του 3<sup>ου</sup> τμήματος περιείχε δύσκολα

ερωτήματα. Σύμφωνα με τον Πίνακα 3, όλοι οι φοιτητές του 1<sup>ου</sup> τμήματος έχουν πολύ καλές επιδόσεις, είτε χρησιμοποίησαν το εικονικό εργαστήριο είτε όχι, και άρα με βάση αυτές τις βαθμολογίες, θα μπορούσαμε να συμπεράνουμε ότι η χρήση του εικονικού εργαστηρίου είναι περιττή. Όμως, παρατηρούμε ότι στους βαθμούς των φοιτητών του 3<sup>ου</sup> τμήματος υπάρχει μία μεγάλη διασπορά. Ενδεικτικά, και χωρίς να μπαίνουμε σε λεπτομέρειες στην παρούσα εργασία, στο 3<sup>ο</sup> τμήμα, ο φοιτητής N<sup>ο</sup>2, με βαθμολογία 9,26 είχε χρησιμοποιήσει το εικονικό εργαστήριο, ενώ στο ίδιο τμήμα ο φοιτητής N<sup>ο</sup>7 και ο N<sup>ο</sup>6, με βαθμολογία 6,30 δεν το είχαν χρησιμοποιήσει. Άρα, με βάση τις βαθμολογίες του 3<sup>ου</sup> τμήματος θα συμπεραίναμε ότι το εικονικό εργαστήριο είναι αναγκαίο. Ποιο από τα δύο συμπεράσματα ισχύει τελικά; Είναι το εικονικό εργαστήριο ένα χρήσιμο εκπαιδευτικό εργαλείο ή όχι; Θα πρέπει να βασιστούμε για να απαντήσουμε σε αυτό το ερώτημα στις βαθμολογίες των φοιτητών του 1<sup>ου</sup> ή του 3<sup>ου</sup> τμήματος; Είναι φανερό ότι με το μοντέλο βαθμολόγησης CTT δεν μπορούμε να έχουμε μία ξεκάθαρη εικόνα.

Για να μπορέσουμε να λάβουμε αξιόπιστες τιμές για την ικανότητα των φοιτητών, τιμές οι οποίες δεν εξαρτώνται από τη δυσκολία των ερωτημάτων των τεστ, χρησιμοποιούμε για την ανάλυσή μας το Rasch μοντέλο. Παρότι αναλύσαμε τις απαντήσεις των φοιτητών στα post-tests και των τριών τμημάτων, στη συνέχεια παρουσιάζουμε ενδεικτικά τα αποτελέσματα από την ανάλυση του post-test του 3<sup>ου</sup> τμήματος των 14 φοιτητών που περιείχε 27 ερωτήματα.

Προκειμένου να χρησιμοποιήσουμε το διχοτομικό Rasch μοντέλο, για κάθε φοιτητή, αντιστοιχίσαμε σε κάθε ένα από τα 27 ερωτήματα την τιμή 1, αν το ερώτημα αυτό απαντήθηκε από τον φοιτητή σωστά, και την τιμή 0 αν απαντήθηκε λάθος. Το αρχείο δεδομένων που δημιουργήσαμε ήταν σε μορφή csv (comma-separated value) και κάθε φοιτητής αναπαραστάθηκε σαν ένα δυαδικό διάνυσμα μεγέθους 27. Άρα, το αρχείο των δεδομένων μας περιελάμβανε συνολικά 14 διανύσματα τέτοιου είδους και για την ανάλυση των δεδομένων αυτών χρησιμοποιήσαμε την ευρέως γνωστή στατιστική γλώσσα προγραμματισμού R. Πιο συγκεκριμένα, κάναμε χρήση του πακέτου TAM (Test Analysis Modules) της R (Kabacoff, 2011). Στο πακέτο αυτό περιέχεται η συνάρτηση tam (Robitzsch, Kiefer, & Wu, 2013) η οποία προσπαθεί να ταιριάξει το Rasch μοντέλο με διχοτομικά βαθμολογημένα δεδομένα χρησιμοποιώντας την μέθοδο της εκτίμησης της μέγιστης πιθανοφάνειας (Maximum Likelihood Estimate-MLE) (Fisher, 1912; Fisher, 1934).

Για να δώσουμε μία σύντομη επεξήγηση της λειτουργίας αυτής της μεθόδου εκτίμησης, αναφέρουμε ότι όταν ο βαθμός δυσκολίας των ερωτημάτων είναι γνωστός, ένας φοιτητής ενδέχεται να απαντήσει σωστά στο 50% των ερωτημάτων που έχουν βαθμό δυσκολίας ίσο με το επίπεδο ικανότητας του φοιτητή. Άρα, γνωρίζοντας τις απαντήσεις του φοιτητή στα ερωτήματα ενός τεστ, μπορούμε να βρούμε την ικανότητά του υπολογίζοντας την πιο πιθανή τιμή της ικανότητας που απαιτείται να έχει ο φοιτητής για να απαντήσει το 50% των ερωτημάτων. Αυτή είναι η έννοια της μεθόδου της εκτίμησης της μέγιστης πιθανοφάνειας.

Υπάρχουν αρκετές διαφορετικές υλοποιήσεις της MLE μεθόδου για την εκτίμηση της παραμέτρου της δυσκολίας των ερωτημάτων καθώς και της ικανότητας των φοιτητών, όπως η εκτίμηση τής από κοινού μέγιστης πιθανοφάνειας (Joint Maximum Likelihood Estimate-JML), η εκτίμηση τής υπό συνθήκης μέγιστης πιθανοφάνειας (Conditional Maximum Likelihood Estimate-CML) και τέλος η εκτίμηση τής περιθωριακής μέγιστης πιθανοφάνειας (Marginal Maximum Likelihood Estimate-MML).

Όλες αυτές οι μέθοδοι εκτίμησης, έχουν ένα επαναληπτικό χαρακτήρα (iterative methods). Οι επαναληπτικές αυτές τεχνικές ξεκινούν από μία αρχική εκτίμηση της

δυσκολίας των ερωτημάτων και προσπαθούν να υπολογίσουν την πιο πιθανή ικανότητα για κάθε φοιτητή. Στη συνέχεια, με δεδομένη αυτήν την πιθανή ικανότητα του κάθε φοιτητή, γίνεται εκ νέου εκτίμηση της δυσκολίας των ερωτημάτων. Η διαδικασία αυτή επαναλαμβάνεται έως ότου η απόσταση στις εκτιμήσεις των τιμών της δυσκολίας των ερωτημάτων, μεταξύ δύο διαδοχικών επαναλήψεων, να είναι μικρή, πράγμα το οποίο υποδηλώνει ότι έχει προκύψει η επιθυμητή σύγκλιση της επαναληπτικής αυτής μεθόδου. Στην έρευνα μας, έχουμε χρησιμοποιήσει τόσο την JML όσο και την MML μέθοδο. Θα πρέπει να αναφέρουμε σε αυτό το σημείο ότι επειδή οι εκτιμήσεις της ικανότητας των φοιτητών που προκύπτουν με την εφαρμογή μεθόδων εκτίμησης τύπων MLE, φαίνεται να παρουσιάζουν μία μεροληψία, για την αφαίρεση αυτής της μεροληψίας τα περισσότερα προγράμματα IRT παρέχουν μία διορθωμένη εκτίμηση της ικανότητας που είναι γνωστή ως εκτίμηση πιθανοφάνειας με βάρη (Weighted Likelihood Estimate- WLE).

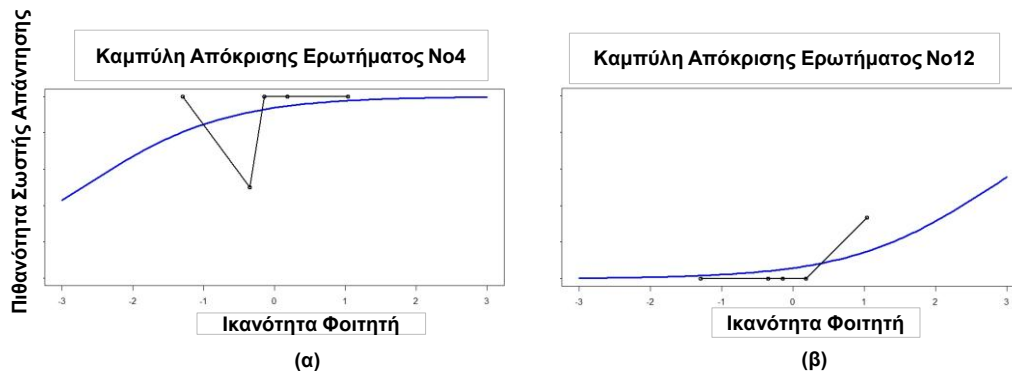
Η εκτέλεση της συνάρτησης tam, στο προγραμματιστικό περιβάλλον της R, έχει σαν αποτέλεσμα τη δημιουργία ενός αντικειμένου της αντικειμενοστραφούς γλώσσας R, μέσα στο οποίο ενθυλακώνονται όλα τα αποτελέσματα. Ένα μέρος των αποτελεσμάτων παρουσιάζονται στον Πίνακα 4.

**Πίνακας 4:** Αποτελέσματα για τη δυσκολία των 27 ερωτημάτων του post-test και για την ικανότητα των 14 φοιτητών του 3<sup>ου</sup> τμήματος.

A/A Ερωτήματος	Δυσκολία Ερωτήματος	A/A Φοιτητή	Ικανότητα Φοιτητή	Αριθμός Σωστών Απαντήσεων (/27)
1	-1.92410557	1	-0.14599193	18
2	0.28803465	2	2.02716522	25
3	-1.00878019	3	0.54230990	21
4	-2.70844823	4	-0.14599193	18
5	-1.41272710	5	0.54230990	21
6	-1.92410557	6	-0.34932498	17
7	-0.65854359	7	-0.34932498	17
8	-0.33603931	8	-0.14599193	18
9	-1.92410557	9	0.29507157	20
10	-0.33603931	10	-0.14599193	18
11	-1.00878019	11	0.06749952	19
12	2.77988303	12	-0.14599193	18
13	-1.92410557	13	-1.11133033	13
14	-2.70844823	14	-1.48690402	11
15	-0.33603931			
16	-0.33603931			
17	-1.41272710			
18	-1.92410557			
19	-2.70844823			
20	-1.41272710			
21	-1.92410557			
22	-0.02490072			
23	-0.33603931			
24	-0.02490072			
25	0.61632357			
26	-1.00878019			
27	0.61632357			

Πιο συγκεκριμένα στη 2<sup>η</sup> στήλη αυτού του πίνακα αποτυπώνεται η δυσκολία κάθε ενός από τα 27 ερωτήματα ενώ στη 4<sup>η</sup> στήλη η ικανότητα του κάθε ενός από τους 14 φοιτητές. Στην τελευταία στήλη εμφανίζεται το σκορ του κάθε φοιτητή το οποίο αντιστοιχεί στον αριθμό των σωστών απαντήσεων που έδωσε ο φοιτητής στο post-test των 27 ερωτήσεων. Από τη στήλη της ικανότητας του φοιτητή προκύπτει ότι ο καλύτερος φοιτητής έχει ικανότητα 2.027 ενώ ο πιο αδύναμος -1.487. Από τη στήλη της δυσκολίας του ερωτήματος παρατηρούμε ότι υπάρχουν τρία ερωτήματα με τη μικρότερη δυσκολία, η οποία είναι -2.70844823. Τα ερωτήματα αυτά είναι τα ερωτήματα N<sup>ο</sup> 4, 14 και 19. Επίσης, το ερώτημα N<sup>ο</sup> 12 παρουσιάζει τη μεγαλύτερη δυσκολία η οποία είναι 2.77988303.

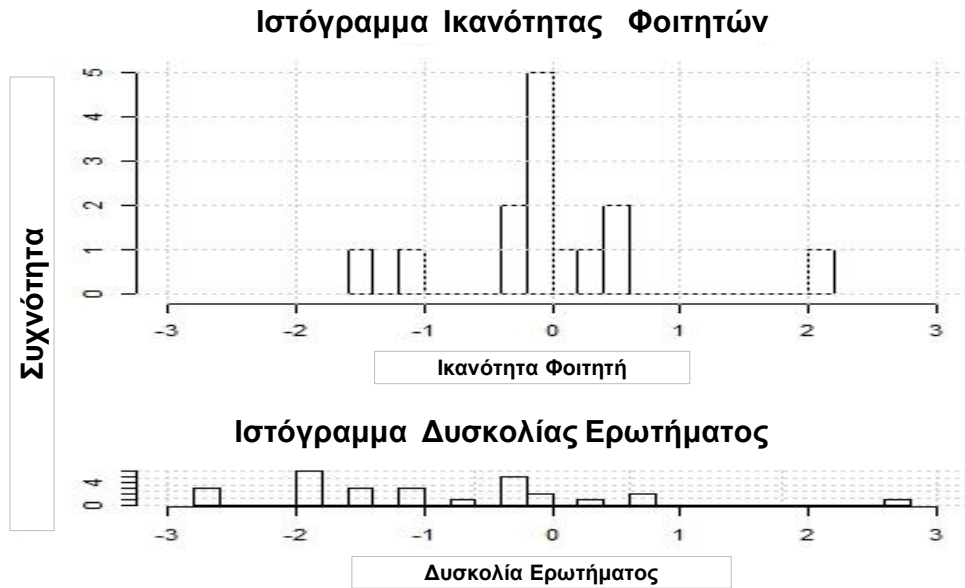
Το αντικείμενο μέσω του αντικειμένου που προκύπτει από την εκτέλεση της συνάρτησης  $tam$  μας δίνει πρόσβαση στις καμπύλες απόκρισης ερωτήματος. Στο Σχήμα 2 παρουσιάζεται (α) η καμπύλη ενός εύκολου ερωτήματος N° 4 και (β) η καμπύλη του πιο δύσκολου ερωτήματος N° 12.



Σχήμα 2: (α) Καμπύλη απόκρισης ερωτήματος του εύκολου ερωτήματος No 4, (β) Καμπύλη απόκρισης ερωτήματος του δύσκολου ερωτήματος N° 12

Σε κάθε ένα από τα γραφήματα του Σχήμα 2 παρουσιάζονται δύο χαρακτηριστικές καμπύλες ερωτημάτων. Η καμπύλη με την μπλε απόχρωση είναι η καμπύλη που υπολογίζεται με βάση την εκτίμηση της μέγιστης πιθανοφάνειας ενώ η καμπύλη που αντιστοιχεί στα συνδεδεμένα γραμμικά κομμάτια αφορά στις πραγματικές απαντήσεις των φοιτητών. Παρατηρούμε στο σχήμα (β) ότι η μαύρη καμπύλη απόκρισης που απεικονίζει τις πραγματικές απαντήσεις των φοιτητών, ταιριάζει περισσότερο με την καμπύλη των θεωρητικά αναμενόμενων αποτελεσμάτων. Η μπλε καμπύλη στο (α) του Σχήμα 2 αντιστοιχεί σε ένα εύκολο ερώτημα που δεν είναι ένα χρήσιμο ερώτημα για τον διαχωρισμό των φοιτητών σε γνωστικά επίπεδα. Είναι φανερό ότι μαθητές με μικρή, με μέτρια αλλά και με μεγάλη ικανότητα έχουν όλοι πολύ μεγάλη πιθανότητα να απαντήσουν σωστά στο συγκεκριμένο ερώτημα. Άρα θα ήταν ασφαλές η ερώτηση αυτή να μην επιλεγεί σε μελλοντικά τεστ.

Στο Σχήμα 3 παρουσιάζονται δύο ιστογράμματα, το ιστόγραμμα της ικανότητας των φοιτητών και το ιστόγραμμα της δυσκολίας των ερωτημάτων. Το σχήμα αυτό, μας δίνει τη δυνατότητα να συγκρίνουμε τις ικανότητες των φοιτητών σε σχέση με τις δυσκολίες των ερωτημάτων. Για παράδειγμα, παρατηρούμε ότι υπάρχουν πέντε φοιτητές των οποίων η ικανότητα βρίσκεται στο διάστημα  $[-0.2, 0]$  και αντίστοιχα πέντε ερωτήσεις με δυσκολία στο διάστημα  $[-2, -1.8]$ . Επίσης από το ιστόγραμμα της ικανότητας των φοιτητών του 3ου τμήματος φαίνεται ότι υπάρχει μικρότερη διασπορά στην ικανότητα των φοιτητών που πήραν το post-test σε σχέση με τη διασπορά της δυσκολίας των ερωτημάτων. Τέλος, παρατηρούμε ότι η κατανομή των ικανοτήτων των φοιτητών προσεγγίζει την κανονική κατανομή, με μέσο όρο που αγγίζει την τιμή μηδέν.



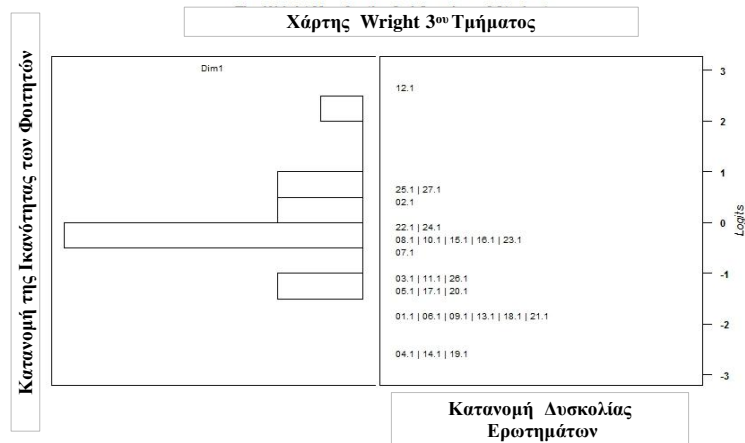
Σχήμα 3: Ιστογράμματα κατανομής ικανότητας φοιτητών και δυσκολίας ερωτημάτων για το 3<sup>ο</sup> τμήμα φοιτητών

Στο **Error! Reference source not found.** παρουσιάζεται ένας χάρτης Wright (WrightMap), ο οποίος χρησιμοποιείται για να αποδώσει ταυτόχρονα και με ακόμα μεγαλύτερη λεπτομέρεια από ότι ο Πίνακας 4, τις κατανομές τόσο της δυσκολίας των ερωτημάτων όσο και των ικανοτήτων των ερωτηθέντων φοιτητών.

Παρατηρούμε ότι ο άξονας της δυσκολίας των ερωτημάτων που είναι ίδιος με τον άξονα της ικανότητας των φοιτητών, είναι πλέον κάθετος και στο κέντρο του χάρτη. Προς το πάνω και δεξιό μέρος του σχήματος εμφανίζονται τα ερωτήματα με την μεγαλύτερη δυσκολία και στο αριστερό μέρος οι φοιτητές με τη μεγαλύτερη ικανότητα. Αντίθετα στο κάτω και δεξιό μέρος εμφανίζονται τα εύκολα ερωτήματα ενώ στο αριστερό οι πιο αδύναμοι φοιτητές.

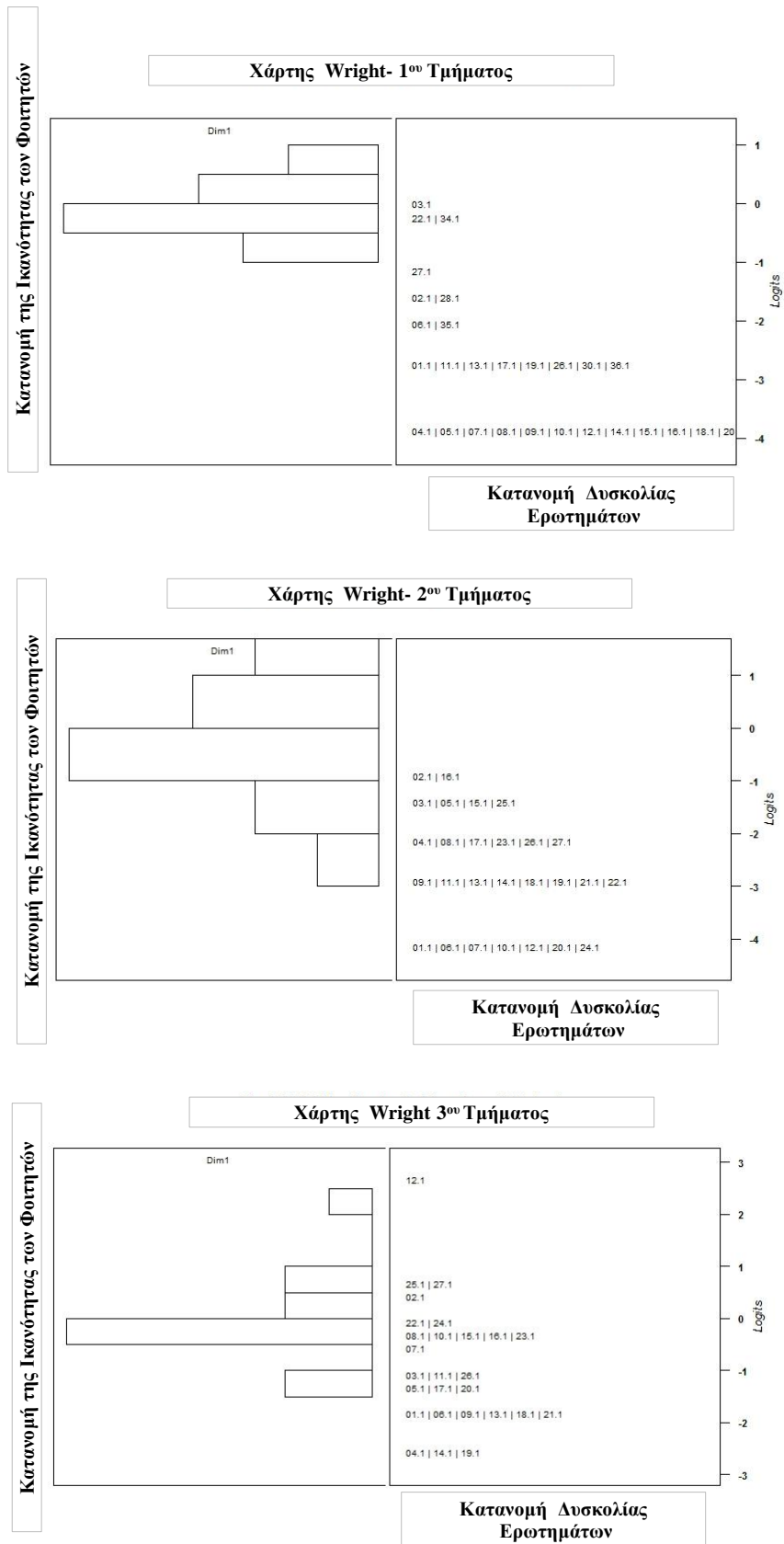
Επαληθεύοντας τα όσα αποτελέσματα μπορούν να προκύψουν από την ανάλυση του Πίνακας 4, στον χάρτη Wright του Σχήμα 4, με ευκολία διακρίνουμε ερωτήματα που έχουν τον ίδιο βαθμό δυσκολίας, π.χ. τα ερωτήματα N<sup>ο</sup>1, 6, 9, 13, 18 και 21 με βαθμό δυσκολίας, περίπου ίσον με -1.9 της κλίμακας logits (log odds units). Άρα, η παρουσία όλων αυτών των ερωτημάτων μέσα στο ίδιο τεστ δεν είναι αναγκαία, και άρα περισσότερα από ένα ερωτήματα της ίδιας δυσκολίας θα μπορούσαν να θεωρηθούν, υπό αυτήν την έννοια, ως περιττά. Σε ότι αφορά στη μέγιστη δυσκολία των ερωτημάτων ανά τεστ, φαίνεται ότι το δυσκολότερο ερώτημα του 1<sup>ου</sup> τεστ είναι κοντά στο 0, του 2<sup>ου</sup> τεστ κοντά στο -1, ενώ του 3<sup>ου</sup> κοντά στο 2.7 της κλίμακας logits. Σε ότι αφορά στην μικρότερη δυσκολία των ερωτημάτων, και στα τρία τεστ τα ευκολότερα ερωτήματα έχουν δυσκολία κοντά στο -4 της κλίμακας logits. Από τον χάρτη Wright προκύπτει ότι το πιο δύσκολο ερώτημα είναι το N<sup>ο</sup>12, ενώ τα τρία πιο εύκολα ερωτήματα είναι τα N<sup>ο</sup> 4, 14 και 19. Τέλος, παρατηρούμε ότι σχετικά μεγάλο πλήθος ερωτημάτων έχει δυσκολία πολύ μικρότερη και από την πιο μικρή ικανότητα των αδύναμων φοιτητών.





**Σχήμα 4:** Χάρτης Wright του 3<sup>ου</sup> τμήματος με απεικονίσεις των κατανομών της ικανότητας των φοιτητών και της δυσκολίας των ερωτημάτων, στην ίδια κλίμακα logits

Στο Σχήμα 4 είναι χρήσιμο να παρατηρήσουμε ότι δεν υπάρχουν ερωτήματα που να αντιστοιχούν σε κάποιες τιμές δυσκολίας του άξονα, π.χ. δεν υπάρχουν ερωτήματα με βαθμό δυσκολίας κοντά στο 1 της κλίμακας logits. Αυτή είναι μία πληροφορία την οποία μπορεί να χρησιμοποιήσει ο εκπαιδευτικός έτσι ώστε να δημιουργήσει μελλοντικά ένα πιο ολοκληρωμένο τεστ με ερωτήματα όλων των βαθμών δυσκολίας. Στο Σχήμα 5, παρουσιάζονται οι χάρτες Wright και για τα τρία τμήματα των φοιτητών. Από την παρατήρηση και των τριών χαρτών προκύπτει ότι οι κατανομές των ικανοτήτων των φοιτητών είναι παρόμοιες, κάτι που είναι αποδεκτό αφού ο διαχωρισμός τους στα τρία τμήματα ήταν τυχαίος και άρα το αναμενόμενο γνωστικό τους επίπεδο είναι θεωρητικά παρόμοιο. Επίσης ο μέσος όρος της δυσκολίας των ερωτημάτων φαίνεται ότι αυξάνεται από το 1<sup>ο</sup> προς το 3<sup>ο</sup> τμήμα, όπως ακριβώς θα έπρεπε να συμβαίνει αφού όπως ήδη έχουμε αναφέρει, το post-test του 1<sup>ου</sup> τμήματος περιείχε αρκετά εύκολες ερωτήσεις, του 2<sup>ου</sup> τμήματος περιείχε ερωτήματα μεσαίας δυσκολίας, ενώ του 3<sup>ου</sup> τμήματος περιείχε δύσκολα ερωτήματα. Από τις παραπάνω παρατηρήσεις μας συμπεραίνουμε ότι από την ανάλυση των απαντήσεων των φοιτητών στα post-tests με το στατιστικό μοντέλο Rasch, οδηγούμαστε σε λογικά συμπεράσματα τα οποία συνάδουν με τα δεδομένα με τα οποία δομήσαμε την έρευνά μας.



Σχήμα 5: Χάρτες Wright για τα τρία τμήματα φοιτητών που συμμετείχαν στην έρευνα

#### 4. Συμπεράσματα

Ο ρόλος και η συνεισφορά του εργαστηρίου ήταν πάντα κεντρική στην εμπέδωση των εννοιών της επιστήμης (Hofstein & Lunetta, 2004). Τα τελευταία χρόνια παρατηρείται ένα τεράστιο ενδιαφέρον για τον εμπλουτισμό των εργαστηριακών δραστηριοτήτων με εργαλεία υψηλής τεχνολογίας προκειμένου να κινητοποιηθούν οι φοιτητές και να εμπλακούν ακόμη περισσότερο στη μαθησιακή διαδικασία. Στην παρούσα έρευνα αξιολογήθηκε η επίδοση ομάδων φοιτητών οι οποίες είτε χρησιμοποίησαν είτε όχι, ένα τρισδιάστατο εικονικό εργαστήριο για την προετοιμασία τους για το πραγματικό εργαστήριο Βιολογίας. Για την αξιολόγηση της επίδοσής τους χρησιμοποιήσαμε τις αποκρίσεις τους στα post-tests. Βασική επιδίωξη μας ήταν να εφαρμόσουμε στα post-tests ένα τέτοιο μοντέλο βαθμολόγησης το οποίο θα μας έδινε αξιόπιστα αποτελέσματα. Η κλασική μέθοδος βαθμολόγησης CTT των post-tests δε μας έδωσε χρήσιμες τιμές για την ικανότητα των φοιτητών, αφού αυτές οι τιμές δεν ήταν ανεξάρτητες από το βαθμό δυσκολίας των ερωτημάτων των τεστ. Υποστηρίζοντας θερμά την άποψη ότι δε θα πρέπει να δίνουμε έμφαση στους βαθμούς αλλά στην ικανότητα του φοιτητών (Winston, 2008), χρησιμοποιήσαμε το μοντέλο Rasch το οποίο δίνει στους εκπαιδευτικούς, αλλά και στους ερευνητές, «πλούσιες» ερμηνείες για τη σχέση της ικανότητας των φοιτητών και των ερωτημάτων ενός τεστ (Pensavalle & Solinas, 2013). Το μοντέλο Rasch που χρησιμοποιήσαμε:

1. Διαχώρισε τη δυσκολία του ερωτήματος  $\beta_j$  από την ικανότητα του φοιτητή  $\theta$
2. Μοντελοποίησε τις αποκρίσεις στα ερωτήματα των φοιτητών ως συνάρτηση της ικανότητάς τους  $\theta$  και της δυσκολίας  $\beta_j$  των ερωτημάτων.
3. Χρησιμοποίησε τη λογαριθμική συνάρτηση για να μετατρέψει μια κλίμακα διάταξης σε συνεχή κλίμακα (λογαριθμικές μονάδες-logits) (Τσιγγίλης, <http://research.edu.uoi.gr/sites/default/files/files/docs/irt.pdf>).

Αφού επεξεργαστήκαμε τις αποκρίσεις των φοιτητών μέσω του μοντέλου Rasch:

1. Λάβαμε σταθερές τιμές για την ικανότητα του φοιτητών ανεξάρτητες της δυσκολίας του δείγματος των ερωτήσεων. Από μία κλίμακα με τιμές από  $-\infty$  έως  $+\infty$ , ο καλύτερος φοιτητής στην έρευνά μας είχε ικανότητα 2,027 και ο πιο αδύναμος -1,487
2. Επιβεβαιώσαμε ότι οι κατανομές των ικανοτήτων των φοιτητών και στα τρία τμήματα είναι παρόμοιες, κάτι το οποίο ήταν λογικό και αναμενόμενο, αφού οι φοιτητές επέλεξαν τυχαία το εργαστηριακό τους τμήμα
3. Πήραμε πληροφορίες για το βαθμό δυσκολίας των ερωτήσεων που χρησιμοποιήσαμε. Σε μία κλίμακα επίσης με τιμές από  $-\infty$  έως  $+\infty$ , η πιο εύκολη ερώτηση είχε βαθμό δυσκολίας -2.708 ενώ η πιο δύσκολη 2.780
4. Διαπιστώσαμε ότι στα post-tests δεν είχαμε συμπεριλάβει ερωτήματα ορισμένου βαθμού δυσκολίας. Άρα, γνωρίζοντας τους βαθμούς δυσκολίας διαφόρων ερωτημάτων μπορούμε να δημιουργήσουμε μία τράπεζα προκειμένου τα post-tests να είναι μελλοντικά πληρέστερα και να απευθύνονται σε τμήματα φοιτητών με διαφορετικό γνωστικό επίπεδο
5. Έχοντας χρησιμοποιήσει τις αποκρίσεις των φοιτητών στα post-tests, δημιουργήσαμε τις καμπύλες απόκρισης όλων των ερωτημάτων της τράπεζας μας, από τις οποίες μπορούμε να υπολογίζουμε την πιθανότητα ένας φοιτητής, γνωστής ικανότητας, να απαντήσει σωστά σε κάθε ένα από αυτά τα ερωτήματα.

Σε επόμενη έρευνά μας, βασιζόμενοι στις ακριβείς τιμές της ικανότητας των φοιτητών της παρούσης εργασίας, θα μπορούσαμε να κάνουμε μία αξιόπιστη αποτίμηση του οφέλους που αποκομίζουν οι φοιτητές οι οποίοι χρησιμοποιούν ένα

τρισετάστατο εικονικό εργαστήριο στις εξ αποστάσεως σπουδές τους. Επίσης με βάση τις ίδιες τιμές θα μπορούσαμε να προσφέρουμε μία πιο εξατομικευμένη εκπαίδευση στους φοιτητές ενός τμήματος ενισχύοντας με επιπλέον διδακτικές ώρες τους πραγματικά και όχι φαινομενικά αδύναμους, και προσφέροντας επιπλέον μαθήματα προχωρημένου επιπέδου στους πολύ ικανούς.

### Βιβλιογραφικές Αναφορές

- Almond, R. G., et al. (2015). *Bayesian Networks in Educational Assessment*. New York: Springer.
- de Ayala, R. (2008). *The theory and Practice of Item Response Theory*. New York: Guilford Press.
- Fisher, R. A. (1912). On an absolute criterion for fitting frequency curves. *Messenger of Mathematics Vol 41*, pp. 155-160.
- Fisher, R. A. (1934). Two New Properties of Mathematical Likelihood. *Proceedings of the Royal Society A Vol 44*, pp. 285-307.
- Gulliksen, H. (1950). *Theory of mental test*. New York: John Willey & Sons.Inc.
- Hambledon, R. K., Swaminathan, H., & Rogers. H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.
- Hofstein, A., & Lunetta, V. N. (2004). The laboratory in Science Education: Foundations for The Twenty-First Century. *Science Education, Vol 88*, (No 1), pp. 28-54.
- Kabacoff, R. I. (2011). *R in Action. Data analysis and graphics with R*. Shelter Island N.Y.: Manning Publications CO.
- Makransky, G., et al. (2016). Simulation based virtual learning environment in medical genetics counseling: an example of bridging the gap between theory and practice in medical education. *BMC Medical Education, Vol 16* (98). doi: 10.1186/s12909-016-0620-6
- Makransky, G., Thisgaard, M. W., & Gadegaard, H. (2016). Virtual Simulations as Preparation for Lab Exercises: Assessing Learning of Key Laboratory Skills in Microbiology and Improvement of Essential Non-Cognitive Skills. *PLOS ONE*. Retrieved September 30, 2017 from <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0155895>
- Mayer, R. E. (1999). Multimedia aids to problem-solving transfer-A dual coding approach. *International Journal of Educational Research Vol 31*, 611-623.
- Pensavalle, C.A., & Sollinas, G. (2013). The Rasch model analysis for understanding mathematics proficiency- A case study: Senior high school Sardinian students. *Creative Education, Vol 4*, (No 12), pp. 767-773.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment test*. Copenhagen: Danmarks Paedagogiske Institut.
- Richardson, M. W. (1936). The relationship between the difficulty and the differential validity of a test. *Psychometrika, Vol 1* (No 2), pp. 33-49.
- Robitzsch, A., Kiefer, T., & Wu, M. (2013). Test Analysis Modules-Package "TAM". *Computer Software*. Retrieved September 30, 2017 from <https://cran.r-project.org/web/packages/TAM/TAM.pdf>
- Shute, J. V., Ventura, M., Bauer, M & Zapata-Rivera, D. (2009). *Serious Games: Mechanisms and Effects*. New York and London: Routledge.
- Thissen, D., & Wainer, H. (2001). *Test Scoring*. New York: Routledge.
- Tucker, L. R. (1946). Maximum validity of a test with equivalent items. *Psychometrika, Vol 11*, pp. 1-3.
- Verhelst, N. D., & Glas, C. A.W. (1995). *The one Parameter Logistic model*. New York: Springer.
- Winston, P. H. (2008). Skills, Big Ideas, and Getting Grades Out of the Way. *MIT Faculty Newsletter Vol 20* (No 4). Retrieved September 30, 2017 from <http://web.mit.edu/fnl/volume/204/winston.html>
- Wu, M., Tam, H. P., & Jen, T-H. (2016). *Educational Measurement for Applied Researchers: Theory into Practice*. Singapore: Spinger.
- Zafeiropoulos, V., Kalles, D., Sgourou, A., & Kameas, A. (2014). Adventure-Style Serious Game for a Science Lab. *9th European Conference on Technology Enhanced Learning, 16-19 September 2014* (pp. 538-541). Austria: Springer.
- Αλεξόπουλος, Δ. (2004). *Ψυχομετρία: Ιστορία, Θεωρίες και Γενικές Αρχές, Τόμος Β'*. Αθήνα: Ελληνικά Γράμματα.
- Λιοναράκης, Α. (2001). Ανοικτή και εξ αποστάσεως πολυμορφική εκπαίδευση. Προβληματισμοί για μια ποιοτική προσέγγιση σχεδιασμού εκπαιδευτικού υλικού. Στο Λιοναράκης, Α. (Επιμ.),

*Απόψεις και Προβληματισμοί για την Ανοικτή και εξ Αποστάσεως Εκπαίδευση.* Αθήνα:  
Προπομπός.

Τσιγγίλης, Ν. *Εφαρμογές της Θεωρίας Απόκρισης Ερωτήματος (Item Response Theory) για την εξέταση των ψυχομετρικών ιδιοτήτων ερωτηματολογίων και κλιμάκων μέτρησης στις κοινωνικές επιστήμες.* Ανακτήθηκε 30 Σεπτεμβρίου, 2017 από <http://research.edu.uoi.gr/sites/default/files/files/docs/irt.pdf>