

ΤΑ ΔΟΚΙΜΙΑ ΕΛΛΗΝΟΜΑΘΕΙΑΣ
«ΜΙΛΑΣ ΕΛΛΗΝΙΚΑ Ι»: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ
ΤΗΝ ΕΡΕΥΝΑ ΕΓΚΥΡΟΠΟΙΗΣΗΣ ΣΤΟ
ΚΥΠΡΙΑΚΟ ΣΥΓΚΕΙΜΕΝΟ

VALIDATION OF GREEK LANGUAGE PROFICIENCY
TESTS “MILAS ELLINIKA I” IN CYPRUS CONTEXT

Αλεξάνδρα Πετρίδου
Κέντρο Εκπαιδευτικής Έρευνας
και Αξιολόγησης
petridou.a@cyearn.pi.ac.cy

Γιασεμίνα Καραγιώργη
Κέντρο Εκπαιδευτικής Έρευνας
και Αξιολόγησης

Abstract

In Cypriot schools the number of pupils with Greek as a Second Language (GSL) is constantly increasing. The lack of a valid and reliable test to assess language proficiency in Greek poses serious difficulties in identifying GSL students' learning needs and offering them Greek language instruction appropriate to their proficiency level. This paper aims to discuss a validation study of the language tests “Milas Ellinika I”. A total of 382 primary school pupils aged 7-11 participated in this study, which collected data on the four dimensions of Weir's (2005) validation framework namely scoring, criterion-related, context and consequential validity, along with the inclusion of content validity. Results support the use of both tests in the Cypriot context, after some minor adaptation revisions.

Key words

Tests, Greek language proficiency, validation study.

Λέξεις κλειδιά

Δοκίμια, ελληνομάθεια, εγκυροποίηση

0. Εισαγωγή

Με την έλευση παλιννοστούτων και μεταναστών και τη συνεπακόλουθη διαφοροποίηση της σύνθεσης της κοινωνίας, ο μαθητικός πληθυσμός στην Κύπρο έχει μετατραπεί τα τελευταία χρόνια από ομοιογενή σε ετερογενή (Οικονομίδου, 2003), με το ποσοστό των αλλόγλωσσων μαθητών σε δημόσια δημοτικά σχολεία τη

σχολική χρονιά 2014-2015 να ανέρχεται σε 14,7% (Υπουργείο Παιδείας και Πολιτισμού, χ.η.).

Η εκπαιδευτική πολιτική που προβλέπεται για την ενσωμάτωση των αλλόγλωσσων μαθητών στο κυπριακό εκπαιδευτικό σύστημα (Υπουργείο Παιδείας και Πολιτισμού, 2014) φαίνεται να συνάδει με το μοντέλο των Christensen και Stanat (2007) για 'ένταξη με συστηματική υποστήριξη στη γλώσσα διδασκαλίας'. Οι μαθητές πρέπει να κατακτήσουν όχι μόνο βασικές επικοινωνιακές γλωσσικές δεξιότητες αλλά και την ακαδημαϊκή γλωσσική ικανότητα (Cummins, 1996), έτσι ώστε να μπορούν να ανταποκριθούν στις απαιτήσεις του σχολείου. Εφόσον στα κυπριακά σχολεία ακολουθείται μονογλωσσικό σύστημα σπουδών (OECD, 2009), τυχόν ελλείψεις στη γνώση και χρήση της ελληνικής γλώσσας ενδέχεται να παρεμποδίσουν τη μαθησιακή επιτυχία των παιδιών.

Παρόλο που στον ελλαδικό χώρο έχουν εκπονηθεί διάφορα δοκίμια ελληνομάθειας, δεν εντοπίζονται αντίστοιχα στον κυπριακό χώρο, με την απουσία κατάλληλων εργαλείων διάγνωσης του επιπέδου ελληνομάθειας των αλλόγλωσσων μαθητών να δυσχεραίνει την εξατομικευμένη και στοχευμένη στήριξη (Βαρλοκώστα και Τριανταφυλλίδου, 2003). Το υφιστάμενο άρθρο επικεντρώνεται στα δοκίμια ελληνομάθειας «Μιλιάς Ελληνικά Ι» που έχουν αναπτυχθεί από το Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης. Η παρούσα εργασία συζητά τα ευρήματα της έρευνας εγκυροποίησης των εν λόγω δοκιμίων στο κυπριακό πλαίσιο, όπως επιχειρήθηκε από το Κέντρο Εκπαιδευτικής Έρευνας και Αξιολόγησης (ΚΕΕΑ) του Υπουργείου Παιδείας και Πολιτισμού, για να στηριχθεί η περαιτέρω αξιοποίησή τους από το εκπαιδευτικό σύστημα της Κύπρου.

1. Η εστίαση της έρευνας

Τα διαγνωστικά δοκίμια «Μιλιάς Ελληνικά» αποτελούν την αναθεωρημένη έκδοση των δοκιμίων «Ας μιλήσουμε Ελληνικά». Αναπτύχθηκαν στο πλαίσιο του Προγράμματος «Εκπαίδευση παλιννοστούτων και αλλοδαπών μαθητών» που αποσκοπούσε στη μέτρηση της ελληνομάθειας των παιδιών της μειονότητας της Θράκης στη βάση μιας αντικειμενικής κλίμακας αξιολόγησης. Τα εν λόγω δοκίμια (Τζεβελέκου κ.ά., 2006, Tzeveleku et al., 2013) βασίστηκαν στην κλίμακα αξιολόγησης του Κοινού Ευρωπαϊκού Πλαισίου Αναφοράς (ΚΕΠΑ) για τις γλώσσες (Συμβούλιο της Ευρώπης, 2001). Η κλίμακα ΚΕΠΑ περιλαμβάνει έξι επίπεδα χειρισμού της γλώσσας (A1-στοιχειώδες, A2-εισαγωγικό, B1-βασικό, B2-επάρκειας, Γ1-προχωρημένο και Γ2-αυτάρκειας) που καθορίζουν τρεις κατηγορίες χρηστών: «αρχάριος» (A1 και A2), «ανεξάρτητος» (B1 και B2) και «αυτάρκης» (Γ1 και Γ2). Στη βάση της γνωσιακής ωριμότητας των παιδιών ηλικίας 6-13 ετών και των επικοινωνιακών καταστάσεων στις οποίες συμμετέχουν, η κλίμακα προσαρμόστηκε μόνο στα τέσσερα πρώτα επίπεδα, εφόσον «οι γλωσσικές πράξεις που ανήκουν σε επίπεδα ανώτερα από το B1, δεν

ταιριάζουν σε παιδιά του δημοτικού, ακόμη και αν είναι φυσικοί ομιλητές της γλώσσας» (Τζεβελέκου κ.ά., 2004:13).

Τα δοκίμια «Μιλιάς Ελληνικά» αποτελούνται από δύο μέρη (I και II), το κάθε ένα από τα οποία διαχωρίζεται σε δύο τεύχη (Α' και Β'). Η υφιστάμενη έρευνα εστίασε μόνο στο πρώτο μέρος («Μιλιάς Ελληνικά I»), εφόσον αυτό απευθυνόταν σε παιδιά 7-11 χρόνων, δηλαδή μαθητές δημοτικού σχολείου. Το τεύχος Α' του δοκιμίου αξιολογεί το πρώτο επίπεδο γλωσσομάθειας (Α1) και περιλαμβάνει τέσσερις ασκήσεις τριών βασικών δεξιοτήτων: της κατανόησης προφορικού λόγου, της κατανόησης γραπτού λόγου και της παραγωγής γραπτού λόγου. Το τεύχος Β' αξιολογεί τα επίπεδα Α2 και Β1 και περιέχει επτά ασκήσεις για αξιολόγηση των τριών βασικών δεξιοτήτων που αναφέρονται πιο πάνω, αλλά και της ευρύτερης χρήσης της γλώσσας (γραμματική και λεξιλόγιο). Τα εν λόγω δοκίμια συνοδεύονταν από οδηγίες χορήγησης, διόρθωσης, καθώς και κατάταξης των μαθητών σε επίπεδα ελληνομάθειας.

2. Το ζήτημα της εγκυρότητας

Η σχετική βιβλιογραφία έχει ασχοληθεί επισταμένα με την έννοια της εγκυρότητας. Παλιότερα επικρατούσε η θέση του Kelley (1927) ότι η εγκυρότητα αφορά κατά πόσον «ένα τεστ πραγματικά μετράει αυτό που δηλώνει ή φαίνεται πως μετράει» (σ. 14). Η παραδοσιακή αυτή άποψη απέδιδε την εγκυρότητα στο ίδιο το δοκίμιο, κάτι που απορρίπτει ο Messick (1989:20), ορίζοντας την ως «αξιολογική κρίση του βαθμού στον οποίο εμπειρικά δεδομένα και θεωρία στηρίζουν την ακρίβεια και την καταλληλότητα των ερμηνειών και των συμπερασμάτων» Στη βάση της νέας αυτής θεώρησης, συντελείται μετατόπιση του ενδιαφέροντος από το εργαλείο μέτρησης στη χρήση του.

Επιπρόσθετα, στη σύγχρονη βιβλιογραφία η εγκυρότητα προσεγγίζεται ως ενιαία και πολυδιάστατη και αφορά λ.χ. στο περιεχόμενο, στην όψη, στη δομή, στη βαθμολόγηση, στις συνέπειες (Messick, 1995, Weir, 2005). Κατά συνέπεια, κατά τη διαδικασία ελέγχου της, ο αριθμός και το είδος των εμπειρικών δεδομένων που συλλέγονται διευρύνεται, ενώ συχνά εξετάζονται διαφορετικές πτυχές των ερμηνειών, των συμπερασμάτων και των χρήσεων των αποτελεσμάτων ενός δοκιμίου (Van der Walt & Steyn, 2008). Παράλληλα, οι διαστάσεις εγκυρότητας διαφοροποιούνται ανάλογα με τον σκοπό και τις εμφάσεις της κάθε ερευνητικής προσπάθειας. Ο Weir (2005) προτείνει τέσσερα θεωρητικά μοντέλα, ένα για κάθε βασική δεξιότητα γλωσσομάθειας, που εξετάζουν πέντε παραμέτρους: το πλαίσιο, το θεωρητικό υπόβαθρο, τη βαθμολόγηση, τις συνέπειες και τα κριτήρια. Επιπρόσθετα, ο Weir (2005) συζητά και τα χαρακτηριστικά των εξεταζομένων (δημογραφικά, ψυχολογικά, εμπειρία), που ενδέχεται να επιδρούν στις επιδόσεις. Τα θεωρητικά αυτά μοντέλα παρουσιάζουν αλληλεπικαλύψεις, εφόσον κάποιες παράμετροι (π.χ. συνέπειες, κριτήρια,

χαρακτηριστικά των εξεταζομένων) παραμένουν οι ίδιες, ενώ οι διαφορές εστιάζουν κυρίως στο θεωρητικό υπόβαθρο και στο πλαίσιο.

Τέλος, λόγω των πολλαπλών διαστάσεων και της πλειάδας των δεδομένων που μπορούν να συλλεγούν, η εγκυρότητα είναι πάντοτε θέμα βαθμού (Messick, 1989).

3. Το μεθοδολογικό πλαίσιο

3.1. Διαστάσεις εγκυρότητας

Η παρούσα έρευνα εγκυροποίησης εστίασε όχι στο σχεδιασμό και την κατασκευή των δοκιμών (a priori validation) αλλά στη χρήση τους (a posteriori validation). Συγκεκριμένα, συνέλεξε δεδομένα ως προς τέσσερις από τις διαστάσεις εγκυρότητας του Weir (2005) -τη βαθμολόγηση, το κριτήριο, το πλαίσιο, τις συνέπειες- ενώ ενσωμάτωσε και τη διάσταση του περιεχομένου.

Για την εγκυρότητα *βαθμολόγησης* (scoring validity), εφαρμόστηκε διπλή διόρθωση για 10% των δοκιμών κάθε τεύχους (49 δοκίμια ή 9,2% για το τεύχος Α' και 36 ή 11% για το τεύχος Β') και διερευνήθηκε ο βαθμός συμφωνίας μεταξύ των διορθωτών με χρήση του Απλού Ποσοστιαίου Βαθμού Συμφωνίας (ΑΠΒΣ) και του στατιστικού δείκτη Cohen's Kappa (K), που κυμαίνεται από το 0 έως το 1 με το 1 να σηματοδοτεί απόλυτη συμφωνία (Cohen, 1960, Fleiss et al., 2003). Σε σύγκριση με τον ΑΠΒΣ, ο δείκτης K έχει δύο σημαντικά πλεονεκτήματα: (α) συνοδεύεται από έλεγχο της στατιστικής σημαντικότητας, και (β) μπορεί να κριθεί στη βάση δημοσιευμένων πινάκων (Altman, 1991). Επίσης, διερευνήθηκε ο βαθμός δυσκολίας των έργων στα δοκίμια, η εξακρίβωση της ποιότητάς τους και η αξιοπιστία των μετρήσεων με το μοντέλο Rasch από τη Σύγχρονη Θεωρία της Μέτρησης. Οι αναλύσεις Rasch αποσκοπούσαν: στην εξακρίβωση της ποιότητας κάθε έργου σύμφωνα με τις βασικές αρχές μέτρησης, στη διακρίβωση του βαθμού δυσκολίας κάθε έργου και στην ανάπτυξη ισοδιαστημικής κλίμακας γλωσσικής ικανότητας. Η τοποθέτηση όλων των έργων και μαθητών στην ίδια κλίμακα επέτρεψε τη σύγκριση των δύο κατανομών, για διαπίστωση κατά πόσον τα δοκίμια ανταποκρίνονται στο αναμενόμενο επίπεδο ελληνομάθειας.

Για την εγκυρότητα *συνεπειών* (consequential validity), διενεργήθηκαν αναλύσεις με το μοντέλο Rasch, για να διερευνηθεί αν και σε ποιον βαθμό υφίστανται δύο σημαντικές απειλές (Messick, 1995): η μη αντιπροσωπευτικότητα του περιεχομένου (construct underrepresentation) και η διακύμανση που δεν οφείλεται στην υπό αξιολόγηση μεταβλητή (construct irrelevant variance).

Όσον αφορά στην εγκυρότητα *κριτηρίου* (criterion-related validity), συλλέχθηκαν δεδομένα για το επίπεδο ελληνομάθειας των μαθητών από μια επιπρόσθετη πηγή, τους εκπαιδευτικούς. Συγκεκριμένα, δόθηκε στους εκπαιδευτικούς σύντομη περιγραφή των

τριών πρώτων επιπέδων του ΚΕΠΑ (Α1, Α2 και Β1) και οι εκπαιδευτικοί κλήθηκαν να κατατάξουν τους μαθητές τους σε ένα από αυτά.

Ως προς την εγκυρότητα *πλαισίου* (context validity), χορηγήθηκε ερωτηματολόγιο στους χορηγητές για τις συνθήκες χορήγησης: το χρόνο συμπλήρωσης, τα αρχεία ήχου, τις απορίες των μαθητών, τις ερωτήσεις που δυσκόλεψαν και άλλες παρατηρήσεις. Συνολικά, συμπληρώθηκαν 30 ερωτηματολόγια χορήγησης για το τεύχος Α' και 23 για το τεύχος Β'.

Τέλος, για την εγκυρότητα *περιεχομένου* συλλέχθηκαν δεδομένα από εκπαιδευτικούς μέσω δύο ερωτηματολογίων για την καταλληλότητα του περιεχομένου (έκταση, ερωτήσεις, εικόνες). Για το τεύχος Α' ανταποκρίθηκαν 186 εκπαιδευτικοί (20% Β' τάξης, 21% Γ' τάξης, 21% Δ' τάξης, 20% Ε' τάξης, 18% ΣΤ' τάξης), ενώ για το τεύχος Β' 188 εκπαιδευτικοί (20% Β' τάξης, 21% Γ' τάξης, 21% Δ' τάξης, 21% Ε' τάξης, 18% ΣΤ' τάξης).

3.2. Ερευνητικές διαδικασίες

Κατά τη διάρκεια της πρώτης φάσης (Οκτώβριος 2014) έγινε χορήγηση του τεύχους Α' στους μαθητές από εξωτερικούς χορηγητές, βάσει των οδηγιών της ομάδας συγγραφής (Διάγραμμα 1). Χορηγήθηκε, επίσης, ερωτηματολόγιο μαθητών από τους εκπαιδευτικούς για συλλογή επιπρόσθετων δεδομένων σχετικά με προσωπικούς (π.χ. φύλο, καταγωγή), οικογενειακούς (π.χ. επάγγελμα γονέων) παράγοντες και τη χρήση της ελληνικής γλώσσας. Στους εκπαιδευτικούς χορηγήθηκαν δύο ερωτηματολόγια. Το Ερωτηματολόγιο Α' αφορούσε στο περιεχόμενο, στο βαθμό δυσκολίας και στην καταλληλότητα του τεύχους Α' για την ηλικία και το επίπεδο ελληνομάθειας των μαθητών. Το Ερωτηματολόγιο Γ' επικεντρώθηκε στο ιστορικό ενίσχυσης των μαθητών, καθώς και στις εκτιμήσεις των εκπαιδευτικών για το επίπεδο ελληνομάθειας των μαθητών τους. Σε κάθε σχολείο συμπληρώθηκε επίσης ερωτηματολόγιο χορηγητών για το χρόνο και τις συνθήκες χορήγησης, καθώς και τις δυσκολίες συμπλήρωσης του δοκιμίου.

Διάγραμμα 1: Μεθοδολογικός σχεδιασμός

ΦΑΣΗ	ΣΥΜΜΕΤΕΧΟΝΤΕΣ	ΔΟΚΙΜΙΟ ΕΛΛΗΝΟΜΑΘΕΙΑΣ	ΆΛΛΑ ΕΡΓΑΣΙΑ
A'	Σύνολο αλλόγλωσσων μαθητών Β'-Στ' τάξης στα επιλεγέντα σχολεία Εκπαιδευτικοί όλων των τάξεων με αλλόγλωσσους μαθητές που λαμβάνουν στήριξη Χορηγητές	Μιλώ Ελληνικά Ι-τεύχος Α'	Ερωτηματολόγιο μαθητή Ερωτηματολόγιο Α' Ερωτηματολόγιο Γ' Ερωτηματολόγιο χορήγησης Τεύχους Α'
B'	Αλλόγλωσσοι μαθητές Β'-Στ' τάξης που συγκέντρωσαν υψηλή βαθμολογία στο Τεύχος Α' Εκπαιδευτικοί όλων των τάξεων με αλλόγλωσσους μαθητές που λαμβάνουν στήριξη. Χορηγητές	Μιλώ Ελληνικά Ι-τεύχος Β'	- Ερωτηματολόγιο Β' Ερωτηματολόγιο χορήγησης Τεύχους Β'

Στη δεύτερη φάση (Νοέμβριος 2014), χορηγήθηκε από εξωτερικούς χορηγητές το τεύχος Β' στους μαθητές που συγκέντρωσαν υψηλή βαθμολογία στο τεύχος Α'. Επίσης, χορηγήθηκε στους εκπαιδευτικούς το Ερωτηματολόγιο Β', με επικέντρωση στο περιεχόμενο, στον βαθμό δυσκολίας και στην καταλληλότητα του τεύχους Β' για την ηλικία και το επίπεδο ελληνομάθειας των μαθητών. Τέλος, όπως και στην πρώτη φάση, συμπληρώθηκε ερωτηματολόγιο χορήγησης.

Η χορήγηση των δοκιμίων και ερωτηματολογίων, η διόρθωση και η καταχώριση των δεδομένων στις δύο φάσεις έγινε σύμφωνα με συγκεκριμένες οδηγίες. Για τον σκοπό αυτόν, αναπτύχθηκαν σχετικοί οδηγοί από το ΚΕΕΑ και οργανώθηκαν ειδικά σεμινάρια για διασφάλιση συνέπειας και ομοιομορφίας στις διαδικασίες. Επίσης, δημιουργήθηκε ηλεκτρονική πλατφόρμα επικοινωνίας διορθωτών για επίλυση αποριών καθ' όλη τη διάρκεια της διαδικασίας. Για τη στατιστική επεξεργασία των ποσοτικών δεδομένων εφαρμόστηκαν διάφορες μέθοδοι ανάλυσης (περιγραφικοί δείκτες, chi-square test, μοντέλα Rasch) στα λογισμικά SPSS και Quest, ενώ για τα ποιοτικά χρησιμοποιήθηκε η μέθοδος της ανάλυσης περιεχομένου.

3.3. Οι συμμετέχοντες

Στην έρευνα συμμετείχαν μαθητές Β'-Στ' τάξης δημοτικού, καταγεγραμμένοι στην πλατφόρμα της Διεύθυνσης Δημοτικής Εκπαίδευσης κατά τη σχολική χρονιά 2014-

2015 για ενισχυτική διδασκαλία στην εκμάθηση της ελληνικής γλώσσας. Στην πρώτη φάση συμμετείχαν 382 μαθητές, από τους οποίους 209 (55%) συμμετείχαν και στη δεύτερη φάση (Πίνακας 1). Οι μαθητές επιλέγηκαν με σκόπιμη δειγματοληψία από 17 σχολεία σε αστικές και αγροτικές περιοχές, από όλες τις επαρχίες. Ωστόσο, στην επιλογή του δείγματος δεν λήφθηκαν υπόψιν άλλες παράμετροι που αφορούσαν μεταβλητές όπως ο χρόνος παραμονής των μαθητών στην Κύπρο ή/και η ηλικία κατά την πρώτη επαφή με τα ελληνικά, καθώς αυτές οι πληροφορίες δεν ήταν διαθέσιμες.

Πίνακας 1: Συμμετέχοντες ανά τάξη

	ΦΑΣΗ Α΄		ΦΑΣΗ Β΄	
	Συχνότητα	Ποσοστό	Συχνότητα	Ποσοστό
Β΄	118	30,9	42	20,1
Γ΄	62	16,2	24	11,5
Δ΄	65	17,0	45	21,5
Ε΄	65	17,0	41	19,6
Στ΄	72	18,8	57	27,3
Σύνολο	382	100,0	209	100,0

Με βάση στοιχεία που συλλέγηκαν στο πλαίσιο της έρευνας, η αντιπροσώπευση των δύο φύλων και στις δύο φάσεις ήταν περίπου ίση (Πίνακας 2). Εφόσον η μεγάλη πλειοψηφία (82%) είχε ως μητρική μία γλώσσα από την κατηγορία των ινδοευρωπαϊκών γλωσσών, εφαρμόστηκε δεύτερη κατηγοριοποίηση με διαχωρισμό των γλωσσών αυτών σε επιμέρους κατηγορίες, από τις οποίες η Νότια Σλαβική, η Δυτική Γερμανική, η Ανατολική Σλαβική και η Ρομανική συγκέντρωσαν ποσοστό γύρω στο 20% έκαστη. Ο συνολικός χρόνος παραμονής των μαθητών στην Κύπρο κυμάνθηκε από 1 μέχρι 12 χρόνια, με μέσο όρο τα 4,6 χρόνια, ενώ για πέραν τους μισούς μαθητές (54%) δεν δόθηκαν από τους εκπαιδευτικούς πληροφορίες για το έτος εισόδου.

Πίνακας 2: Δημογραφικά στοιχεία συμμετεχόντων

	Φάση Α΄	Φάση Β΄
<i>Φύλο</i>		
Αγόρια	50,0%	45,0%
Κορίτσια	50,0%	55,0%
<i>Μητρική γλώσσα</i>		
Ινδοευρωπαϊκή	81,5%	81,7%
<i>Δυτική Γερμανική</i>	19,7%	16,8%
<i>Νότια Σλαβική</i>	19,4%	19,8%
<i>Ανατολική Σλαβική</i>	19,9%	23,9%
<i>Ρομανική</i>	17,7%	18,3%
<i>Άλλη ινδοευρωπαϊκή</i>	4,8%	2,9%
Αφροασιατική	14,0%	11,7%
Άλλη	4,5%	6,6%

4. Τα κύρια αποτελέσματα της έρευνας

4.1. Εγκυρότητα βαθμολόγησης

Ο Απλός Ποσοστιαίος Βαθμός Συμφωνίας (ΑΠΒΣ) μεταξύ των διορθωτών για το τεύχος Α΄ κυμάνθηκε από 78 έως 100%, με μέσο όρο συμφωνίας 97% και για το τεύχος Β΄ από 60 έως 100%, με μέσο όρο συμφωνίας 93%. Σε έξι δοκίμια του τεύχους Β΄ το ποσοστό συμφωνίας ήταν κάτω από 90%, στοιχείο που προβληματίζει, καθώς οι αποκλίσεις μπορεί να θεωρηθούν μεγάλες. Οι αποκλίσεις αυτές παρατηρήθηκαν κυρίως στην αξιολόγηση της παραγωγής γραπτού λόγου, γεγονός που χρήζει περαιτέρω διερεύνησης με σκοπό να διακριβωθεί αν αυτές οι αποκλίσεις οφείλονται σε ασάφειες της κλειδας διόρθωσης ή στη μη κατανόηση της κλειδας από συγκεκριμένους διορθωτές. Παράλληλα, ο δείκτης K για τα δοκίμια του τεύχους Α΄ κυμάνθηκε από 0,90 έως 1,00 και για τα δοκίμια του τεύχους Β΄ από 0,43 έως 1,00, ενώ σε όλες τις περιπτώσεις ήταν στατιστικά σημαντικός. Σε τρεις περιπτώσεις σημειώθηκαν τιμές που κατέδειξαν μέτριο βαθμό συμφωνίας, ενώ στις υπόλοιπες ο βαθμός συμφωνίας ήταν υψηλός έως πολύ υψηλός.

Ακολούθως, μέσω αναλύσεων με το μοντέλο Rasch, εξετάστηκε ο βαθμός στον οποίο τα δεδομένα ικανοποιούσαν τις προαπαιτούμενες παραδοχές του μοντέλου με τη χρήση των δύο δεικτών Infit Mean-Square (MNSQ) και Outfit Mean-Square (MNSQ). Τα δεδομένα φάνηκε να ικανοποιούν τις παραδοχές του μοντέλου σε μεγάλο βαθμό με τους δύο δείκτες και στις δύο περιπτώσεις να είναι κοντά στην προσδοκώμενη τιμή,

δηλαδή τη μονάδα (Πίνακας 3). Τα έργα των δύο τευχών 'συμπεριφέρονταν' όπως αναμενόταν, ενώ φάνηκε να συμβάλλουν στη μέτρηση μιας διάστασης (δηλαδή, της γλωσσικής ικανότητας) και συμβάλλοντας στη διάκριση ατόμων διαφορετικών ικανοτήτων. Συνεπώς, για την ανάλυση των δεδομένων χρησιμοποιήθηκε ένα έγκυρο μοντέλο μέτρησης. Ο Πίνακας 3 παρουσιάζει επίσης περιγραφικά στοιχεία των δύο βασικών παραμέτρων του μοντέλου, δηλαδή του βαθμού δυσκολίας των έργων και του βαθμού ικανότητας των μαθητών. Ο μέσος όρος του βαθμού δυσκολίας των έργων προκαθορίζεται από το λογισμικό Quest στα 0 logits (default setting). Για τα άτομα που συμπλήρωσαν το τεύχος Α' η μέτρηση του βαθμού ικανότητας των μαθητών (1,9 logits) ήταν πολύ υψηλότερη από τον αντίστοιχο μέσο όρο του βαθμού δυσκολίας των έργων, συνεπώς τα έργα θεωρήθηκαν 'εύκολα' για την πλειοψηφία. Αυτό επιβεβαιώνεται και από τον μεγάλο αριθμό ατόμων (95) που συγκέντρωσαν την υψηλότερη δυνατή βαθμολογία (perfect scores), δίνοντας μόνο σωστές απαντήσεις. Παρόμοια μέτρηση για το τεύχος Β' με τον μέσο όρο του βαθμού ικανότητας των ατόμων στα 1,58 logits καταδεικνύει δοκίμιο πιο 'δύσκολο' από το Α', ενώ μόνο ένα άτομο εξασφάλισε την υψηλότερη δυνατή βαθμολογία. Η αξιοπιστία των μετρήσεων του βαθμού δυσκολίας των έργων και της ικανότητας των ατόμων (που ερμηνεύεται σε κλίμακα από το μηδέν μέχρι το ένα) προέκυψε αρκετά υψηλή και στις δύο περιπτώσεις (>0,85).

Πίνακας 3: Περιγραφικά στοιχεία για τα έργα των δοκιμίων «Μιλιάς Ελληνικά Ι» και τους μαθητές, με βάση το μοντέλο Rasch

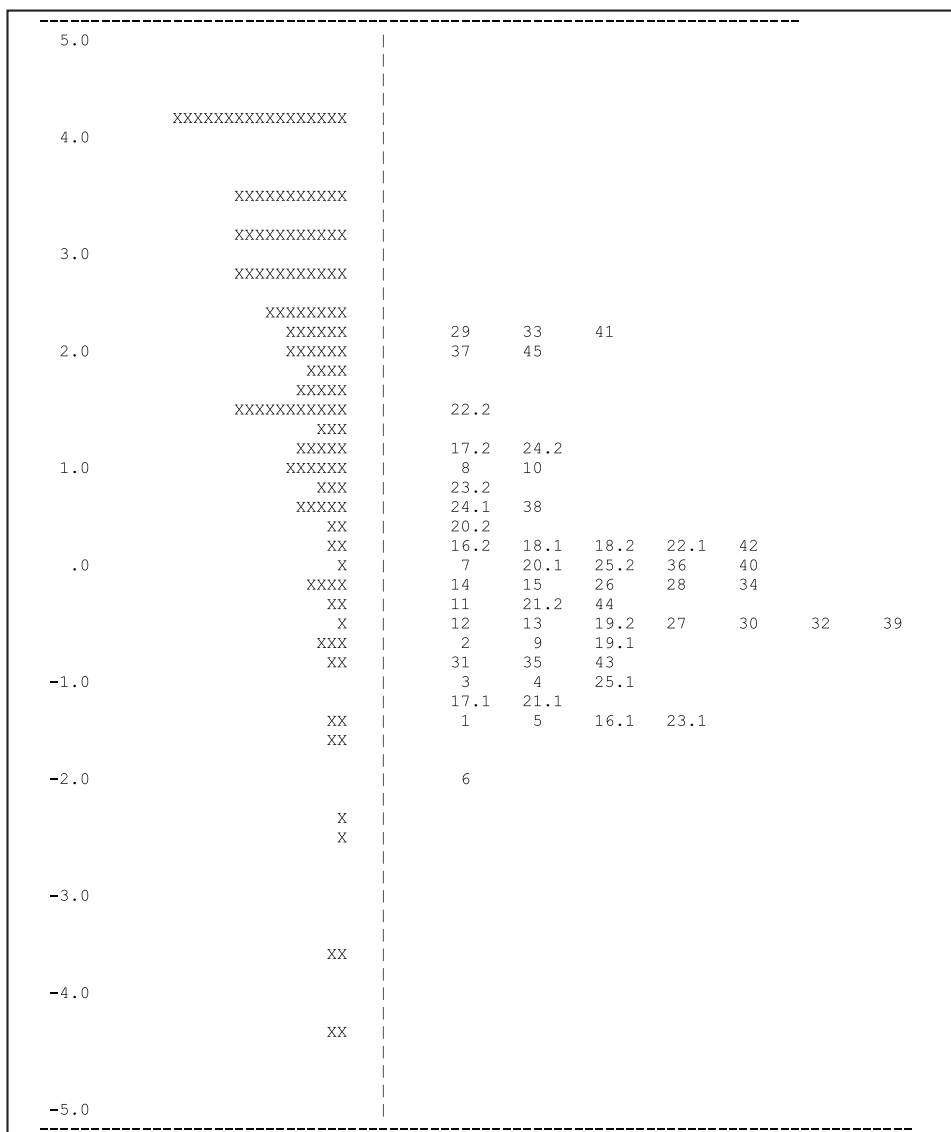
	Τεύχος Α'		Τεύχος Β'	
	Έργα	Άτομα	Έργα	Άτομα
Βαθμός δυσκολίας/ικανότητας Μέσος Όρος (Τ.Α.*)	0,00 (0,99) logits	1,90 (1,80) logits	0,00 (1,48) logits	1,58 (1,32) logits
Infit MNSQ (Τ.Α.)	1,00 (0,21)	1,01 (0,25)	1,00 (0,12)	1,00 (0,32)
Outfit MNSQ (Τ.Α.)	1,04 (0,53)	1,04 (0,81)	0,93 (0,36)	0,93 (0,64)
Αξιοπιστία μετρήσεων	0,93	0,89	0,95	0,88

*Τ.Α.= Τυπική Απόκλιση

Τα Διαγράμματα 2 και 3 παρουσιάζουν την ισοδιαστημική κλίμακα ελληνομάθειας με κοινή μονάδα μέτρησης (logits) για τα τεύχη Α' και Β' αντίστοιχα. Η τοποθέτηση όλων των έργων και μαθητών πάνω στην ίδια κλίμακα επέτρεψε την άμεση σύγκριση

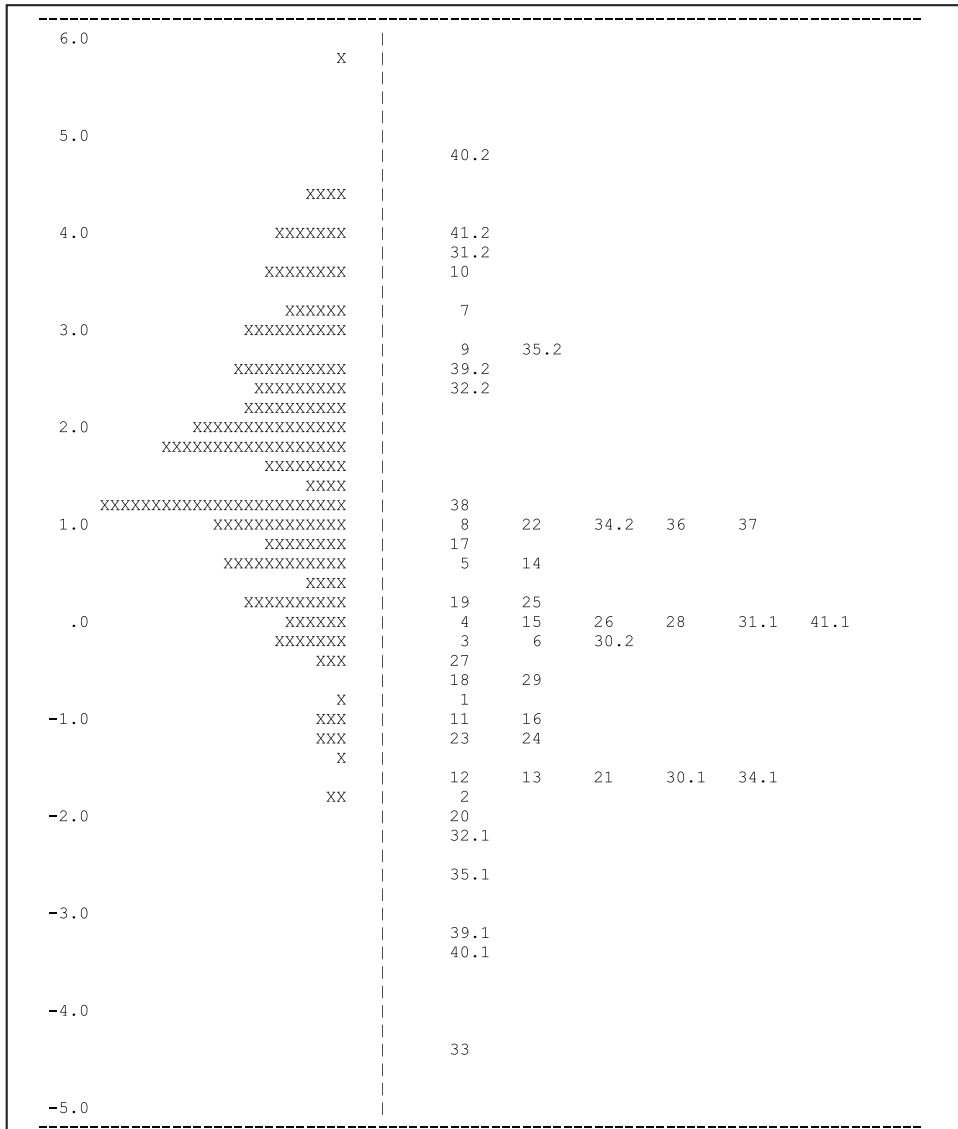
των δύο κατανομών (μαθητών και έργων), για να διερευνηθεί αν τα δοκίμια ανταποκρίνονται στο αναμενόμενο επίπεδο ελληνομάθειας. Άτομα και έργα στο άνω άκρο της κλίμακας είχαν υψηλό βαθμό ικανότητας και δυσκολίας αντίστοιχα, και αντίστροφα.

Διάγραμμα 2: Ισοδιαστημική κλίμακα ελληνομάθειας, με βάση το δοκίμιο «Μιλιάς Ελληνικά Ι-τεύχος Α΄»



Κάθε X αντιπροσωπεύει 2 μαθητές

Διάγραμμα 3: Ισοδιαστημική κλίμακα ελληνομάθειας με βάση το δοκίμιο «Μιλάς Ελληνικά Ι-τεύχος Β΄»



Κάθε X αντιπροσωπεύει 1 μαθητή

Η πλειοψηφία των έργων του τεύχους Α΄ συγκεντρώθηκε απέναντι από άτομα που βρίσκονται στη μέση και προς το κάτω άκρο της κατανομής του βαθμού γλωσσικής ικανότητας των ατόμων (Διάγραμμα 2). Συνεπώς, το επίπεδο ελληνομάθειας

των ατόμων στο κάτω άκρο της κατανομής μπορούσε να μετρηθεί με μεγάλη ακρίβεια, καθώς το δοκίμιο φάνηκε να απευθύνεται σε άτομα με χαμηλό επίπεδο ελληνομάθειας (A1). Αντίθετα, το τεύχος Β' φάνηκε να στοχεύει σε άτομα διαφόρων επιπέδων ελληνομάθειας, καλύπτοντας ολόκληρη την κατανομή του βαθμού γλωσσικής ικανότητας (Διάγραμμα 3) σε συνάφεια με τον προσανατολισμό του δοκιμίου που απευθύνεται σε εύρος επιπέδων (A2-B1). Η πλειοψηφία των έργων συγκεντρώθηκε απέναντι από άτομα που βρίσκονται στη μέση και προς τα κάτω της κατανομής του βαθμού γλωσσικής ικανότητας, ενώ το επίπεδο ελληνομάθειας των ατόμων στο κάτω μέρος και προς το μέσο της κατανομής μπορούσε να μετρηθεί με μεγαλύτερη ακρίβεια.

4.2. Εγκυρότητα συνεπειών

Η διερεύνηση για πιθανό σχηματισμό κενών μεταξύ των έργων σε μια ισοδιαστημική κλίμακα ικανότητας θεωρείται σημαντική, καθώς διασφαλίζει επαρκή κάλυψη όλων των πτυχών της υπό διερεύνηση μεταβλητής (αντιπροσωπευτικότητα περιεχομένου). Ο σχηματισμός κενών στα έργα του τεύχους Α' εντοπίστηκε στο άνω άκρο της κλίμακας όπως αναμενόταν (Διάγραμμα 2), ενώ για το τεύχος Β' στο διάστημα 1,5-2,0 logits στην κλίμακα (Διάγραμμα 3), επομένως το τεύχος Β' θα μπορούσε να ενισχυθεί με έργα μέτριου- υψηλού βαθμού δυσκολίας.

Η διακύμανση στην επίδοση που δεν οφείλεται στην υπό αξιολόγηση μεταβλητή (construct irrelevant variance) αποτελεί σημαντική απειλή για ένα δοκίμιο, αλλά εφόσον οι δύο δείκτες Infit Mean-Square (MNSQ) και Outfit Mean-Square (MNSQ) είχαν τιμές κοντά στη μονάδα (Πίνακας 3), τα έργα των δύο δοκιμίων δεν φάνηκε να αξιολογούν 'ανεπιθύμητες' διαστάσεις.

4.3. Εγκυρότητα κριτηρίου

Για μεγάλο αριθμό μαθητών (37% τεύχος Α', 34% τεύχος Β'), οι εκπαιδευτικοί δεν έδωσαν εκτιμήσεις για το επίπεδο ελληνομάθειας τους. Σε 162 περιπτώσεις (67%) συμφώνησαν με τις μετρήσεις του τεύχους Α', σε 61 (25%) οι εκτιμήσεις τους υπερέβησαν αυτές του δοκιμίου και σε 19 (8%) ήταν χαμηλότερες (Πίνακας 4). Για το τεύχος Β' οι εκπαιδευτικοί συμφώνησαν με τις μετρήσεις του δοκιμίου σε 131 περιπτώσεις (62%), σε 10 (5%) οι εκτιμήσεις τους υπερέβησαν αυτές του δοκιμίου, ενώ σε 70 (33%) ήταν χαμηλότερες. Συνολικά, οι εκπαιδευτικοί στη μεγάλη πλειοψηφία τους έτειναν να συμφωνούν με τα αποτελέσματα των δοκιμίων, ενώ για σχετικά μεγάλο αριθμό περιπτώσεων (33% και 38% για τεύχη Α' και Β', αντίστοιχα) εντοπίστηκαν αποκλίσεις. Οι αποκλίσεις αυτές προβληματίζουν σχετικά με το αν είναι σε θέση οι εκπαιδευτικοί να αξιολογήσουν το επίπεδο ελληνομάθειας των μαθητών τους στην κλίμακα αξιολόγησης του ΚΕΠΑ.

Πίνακας 4: Επίπεδο ελληνομάθειας με βάση τις απόψεις των εκπαιδευτικών και του δοκιμίου «Μιλιάς Ελληνικά-τεύχος Α΄

		Επίπεδο Ελληνομάθειας με βάση τις απόψεις των εκπαιδευτικών				
		<A1	A1	A2	B1	Σύνολο
Επίπεδο Ελληνομάθειας με βάση το τεύχος Α΄	<A1	3	20	2	1	26
	A1	4	33	29	9	75
	>A1	1	14	59	67	141
Σύνολο		8	67	90	77	242
Επίπεδο Ελληνομάθειας με βάση το τεύχος Β΄	<A2	0	1	1	1	3
	A2	1	13	30	8	52
	B1	0	5	51	100	156
Σύνολο		1	19	82	109	211

Σημείωση: Έντονη: Εκτιμήσεις εκπαιδευτικών = Εκτιμήσεις δοκιμίου, Ανοικτή γκρίζα σκίαση: Εκτιμήσεις εκπαιδευτικών < Εκτιμήσεις δοκιμίου, Σκούρα γκρίζα σκίαση: Εκτιμήσεις εκπαιδευτικών < Εκτιμήσεις δοκιμίου.

4.4. Εγκυρότητα πλαισίου

Ο χρόνος χορήγησης των 90 λεπτών που εισηγείται η ομάδα συγγραφής, με βάση τις απαντήσεις των χορηγητών, κρίθηκε ικανοποιητικός σε όλες τις περιπτώσεις. Εφόσον οι χορηγητές σημείωναν σε κάθε δοκίμιο και την ώρα παράδοσής του, διαφάνηκε ότι κανένας μαθητής δεν εξάντλησε τον χρόνο. Σχετικά με τα αρχεία ήχου δεν έγινε οποιοδήποτε σχόλιο και όλοι οι χορηγητές απάντησαν ότι αυτά ακούγονταν καθαρά. Στις περισσότερες ομάδες χορήγησης του τεύχους Α΄ (77%) και Β΄ (60%), οι χορηγητές δέχθηκαν απορίες, που αφορούσαν κυρίως στην επεξήγηση των οδηγιών. Για το τεύχος Α΄ δεν αναφέρθηκε κάποια άσκηση που δυσκόλεψε τους μαθητές, ενώ για το τεύχος Β΄ η πλειοψηφία (60%) εντόπισε την άσκηση 2 (κατανόηση προφορικού λόγου) ως δύσκολη για δύο κυρίως λόγους: (α) τα αρχείου ήχου ακούγονταν πολύ γρήγορα και (β) το περιεχόμενό της δεν ήταν κατάλληλο για μαθητές μικρότερων τάξεων. Οι χορηγητές δεν ανέφεραν οποιαδήποτε άλλη παρατήρηση.

4.5. Εγκυρότητα περιεχομένου

Η πλειοψηφία των εκπαιδευτικών θεώρησε το τεύχος Α΄ κατάλληλο από άποψη έκτασης, διατύπωσης των ερωτήσεων και εικόνων, τόσο για το συγκεκριμένο επίπεδο ελληνομάθειας (A1) όσο και για την ηλικία των μαθητών (Πίνακας 5). Σε στατιστικό έλεγχο (chi-square test) για διερεύνηση των απαντήσεων των εκπαιδευτικών

για το περιεχόμενο του δοκιμίου ως προς την τάξη που διδάσκουν, μόνο σε μία ερώτηση σημειώθηκε διαφοροποίηση σε στατιστικά σημαντικό βαθμό. Συγκεκριμένα, σημαντικό ποσοστό (30-40%) των εκπαιδευτικών που διδάσκουν σε μεγαλύτερες τάξεις (Δ'-ΣΤ') θεώρησε την έκταση του δοκιμίου από καθόλου έως λίγο κατάλληλη για τους μαθητές της τάξης τους, σε αντίθεση με τους εκπαιδευτικούς μικρότερων τάξεων (10%). Αυτό θα μπορούσε να είναι αναμενόμενο, εφόσον οι μαθητές μεγαλύτερης ηλικίας εργάζονται συνήθως σε μεγαλύτερης έκτασης δοκίμια.

Πίνακας 5: Αξιολόγηση περιεχομένου Τεύχους Α' από τους εκπαιδευτικούς

	Καθόλου (%)	Λίγο (%)	Αρκετά (%)	Πολύ (%)
Η έκταση του δοκιμίου είναι:				
- κατάλληλη για μαθητές με επίπεδο ελληνομάθειας Α1.	-	10,8	55,1	34,1
- κατάλληλη για την ηλικία των μαθητών της τάξης μου.	8,1	17,8	34,6	39,5
Η διατύπωση των ερωτήσεων:				
- είναι σαφής.	1,6	12,0	40,2	46,2
- περιέχει δύσκολες λέξεις, όσον αφορά στην κατανόηση από μαθητές με επίπεδο ελληνομάθειας Α1.	34,6	40,2	16,8	8,4
- περιέχει δύσκολες λέξεις, όσον αφορά στην κατανόηση για την ηλικία των μαθητών της τάξης μου.	53,5	30,3	12,4	3,8
Οι εικόνες του δοκιμίου:				
- έχουν άμεση σχέση και είναι σε πλήρη αντιστοιχία με το λεκτικό κομμάτι των ερωτήσεων.	-	3,8	32,3	64,0
- είναι υποβοηθητικές, όσον αφορά στην κατανόηση των ερωτήσεων από μαθητές με επίπεδο ελληνομάθειας Α1.	0,5	6,5	35,5	57,5
- είναι κατάλληλες για την ηλικία των μαθητών της τάξης μου.	4,3	12,4	27,4	55,9

Σε σχέση με το τεύχος Β', η πλειοψηφία των εκπαιδευτικών το θεώρησε κατάλληλο

από άποψη έκτασης, διατύπωσης των ερωτήσεων και εικόνων και για τα δύο επίπεδα ελληνομάθειας (A2 και B1), αλλά και για την ηλικιακή ομάδα των μαθητών (Πίνακας 6).

Πίνακας 6: Αξιολόγηση περιεχομένου Τεύχους Β' από εκπαιδευτικούς

	Καθόλου (%)	Λίγο (%)	Αρκετά (%)	Πολύ (%)
Η έκταση του δοκιμίου είναι:				
– κατάλληλη για μαθητές με επίπεδο ελληνομάθειας A2	11,3	34,4	46,2	8,1
– κατάλληλη για μαθητές με επίπεδο ελληνομάθειας B1	2,2	15,1	57,0	25,8
– κατάλληλη για την ηλικία των μαθητών της τάξης μου	8,2	25,7	39,3	26,8
Η διατύπωση των ερωτήσεων:				
– είναι σαφής	-	9,1	53,2	37,6
– περιέχει δύσκολες λέξεις, όσον αφορά στην κατανόηση από μαθητές με επίπεδο ελληνομάθειας A2	9,1	49,2	33,2	8,6
– περιέχει δύσκολες λέξεις, όσον αφορά στην κατανόηση από μαθητές με επίπεδο ελληνομάθειας B1	35,1	43,6	18,6	2,7
– περιέχει δύσκολες λέξεις, όσον αφορά στην κατανόηση για την ηλικία των μαθητών της τάξης μου	34,4	36,6	22,0	7,0
Οι εικόνες του δοκιμίου:				
– έχουν άμεση σχέση και είναι σε πλήρη αντιστοιχία με το λεκτικό κομμάτι των ερωτήσεων	-	5,3	52,4	42,2
– είναι υποβοηθητικές, όσον αφορά στην κατανόηση των ερωτήσεων από μαθητές με επίπεδο ελληνομάθειας A2	0,5	17,1	52,4	29,9
– είναι υποβοηθητικές, όσον αφορά στην κατανόηση των ερωτήσεων από μαθητές με επίπεδο ελληνομάθειας B1	-	5,9	62,9	31,2
– είναι κατάλληλες για την ηλικία των μαθητών της τάξης μου	1,6	12,4	45,7	40,3

Σε τρεις ερωτήσεις σημειώθηκαν στατιστικά σημαντικές διαφοροποιήσεις (chi-square test) στις απαντήσεις, ανάλογα με την τάξη. Για την καταλληλότητα της έκτασης του δοκιμίου η πλειοψηφία των εκπαιδευτικών της Β' τάξης απάντησε 'λίγο', ενώ των τάξεων Γ'-Ε' 'αρκετά' και της ΣΤ' τάξης 'πολύ'. Για την παρουσία δύσκολων λέξεων, η πλειοψηφία των Β', Δ' και Ε' τάξεων απάντησε 'λίγο', ενώ των Γ' και ΣΤ' τάξεων 'καθόλου'. Τέλος, για το κατά πόσον οι εικόνες είναι υποβοηθητικές στην κατανόηση των ερωτήσεων από μαθητές επιπέδου Β1, η συντριπτική πλειοψηφία (60-75%) των εκπαιδευτικών Β', Γ' και Δ' τάξης απάντησε 'αρκετά', ενώ των τάξεων Ε' και ΣΤ' διχάστηκαν ανάμεσα στο 'αρκετά' και στο 'πολύ'. Τα αποτελέσματα αυτά δεν προκαλούν ιδιαίτερη έκπληξη, εφόσον το τεύχος Β' αφορά στα δύο ανώτερα επίπεδα ελληνομάθειας, συνεπώς ανταποκρίνεται καλύτερα σε μεγαλύτερης ηλικίας μαθητές.

5. Συζήτηση αποτελεσμάτων – συμπεράσματα

Η απουσία αντικειμενικού τρόπου διάγνωσης του επιπέδου ελληνομάθειας των μαθητών που δεν έχουν ως μητρική γλώσσα την ελληνική και φοιτούν σε δημόσια δημοτικά σχολεία στην Κύπρο κατέστησε τη διεξαγωγή της παρούσας έρευνας εγκυροποίησης αναγκαία. Συμπερασματικά, η διερεύνηση κατέδειξε τα δύο τεύχη του δοκιμίου («Μιλάς Ελληνικά Ι») κατάλληλα για διάγνωση του επιπέδου ελληνομάθειας μαθητών δημοτικού σχολείου (ηλικίας 7-11 χρόνων) στην Κύπρο, με ανάγκη για μικρές προσαρμογές, ιδιαίτερα σε σχέση με το τεύχος Β'.

Για το τεύχος Α', η διπλή διόρθωση δοκιμίων κατέδειξε υψηλό βαθμό συμφωνίας μεταξύ των διορθωτών (μέσος ΑΠΒΣ 97%, δείκτης Κ 0,90-1), υποδεικνύοντας ότι η κλειδα διόρθωσης ήταν κατανοητή και εφαρμόστηκε με συνέπεια (εγκυρότητα βαθμολόγησης). Επίσης, τα δεδομένα ικανοποιούσαν τις παραδοχές του μοντέλου Rasch (εγκυρότητα συνεπειών). Ο βαθμός δυσκολίας των έργων ήταν ο αναμενόμενος, καθώς η πλειοψηφία τους φάνηκε να απευθύνεται σε άτομα με χαμηλό επίπεδο ελληνομάθειας. Παράλληλα, η αξιοπιστία των μετρήσεων ήταν αρκετά υψηλή. Εφόσον το δοκίμιο μετρούσε μόνο μία διάσταση (ελληνομάθεια), η χρήση μίας συνολικής βαθμολογίας στο δοκίμιο για κάθε μαθητή ήταν δικαιολογημένη. Επίσης, διασφαλίστηκαν οι δύο παράμετροι εγκυρότητας των αποτελεσμάτων (Messick, 1995) για την αντιπροσωπευτικότητα του περιεχομένου και τη διακύμανση στην επίδοση που οφείλεται μόνο στην υπό αξιολόγηση μεταβλητή. Στη μεγάλη πλειοψηφία των περιπτώσεων (67%), οι εκπαιδευτικοί συμφωνούσαν με τα αποτελέσματα του δοκιμίου (εγκυρότητα κριτηρίου), ενώ οι διαφωνίες ενδέχεται να σχετίζονται με το βαθμό εξοικείωσης των εκπαιδευτικών με την κλίμακα αξιολόγησης ΚΕΠΑ. Για έναν αρκετά μεγάλο αριθμό περιπτώσεων (37%) οι εκπαιδευτικοί δεν έδωσαν εκτιμήσεις για το επίπεδο ελληνομάθειας των μαθητών τους, λόγω ενδεχόμενης δυσκολίας για

τέτοια αποτίμηση. Τα αποτελέσματα αυτά καταδεικνύουν την ανάγκη για πιο 'αντικειμενικό' μηχανισμό διάγνωσης του βαθμού γνώσης της ελληνικής γλώσσας, εφόσον αυτή δεν μπορεί να επαφίεται αποκλειστικά στις εκτιμήσεις των εκπαιδευτικών. Ο χρόνος των 90 λεπτών κρίθηκε ικανοποιητικός (*εγκυρότητα πλαισίου*). Γενικά, δεν προέκυψαν ζητήματα κατά τη διάρκεια της χορήγησης, ενώ η μόνη δυσκολία αφορούσε στην κατανόηση των οδηγιών των ασκήσεων. Πρέπει να αναφερθεί όμως ότι οι χορηγητές είχαν την ευχέρεια επεξήγησης των οδηγιών. Τέλος, για την *εγκυρότητα περιεχομένου* η πλειοψηφία των εκπαιδευτικών θεώρησε το τεύχος κατάλληλο από άποψη έκτασης, διατύπωσης ερωτήσεων και εικόνων, τόσο για το επίπεδο ελληνομάθειας (A1) όσο και για την ηλικιακή ομάδα των μαθητών του δημοτικού σχολείου, ενώ εκπαιδευτικοί μεγαλύτερων τάξεων (Ε' - ΣΤ') έκριναν την έκταση σχετικά μικρή, όπως αναμενόταν. Οι θετικές απόψεις των εκπαιδευτικών θεωρούνται πολύ σημαντικές, καθώς αυτοί θα χρησιμοποιήσουν το εν λόγω δοκίμιο.

Για το τεύχος Β', ο βαθμός συμφωνίας μεταξύ των διορθωτών σύμφωνα με τους δύο δείκτες ήταν ικανοποιητικός (*εγκυρότητα βαθμολόγησης*), ενώ παρατηρήθηκαν κάποιες αποκλίσεις. Σε αυτές τις περιπτώσεις χρειάζεται έλεγχος της συνέπειας διόρθωσης σε συστηματική βάση, έτσι ώστε να αξιολογούνται ξανά από τρίτο ανεξάρτητο διορθωτή ή να αποκλείονται συγκεκριμένοι διορθωτές από τη διαδικασία. Τα δεδομένα ικανοποιούσαν γενικά τις παραδοχές του μοντέλου Rasch, εφόσον δεν υφίστανται οι δύο σημαντικές απειλές για την εγκυρότητα των αποτελεσμάτων (*εγκυρότητα συνεπειών*). Τα έργα στόχευσαν σε άτομα με διάφορα επίπεδα ελληνομάθειας, κάτι που συγκλίνει με τον προσανατολισμό του δοκιμίου, εφόσον καλύπτει εύρος επιπέδων ελληνομάθειας (A2-B1). Επίσης, το δοκίμιο φάνηκε να έχει μεγαλύτερο βαθμό δυσκολίας από το τεύχος Α'. Ο χρόνος συμπλήρωσης των 90 λεπτών κρίθηκε ικανοποιητικός (*εγκυρότητα πλαισίου*). Οι μόνες δυσκολίες κατά τη χορήγηση αφορούσαν στην κατανόηση των οδηγιών των ασκήσεων, αλλά και στο ηχητικό αρχείο της Άσκησης 2 που ακουγόταν πολύ γρήγορα. Για την *εγκυρότητα κριτηρίου*, οι εκπαιδευτικοί έτειναν να συμφωνούν στην αξιολόγηση του επιπέδου ελληνομάθειας σε 62% των περιπτώσεων. Και πάλι, για έναν σχετικά μεγάλο αριθμό περιπτώσεων (34%), δεν έδωσαν τις εκτιμήσεις τους αναφορικά με το επίπεδο ελληνομάθειας των μαθητών τους. Η πλειοψηφία θεώρησε το δοκίμιο κατάλληλο από άποψη έκτασης, διατύπωσης των ερωτήσεων και εικόνων για τα δύο επίπεδα ελληνομάθειας (A2 και B1), και για την ηλικιακή ομάδα των μαθητών (*εγκυρότητα περιεχομένου*). Παρόλο που η έκταση κρίθηκε αρκετά μεγάλη από εκπαιδευτικούς μικρότερων τάξεων, αυτή ενδεχομένως επηρεάστηκε από την εστίαση σε δύο επίπεδα ελληνομάθειας. Επίσης, σύμφωνα με κάποιους εκπαιδευτικούς, οι οδηγίες κρίθηκαν πολύπλοκες, με αποτέλεσμα να υπάρξουν κάποιες δυσκολίες στην προσέγγισή τους.

6. Καταληκτικά

Η διεθνής βιβλιογραφία τονίζει την ανάγκη εγκυροποίησης εργαλείων, όταν αυτά χρησιμοποιούνται σε διαφορετικό συγκείμενο από αυτό που αναπτύχθηκαν. Τα δοκίμια «Μιλιάς Ελληνικά Ι» κατασκευάστηκαν στον ελλαδικό χώρο για αξιολόγηση του βαθμού κατάρκτησης της ελληνικής γλώσσας ως δεύτερης. Η καταλληλότητά τους για χρήση στο κυπριακό πλαίσιο δεν θεωρήθηκε δεδομένη, λόγω και της γλωσσικής ιδιαιτερότητας της Κύπρου, εφόσον γίνεται χρήση τόσο της κοινής νεοελληνικής γλώσσας όσο και της κυπριακής διαλέκτου, καθιστώντας τη διενέργεια της παρούσας έρευνας απαραίτητη. Για το σκοπό αυτό, συλλέγηκαν δεδομένα από μαθητές, εκπαιδευτικούς και χορηγητές, τα οποία έτυχαν ποσοτικών και ποιοτικών αναλύσεων. Η διερεύνηση αφορούσε σε πέντε διαστάσεις (βαθμολόγηση, συνέπειες, πλαίσιο, κριτήριο, περιεχόμενο), όπως αυτές εννοιολογήθηκαν και συζητήθηκαν ενδελεχώς στο πλαίσιο της παρούσας εργασίας. Ενδεχομένως και άλλες διαστάσεις εγκυρότητας χρήζουν διερεύνησης στο μέλλον, όπως για παράδειγμα η επίδραση εξωτερικών παραγόντων στην επίδοση των μαθητών στα δύο δοκίμια, όπως το φύλο, η ηλικία, η μητρική γλώσσα κ.λπ. Όπως υποστηρίζει, ωστόσο, και ο Messick (1989), μια έρευνα εγκυροποίησης είναι αδύνατον να καλύψει όλες τις διαστάσεις που σχετίζονται με την ερμηνεία και χρήση των αποτελεσμάτων ενός εργαλείου, εφόσον πρόκειται για μια διαδικασία χωρίς τέλος.

Βιβλιογραφία

- Altman D.G. (1991) *Practical statistics for medical research*. London: Chapman and Hall.
- Bartlett, M.S. (1954) A note on the multiplying factors for various χ^2 approximation. *Journal of the Royal Statistical Society, Series B16*, 296-298.
- Brown, H.D. (1977) *Cognitive and affective characteristics of good language learners*. Paper presented at Los Angeles Second Language Acquisition, Research Forum, UCLA, February 1977.
- Βαρλοκώστα, Σ. & Λ. Τριανταφυλλίδου (2003) *Επίπεδα Γλωσσομάθειας στην Ελληνική ως Δεύτερη Γλώσσα*. Αθήνα: Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.
- Βαρλοκώστα, Σ., Μ. Μάρκου & Λ. Τριανταφυλλίδου (2001) Γλωσσομαθησιακή πορεία των παλιννοστούντων και αλλοδαπών μαθητών της δευτεροβάθμιας εκπαίδευσης: Εργαλεία αξιολόγησης της Ελληνικής ως δεύτερης γλώσσας. Στο Μ. Δαμανάκης (επιμ.), *Επίπεδα και Κριτήρια Διαπίστωσης/Πιστοποίησης της Ελληνομάθειας: Αποτελέσματα και Προβληματισμοί Ερευνητικών Προγραμμάτων*. Ρέθυμνο: Ε.ΔΙΑ.Μ.ΜΕ., Πανεπιστήμιο Κρήτης, 127-143.

- Christensen, G. & P. Stanat (2007) *Language, migration and social inequalities; Language policies and practices for helping immigrants and second-generation students succeed*. Washington, DC: Migration Policy Institute. Available at: www.migrationpolicy.org/pubs/ChristensenEducation091907.pdf [Ημερομηνία πρόσβασης 20/7/2015]
- Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20: 37-46.
- Cummins, J. (1996) *Negotiating Identities: Education for Empowerment in a Diverse Society*. Los Angeles: California Association for Bilingual Education.
- Ellis, R. (1994) *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Fleiss J.L., B. Levin & M.C. Paik (2003) *Statistical methods for rates and proportions*. Hoboken: John Wiley & Sons.
- Kaiser, H.F. (1974) An index of factorial simplicity. *Psychometrika*, 39: 31-36.
- Kelley T.L. (1927) *Interpretation of educational measurements*. Yonkers, NY, World Book Company.
- Larsen-Freeman, D. & M. Long (1991) *An Introduction to Second Language Acquisition Research*. London: Longman.
- Marcos, K. M. (1998) Learning a Second Language: What Parents Need to Know. «*National PTA Magazine*», August/September, 32-33.
- Messick, S. (1989) Validity. In R.L. Linn (Ed.), *Educational measurement*. New York: MacMillan, 13-103.
- Messick, S. (1995) Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50 (9): 741-749.
- Met, M. & N. Rhodes (1990) Priority: Instruction. Elementary school foreign language instruction: Priorities for the 1990s. *Foreign Language Annals*, 25, 433-43.
- O'Malley, J. M. & A. U. Chamot (1990) *Learning strategies in second language acquisition*. Cambridge: Cambridge: University Press.
- OECD (2009) *PISA 2009 Results: Executive Summary*. Available at: <http://www.oecd.org/pisa/pisaproducts/46619703.pdf> [Ημερομηνία πρόσβασης 9/7/2015].
- Oikonomidou, M. (2003) Multicultural education in Cyprus. *Parousia (Educational magazine of Greek scientists' association in secondary education)*, 13, 44-51.
- Shoebottom, P. (n.d.) *The factors that influence the acquisition of a second language*. Available at: <http://esl.fis.edu/teachers/support/factors.htm> [Ημερομηνία πρόσβασης 11/7/2015]

- Συμβούλιο της Ευρώπης (2001) *Κοινό πλαίσιο αναφοράς για τη γλώσσα: εκμάθηση, διδασκαλία και αξιολόγηση*. Ανάκτηση από: http://www.pi-schools.gr/lessons/english/pdf/cef_gr.pdf [Ημερομηνία πρόσβασης 20/1/2016]
- Tzeveleku M., M. Giagkou, V. Kantzou, S. Stamouli, S. Varlokosta, I. Mitzias & D. Papadopoulou (2013) Second language assessment in the Greek educational system: the case of Reception Classes. *Γλωσσολογία*, 21: 75-89.
- Τζεβελέκου, Μ., Β. Κάντζου, Σ. Σταμούλη, Β. Χονδρογιάννη, Σ. Βαρλοκώστα, Μ. Ιακώβου, Β. Λύτρα & Γ. Παπαγεωργακόπουλος (2004) *Η ελληνική ως δεύτερη γλώσσα: Κλίμακα αξιολόγησης και μέτρηση του επιπέδου γνώσης της ελληνικής στα μειονοτικά δημοτικά σχολεία της Θράκης*. Ανάκτηση από: http://www.ilsp.gr/files/web_erevna.pdf.
- Τζεβελέκου, Μ., Β. Κάντζου, Σ. Σταμούλη, Β. Χονδρογιάννη, Σ. Βαρλοκώστα, Μ. Ιακώβου, Β. Λύτρα, Α. Σιούπη & Α. Τερζή (2006) Μελέτη της ελληνομάθειας στα μειονοτικά δημοτικά σχολεία της Θράκης. *Νέες κατευθύνσεις στην εφαρμοσμένη γλωσσολογία. Πρακτικά του 13ου Διεθνούς Συνεδρίου της Ελληνικής Εταιρείας Εφαρμοσμένης Γλωσσολογίας*. Θεσσαλονίκη: University Studio Press, τόμος 9, 343-357.
- Τριάρχη-Herrmann, B. (2000) *Η διγλωσσία στην Παιδική Ηλικία - μια ψυχολογική προσέγγιση*. Αθήνα: Gutenberg.
- Van der Walt, J. L. & F. Steyn (2008) The validation of language tests. *Papers in Linguistics*, 38, 191-204.
- Weir, C. J. (2005) *Language testing and validation: An evidence-based approach*. Series: Research and Practice in Applied Linguistics. Palgrave: Macmillan.
- Wright, B. D. & J. M. Linacre (1994) Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8 (3): 370.
- Υπουργείο Παιδείας και Πολιτισμού. (χ.η.) *Διαπολιτισμική Αγωγή και Εκπαίδευση*. Ανάκτηση από: <http://www.moec.gov.cy/dde/diapolitismiki.html> [Ημερομηνία πρόσβασης 15/6/2015]
- Υπουργείο Παιδείας και Πολιτισμού. (2014) *Ενισχυτική διδασκαλία για αλλόγλωσσους μαθητές για τη σχολική χρονιά 2014-2015*. Ανάκτηση από: <http://egkyklioi.moec.gov.cy/Data/dde4153a.pdf> [Ημερομηνία πρόσβασης 15/6/2015]
- Zhuanglin, H. (1989) *Linguistics: An Introduction*. Beijing: Peking University Press.