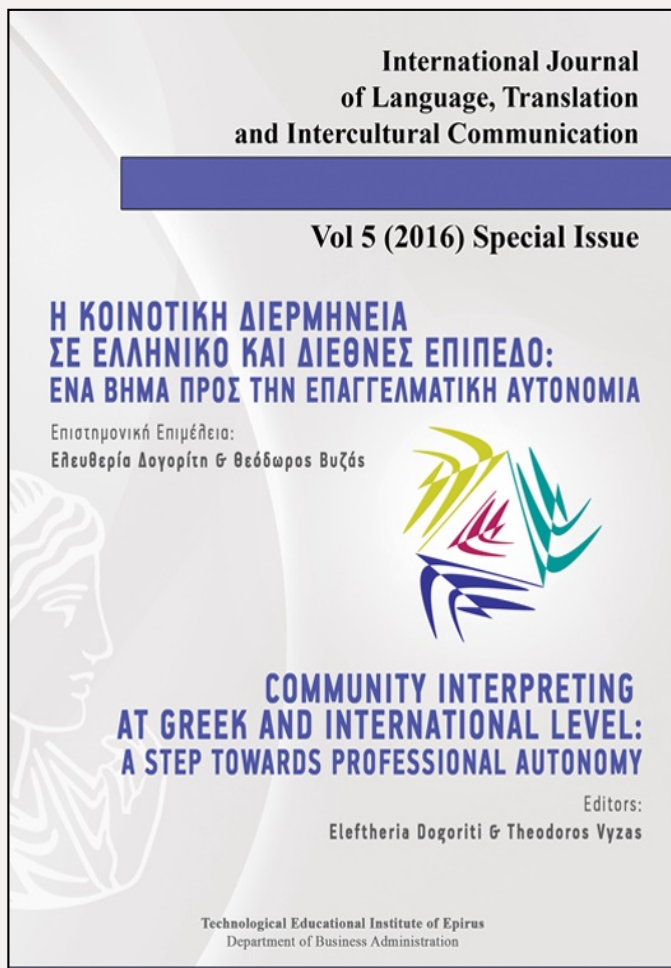


# International Journal of Language, Translation and Intercultural Communication

Vol. 5, 2016



Στατιστικές μέθοδοι ελέγχου εγκυρότητας και αξιοπιστίας ερωτηματολογίων. Η περίπτωση του CiGreece.

Iseris Georgios TEI of Epirus  
<http://dx.doi.org/10.12681/ijltic.10665>

Copyright © 2016 Georgios Iseris



## To cite this article:

Iseris, G. (2016). Στατιστικές μέθοδοι ελέγχου εγκυρότητας και αξιοπιστίας ερωτηματολογίων. Η περίπτωση του CiGreece.. *International Journal of Language, Translation and Intercultural Communication*, 5, 175-189.  
doi:<http://dx.doi.org/10.12681/ijltic.10665>

## Στατιστικές μέθοδοι ελέγχου εγκυρότητας και αξιοπιστίας ερωτηματολογίων. Η περίπτωση του CiGreece.

Γεώργιος Ίσερης  
ΤΕΙ Ηπείρου  
ginocorfatiis@gmail.com

### Περίληψη

Η συστηματική χρήση των ερωτηματολογίων σε όλων των ειδών τις έρευνες καθιστά αναγκαίο τον συνδυασμό στατιστικών ερευνητικών μεθόδων που να επιβεβαιώνουν την εγκυρότητα και την αξιοπιστία αυτών των τύπων κλιμάκων. Ο σκοπός αυτής της εργασίας είναι να παρουσιάσει τμήμα της πρότασής μας για το μεθοδολογικό σχέδιο της διεξαγωγής της ερευνητικής προσπάθειας CiGreece που έχει ως στόχο να χαρτογραφήσει για πρώτη φορά την κατάσταση στην Ελλάδα αναφορικά με την παροχή υπηρεσιών κοινοτικής διερμηνείας, να αναδείξει τις ελλείψεις τέτοιων υπηρεσιών σε ευαίσθητα περιβάλλοντα, να αυξήσει την ευαισθησία αλλά και να προμηθεύσει με τις απαραίτητες πληροφορίες που θα βοηθήσουν τις αρμόδιες κρατικές αρχές να λάβουν τα όσα και όποια πρόσφορα μέτρα για τη βελτίωση της κατάστασης παροχής υπηρεσιών κοινοτικής διερμηνείας στην Ελλάδα. Στην εργασία μας δίνουμε έμφαση στην περιγραφή των προϋποθέσεων για την εφαρμογή τεχνικών στάθμισης του ερωτηματολογίου της έρευνας μέσω της εξαγωγής του δείκτη αξιοπιστίας Cronbach's  $\alpha$  αλλά και της Διερευνητικής και Επιβεβαιωτικής ανάλυσης παραγόντων για τον έλεγχο της δομικής εγκυρότητας του. Υπογραμμίζουμε πως η συνδυαστική χρήση των περιγραφόμενων στο κύριο μέρος της παρούσας εργασίας στατιστικών μεθόδων είναι σε θέση να αποκαλύψει την ύπαρξη ασθενούς δομικής εγκυρότητας ενός ερωτηματολογίου, παρά την επιτυχή έκβαση μεμονωμένων στατιστικών ελέγχων, γεγονός που υποδεικνύει την αναγκαιότητα εφαρμογής πολλαπλών στατιστικών μεθόδων για τον έλεγχο εγκυρότητας και αξιοπιστίας των κλιμάκων μέτρησης.

*Λέξεις-κλειδιά:* Διερευνητική παραγοντική ανάλυση, Επιβεβαιωτική παραγοντική ανάλυση, Συντελεστής σωτηρικής συνέπειας Cronbach's Alpha

### Εισαγωγή

Μια πρώτη προσπάθεια χαρτογράφησης της κατάστασης που αφορά στην παροχή υπηρεσιών κοινοτικής διερμηνείας στην Ελλάδα ανελήφθη, στο πλαίσιο του προγράμματος CiGreece, από το Τμήμα Διοίκησης Επιχειρήσεων του ΤΕΙ ΗΠΕΙΡΟΥ και συγκεκριμένα από την επιστημονική ομάδα της κατεύθυνσης Εφαρμοσμένων ξένων γλωσσών στη διοίκηση και την οικονομία, με στόχο να διαπιστωθούν οι πιθανές ελλείψεις υπηρεσιών κοινοτικής διερμηνείας σε περιβάλλοντα όπως στα κέντρα φιλοξενίας προσφύγων, στις υγειονομικές υπηρεσίες, στις διοικητικές υπηρεσίες, στα σχολικά περιβάλλοντα, στα σώματα ασφαλείας, κ.α. και να συλλεχθούν οι απαραίτητες πληροφορίες που θα βοηθήσουν τις αρμόδιες κρατικές αρχές να λάβουν τα όσα και όποια πρόσφορα μέτρα για τη βελτίωση της κατάστασης.

Στο πλαίσιο αυτό συντάχθηκε και η παρούσα εργασία η οποία αποτελεί τμήμα, λόγω του επιτρεπόμενου εκ του σκοπού της σύνταξής της εύρους, του μεθοδολογικού σχεδίου που προτάθηκε για τη διεξαγωγή της ανωτέρω ερευνητικής προσπάθειας και αναφέρεται πρωτίστως στα χαρακτηριστικά και τις μεθόδους ελέγχου της αξιοπιστίας και εγκυρότητας του ερωτηματολογίου που αφορά στους αλλοδαπούς, ένα από τα δύο υπό δημιουργία ερευνητικά εργαλεία συλλογής των δεδομένων της έρευνας.

## 1. Περί ερωτηματολογίων ο λόγος

Είναι παγκοίως γνωστό πως τα τελευταία χρόνια σε ευρεία χρήση ως βασικά ερευνητικά εργαλεία συλλογής δεδομένων υιοθετούνται τα ερωτηματολόγια είτε σε αυτοσυμπληρούμενη εκδοχή είτε ως βάση για δομημένη συνέντευξη. Μία από τις πρώτες αποφάσεις που θα πρέπει να λάβει ο ερευνητής κατά την κατασκευή ενός ερωτηματολογίου είναι εκείνη της επιλογής του ανοικτού ή κλειστού τύπου των ερωτημάτων (ή συνδυασμούς των). Αν δηλαδή θα δίνεται η δυνατότητα στον ερωτώμενο να απαντήσει ελεύθερα ή να επιλέξει από μια σειρά προσφερόμενων απαντήσεων. Η πρώτη επιλογή, παρά το πλήθος των θετικών χαρακτηριστικών της (Bradburn, Sudman & Wansink, 2004: 154), δεν είναι ελκυστική τόσο για τον ερωτώμενο, ένεκα των χρονοβόρων απαιτήσεών της για να απαντηθούν τέτοιου τύπου ερωτήματα, όσο και για τον ερευνητή συνεπεία του περιοριστικού της χαρακτήρα σχετικά με το εύρος των θεμάτων που ένα ερωτηματολόγιο απαρτιζόμενο από παρόμοιου τύπου ερωτήσεις μπορεί να συμπεριλάβει αλλά και των δυσχερειών κωδικοποίησής τους (Dornyei, 2003: 47· Bradburn, Sudman & Wansink, 2004: 154). Στη δεύτερη περίπτωση, του κλειστού τύπου ερωτημάτων, ο ερευνητής, με συγκεκριμένο αριθμό εκ των προτέρων κωδικοποιημένων ερωτημάτων, ευελπιστεί, στις περισσότερες των περιπτώσεων, να ποσοτικοποιήσει διάφορες εξωτερικεύσεις του ανθρώπινου ψυχισμού. Το υποκείμενο της έρευνας καλείται να συγκρίνει τη συμπεριφορά ή τις συνήθειές του με αυτές που περιγράφονται στα ερωτήματα και να εκφράσει μια αξιολογική κρίση. Σε πολλά ερωτήματα τοποθετείται καταφατικά ή αρνητικά (διωνυμικού τύπου ερωτήματα) ως προς το εάν τα συγκεκριμένα ερωτήματα τον εκφράζουν ή περιγράφουν επακριβώς τη στάση του ως προς το υπό εκάστου ερωτήματος περιγραφόμενο θέμα. Συχνότερα, βεβαίως, στα ερωτηματολόγια συναντώνται ερωτήματα στα οποία ο αποκρινόμενος προτρέπει να δηλώσει τον βαθμό συμφωνίας του ως προς μια δήλωση που περιγράφει μια συγκεκριμένη συμπεριφορά (του) με ένα περιγραφικό λήμμα ή με ένα αριθμητικό συχνότητας (κλίμακες τύπου Likert 1932). Τα ερωτήματα αυτού του τύπου, λόγω των περισσότερων αριθμητικά αλλά, κυρίως, των αντιπροσωπευτικότερων, ως προς τις απαντήσεις, προσφερόμενων επιλογών, τείνουν να περιθωριοποιήσουν τα διωνυμικού τύπου ερωτήματα.

Στις πλείστες των περιπτώσεων μια κλίμακα μέτρησης, όπως το ερωτηματολόγιο, αποτελείται από ερωτήματα τα οποία συνενούνται σε μικρότερες ομάδες, εκάστη δε εξ αυτών περιγράφει πολυπλευρώς ένα χαρακτηριστικό της ανθρώπινης συμπεριφοράς. Συνεπεία δε του πολυδιάστατου των ανθρώπινων χαρακτηριστικών, για την ει δυνατόν πλέον αξιόπιστη μέτρησή τους, απαιτούνται και κλίμακες πολλαπλών διαστάσεων. Επειδή, βεβαίως, ο ανθρώπινος ψυχισμός αλλά και η εξωτερικεύσή του υφίστανται επιδράσεις τόσο εκ του εγγύς όσο και του ευρύτερου περιβάλλοντος, οφείλει ο ερευνητής να προβαίνει σε έλεγχο της αξιοπιστίας και της εγκυρότητας του ανά περίπτωση δημιουργηθέντος ερωτηματολογίου. Για τον ίδιο λόγο, καθώς μεταξύ των κρατών (ή/και εθνικοτήτων) οι κοινωνικο-οικονομικές και πολιτισμικές συνθήκες αλλά και η ίδια κουλτούρα είναι διαφορετικές (ή/και ασκούν διαφορετικές επιρροές), είναι προφανές πως η επαναχρησιμοποίηση ενός ερωτηματολογίου σε δείγμα πληθυσμού με διαφορετικά χαρακτηριστικά από αυτόν για τον οποίο δημιουργήθηκε, προϋποθέτει την ανάλογη τροποποίησή του και τον έλεγχο καλής εφαρμογής του, ώστε να δημιουργηθεί μια αξιόπιστη και έγκυρη μεταφορά του στον νέο πληθυσμό (Riva et al., 2015: 599· Brancato et al., 2006). Στην παρούσα μελέτη θα μας απασχολήσουν οι έννοιες της αξιοπιστίας και της εγκυρότητας ως δύο από τις βασικές ιδιότητες για τις οποίες κάθε εργαλείο μέτρησης πρέπει να διακρίνεται.

### 1.1 Το ερωτηματολόγιο των αλλοδαπών

Το ερωτηματολόγιο των αλλοδαπών, με δεδομένο ότι το δείγμα θα αποτελείται από άτομα διαφορετικής εθνικότητας και ομιλούντα ένα ευρύ φάσμα γλωσσών και διαλέκτων, για να αποφευχθούν προβλήματα αξιοπιστίας, λόγω της αμφίβολης πιστότητας των μεταφράσεων σε λιγότερο συχνές γλώσσες, θα πρέπει ως κύρια γλώσσα σύνταξής του να χρησιμοποιηθεί η Ελληνική με παράλληλη μετάφρασή του στην Αγγλική.

Θα αποφύγουμε εδώ, λόγω των περιορισμών της έκτασης της εργασίας, να παρουσιάσουμε τις προτάσεις μας για τις τεχνικές σχεδίασης που καθιστούν ένα ερωτηματολόγιο ελκυστικό προς συμπλήρωση θεωρώντας τες ως εκ των ων ουκ άνευ στη δόμηση του περιού ο λόγος ερωτηματολογίου και θα περιοριστούμε στα χαρακτηριστικά που συνάδουν με τους σκοπούς του άρθρου, ήτοι σε όλα εκείνα που θα πρέπει να πληροί μια κλίμακα μέτρησης προκειμένου να λάβουν χώρα οι εδώ αναφερόμενοι στατιστικοί έλεγχοι αξιοπιστίας και εγκυρότητας.

Από τα όσα μέχρις εδώ έχουν συζητηθεί και από όσα θα ακολουθήσουν προκύπτει η ανάγκη δημιουργίας ενός ερωτηματολογίου το οποίο θα πρέπει να απαρτίζεται από τα παρακάτω τμήματα ώστε, με την αναγκαία ποσοτική κωδικοποίηση των υπό εξέταση μεταβλητών, να παρέχεται η δυνατότητα των απαραίτητων στατιστικών συσχετισμών και να συλλέγονται πληροφορίες σχετικά με τις παρακάτω κατηγορίες πληροφοριών (σε παρένθεση η δομή του ερωτήματος):

#### A) Δημογραφικά χαρακτηριστικά

- το φύλο (δίτιμη μεταβλητή Άντρας, Γυναίκα),
- την ηλικία, τον τόπο καταγωγής και γέννησης, τις γλώσσες που γνωρίζουν και χρησιμοποιούν (σύντομη αριθμητική ή αλφαριθμητική απάντηση),
- το επίπεδο εκπαίδευσης (θα δίνονται οι επιλογές: Αναλφάβητος, Πρωτοβάθμια εκπαίδευση, Δευτεροβάθμια εκπαίδευση, Τεχνική εκπαίδευση, Κολέγιο, Πανεπιστήμιο και πρόβλεψη για άλλη ανοικτή απάντηση),

#### B) Πληροφοριακά στοιχεία

- την ημερομηνία εισόδου για πρώτη φορά στην Ελλάδα (σύντομη αριθμητική απάντηση),
- συνθήκες παραμονής στη χώρα (θα δίνονται επιλογές: Αιτών άσυλο, Κάτοχος άδειας παραμονής, Σε καθεστώς μετάβασης σε άλλη ευρωπαϊκή χώρα, Προσωρινός επισκέπτης (τουρίστας), Φοιτητής και πρόβλεψη για άλλη ανοικτή απάντηση),
- τις δημόσιες υπηρεσίες που έχουν επισκεφτεί (θα δίνονται επιλογές: Υπηρεσίες υγείας, Δημοτικές υπηρεσίες, ΚΕΠ, Υπηρεσίες ασφάλισης, Εφορία, Σχολείο, Τράπεζες, και πρόβλεψη για άλλη ανοικτή απάντηση). Για κάθε μία επιλογή θα πρέπει να δίδεται η δυνατότητα να εκφράσουν την ικανοποίησή τους από τις υπηρεσίες διερωτώντας σε πεντάβαθμη κλίμακα Likert (1=πολύ απογοητευμένος έως 5=πολύ ικανοποιημένος),
- στον τρόπο που επικοινωνούν με τους υπαλλήλους (με συνοδό ή όχι) (θα δίνονται επιλογές: Μόνος/η μου, Συμπατριώτης που μιλά Ελληνικά, Μέλος της Οικογένειας που μιλά Ελληνικά, Έλληνας φίλος, Μέλος ΜΚΟ και πρόβλεψη για

άλλη ανοικτή απάντηση). Για κάθε μία επιλογή θα πρέπει να δίνεται η δυνατότητα να αναφέρουν και την υπηρεσία γι αυτό και θα δίνονται εκ νέου οι επιλογές: Υπηρεσίες υγείας, Δημοτικές υπηρεσίες, ΚΕΠ, Υπηρεσίες ασφάλισης, Εφορία, Σχολείο, Τράπεζες, και πρόβλεψη για άλλη ανοικτή απάντηση).

### Γ) Η διαδικασία της διερμηνείας

Το τρίτο τμήμα του ερωτηματολογίου θα πρέπει να περιλαμβάνει ερωτήματα που να συσχετίζονται με τους παρακάτω τέσσερις (4) παράγοντες που μπορούν να θεωρηθούν βασικοί συντελεστές περιγραφής της κατάστασης της κοινοτικής διερμηνείας:

- *Επιτυχής παροχή διερμηνείας*: βαθμός ευκολίας εύρεσης και επιλογής του διερμηνέα από τους ίδιους, εάν μιλούσε και πόσο καλά τη γλώσσα τους.
- *Επαγγελματισμός διερμηνείας*: σε ποιο βαθμό ο διερμηνέας κρατούσε σημειώσεις, διερμήνευε όλα όσα λέγονταν, ήταν ουδέτερος.
- *Επίπεδο μεσολάβησης διερμηνείας*: σε ποιο βαθμό εξηγούσε πολιτισμικές διαφορές, βοηθούσε σε άλλου είδους ζητήματα (συμπλήρωση αιτήσεων, κλπ), εξηγούσε σημαντικά σημεία της διαδικασίας.
- *Ικανοποίηση από τη διερμηνεία*: βαθμός κατανόησης του διερμηνέα, ικανοποίηση από τη διερμηνεία, αξιολόγηση των υπηρεσιών διερμηνείας.

Τα ερωτήματα, όπως θα εξηγηθεί περαιτέρω στις οδηγίες στάθμισης του ερωτηματολογίου, δεν θα πρέπει να είναι λιγότερα των τριών ανά παράγοντα, ενώ και ο συνολικός αριθμός να είναι προσηκόντως ικανός ώστε, εάν απαιτηθεί από τον έλεγχο αξιοπιστίας και εγκυρότητας, να μπορούν να αφαιρεθούν ερωτήματα, χωρίς να τίθεται εν αμφιβόλω η αρτιότητα της κλίμακας. Για κάθε ένα από τα ερωτήματα του τρίτου τμήματος του ερωτηματολογίου θα πρέπει να δίδεται η δυνατότητα στα υποκείμενα της έρευνας να εκφράσουν τη γνώμη τους σε πεντάβαθμη ή ανώτερη κλίμακα Likert που να συνάδει με το περιεχόμενο της ερώτησης.

## 2. Αξιοπιστία (Reliability)

Στις επιστημονικές έρευνες ποσοτικού τύπου στενά συνυφασμένη με την αξιοπιστία μιας μέτρησης είναι η έννοια της συνέπειας (consistency) οριζόμενη με τη σειρά της μέσω της επαναληπτικότητας (repeatability) και της αναπαραγωγιμότητας (reproducibility) του αποτελέσματος της μέτρησης, ήτοι της επανάληψης της μέτρησης υπό αντίστοιχες συνθήκες, αν και όχι όμοιες, και της αναπαραγωγής αποτελεσμάτων με μικρή μεταβλητότητα (Cohen, Manion & Morrison, 2007: 146). Οι έρευνες ποιοτικού τύπου, αν και δεν είναι υποχρεωμένες να ακολουθήσουν τους κανόνες ελέγχου αξιοπιστίας των ποσοτικών ερευνών, δεν σημαίνει πως δεν πρέπει να παρουσιάζουν παρόμοια χαρακτηριστικά τα οποία να διασφαλίζουν τη φερεγγυότητα των αποτελεσμάτων (Cohen, Manion & Morrison, 2007: 148-149).

### 2.1 Συντελεστές αξιοπιστίας (Reliability Coefficients)

#### 2.1.1. Συντελεστής αξιοπιστίας επαναληπτικών μετρήσεων (Test-Retest Reliability Coefficient)

Η εκτίμηση του μεγέθους της αξιοπιστίας των επαναληπτικών μετρήσεων γίνεται μέσω των δεικτών συνάφειας, με προεξάρχοντα τον Pearson  $r$ , ο οποίος με τη σειρά του απαιτεί

προϋποθέσεις γραμμικής συνάφειας των μεταβλητών, αναλογικών κλιμάκων μέτρησης ή κατ' ελάχιστον ισοδιαστημικών και κανονικής κατανομής (Pagano, 2009: 131-132). Πρόβλημα αποτελεί η πιθανή μη διαθεσιμότητα της ίδιας ομάδας ατόμων για επανάληψη των μετρήσεων, αλλά και η επίδραση των προηγούμενων μετρήσεων στις επόμενες, ακόμα και αν μεσολαβεί ικανό μεταξύ τους διάστημα, που με τη σειρά του δεν θα πρέπει να είναι μεγάλο, διότι υπάρχει κίνδυνος να υπεισέλθουν άλλοι εξελικτικοί παράγοντες οι οποίοι θα ασκήσουν επίδραση στις εξεταζόμενες μεταβλητές (Drost, n.d.:108).

#### 2.1.2 Αξιοπιστία εναλλακτικών τύπων ή ισοδύναμων τύπων (Parallel or Equivalent forms Reliability)

Προσπαθεί να θεραπεύσει τον θόρυβο που προκαλείται από τη συγκεκριμένη θεματολογία στην οποία αναφέρεται το τεστ. Στην περίπτωση αυτή θα πρέπει να διαθέτουμε δύο ισοδύναμους τύπους του ίδιου εργαλείου μέτρησης (Webb, Shavelson & Haertel, 2006:4). Σημαντικό πρόβλημα στην περίπτωση αυτή είναι να αποδειχθεί η ισοδυναμία των δύο τύπων και η δυσκολία δημιουργίας δύο συνόλων ερωτημάτων για τη μέτρηση της ίδιας έννοιας.

#### 2.1.3 Αξιοπιστία των ημικλάστων (Split-half Reliability)

Αντί να έχουμε δύο συγκρινόμενους τύπους του ίδιου τεστ, διαιρούμε το τεστ και συγκρίνουμε μεταξύ τους τα δύο τμήματά του. Ο δείκτης συνάφειας μεταξύ των δύο τμημάτων είναι στην ουσία ο δείκτης αξιοπιστίας των ημικλάστων του τεστ (Webb, Shavelson & Haertel, 2006: 6). Πρόβλημα αποτελεί ο μικρός αριθμός ερωτημάτων μιας ψυχομετρικής κλίμακας. Ως εναλλακτική λύση προτείνεται ο δείκτης αξιοπιστίας ημικλάστων του Guttman ο οποίος δεν υποθέτει ότι τα δύο μισά έχουν την ίδια αξιοπιστία ή την ίδια διασπορά και παρακάμπτει έτσι το θέμα της διόρθωσης για την αρχική έκταση της κλίμακας (Benton, 2015: 302). Καταλήγουμε όμως με δύο δύσχρηστους δείκτες αξιοπιστίας ημικλάστων, έναν για το κάθε μισό.

#### 2.1.4 Αξιοπιστία εσωτερικής συνέπειας (Internal consistency Reliability) - συντελεστής $\alpha$ του Cronbach (Cronbach's Alpha Coefficient)

Η ασκηθείσα έντονη κριτική στον συντελεστή του ημίκλαστου οδήγησε τους Kuder και Richardson, δύο από τους επικριτές του, να αντιπροτείνουν τη χρήση του τύπου KR-20 για τον υπολογισμό της αξιοπιστίας με μοναδική διαφορά τη μη χρήση συσχετίσεων μεταξύ των ερωτημάτων διχοτομικού τύπου ενός ερωτηματολογίου (Cronbach, 1951: 298-299).

Ο δείκτης αξιοπιστίας εσωτερικής συνέπειας προκύπτει είτε από τον τύπο KR-20 (Kuder-Richardson), όταν τα ερωτήματα επιδέχονται μόνο δύο απαντήσεις, τα στοιχεία δηλαδή που αναλύονται είναι διωνυμικού τύπου (δίτιμες μετρήσεις), είτε μέσω του τύπου Cronbach's  $\alpha$ , όπου το  $\alpha$  είναι ο δείκτης αξιοπιστίας, όταν τα ερωτήματα μιας κλίμακας επιδέχονται βαθμολόγηση με περισσότερες από δύο βαθμίδες (Cronbach, 1951: 299).

Μια επιπλέον προϋπόθεση είναι ότι τα ερωτήματα της κλίμακας πρέπει να είναι τουλάχιστον τρία σε αριθμό (Eisinga, Grotenhuis & Pelzer, 2012). Ο τύπος αναφέρεται σε διακυμάνσεις τιμών και σε αθροίσματα διακυμάνσεων, και κατ' ουσίαν βασίζεται στον πίνακα διακυμάνσεων μεταξύ όλων των πιθανών ζευγών των ερωτημάτων μιας κλίμακας μέτρησης. Οι δείκτες συνάφειας πρέπει να είναι ομοιογενώς υψηλοί, αλλά ταυτόχρονα ο πίνακας συναφειών δεν θα πρέπει έχει πολύ υψηλές συνάφειες σε όλο του το εύρος, αλλιώς τα ερωτήματα σημαίνει πως αλληλεπικαλύπτονται και συνεπώς δεν χρειάζονται όλα, ενώ ο

δείκτης θα καταλήξει να είναι υπερβολικά υψηλός ή εκτός ορίων (-1 έως +1). Οι τιμές της συνάφειας ερμηνεύονται όπως ο Pearson  $r$  χωρίς να μας ενδιαφέρει το πρόσημο. Εάν ο δείκτης είναι εκτός ορίων, θα σημαίνει ότι κατά την υπολογιστική διαδικασία παραβιάστηκαν οι στατιστικές παραδοχές. Εάν ο δείκτης αξιοπιστίας είναι μικρότερος από την απόλυτη τιμή 0,70, τότε ίσως θα πρέπει να επανεξετάσουμε το θέμα της εσωτερικής συνέπειας των ερωτημάτων της κλίμακάς μας.

Όσο μεγαλύτερο το πλήθος του δείγματος (N), τόσο μικρότερο το τυπικό σφάλμα για κάθε ερώτημα, επομένως τόσο μικρότερη και η διασπορά του. Μικρότερο το N μεγιστοποιεί το τυπικό σφάλμα μέτρησης και ελαχιστοποιεί τον τελικό δείκτη. Ο μεγάλος αριθμός ερωτημάτων, επίσης, επιφέρει συνολικά μεγάλη διασπορά τιμών. Ο κανόνας είναι να χρησιμοποιούμε τόσα ερωτήματα και τόσοι συμμετέχοντες, όσα και όσους θα χρησιμοποιούσαμε, αν το αντικείμενό μας ήταν μια ερευνητική διαδικασία μέσω του εργαλείου αυτού και όχι η προσπάθεια ανεύρεσης της αξιοπιστίας του. Να πληρούμε δηλαδή τις μεθοδολογικές και στατιστικές απαιτήσεις και παραδοχές.

#### 2.1.5 Αξιοπιστία μεταξύ βαθμολογητών (Reliability across raters)

Όλοι οι προηγούμενοι δείκτες αξιοπιστίας αναφέρονται σε περιπτώσεις όπου οι απαντήσεις του εξεταζόμενου βαθμολογούνται με βάση αντικειμενικά, ή έστω κοινά αποδεκτά και συμφωνημένα κριτήρια. Όταν τίθεται ζήτημα υποκειμενικής βαθμολόγησης των απαντήσεων των συμμετεχόντων σε ένα τεστ, είναι ανάγκη να υπολογίσουμε την αξιοπιστία μεταξύ των βαθμολογητών. Καταγράφοντας τις αρχικές τιμές από τουλάχιστον δύο βαθμολογητές και συσχετίζοντας τις βαθμολογίες μπορούμε να βρούμε τον βαθμό συμφωνίας αυτών των βαθμολογητών υπολογίζοντας και πάλι τον δείκτη συνάφειας Pearson  $r$  ή άλλους δείκτες συνάφειας, εάν οι μετρήσεις δεν είναι αριθμητικές (Sherman, Brooks, Iverson, Slick & Strauss, 2011: 880-881)

### 3. Εγκυρότητα (Validity)

Η εγκυρότητα των εργαλείων μέτρησης εμφανίστηκε ως όρος σχετικός με την εγκυρότητα των ψυχομετρικών κλιμάκων ύστερα από πρωτοβουλία της Επιτροπής του Αμερικανικού Συνδέσμου Ψυχολόγων APA η οποία προσδιόρισε τα προς αξιολόγηση χαρακτηριστικά πριν από τη δημοσίευση μιας κλίμακας και τους τύπους εγκυρότητας. Με προεξάρχουσα την *εγκυρότητα εννοιολογικής κατασκευής (construct validity)* προσδιορίστηκαν αρχικά οι παρακάτω τύποι εγκυρότητας (Cronbach & Meehl, 1955: 281):

- *Η προβλεπτική ή προγνωστική εγκυρότητα (predictive validity)*
- *Η συντρέχουσα εγκυρότητα (concurrent validity)*
- *Η εγκυρότητα περιεχομένου (content validity)*
- *Η εγκυρότητα εννοιολογική κατασκευής (construct validity)*

#### 3.1 Εγκυρότητα βάση κριτηρίου (Criterion Validity)

##### 3.1.1 Η προβλεπτική ή προγνωστική εγκυρότητα (predictive validity)

Μια κλίμακα μέτρησης διακρίνεται για την προβλεπτική της εγκυρότητα στο βαθμό που είναι σε θέση, με βάση τις σημερινές αξιολογήσεις, να προβλέψει μελλοντικές παραμέτρους, που

αποτελούν ειδοποιά χαρακτηριστικά της σημερινής εννοιολογικής κατασκευής (Gignac, 2009: 27).

### 3.1.2 Η συντρέχουσα εγκυρότητα (concurrent validity)

Μια κλίμακα μέτρησης διακρίνεται για τη συντρέχουσα εγκυρότητά της, όταν οι τιμές διαφοροποιούνται προς την αναμενόμενη κατεύθυνση μεταξύ ομάδων ατόμων, οι οποίες αντικειμενικά διαφέρουν ως προς την αξιολογούμενη εννοιολογική κατασκευή. Ο έλεγχος της συντρέχουσας εγκυρότητας γίνεται στη βάση της συνάφειας της αξιολογούμενης κλίμακας με ένα αντικειμενικό κριτήριο της εννοιολογικής κατασκευής για την οποία τη χρησιμοποιούμε. Βέβαια, ο έλεγχος μπορεί να γίνει και στη βάση της συνάφειας της αξιολογούμενης κλίμακας με μια άλλη σταθμισμένη κλίμακα τεκμηριωμένης αξιοπιστίας και εγκυρότητας, που αποτυπώνει την ίδια εννοιολογική κατασκευή με την αξιολογούμενη κλίμακα. Έτσι ο συντελεστής συνάφειας εκλαμβάνεται ως δείκτης της συγκλίνουσας εγκυρότητας της αξιολογούμενης κλίμακας (Gignac, 2009: 30).

### 3.2 Εγκυρότητα σχετική με το περιεχόμενο

Μία κλίμακα μέτρησης αξιολογείται ως προς το περιεχόμενό της από δύο κατηγορίες ατόμων. Αφενός, από τα άτομα στα οποία εφαρμόζεται, δηλαδή από δείγμα ατόμων του πληθυσμού ενδιαφέροντος. Αφετέρου, αξιολογείται από τους ειδικούς στον εκάστοτε επιστημονικό χώρο στον οποίο ανήκει. Έτσι, η εγκυρότητα σχετικά με το περιεχόμενο ενός ερωτηματολογίου, περιλαμβάνει δύο μορφές: την φαινομενική εγκυρότητα (Face Validity) και την εγκυρότητα περιεχομένου (Content Validity). Αναλυτικότερα έχουμε:

#### 3.2.1 Φαινομενική εγκυρότητα (Face validity)

Ένα ερωτηματολόγιο υποστηρίζουμε πως χαρακτηρίζεται από φαινομενική εγκυρότητα, όταν δείγμα ατόμων από τον πληθυσμό στόχο εκτιμά πως αποτελείται από ερωτήματα που συνάδουν με το προς μέτρηση χαρακτηριστικό. Παρά το γεγονός ότι το είδος αυτό της εγκυρότητας δεν βασίζεται σε επιστημονικά κριτήρια αλλά σε γνώμη ατόμων μη ειδικών με το αντικείμενο, είναι εξίσου σημαντικό, διότι στην περίπτωση που θεωρηθεί ως μη έγκυρο, επιφέρει ελλιπή απόκριση στα ερωτήματα ή παραπλανητικές απαντήσεις (Gignac, 2009: 23).

#### 3.2.2 Εγκυρότητα περιεχομένου (Content validity)

Σε αντίθεση με τον προηγούμενο τύπο, η εγκυρότητα περιεχομένου ενός ερωτηματολογίου αξιολογείται από έναν αριθμό επιστημόνων (που δεν θα είναι λιγότεροι των τριών) η ειδίκευση των οποίων είναι σχετική με το γνωστικό αντικείμενο που πραγματεύεται το ερωτηματολόγιο και οι οποίοι έχουν συνεπώς τη γνωστική επάρκεια να αξιολογήσουν το εργαλείο αλλά και να αντιπροτείνουν μεθόδους βελτίωσής του. Η βαθμολογία γίνεται με τον Δείκτη Εγκυρότητας Περιεχομένου (Content Validity Index (CVI)) με τη χρήση μιας κλίμακας τεσσάρων βαθμών. Ο CVI, αν και βαθμολογείται με τρεις τρόπους, είναι με απλά λόγια το ποσοστό των ερωτημάτων που βαθμολογούνται με βαθμούς 3 και 4 από τους ειδικούς (Polit & Tatano Beck, 2006: 491-493).

### 3.3 Εγκυρότητα εννοιολογικής κατασκευής (Construct validity)

Η εγκυρότητα εννοιολογικής κατασκευής ενός ερωτηματολογίου αναφέρεται στο εάν αυτό αντιπροσωπεύει επακριβώς τις διαστάσεις ενός φαινομένου που σκοπεύει να περιγράψει, να



εξηγήσει ή να διαμορφώσει σε θεωρητικό επίπεδο. Στις περισσότερες αναλύσεις, η εγκυρότητα σχετίζεται με την αξιοπιστία. Η εγκυρότητα συνίσταται στο κατά πόσο η κλίμακα μετρά αυτό που επικαλείται ότι μετρά. Το βασικό κριτήριο για την επιλογή μιας κλίμακας είναι η εγκυρότητά της. Το εάν μια κλίμακα μέτρησης, εδώ ένα ερωτηματολόγιο, μετράει στην πραγματικότητα αυτό για το οποίο έχει δημιουργηθεί μπορεί να εκτιμηθεί με τη χρήση στατιστικών μεθόδων της Ανάλυσης Παραγόντων οι οποίες βασίζονται στην ανάλυση των συσχετίσεων μεταξύ των ερωτημάτων, ώστε να διαπιστωθεί εάν τα ερωτήματα που ανήκουν στην ίδια διάσταση αποτελούν έναν κοινό παράγοντα (Sherman et al., 2011: 889). Εάν οι διαστάσεις ενός ερωτηματολογίου και συνεπώς η παραγοντική δομή είναι άγνωστη τότε εφαρμόζεται η Διερευνητική Ανάλυση Παραγόντων (Exploratory Factor Analysis (EFA)), ενώ εάν η παραγοντική δομή βασίζεται σε κάποιο θεωρητικό μοντέλο και επιθυμούμε να επιβεβαιώσουμε εάν οι μετρήσεις προσαρμόζονται σε αυτή τη δομή, χρησιμοποιούμε την Επιβεβαιωτική Ανάλυση Παραγόντων (Confirmatory Factor Analysis (CFA)) (Fabrigar & Wegener, 2012: 4). Εάν μετά τη CFA η συνολική εικόνα προσαρμογής είναι πολύ φτωχή και μη αποδεκτή μπορεί να αποφασιστεί η πραγματοποίηση EFA ώστε να φανούν οι πραγματικές διαστάσεις του ερωτηματολογίου και εκ νέου CFA ώστε να ελεγχθεί και πάλι η καλή προσαρμογή του.

#### 4. Πιλοτική εφαρμογή

Για τη συλλογή των απαραίτητων δεδομένων για τη πραγματοποίηση του ελέγχου της εγκυρότητας και της αξιοπιστίας του ερωτηματολογίου δεν θα μπορούσαμε να κάνουμε τίποτε λιγότερο από το να συστήσουμε τη διεξαγωγή πιλοτικής μελέτης, αφού θεωρείται ακρογωνιαίος λίθος ενός καλού ερευνητικού σχεδιασμού (Hazzi & Maldaon, 2015: 53). Αυτονόητη επίσης είναι η σύσταση για συλλογή δεδομένων από όλες τις επιλογείς για τη διενέργεια της κύριας έρευνας περιοχές, ήτοι την Αττική (πολεοδομικό συγκρότημα Αθηνών - Πειραιώς), τη Θεσσαλονίκη, τα νησιά του Ανατολικού Αιγαίου καθώς και τη Θράκη. Με δεδομένο δε ότι η πιλοτική εφαρμογή ορίζεται ως η «σε μικρότερη κλίμακα δοκιμή των χρησιμοποιούμενων σε ευρεία κλίμακα μεθόδων και των διαδικασιών...» (Porta, 2008: 215) θα πρέπει να τηρηθούν επιπλέον απαρεγκλίτως και τα όσα θα αναφερθούν ακολούθως κατά την περιγραφή των στατιστικών ελέγχων αξιοπιστίας και εγκυρότητας και σχετίζονται τόσο με το απαιτούμενο μέγεθος του δείγματος και τη φύση των δεδομένων όσο και με τη δομή του ίδιου του οργάνου μέτρησης.

Η στατιστική ανάλυση μπορεί να διεξαχθεί με τη βοήθεια του λογισμικού στατιστικής επεξεργασίας IBM SPSS για τον έλεγχο αξιοπιστίας και τη Διερευνητική Ανάλυση Παραγόντων, ενώ για την Επιβεβαιωτική Ανάλυση Παραγόντων μπορεί να χρησιμοποιηθεί το SSI LISREL.

#### 5. Έλεγχος Αξιοπιστίας ερωτηματολογίου

Ο πιο συχνά χρησιμοποιούμενος δείκτης αξιοπιστίας στις έρευνες κοινωνικών επιστημών είναι ο Cronbach's  $\alpha$  για τον οποίο υπάρχει συμφωνία τόσο για την ερμηνεία του όσο και για το αποδεκτό εύρος των τιμών του. Στις μέρες μας ο υπολογισμός του Cronbach's  $\alpha$  γίνεται ταχύτατα με χρήση υπολογιστή με ειδικό λογισμικό στατιστικής επεξεργασίας όπως το IBM SPSS.

### 5.1 Προϋποθέσεις εφαρμογής

Πριν υπολογισθεί ο δείκτης αξιοπιστίας Cronbach's  $\alpha$ , εάν υπάρχουν ερωτήματα με αρνητική βαθμολογία αυτά θα πρέπει να αντιστραφούν (Field, 2009), ώστε η εξ αυτών προκληθείσα αρνητική συνδιακύμανση να μην επιφέρει μείωση του συνολικού δείκτη, αφού ο αριθμητής της εξίσωσης του Cronbach's  $\alpha$  ενσωματώνει τον μέσο όρο όλων των συνδιακυμάνσεων μεταξύ των ερωτημάτων.

Πολύ συχνά στα ερωτηματολόγια συναντούμε δέσμες ερωτημάτων που σχετίζονται με συγκεκριμένες υπερμεταβλητές – παράγοντες. Σε αυτή την περίπτωση αναφερόμαστε σε πολυδιάστατα ερωτηματολόγια στα οποία κάθε διάσταση συμπίπτει με έναν παράγοντα. Παραδείγματος χάριν σε κάποιο ερωτηματολόγιο για τη μέτρηση της νοημοσύνης μια δέσμη ερωτημάτων μπορεί αναφέρεται στη μέτρηση της *αριθμητικής ικανότητας*, άλλη δέσμη ερωτημάτων στη μέτρηση της *λεκτικής ικανότητας*, κ.ο.κ. Σε μια τέτοια περίπτωση θα πρέπει να υπολογιστεί για κάθε μία από τις διαστάσεις του οργάνου μέτρησης ο δείκτης αξιοπιστίας Cronbach's  $\alpha$ .

Θετική επίδραση στον δείκτη αξιοπιστίας φαίνεται να ασκεί, επίσης, το μεγάλο για κάθε ερώτημα εύρος των εναλλακτικών απαντήσεων της κλίμακας Likert που χρησιμοποιείται στην έρευνα με ένα μέσο αποδεκτό εύρος  $7 \pm 2$  εναλλακτικών απαντήσεων, εύρος το οποίο το ανθρόπινο μυαλό είναι σε θέση να διακρίνει (Miller, 1955).

Για να χαρακτηρίσουμε ένα ερωτηματολόγιο αξιόπιστο θα πρέπει ο δείκτης αξιοπιστίας Cronbach's  $\alpha$  να είναι μεγαλύτερος του 0,70, ενώ άλλοι μελετητές υποστηρίζουν πως για τα αρχικά στάδια κάποιας μελέτης αρκεί ένας συντελεστής Cronbach's  $\alpha$  μεταξύ του 0,5 και 0,6 αλλά οπωσδήποτε για την εξαγωγή σημαντικών συμπερασμάτων τουλάχιστον  $\alpha = 0,9$  ή καλύτερα  $\alpha = 0,95$  (Nunnally & Bernstein, 1994· Tavakol & Dennick, 2011: 54· Vaske, Beaman & Sponarski, n.d.: 5-6· Pallant, 2005: 90).

Επειδή μεταξύ της συνολικής τιμής του συντελεστή αξιοπιστίας Cronbach's  $\alpha$  του ερωτηματολογίου και των επιμέρους συντελεστών Cronbach's  $\alpha$  των παραγόντων του υπάρχει υψηλή συνάφεια θα πρέπει να εξεταστούν αρχικά οι δείκτες αξιοπιστίας των επιμέρους διαστάσεων και στη συνέχεια ο συνολικός δείκτης. Συνεπώς στην περίπτωση του ερωτηματολογίου των αλλοδαπών θα πρέπει να υπολογιστεί πρώτα ο δείκτης αξιοπιστίας για καθεμιά από τις διαστάσεις:

- Επιτυχής παροχή διερμηνεία.
- Επαγγελματισμός διερμηνεία.
- Επίπεδο μεσολάβησης διερμηνεία.
- Ικανοποίηση από τη διερμηνεία.

Εάν για κάποια διάσταση του ερωτηματολογίου προκύπτει δείκτης αξιοπιστίας μικρότερος του 0,70 ( $\alpha < 0,70$ ), θα πρέπει να αναζητήσουμε τα ερωτήματα που επηρεάζουν το μέγεθος του δείκτη. Αυτό είναι εφικτό από τον πίνακα *Item-Total Statistics* και συγκεκριμένα από τη στήλη *Corrected Item-Total Correlation* από την οποία θα διαπιστώσουμε ποια ερωτήματα έχουν πολύ χαμηλό συντελεστή συνάφειας Pearson  $r$ , αλλά και από τη στήλη *Cronbach's Alpha if Item deleted*, η οποία θα μας πληροφορεί κατά πόσο η μειωμένη τιμή του δείκτη μπορεί να βελτιωθεί, εάν απαλειφθούν από το ερωτηματολόγιο τα προαναφερόμενα ερωτήματα. Εάν θεωρήσουμε πως τα συγκεκριμένα ερωτήματα δεν συνεισφέρουν σημαντικά και δεν συνιστούν μη εξαιρετικά στοιχεία στην έρευνά μας,

μπορούμε να τα αφαιρέσουμε, ειδάλλως θα χρειαστεί επαναδιατύπωσή τους, εκ νέου πιλοτική εφαρμογή και επανέλεγχος.

Η ίδια διαδικασία θα πρέπει να επαναληφθεί και για το συνολικό ερωτηματολόγιο.

## 6. Διερευνητική Παραγοντική Ανάλυση

Στα πλαίσια του άρθρου αυτού θα μας απασχολήσει η μέθοδος των κύριων συνιστωσών ως πρώτο βήμα για την περαιτέρω ανάλυση παραγόντων και η πλάγια περιστροφή των αξόνων κατά τη φάση της παραγοντικής ανάλυσης για τους λόγους που θα εξηγηθούν περαιτέρω (Fabrigar & Wegener, 2012: 43, 73).

### 6.1 Προϋποθέσεις εφαρμογής

Για την εκτέλεση μιας αξιόπιστης παραγοντικής ανάλυσης θα πρέπει να τηρούνται μια σειρά από προϋποθέσεις:

1. Είναι σημαντικό να προηγηθεί *εξέταση της κανονικότητας της κατανομής* είτε εποπτικά με Q-Q plot ή ιστογράμματα είτε/και στατιστικά με τα tests of Normality Kolmogorov-Smirnov (για δείγματα >50) και Shapiro-Wilk (και για δείγματα <50), αν και χρειάζεται προσοχή στην ερμηνεία τους, αφού είναι δοκιμασίες στενά εξαρτημένες από το μέγεθος του δείγματος (DeCoster, 2001:10). Τα tests αυτά δε θα πρέπει να απορρίπτουν τη μηδενική υπόθεση η οποία ισοδυναμεί με τη δήλωση πως το δείγμα μας δεν είναι στατιστικώς σημαντικά διαφορετικό από τον κανονικό πληθυσμό. Είναι σημαντικό εδώ να τονιστεί πως αν παραβιάζεται η προϋπόθεση της κανονικής κατανομής, τότε εάν δεν γίνει με μία από τις ενδεικνύμενες μεθόδους (Log10, LN, Sqrt) η κανονικοποίηση της κατανομής (DeCoster, 2001: 11-12), προτείνεται η μέθοδος των κύριων συνιστωσών για την εκτέλεση της EFA (Fabrigar, Wegener, MacCallum & Strahan, 1999: 283).
2. Το *μέγεθος του δείγματος* θα πρέπει να είναι επαρκές, αφού στα μικρά δείγματα επηρεάζονται οι συντελεστές συνάφειας πολύ περισσότερο από όσο στα μεγαλύτερα δείγματα. Παρόλο που δεν υπάρχει απόλυτη συμφωνία για τον αριθμό των υποκειμένων του δείγματος, κοινός κανόνας είναι πως ο ερευνητής χρειάζεται μια αναλογία 10 – 15 συμμετεχόντων ανά ερώτημα και ποτέ αυτή να μην είναι μικρότερη από 5, ή τουλάχιστον ο αριθμός των υποκειμένων του δείγματος να είναι 100. Άλλοι ανεβάζουν τον αριθμό σε 200 ανεξάρτητα από την αναλογία ατόμων προς τα ερωτήματα. Εναλλακτικά η επάρκεια του μεγέθους της δειγματοληψίας (sampling adequacy) εξετάζεται με τον στατιστικό δείκτη καταλληλότητας του δείγματος για ανάλυση παραγόντων των Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) με αποδεκτή τιμή  $KMO > 0,6$  και κυμαίνεται μεταξύ του 0 και του 1 ( $0 \leq KMO \leq 1$ ) (Beavers, Lounsbury, Richards, Huck, Skolits & Esquivel, 2013).
3. Το ερωτηματολόγιο, θα πρέπει να έχει ως κατώτατο όριο τα 10 ερωτήματα, χωρίς και πάλι να είμαστε σίγουροι για την ασφάλεια των αποτελεσμάτων της ανάλυσης, αφού στην αρχική φάση της δημιουργίας ενός ερωτηματολογίου θα πρέπει να υπάρχουν εναλλακτικά ερωτήματα από τα οποία ο ερευνητής θα μπορεί να επιλέξει κάποια που να χαρακτηρίζονται από υψηλή αξιοπιστία. Πάντως στην ανάλυση παραγόντων θα πρέπει να περιλαμβάνονται τουλάχιστον τρεις (3) μεταβλητές ανά παράγοντα (Yong & Pearce, 2013: 80).

4. Οι μεταβλητές μας θα πρέπει να μετριοούνται με κλίμακα μέτρησης αναλογική ή τουλάχιστον ισοδιαστημική (πεντάβαθμες και άνω κλίμακες Likert) και η σχέση μεταξύ τους να είναι ευθύγραμμη και όχι καμπυλόγραμμη, ως δύο από τις βασικές προϋποθέσεις για τη χρήση του δείκτη συνάφειας Pearson  $r$  στον πίνακα συνάφειας.
5. Ο δείκτης σφαιρικότητας (Bartlett's test of sphericity) αναφέρεται στην ομοιογένεια του δείγματος και εξετάζει την πιθανότητα μοναδιαίας συμμετρίας (compound symmetry) στον πίνακα διασποράς-συνδιασποράς των μετρήσεων (variance-covariance matrix). Αν ο δείκτης προκύψει στατιστικά σημαντικός, αποδεικνύεται πως υπάρχει συνάφεια μεταξύ των μεταβλητών μας και σε συνδυασμό με τον δείκτη MKO και την ορίζουσα (determinant) του πίνακα συνάφειας συμπεραίνουμε την επαρκή συνάφεια και συνεπώς καταλληλότητα των δεδομένων μας.
6. Τα ερωτήματα πρέπει να χαρακτηρίζονται από μέτρια έως υψηλή συνάφεια μεταξύ τους (ενδοσυνάφεια) ανεξάρτητα από το πρόσημο. Τιμές συνάφειας χαμηλότερες του  $\pm 0,3$  ή υψηλότερες του  $\pm 0,8$  δείχνουν ερωτήματα ασυσχέτιστα μεταξύ τους ή επικαλυπτόμενα. (Field, 2009).
7. Πριν την εξαγωγή των παραγόντων ο ερευνητής μπορεί να επιλέξει, εάν οι θεωρητικές ή μεθοδολογικές επιλογές τού το υποδεικνύουν, συγκεκριμένο αριθμό παραγόντων προς εξαγωγή. Επιλέξιμοι είναι οι παράγοντες με ιδιοτιμή (Eigenvalue) μεγαλύτερη του 1 σύμφωνα με το κριτήριο του Kaiser το οποίο σε συνδυασμό με το scree test, παρά τις όποιες αδυναμίες επακριβούς ανίχνευσης των παραγόντων (Ledesma & Valero-Mora, 2007: 2-3· Nunnally & Bernstein, 1994: 482), μπορούν να επιβεβαιώσουν την αρχική επιλογή η οποία μπορεί να στηρίζεται από το θεωρητικό μοντέλο ή να υποδείξουν συγκεκριμένο αριθμό προς εξαγωγή.
8. Είναι πολύ σημαντικό στη φάση της εξαγωγής να μελετηθεί ο πίνακας των αποτελεσμάτων της κοινότητας διακύμανσης των παραγόντων (Communalities) ο οποίος μας δίνει πληροφορίες για το ποσοστό της διακύμανσης κάθε μεταβλητής (ερωτήματος) που εξηγείται από τους τρεις παράγοντες. Ικανοποιητική θεωρείται τιμή μεγαλύτερη του 0,5. Εάν υπάρξει κάποια μεταβλητή που να έχει μικρότερη τιμή, γεγονός που υποδεικνύει μη συνάφεια με κάποιο παράγοντα, πριν λάβουμε απόφαση απομάκρυνσής της, θα πρέπει να δούμε εάν συμφωνεί με την παραγοντική δομή (Fabrigar & Wegener, 2012: 134).
9. Το είδος της στρέψης των παραγόντων που θα εφαρμοστεί είναι στενά συνδεδεμένο με τη συνάφεια ή μη των παραγόντων. Η ορθόγωνη στρέψη (orthogonal rotation) με κυριότερες τεχνικές τις Quartimax και Varimax προτιμάται για ασυσχέτιστους παράγοντες, ενώ η πλάγια στρέψη (oblique rotation) με κυριότερες τεχνικές τις Direct Oblimin και Promax ενδείκνυται όταν υπάρχει υπόνοια ή το μεθοδολογικό πλαίσιο υποστηρίζει την ύπαρξη συνάφειας μεταξύ των παραγόντων (Yong & Pearce, 2013: 84· Zygmont & Smith, 2014: 46). Στην περίπτωση της έρευνάς μας προτείνεται η εφαρμογή της Promax αφού υποθέτουμε πως υπάρχει συσχέτιση μεταξύ των παραγόντων.
10. Οι φορτίσεις (loadings) των μεταβλητών στους παράγοντες υποδεικνύουν τη συνάφεια μεταξύ τους. Κυμαίνονται από -1 έως 1 και φορτίσεις γύρω στο μηδέν δείχνουν χαμηλή επίδραση στη μεταβλητή. Όσο πιο ψηλές είναι οι φορτίσεις τόσο ευκολότερα καταδεικνύεται από ποιες μεταβλητές ερμηνεύεται ο παράγοντας. Αποδεκτές τιμές των φορτίσεων θεωρούνται οι άνω των 0,30 (στην έρευνά μας το όριο που θα πρέπει να τεθεί είναι το 0,40).

11. Ο πίνακας Component Correlation Matrix περιλαμβάνει τις ενδοσυσχετίσεις των εξαχθέντων παραγόντων και της συνάφειας μεταξύ των ζευγών.

### 7. Επιβεβαιωτική παραγοντική ανάλυση (Confirmatory Factor Analysis (CFA))

Για να επιβεβαιωθεί η προκύψασα εκ της ανωτέρω διερευνητικής παραγοντικής ανάλυσης (EFA) παραγοντική δομή και να αξιολογηθεί η καλή προσαρμογή του μοντέλου στα δεδομένα θα πρέπει να πραγματοποιηθεί επιβεβαιωτική παραγοντική ανάλυση (CFA). Η ανάλυση αυτή μπορεί να γίνει με τη χρήση του λογισμικού SSI LISREL.

Η μέθοδος υπολογισμού των στατιστικών μοντέλων της ανάλυσης που εφαρμόζεται εξ' ορισμού από το πρόγραμμα είναι εκείνη της Μεγίστης Πιθανοφάνειας (Maximum Likelihood) η οποία ανέχεται μικρές αποκλίσεις από την κανονική κατανομή. Εάν υπάρχουν απύσες τιμές (missing values) μπορούν να αντιμετωπισθούν με τη μέθοδο αποκλεισμού Listwise pairs deletion μία από τις πιο συχνά χρησιμοποιούμενες (Soley-Bori, 2013: 6· Kamakura & Wedel, 2000).

Στα αποτελέσματα της ανάλυσης οι δείκτες που μας ενδιαφέρουν και δείχνουν επάρκεια ή μη προσαρμογής του μοντέλου στα δεδομένα είναι:

- Το στατιστικό κριτήριο  $\chi^2$  το οποίο μπορεί να χρησιμοποιηθεί εάν το μέγεθος του δείγματος είναι μεταξύ των 100 και 200 περιπτώσεων με αποδεκτή τιμή να πλησιάζει στο μηδέν και επίπεδο σημαντικότητας  $\geq 0,05$ , ενώ
- εάν το μέγεθος του δείγματος είναι μεγαλύτερο (όπως θα πρέπει να είναι στη περίπτωση μας) τότε θα πρέπει να χρησιμοποιηθεί ο λόγος  $CMIN/df$  ( $\chi^2/df$  ratio) με αποδεκτή τιμή  $\leq 2$  (ή  $\leq 5$ ),
- ο δείκτης καλής προσαρμογής GFI (Goodness of fit index) με αποδεκτή τιμή  $\geq 0,90$  (εύρος τιμών από 0 (πολύ φτωχή προσαρμογή) έως 1 (τέλεια προσαρμογή),
- ο διορθωμένος δείκτης καλής προσαρμογής AGFI (Adjusted goodness of fit index) με αποδεκτή τιμή  $\geq 0,90$ , (εύρος τιμών από 0 (πολύ φτωχή προσαρμογή) έως 1 (τέλεια προσαρμογή),
- ο συγκριτικός δείκτης προσαρμογής CFI (Comparative Fit Index) με αποδεκτή τιμή  $\geq 0,95$ ,
- η ρίζα των μέσων τετραγώνων του σφάλματος εκτίμησης RMSEA (Root Mean Square Error of Approximation) με αποδεκτή τιμή  $\leq 0,08$  (Hoe, 2008: 78· Byrne, 2010· Bentler, 2006: 355).

Ως μέθοδοι θεραπείας του προβλήματος της τυχόν μη καλής προσαρμογής του μοντέλου στα δεδομένα, στην περίπτωση που η κατανομή του δείγματος αποκλίνει από την κανονική κατανομή, προτείνονται στη CFA εναλλακτικές μέθοδοι υπολογισμού των στατιστικών μοντέλων της ανάλυσης όπως εκείνες των Σταθμισμένων Ελαχίστων Τετραγώνων (Weighted Least Squares (WLS)), και Εύρωστης Μεγίστης Πιθανοφάνειας Robust Maximum Likelihood (RML) (Kline, 2011: 180-181· Hair et al., 2009: 638). Μια περαιτέρω επιλογή είναι εκείνη της εφαρμογής της CFA σε τιμές Z αντί των πρωτογενών δεδομένων.

Εάν δεν ευδοκιμήσουν οι ανωτέρω αναφερόμενες μέθοδοι επαναχρησιμοποιούμε την EFA επιχειρώντας, κάθε φορά ανάλογα με τα αποτελέσματα, διαφορετικό είδος στρέψης των

αξόνων, ώστε να αποκαλυφθούν περισσότεροι παράγοντες που να καλύπτουν το μεγαλύτερο ποσοστό της συνολικής διακύμανσης και στη συνέχεια να επανεφαρμοσθεί η CFA.

### Συμπεράσματα

Στην παρούσα εργασία επιχειρήθηκε μια αναφορά στο μεθοδολογικό σχέδιο που προτάθηκε για τη διεξαγωγή της ερευνητικής προσπάθειας καταγραφής της κοινοτικής διερμηνείας στην Ελλάδα, στο πλαίσιο του προγράμματος CiGreece, δίνοντας έμφαση στην περιγραφή των προϋποθέσεων για την εφαρμογή τεχνικών στάθμισης ενός ερωτηματολογίου μέσω της εξαγωγής του δείκτη αξιοπιστίας Cronbach's  $\alpha$  αλλά και της Διερευνητικής και Επιβεβαιωτικής ανάλυσης παραγόντων για τον έλεγχο της δομικής εγκυρότητάς του. Η σειρά των τεχνικών στάθμισης ενός ερωτηματολογίου όπως παρουσιάζονται εδώ συνάδει με τους σκοπούς του παρόντος άρθρου, και δεν απεικονίζει τη δέουσα, αφού είναι αναμφίβολο ότι πρωτίστως θα πρέπει να ελέγχεται η εγκυρότητα ενός οργάνου μέτρησης, ήτοι ότι αυτό μετρά πράγματι την έννοια την οποία διατείνεται ότι μετρά, και ακολούθως, εάν αυτό κριθεί αποδεκτό, να ελέγχεται η αξιοπιστία του, δηλαδή εάν η μέτρηση αυτή παράγει σταθερά και επαναλαμβανόμενα αποτελέσματα.

Η συνδυαστική χρήση των αναφερόμενων στο κύριο μέρος της παρούσας εργασίας στατιστικών μεθόδων μπορεί να αναδείξει την πιθανότητα ύπαρξης ασθενούς παραγοντικής δομής ενός μοντέλου παρά την επιτυχή και ταυτιζόμενη με το θεωρητικό ή/και μεθοδολογικό μοντέλο εξαγωγή των παραγόντων της κλίμακας από τη Διερευνητική Παραγοντική Ανάλυση και την υψηλή τιμή του δείκτη αξιοπιστίας, γεγονός που υποδεικνύει την αναγκαιότητα εφαρμογής πολλαπλών στατιστικών μεθόδων για τον έλεγχο εγκυρότητας και αξιοπιστίας των κλιμάκων μέτρησης.

Σίγουρα στα πλαίσια του παρόντος άρθρου δεν θα μπορούσε να πραγματοποιηθεί μια εξαντλητική παρουσίαση, αλλά μια μόνο ενδεικτική αναφορά ως προτροπή για να ακολουθούνται και μέθοδοι που προσφέρουν ποικίλες δυνατότητες ελέγχου και αξιοποίησης των ερευνητικών δεδομένων, χωρίς βεβαίως να αυτοπαγιδεύεται κανείς από στατιστικά έγκυρα αποτελέσματα ή παραγοντικές δομές που πόρρω απέχουν από την πραγματικότητα η οποία με τη σειρά της επιβάλλει τη δική της βέλτιστη περιοριστική λύση.

### ΒΙΒΛΙΟΓΡΑΦΙΑ

- Beavers, A., Lounsbury, J., Richards, J., Huck, S., Skolits, G. & Esquivel, S. (2013). Practical Considerations for Using Exploratory Factor Analysis in Educational Research. *Practical Assessment, Research & Evaluation* 18 (6), pp. 1-13.
- Bentler, P. (2006). *EQS 6 Structural Equations Program Manual*. Encino, CA: Multivariate Software, Inc.
- Benton, T. (2015). An Empirical Assessment of Guttman's Lambda 4 Reliability Coefficient. In R. E. Millsap (ed.) *Quantitative Psychology Research, Springer Proceedings in Mathematics & Statistics* (pp. 301-310). Springer International Publishing Switzerland.
- Bradburn, N., Sudman, S. & Wansink, B. (2004). *Asking Questions, The Definitive Guide to Questionnaire Design - For Market Research, Political Polls and Social and Health Questionnaires*. San Francisco: Jossey-Bass.

- Brancato, S., Macchia, M., Signore, G., Simeoni, G., Blanke, K., Koerner, T. et al. (2006). *Handbook of Recommended Practises for Questionnaire Development and Testing in the European Statistical System*. ISTAT.
- Byrne, B. (2010). *Structural Equation Modelling with AMOS, Basic Concepts, Applications, and Programming* (2<sup>nd</sup> ed.). New York: Routledge.
- Cohen, L., Manion, L. & Morrison, K. (2007). *Research Methods in Education* (6<sup>th</sup> ed.). London: Routledge.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16(3), pp. 297-334.
- Cronbach, L. & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin* 52, pp. 281-302.
- DeCoster, J. (2001). Transforming and Restructuring Data. Ανάρτηση 10/6/2016, από <http://www.stat-help.com/notes.html>
- Dornyei, Z. (2003). *Questionnaires in Second Language Research, Construction, Administration, and Processing*. London: Lawrence Erlbaum Associates, Publishers.
- Drost, E. (n.d.). Validity and Reliability in Social Science Research. *Education Research and Perspectives*, 38(1), pp. 105-123.
- Eisinga, R., Grotenhuis, M. & Pelzer, B. (2012). The reliability of a two-item scale: Pearson, Cronbach or Spearman-Brown? Ανάρτηση 14/6/2016, από *Int J Public Health*: <http://robeisinga.ruhosting.nl/MyOpenAccess/The%20reliability%20of%20a%20two-item%20scale.pdf>
- Fabrigar, L. & Wegener, D. (2012). *Exploratory Factor Analysis*. New York: Oxford University Press.
- Fabrigar, L., Wegener, D., MacCallum, R. & Strahan, E. (1999). Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods* 4(3), pp. 272-299.
- Field, A. (2009). *Discovering Statistics Using SPSS* (3<sup>rd</sup> ed.). London: Sage.
- Gignac, G. (2009). Psychometrics and the Measurement of Emotional Intelligence. In *Assessing Emotional Intelligence, The Springer Series on Human Exceptionality* (pp. 9-40). Springer Science+Business Media.
- Hair, J., Black, W., Babin, B. & Anderson, R. (2009). *Multivariate Data Analysis* (7<sup>th</sup> ed.). New York: Pearson.
- Hazzi, O. & Maldaon, I. (2015). A pilot study: vital Methodological issues. *Verslas:Teorija ir praktika/business:Theory and Practice*, 16(1), pp. 53-62.
- Hoe, S. L. (2008). Issues and Procedures in Adoping Structural Equation Modelin Technique. *Journal of Applied Quantitative Methods* 3(1), pp. 76-83.
- Kamakura, W., & Wedel, M. (2000). Factor Analysis and Missing Data. *Journal of Marketing Research* XXXVII, pp. 490-498.
- Kline, R. (2011). *Principles and Practice of Structural Equation Modeling* (3<sup>rd</sup> ed.). New York: The Guilford Press.
- Ledesma, R. & Valero-Mora, P. (2007). Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis. Ανάρτηση 10/6/2016, από <http://pareonline.net/getvn.asp?v=12&n=2>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology* 22(140), pp. 5-55.

- Miller, G. (1955). The Magical Number Seven, Plus or Minus Two, Some Limits on Our Capacity for Processing Information. *Psychological Review* 101(2), pp. 343-352.
- Nunnally, J. & Bernstein, I. (1994). *Psychometric Theory* (3<sup>rd</sup> ed.). New York: McGraw-Hill, Inc.
- Pagano, R. (2009). *Understanding Statistics in the Behavioral Sciences*. Belmont: Wadsworth Cengage Learning.
- Pallant, J. (2005). *Spss Survival Manual*. Sydney: Allen & Unwin.
- Polit, D. & Tatano Beck, C. (2006). The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations. *Research in Nursing & Health* 26, pp. 489-497.
- Porta, M. (2014). *A Dictionary of Epidemiology* (6<sup>th</sup> ed.). New York: Oxford University Press.
- Riva, S., Gorini, A., Cutica, I., Mazzocco, K. & Pravettoni, G. (2015). Translation, cross-cultural adaptation, and reliability, of the Italian version of the Passive Risk Taking (PRT) Scale. *Judgement and Decision Making* 10(6), pp. 597-604.
- Sherman, E., Brooks, B., Iverson, G., Slick, D. & Strauss, E. (2011). Reliability and Validity in Neuropsychology. In *The Little Black Book of Neuropsychology: A Syndrome-Based Approach*. Springer Science+Business Media.
- Soley-Bori, M. (2013). Dealing with missing data: Key assumptions and methods for applied analysis. Technical Report, pp. 1-19. Ανάκτηση 12/6/2016, από <http://www.bu.edu/sph/files/2014/05/Marina-tech-report.pdf>
- Tavakol, M. & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education* 2, pp. 53-55.
- Vaske, J., Beaman, J. & Sponarski, C. (n.d.). Rethinking Internal Consistency in Cronbach's alpha. Ανάκτηση 12/6/2016, από [http://warnercnr.colostate.edu/~jerryv/Cronbach/consistency\\_in\\_alpha.pdf](http://warnercnr.colostate.edu/~jerryv/Cronbach/consistency_in_alpha.pdf)
- Webb, N., Shavelson, R. & Haertel, E. (2006). Reliability Coefficients and Generalizability Theory. *Handbook of Statistics* 26, pp. 1-44.
- Yong, A. G. & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Method for Psychology* 9(2), pp. 79-94.
- Zygmunt, C. & Smith, M. (2014). Robust factor analysis in the presence of normality violations, missing data, and outliers: Empirical questions and possible solutions. *The Quantitative Methods for Psychology* 10(1).