

ΙΣΟΔΥΝΑΜΙΑ ΤΗΣ ΜΕΤΡΗΣΗΣ ΣΤΗ ΣΥΓΚΡΙΤΙΚΗ ΨΥΧΟΛΟΓΙΚΗ ΕΡΕΥΝΑ

Συμεών Π. Βλαχόπουλος

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Περίληψη: Σημαντικά ερωτήματα τα οποία γεννιούνται στη συγκριτική ψυχολογική έρευνα περιλαμβάνουν το εάν συμμετέχοντες από διάφορες ομάδες που αντιπροσωπεύουν διαφορετικούς πληθυσμούς ερμηνεύουν τις ερωτήσεις των ερωτηματολογίων με τον ίδιο τρόπο και, επίσης, εάν ατομικές διαφορές λόγω διαφορετικής εθνικότητας ή φύλου ή άλλης κατηγοριοποίησης αποκλείουν την παρόμοια ερμηνεία των τιμών στα ερωτηματολόγια. Σε αυτό το άρθρο παρουσιάζεται το θέμα της ισοδυναμίας της μέτρησης στο πλαίσιο της επιβεβαιωτικής παραγοντικής ανάλυσης και η εφαρμογή του στην ψυχολογική έρευνα. Συγκεκριμένα, (α) συζητείται γιατί η αξιολόγηση της ισοδυναμίας της μέτρησης σε διαφορετικούς πληθυσμούς είναι σημαντική στην ψυχολογική έρευνα, (β) περιγράφεται η διαδικασία της αξιολόγησης της ισοδυναμίας της μέτρησης, (γ) συζητούνται θεωρητικές ερμηνείες και πρακτικές εφαρμογές των διάφορων βαθμών της ισοδυναμίας της μέτρησης, και (δ) παρουσιάζεται ένα παράδειγμα της αξιολόγησης της ισοδυναμίας της μέτρησης από την ψυχολογική έρευνα.

Λέξεις κλειδιά: Διαπολιτισμική εγκυρότητα, Διομαδική σύγκριση, Επιβεβαιωτική παραγοντική ανάλυση πολλαπλών ομάδων.

Σημείωση: Εκφράζονται θερμές ευχαριστίες προς την κ. Αναστασία Ευκλείδη για τη συνεισφορά της στη μετάφραση της σχετικής στατιστικής ορολογίας του άρθρου.

Διεύθυνση: Συμεών Π. Βλαχόπουλος, Εργαστήριο Κοινωνικής Έρευνας στη Φυσική Δραστηριότητα, Τμήμα Επιστήμης Φυσικής Αγωγής και Αθλητισμού Σερρών, Άγιος Ιωάννης, 621 10 Σέρρες. Τηλ.: 2310-991045. Fax: 23210-64806. E-mail: vlachop@phed-sr.auth.gr

ΕΙΣΑΓΩΓΗ

Η συγκριτική έρευνα εξετάζει ποσοτικά τις ομοιότητες και τις διαφορές συγκεκριμένων πληθυσμιακών ομάδων και επιχειρεί να ερμηνεύσει αυτές τις διαφορές. Διάφοροι κίνδυνοι που ελλοχεύουν στη συγκριτική έρευνα εκφράζονται με ερωτήματα όπως: (α) Αλλάζει η συμμετοχή σε μια συγκεκριμένη διαδικασία (π.χ., η συμμετοχή σε ένα πρόγραμμα παρέμβασης) την οπτική μιας ομάδος ατόμων με βάση την οποία απαντά σε κάποιο όργανο μέτρησης διαχρονικά; (β) Τα άτομα που ανήκουν σε διαφορετικές κουλτούρες θα ερμηνεύσουν νοηματικά το ίδιο ερωτηματολόγιο με τον ίδιο τρόπο; (γ) Οι διαφορές ως προς το φύλο, την εθνικότητα ή και άλλες ατομικές διαφορές εμποδίζουν τα άτομα να απαντήσουν στο ίδιο ερωτηματολόγιο με τον ίδιο τρόπο; (Vandenberg & Lance, 2000).

Ένας κίνδυνος για τη συναγωγή έγκυρων συγκριτικών συμπερασμάτων ανάμεσα σε διαφορετικούς πληθυσμούς σε αυτό το είδος της έρευνας ελλοχεύει και στην καταλληλότητα των οργάνων της μέτρησης (Gregorich, 2006. Little, 1997. Meredith, 1993. Yoo, 2002). Αυτός ο κίνδυνος γίνεται εντονότερος όταν η αξιολόγηση πραγματοποιείται με τη μέθοδο των αυτο-αναφορών και αφορά χαρακτηριστικά τα οποία δεν είναι ευθέως παρατηρήσιμα όπως οι προθέσεις, τα κίνητρα, και οι καταστάσεις της διάθεσης. Η αντιμετώπιση αυτού του κινδύνου γίνεται ακόμη πιο επιτακτική όταν οι ομάδες των ατόμων που συγκρίνονται ανήκουν σε διαφορετικές κουλτούρες ή έθνη ή μιλούν διαφορετική γλώσσα και τα όργανα μέτρησης που χρησιμοποιούνται είναι μεταφρασμένα στις αντίστοιχες γλώσσες (Janssens, Brett, & Smith, 1995. Reise, Widaman, & Pugh, 1993. Rioridan & Vandenberg, 1994. Steenkamp & Baumgartner, 1998). Άλλα παραδείγματα ομάδων σύγκρισης περιλαμβάνουν άτομα με διαφορετικά επίπεδα ακαδημαϊκής επίτευξης (Byrne, Shavelson, & Muthen, 1989), διαφορετικού φύλου (Byrne, 1994) και άτομα που ανήκουν στην πειραματική και στην ομάδα ελέγχου στα πλαίσια της πειραματικής έρευνας (Pentz & Chou, 1994).

Οι κλίμακες που χρησιμοποιούνται για τη μέτρηση τέτοιων χαρακτηριστικών αποτελούνται από ένα αριθμό θεμάτων (ή ερωτήσεων) όπου κάθε ένα θέμα θεωρείται ένας έμμεσος δείκτης του χαρακτηριστικού που αξιολογείται. Αυτό συμβαίνει διότι το χαρακτηριστικό δεν είναι άμεσα παρατηρήσιμο δεδομένου ότι αποτελεί μια υποθετική έννοια (π.χ., κίνητρο). Τα χαρακτηριστικά τα οποία δεν είναι άμεσα παρατηρήσιμα ονομάζονται

λανθάνουσες μεταβλητές¹. Στο πλαίσιο αυτής της φιλοσοφίας, η κατάλληλη μέθοδος για να εξεταστεί εάν τα θέματα μιας κλίμακας αποτελούν έμμεσους δείκτες αξιολόγησης μιας συγκεκριμένης λανθάνουσας μεταβλητής είναι η *επιβεβαιωτική παραγοντική ανάλυση*² (ΕΠΑ. Bollen, 1989). Επίσης, η *ΕΠΑ πολλαπλών ομάδων*³ είναι η κατάλληλη μέθοδος για να εξεταστεί: (α) εάν τα θέματα της κλίμακας, ως έμμεσοι δείκτες της έννοιας που μετριέται, είναι νοηματικώς ισοδύναμα ανά πληθυσμιακή ομάδα και (β) εάν η σύγκριση των δειγμάτων που αντιπροσωπεύουν τον κάθε πληθυσμό αντανakλά αληθινές πληθυσμιακές διαφορές και δεν επηρεάζεται από χαρακτηριστικά των ομάδων τα οποία είναι άσχετα με το χαρακτηριστικό που αξιολογείται μέσω της κλίμακας (Cheung & Rensvold, 2000. Little, 2000. Meredith & Teresi, 2006. Steenkamp & Baumgartner, 1998).

Ο όρος *ισοδυναμία της μέτρησης*⁴ είναι ένας γενικός όρος και αυτή η ισοδυναμία μπορεί να εξεταστεί αναφορικά με διάφορες παραμέτρους των μοντέλων επιβεβαιωτικής παραγοντικής ανάλυσης. Η εξέταση της ισοδυναμίας της μέτρησης περιλαμβάνει έναν αριθμό από συγκεκριμένες δοκιμασίες, οι οποίες αντιστοιχούν σε διαφορετικούς τύπους ισοδυναμίας της μέτρησης. Αυτοί οι διαφορετικοί τύποι ισοδυναμίας της μέτρησης αντανακλούν ισοδυναμία συγκεκριμένων παραμέτρων των ψυχομετρικών ιδιοτήτων των θεμάτων της κλίμακας (Meredith, 1993. Meredith & Teresi, 2006). Υπάρχουν διαφορετικοί τύποι απουσίας της ισοδυναμίας της μέτρησης και κάθε ένας από αυτούς αντιστοιχεί σε μία συγκεκριμένη απειλή εγκυρότητας ως προς την ποσοτική σύγκριση συγκεκριμένων πληθυσμιακών ομάδων (Gregorich, 2006). Η πλέον σύγχρονη πρόταση για τον έλεγχο της ισοδυναμίας της μέτρησης περιγράφεται από τον Gregorich (2006). Οι τύποι της ισοδυναμίας της μέτρησης που προτείνονται από τον Gregorich (2006) είναι η *διαστασιακή ισοδυναμία*⁵ (δηλαδή ισοδυναμία του αριθμού των παραγόντων), η *συμμορφική ισοδυναμία*⁶ (δηλαδή ισοδυναμία της μορφής των παραγόντων), η *μετρική ισοδυναμία*⁷ (δηλαδή ισοδυναμία της

¹ Latent variables.

² Confirmatory Factor Analysis.

³ Multi-group Confirmatory Factor Analysis.

⁴ Measurement invariance.

⁵ Dimensional invariance.

⁶ Configural invariance.

⁷ Metric invariance.

φόρτισης των παραγόντων), η *ισχυρή παραγοντική ισοδυναμία*⁸, και η *αυστηρή παραγοντική ισοδυναμία*⁹.

Οι δοκιμασίες που χρησιμοποιούνται για την εξέταση της ισοδυναμίας της μέτρησης έχουν συζητηθεί εκτενώς στη βιβλιογραφία (Bollen, 1989. Byrne et al., 1989. Cheung & Rensvold, 2000. Drasgow & Kanfer, 1985. Little, 1997. Steenkamp & Baumgartner, 1998. Vandenberg & Lance, 2000). Υπάρχουν δύο γενικές κατηγορίες δοκιμασιών της παραγοντικής ισοδυναμίας (Little, 1997). Στην Κατηγορία I ανήκουν οι δοκιμασίες ισοδυναμίας που επικεντρώνονται στις ψυχομετρικές ιδιότητες των τιμών των οργάνων μέτρησης και περιλαμβάνουν τη συμμορφική ισοδυναμία (Buss & Royce, 1975. Irvine, 1969. Suzuki & Rancer, 1994), τη μετρική ισοδυναμία (Horn & McArdle, 1992), την *ισοδυναμία του σφάλματος μέτρησης*¹⁰ (Mullen, 1995. Singh, 1995), και την *ισοδυναμία της κλίμακας μέτρησης*¹¹ (Meredith, 1993. Steenkamp & Baumgartner, 1998. Vandenberg & Lance, 2000). Στην Κατηγορία II ανήκουν οι δοκιμασίες ισοδυναμίας που σχετίζονται με διομαδικές διαφορές στους λανθάνοντες μέσους όρους και στην παραγοντική διακύμανση και συνδιακύμανση.

Σε γενικές γραμμές, οι δοκιμασίες ισοδυναμίας της Κατηγορίας I αποτελούν προϋπόθεση για τη διεξαγωγή των δοκιμασιών της ισοδυναμίας της Κατηγορίας II, οι οποίες αντανακλούν διαφορές θεωρητικού και περισσότερο ουσιαστικού ενδιαφέροντος συγκριτικά με τις δοκιμασίες της Κατηγορίας I. Τα αποτελέσματα κάθε δοκιμασίας της ισοδυναμίας της μέτρησης παρέχουν πληροφορίες σχετικά με το είδος της υπόθεσης της διομαδικής σύγκρισης, η οποία μπορεί έγκυρα να εξεταστεί περαιτέρω από τους ερευνητές. Για παράδειγμα, αν υποστηρίζεται από τον έλεγχο που έγινε η μετρική ισοδυναμία για συγκεκριμένα θέματα της κλίμακας, οι ερευνητές μπορούν να χρησιμοποιήσουν αυτά τα θέματα σε ένα μοντέλο δομικών εξισώσεων με λανθάνουσες μεταβλητές ώστε να συγκρίνουν διομαδικά ένα συγκεκριμένο συντελεστή παλινδρόμησης. Αυτή η σύγκριση μπορεί να γίνει μόνο για τις λανθάνουσες μεταβλητές των οποίων οι παραγοντικές φορτίσεις αποδείχτηκαν ισοδύναμες στον προηγούμενο έλεγχο της μετρικής ισοδυναμίας τους.

⁸ Strong factorial invariance.

⁹ Strict factorial invariance

¹⁰ Measurement error invariance.

¹¹ Scalar invariance.

Ακολουθεί η περιγραφή του ελέγχου κάθε τύπου ισοδυναμίας της μέτρησης και τα είδη των διομαδικών συγκρίσεων που υποστηρίζονται δεδομένης της ισοδυναμίας των παραμέτρων που εξετάζονται στο αντίστοιχο βήμα. Οι έλεγχοι της ισοδυναμίας της μέτρησης μαζί με τις υποθέσεις που εξετάζονται και τις διομαδικές συγκρίσεις που υποστηρίζονται παρουσιάζονται στον Πίνακα 1. Σε γενικές γραμμές, οι ερευνητές πρέπει να ακολουθήσουν την προτεινόμενη ιεραρχική σειρά στους τύπους των ελέγχων ισοδυναμίας που πρόκειται να διενεργήσουν και προϋπόθεση για τον έλεγχο ισοδυναμίας ενός επιπέδου είναι η επίτευξη ισοδυναμίας στο αμέσως προηγούμενο ιεραρχικά επίπεδο στα συγκεκριμένα θέματα του οργάνου μέτρησης.

Πίνακας 1. Σύνοψη των παραγοντικών μοντελών πολλαπλών ομάδων για την εξέταση της ισοδυναμίας της μέτρησης

Έλεγχος διαστασιακής ισοδυναμίας	Έλεγχος συμμορφικής ισοδυναμίας	Έλεγχος μετρικής ισοδυναμίας	Έλεγχος ισχυρής ισοδυναμίας	Έλεγχος αυστηρής ισοδυναμίας
Ισοδυναμία της μέτρησης				
Ισοδυναμία του αριθμού των παραγόντων	+ Ισοδυναμία της μορφής των παραγόντων	+ Ισοδυναμία της φόρτισης των παραγόντων	+ Ισοδυναμία των σταθερών των θεμάτων	+ Ισοδυναμία των υπολοίπων σφάλματος των θεμάτων
Διομαδικές συγκρίσεις που υποστηρίζονται				
Καμία	Καμία	Διακυμάνσεις / συνδιακυμάνσεις των λανθανόντων παραγόντων και συντελεστές παλινδρόμησης λανθανόντων παραγόντων	Μέσοι όροι λανθανόντων και παρατηρούμενων παραγόντων	Παρατηρούμενες διακυμάνσεις / συνδιακυμάνσεις θεμάτων και (υπο)κλιμάκων

ΤΥΠΟΙ ΙΣΟΔΥΝΑΜΙΑΣ ΤΗΣ ΜΕΤΡΗΣΗΣ

Διαστασιακή ισοδυναμία

Κατά τον έλεγχο της διαστασιακής ισοδυναμίας, δηλαδή της ισοδυναμίας του αριθμού των παραγόντων, εξετάζεται εάν η κλίμακα (π.χ., ερωτηματολόγιο) αντανakλά τον ίδιο αριθμό παραγόντων ανά πληθυσμιακή ομάδα. Αυτός ο έλεγχος επικεντρώνεται μόνο στον αριθμό των παραγόντων και όχι στη διαμόρφωση ή αλλιώς σύσταση των παραγόντων με βάση τα θέματα που λειτουργούν ως έμμεσοι δείκτες του κάθε παράγοντα. Στην ποσοτική συγκριτική έρευνα είναι πολύ σημαντικό να υπάρχει ο ίδιος αριθμός παραγόντων στην κλίμακα που χρησιμοποιείται ανά πληθυσμιακή ομάδα, διότι, όταν η μέτρηση του ίδιου χαρακτηριστικού στηρίζεται σε διαφορετικό αριθμό παραγόντων ανά πληθυσμιακή ομάδα, αυτό σημαίνει ότι τουλάχιστον ένας ή και περισσότεροι παράγοντες διαφέρουν ανά πληθυσμιακή ομάδα στην ποιοτική τους διάσταση. Παρά το ότι η ισοδυναμία του αριθμού των παραγόντων είναι απαραίτητη και επιθυμητή, από μόνη της δε διασφαλίζει την έγκυρη ποσοτική σύγκριση των πληθυσμιακών ομάδων.

Συμμορφική ισοδυναμία

Με δεδομένο ότι έχει υποστηριχθεί από τον έλεγχο η διαστασιακή ισοδυναμία, ελέγχεται στη συνέχεια η συμμορφική ισοδυναμία, δηλαδή εάν σε κάθε πληθυσμιακή ομάδα το ίδιο ακριβώς σύνολο θεμάτων της κλίμακας ορίζει τον ίδιο παράγοντα, ο οποίος έχει προσδιορισθεί θεωρητικώς. Στην περίπτωση της συμμορφικής ισοδυναμίας δεν απαιτείται η ισοδυναμία καμίας παραμέτρου των ψυχομετρικών ιδιοτήτων των θεμάτων της κλίμακας ανά ομάδα (π.χ., μετρική ισοδυναμία, ισοδυναμία του σφάλματος μέτρησης ανά θέμα, ισοδυναμία παραγοντικής διακύμανσης, κ.ά.). Το μόνο που απαιτείται είναι η ομοιότητα του συνόλου των θεμάτων που αντανakλούν και προσδιορίζουν τον παράγοντα. Κατά τον καθορισμό του παραγοντικού μοντέλου που θα εξεταστεί στην ΕΠΑ καθορίζεται το σύνολο των θεμάτων που λειτουργούν ως δείκτες του συγκεκριμένου παράγοντα και έτσι μπορεί να διενεργηθεί ο έλεγχος της υπόθεσης της συμμορφικής ισοδυναμίας ανά πληθυσμιακή ομάδα. Στήριξη αυτής της υπόθεσης σημαίνει ότι τα άτομα των διαφορετικών ομάδων αντιλαμβάνονται την έννοια που μετρείται με τον ίδιο τρόπο (Riordan & Vandenberg, 1994). Αντιθέτως, η υπόθεση αυτή μπορεί να μην υποστηριχθεί όταν για παράδειγ-

μα η έννοια είναι τόσο αφηρημένη ώστε η αντίληψη των ατόμων για την έννοια να επηρεάζεται από το πολιτισμικό πλαίσιο στο οποίο ανήκουν (Tayeb, 1994), ή όταν τα άτομα των διαφορετικών ομάδων χρησιμοποιούν διαφορετική οπτική και δίνουν διαφορετικό νόημα στην έννοια (Millsap & Ever-son, 1991. Millsap & Hartog, 1988. Riordan & Vandenberg, 1994). Επίσης, η υπόθεση μπορεί να μην υποστηριχθεί εξαιτίας προβλημάτων με τη συλλογή των δεδομένων, λάθη στη μετάφραση των θεμάτων, κ.τ.λ. (Cheung & Rensvold, 2002). Ο έλεγχος της συμμορφικής ισοδυναμίας αποτελεί ισχυρότερη δοκιμασία συγκριτικά με τον έλεγχο της διαστασιακής ισοδυναμίας. Ωστόσο, ακόμη και αυτός ο έλεγχος δεν επαρκεί για τη διασφάλιση της έγκυρης ποσοτικής σύγκρισης των πληθυσμιακών ομάδων.

Μετρική ισοδυναμία

Με δεδομένη τη συμμορφική ισοδυναμία, ελέγχεται στη συνέχεια η μετρική ισοδυναμία, δηλαδή η ισοδυναμία της φόρτισης των παραγόντων ανά πληθυσμιακή ομάδα. Κάτι τέτοιο θα σήμαινε ότι το περιεχόμενο των θεμάτων ανά παράγοντα εκλαμβάνεται με τον ίδιο τρόπο από τα άτομα της κάθε πληθυσμιακής ομάδας και ότι ο συγκεκριμένος παράγοντας έχει το ίδιο νόημα σε όλες τις πληθυσμιακές ομάδες που μελετούνται. Με άλλα λόγια, όταν ένας παράγοντας έχει το ίδιο νόημα ανά ομάδα, ισοδύναμη αναμένεται να είναι και η σχέση του κάθε ενός θέματος που προσδιορίζει τον παράγοντα (δηλ. η παραγοντική φόρτιση) με τον παράγοντα αυτόν, ανά ομάδα. Η εξέταση της μετρικής ισοδυναμίας ενός ή και περισσότερων παραγόντων ταυτοχρόνως για δύο ή και περισσότερες ομάδες ατόμων εκτελείται με τη μέθοδο της ΕΠΑ πολλαπλών ομάδων.

Συγκεκριμένα, η μετρική ισοδυναμία ελέγχεται με την επιβολή *περιορισμών ισότητας*¹² στις παραγοντικές φορτίσεις των θεμάτων που προσδιορίζουν τον παράγοντα ανά ομάδα. Εάν το θεωρητικό παραγοντικό μοντέλο με τους περιορισμούς ισότητας στις παραγοντικές φορτίσεις ταιριάζει με τα δεδομένα, αυτό σημαίνει ότι οι περιορισμοί ισότητας ισχύουν στα συγκεκριμένα δεδομένα και άρα οι συγκεκριμένες παραγοντικές φορτίσεις είναι ισοδύναμες ανά ομάδα. Ένας συμπληρωματικός τρόπος εξέτασης της μετρικής ισοδυναμίας είναι η σύγκριση του

¹² Equality constraints.

μοντέλου της συμμορφικής ισοδυναμίας με το μοντέλο της μετρικής ισοδυναμίας, το οποίο περιλαμβάνει, επιπλέον του πρώτου, και τους περιορισμούς ισότητας των παραγοντικών φορτίσεων. Η σύγκριση γίνεται με τη δοκιμασία του χ^2 και των δεικτών καλής προσαρμογής¹³. Αυτό γίνεται χρησιμοποιώντας τα δεδομένα και από τους δύο πληθυσμούς. Εάν η επάρκεια του μοντέλου της μετρικής ισοδυναμίας αποδειχθεί χειρότερη συγκριτικά με το μοντέλο της συμμορφικής ισοδυναμίας σημαίνει ότι η μετρική ισοδυναμία δεν υποστηρίζεται. Με άλλα λόγια, οι περιορισμοί ισότητας στις παραγοντικές φορτίσεις δεν υποστηρίζονται από τα δεδομένα και, κατά συνέπεια, τα θέματα που φορτίζουν στους παράγοντες δεν είναι ισοδύναμα ανά ομάδα. Έτσι, οι ερευνητές δεν μπορούν να προχωρήσουν σε περαιτέρω ελέγχους άλλων μοντέλων που να προϋποθέτουν τη μετρική ισοδυναμία (π.χ., σύγκριση της διακύμανσης του παράγοντα ανά ομάδα).

Υποστήριξη της μετρικής ισοδυναμίας επιτρέπει στους ερευνητές να προχωρήσουν σε έγκυρη ποσοτική σύγκριση της *διακύμανσης του λανθάνοντα παράγοντα*¹⁴ ή/και *συνδιακύμανσης του λανθάνοντα παράγοντα*¹⁵ ανά ομάδα καθώς και των *δομικών συντελεστών*¹⁶ κατά την *ανάλυση παλινδρόμησης*¹⁷ ανάμεσα σε δύο *λανθάνουσες μεταβλητές*¹⁸ στα πλαίσια της *ανάλυσης διαδρομών*¹⁹ (Gregorich, 2006). Τέτοιες συγκρίσεις είναι έγκυρες διότι (α) οι αντίστοιχοι παράγοντες θεωρούνται νοηματικά ισοδύναμοι ανά ομάδα, και (β) στο μοντέλο της ΕΠΑ η παρατηρούμενη *συνολική διακύμανση*²⁰ των θεμάτων ανά λανθάνοντα παράγοντα διασπάται στη *διακύμανση που αποδίδεται στο λανθάνοντα παράγοντα* (δηλαδή την *αληθή διακύμανση*²¹) και στη *διακύμανση του σφάλματος των θεμάτων*²² (δηλαδή, τη διακύμανση που δεν οφείλεται στην επίδραση του χαρακτηριστικού το οποίο μετριέται αλλά μπορεί να οφείλεται

¹³ Goodness-of-fit indexes.

¹⁴ Latent factor variance.

¹⁵ Latent factor covariance.

¹⁶ Structure coefficients.

¹⁷ Regression analysis.

¹⁸ Latent variables.

¹⁹ Path analysis.

²⁰ Total variance.

²¹ True variance.

²² Residual variance.

ταυτοχρόνως και σε κάποιον άλλο συστηματικό παράγοντα και στο σφάλμα μέτρησης). Αυτό σημαίνει ότι οι διαφορές στην αληθή διακύμανση και συνδιακύμανση των παραγόντων ανά ομάδα δεν επηρεάζονται από πιθανές διομαδικές διαφορές στη διακύμανση του σφάλματος των θεμάτων. Σε αντίθεση με τους παραπάνω ελέγχους, όμως, η μετρική ισοδυναμία δεν μπορεί να υποστηρίξει διομαδικές συγκρίσεις των διακυμάνσεων και συνδιακυμάνσεων που αναφέρονται σε παρατηρούμενες τιμές²³ (π.χ., τιμές διακύμανσης μεμονωμένων θεμάτων ή μέσων όρων υποκλιμάκων), αλλά μόνο συγκρίσεις στους λανθάνοντες παράγοντες. Αυτό συμβαίνει διότι διομαδικές διαφορές σε παρατηρούμενες διακυμάνσεις και συνδιακυμάνσεις δεν αντανακλούν υποχρεωτικά και αντίστοιχες διαφορές στις λανθάνουσες διακυμάνσεις και συνδιακυμάνσεις των παραγόντων μόνο.

Πώς, όμως, ερμηνεύεται το φαινόμενο της απουσίας της μετρικής ισοδυναμίας; Στην περίπτωση που δεν υποστηρίζεται η μετρική ισοδυναμία υπάρχουν δύο εναλλακτικές ερμηνείες. Η πρώτη ερμηνεία είναι ότι ο παράγοντας ή τουλάχιστον κάποια θέματα που προσδιορίζουν τον παράγοντα δεν έχουν το ίδιο νόημα ανά ομάδα. Η δεύτερη είναι ότι ένας αριθμός των παραγοντικών φορτίσεων έχει επηρεαστεί από το *ύφος ακραίων απαντήσεων*²⁴, το οποίο μπορεί να εμφανιστεί με δύο μορφές (Baumgartner & Steenkamp, 2001. Cheung & Rensvold, 2000). Πρώτον, με ακραίες απαντήσεις στις κλίμακες απαντήσεων από ομάδες των οποίων τα μέλη δίνουν αξία στο χαρακτηριστικό της αποφασιστικότητας ή/και της βεβαιότητας και, δεύτερον, με την αποφυγή ακραίων απαντήσεων (δηλ. την επιλογή απαντήσεων κοντά στο μέσο της κλίμακας). Η δεύτερη περίπτωση μπορεί να εμφανιστεί σε ομάδες που προσδίδουν αξία στο χαρακτηριστικό της μετριοπάθειας (Gregorich, 2006). Επειδή το *ύφος ακραίων απαντήσεων* μπορεί να επηρεάσει τις συσχετίσεις των θεμάτων μεταξύ τους και, κατά συνέπεια, τις παραγοντικές φορτίσεις και τις κατανομές των τιμών ανά θέμα, έγκειται στους ερευνητές να αποφασίσουν εάν η μετρική ισοδυναμία δεν υποστηρίχθηκε εξαιτίας του διαφορετικού νοήματος του παράγοντα ανά ομάδα ή του διαφορετικού *ύφους ακραίων απαντήσεων* ανά ομάδα (Gregorich, 2006).

²³ Observed values.

²⁴ Extreme response style.

Ισχυρή παραγοντική ισοδυναμία

Με δεδομένο ότι υποστηρίζεται η μετρική ισοδυναμία, τότε οι ερευνητές πρέπει να εξετάσουν την ύπαρξη ισχυρής παραγοντικής ισοδυναμίας²⁵, δηλαδή την πιθανότητα οι απαντήσεις στα θέματα του παράγοντα να επηρεάζονται με τέτοιο τρόπο ώστε οι τιμές των θεμάτων να καταλήγουν να είναι συστηματικά υψηλότερες ή χαμηλότερες για τη μία ομάδα συγκριτικά με τις άλλες ομάδες εξαιτίας επιρροών οι οποίες δε σχετίζονται με τους λανθάνοντες παράγοντες αυτούς καθαυτούς. Αυτός ο κίνδυνος στην ισοδυναμία της μέτρησης αντανακλά το *ύφος διαφορικής συμφωνίας στις απαντήσεις*²⁶, δηλαδή την τάση των ατόμων να συμφωνούν περισσότερο ή λιγότερο στα ίδια θέματα ανάλογα με την ομάδα στην οποία ανήκουν εξαιτίας, παραδείγματος χάρη, των πολιτισμικών χαρακτηριστικών τους (Baumgartner & Steenkamp, 2001. Cheung & Rensvold, 2000). Η επίδραση αυτής της τάσης είναι συσσωρευτική και επηρεάζει τους παρατηρούμενους μέσους όρους αλλά όχι τη διακύμανση των απαντήσεων. Αν οι επιδράσεις στα επίπεδα των τιμών των απαντήσεων δεν είναι ισοδύναμες ανά ομάδα τότε θα διαφοροποιήσουν τους μέσους όρους των παραγόντων ανά ομάδα. Οι *σταθερές των θεμάτων*²⁷ αντιπροσωπεύουν την τιμή του μέσου όρου του κάθε θέματος, η οποία αντιστοιχεί στην τιμή του μηδενός για το μέσο όρο του λανθάνοντα παράγοντα. Με άλλα λόγια, η διαφοροποίησή τους δίνει απάντηση στο ερώτημα σχετικά με το εάν οι κλίμακες μέτρησης μοιράζονται τον ίδιο λειτουργικό ορισμό ανά ομάδα (δηλαδή, εάν έχουν τα ίδια διαστήματα και σημεία εκκίνησης. Βλ. Cheung & Rensvold, 2002). Όταν υπάρχει ισχυρή παραγοντική ισοδυναμία θεωρείται ότι (α) οι διαφορές στους μέσους όρους του λανθάνοντα παράγοντα ανά ομάδα δεν επηρεάζονται από την τάση των ατόμων να συμφωνούν περισσότερο ή λιγότερο στα ίδια θέματα ανάλογα με την ομάδα στην οποία ανήκουν και (β) οι διαφορές στους *παρατηρούμενους μέσους όρους*²⁸ θα έχουν άμεση σχέση με τις αντίστοιχες διαφορές στους *μέσους όρους των λανθανόντων παραγόντων*²⁹ χωρίς να επηρεάζονται από τον κίνδυνο που προαναφέρθηκε.

²⁵ Η ισχυρή παραγοντική ισοδυναμία αναφέρεται αλλού και ως ισοδυναμία της κλίμακας μέτρησης (scalar invariance. Βλ. Mullen, 1995).

²⁶ Differential acquiescence response style.

²⁷ Item intercepts.

²⁸ Observed means.

²⁹ Latent factor means.

Η ισχυρή παραγοντική ισοδυναμία μπορεί να εξεταστεί με την επιβολή περιορισμών ισότητας στις σταθερές των θεμάτων εκείνων των οποίων οι παραγοντικές φορτίσεις αποδείχθηκαν ισοδύναμες στον έλεγχο του μοντέλου της μετρικής ισοδυναμίας. Εάν οι δείκτες προσαρμογής οδηγούν στο συμπέρασμα ότι το μοντέλο της ισχυρής παραγοντικής ισοδυναμίας είναι συμβατό με τα δεδομένα τότε αυτό σημαίνει ότι γενικά στο μοντέλο οι περιορισμοί ισότητας που έχουν προστεθεί σε κάποιες ή σε όλες τις σταθερές είναι και αυτοί συμβατοί με τα δεδομένα και έτσι αυτές οι σταθερές είναι ισοδύναμες ανά ομάδα. Ταυτοχρόνως, πρέπει να γίνει και σύγκριση του βαθμού στον οποίο το παραγοντικό μοντέλο που αντιστοιχεί στην υπόθεση της ισχυρής παραγοντικής ισοδυναμίας διαφοροποιείται στατιστικώς από το μοντέλο που αντιστοιχεί στην υπόθεση της μετρικής ισοδυναμίας. Εάν το πρώτο μοντέλο βρεθεί να είναι στατιστικώς χειρότερο από το δεύτερο σημαίνει ότι η υπόθεση της ισχυρής παραγοντικής ισοδυναμίας δεν υποστηρίζεται.

Αυστηρή παραγοντική ισοδυναμία

Στις βασικές εφαρμογές της ΕΠΑ η συνολική παρατηρούμενη διακύμανση του κάθε θέματος ανά λανθάνοντα παράγοντα διασπάται στη διακύμανση που αποδίδεται στο λανθάνοντα παράγοντα (δηλαδή την αληθή διακύμανση) και στη διακύμανση του σφάλματος των θεμάτων (δηλαδή τη διακύμανση που δεν οφείλεται στην επίδραση του χαρακτηριστικού το οποίο μετριέται). Εάν ο στόχος της έρευνας είναι η σύγκριση ανά ομάδα των διακυμάνσεων των παρατηρούμενων παραγόντων και όχι των λανθανόντων, η σύγκριση αυτή πρέπει να αντανakλά εξ ολοκλήρου διαφορές στη διακύμανση που αποδίδεται στο λανθάνοντα παράγοντα, χωρίς να επηρεάζεται από πιθανές διαφορές στις διακυμάνσεις των υπολοίπων σφάλματος των θεμάτων. Έτσι, λοιπόν, για να προχωρήσουν οι ερευνητές σε έγκυρη σύγκριση της παρατηρούμενης διακύμανσης είτε του παράγοντα είτε συγκεκριμένων θεμάτων πρέπει να ισχύει μία επιπρόσθετη μορφή ισοδυναμίας της μέτρησης, δηλαδή η ισοδυναμία της διακύμανσης των υπολοίπων σφάλματος των θεμάτων. Επειδή η ισοδυναμία της διακύμανσης των υπολοίπων σφάλματος δεν αποτελεί προϋπόθεση για την έγκυρη σύγκριση των μέσων όρων των λανθανόντων παραγόντων και επειδή είναι αυτές οι συγκρίσεις που ενδιαφέρουν συχνότερα τους ερευνητές, ο έλεγχος της ισοδυναμίας της διακύμανσης των υπολοίπων σφάλματος των θεμάτων θεω-

ρείται μειωμένης πρακτικής αξίας. Όμως, όταν στόχος είναι η έγκυρη σύγκριση των παρατηρούμενων διακυμάνσεων ανά ομάδα, τότε η υποστήριξη της αυστηρής παραγοντικής ισοδυναμίας κρίνεται απαραίτητη (Gregorich, 2006). Στον καθορισμό του παραγοντικού μοντέλου της αυστηρής παραγοντικής ισοδυναμίας διατηρούνται οι περιορισμοί ισότητας που ισχύουν στο μοντέλο της ισχυρής παραγοντικής ισοδυναμίας (δηλ. στις αντίστοιχες παραγοντικές φορτίσεις και στις σταθερές των θεμάτων) και προστίθενται περιορισμοί και στα σφάλματα των θεμάτων.

Εάν υποστηριχθεί η αυστηρή παραγοντική ισοδυναμία, τότε οι διαφορές στις διακυμάνσεις των παρατηρούμενων παραγόντων ισοδυναμούν με τις διαφορές στις διακυμάνσεις των λανθανόντων παραγόντων. Επιπροσθέτως, η εξέταση της ισοδυναμίας των υπολοίπων σφάλματος των θεμάτων παρέχει πληροφορίες σχετικά με το εάν τα θέματα μετρούν το λανθάνον χαρακτηριστικό με τον ίδιο βαθμό σφάλματος (Cheung & Rensvold, 2002). Η ισοδυναμία σφάλματος μπορεί να μην υποστηριχθεί, για παράδειγμα, όταν τα άτομα της μιας ομάδας είναι λιγότερο εξοικειωμένα με το όργανο μέτρησης κατά τη διαδικασία της αξιολόγησης από τα άτομα των άλλων ομάδων και έτσι απαντούν με διαφορετικό βαθμό συνέπειας (Mullen, 1995). Επίσης, το ίδιο φαινόμενο μπορεί να προκληθεί από διαφορές στο λεξιλόγιο, ιδιοματισμούς, τη γραμματική και τη σύνταξη των θεμάτων, καθώς και τις κοινές εμπειρίες των ατόμων στην ίδια κουλτούρα (Malpass, 1977).

Μερική ισοδυναμία της μέτρησης

Εκτός από την *πλήρη ισοδυναμία της μέτρησης*³⁰ κάθε τύπου που παρουσιάστηκε προηγουμένως υπάρχει και η περίπτωση της *μερικής ισοδυναμίας της μέτρησης*³¹ για κάθε έναν από τους τύπους ισοδυναμίας που περιγράφηκαν. Για παράδειγμα, στην περίπτωση μιας μονοδιάστατης κλίμακας με εννέα ερωτήσεις, όταν μόνον οι έξι παραγοντικές φορτίσεις βρεθούν να είναι ισοδύναμες ανά ομάδα έχουμε την περίπτωση της *μερικής μετρικής ισοδυναμίας*³². Η έννοια της μερικής ισοδυναμίας της μέτρησης αρχικά πα-

³⁰ Full measurement invariance.

³¹ Partial measurement invariance.

³² Partial metric invariance.

ρουσιάστηκε από τους Byrne et al. (1989) και τους Marsh και Hocevar (1985) και αργότερα αναλύθηκε περαιτέρω από άλλους συγγραφείς (π.χ., Steenkamp & Baumgartner, 1998). Σύμφωνα με τους παραπάνω συγγραφείς, ακόμη και στην περίπτωση της μερικής ισοδυναμίας της μέτρησης για κάθε έναν από τους διαφορετικούς τύπους της που προαναφέρθηκαν, η έγκυρη σύγκριση των ομάδων στις παραμέτρους που υποστηρίζονται από κάθε τύπο ισοδυναμίας είναι εφικτή. Όμως, κατά τη σύγκριση των ομάδων, πρέπει να χρησιμοποιούνται μόνον εκείνα τα θέματα των οποίων οι ψυχομετρικές ιδιότητες κρίθηκαν ισοδύναμες ανά ομάδα ανάλογα με τον τύπο της ισοδυναμίας της μέτρησης που εξετάστηκε (δηλαδή, παραγοντικές φορτίσεις, σφάλματα των θεμάτων, κ.ά.). Έτσι, για παράδειγμα, στην περίπτωση της σύγκρισης των λανθανόντων μέσων όρων πρέπει να χρησιμοποιηθούν μόνο εκείνα τα θέματα των οποίων οι παραγοντικές φορτίσεις και οι σταθερές κρίθηκαν ισοδύναμες κατά την εξέταση της υπόθεσης της ισχυρής παραγοντικής ισοδυναμίας. Όμως, για να λάβουν μια τέτοια απόφαση για συγκεκριμένα θέματα μίας κλίμακας, οι ερευνητές θα πρέπει να λάβουν υπόψη τους σχετικά αποτελέσματα ενός αριθμού ερευνών που εξέτασαν το συγκεκριμένο τύπο ισοδυναμίας της μέτρησης και φάνηκε ότι στις περισσότερες έρευνες, για τα συγκεκριμένα θέματα, οι συγκεκριμένες ψυχομετρικές τους ιδιότητες δεν αποδείχθηκαν ισοδύναμες ανά ομάδα (π.χ., παραγοντική φόρτιση, σταθερές των θεμάτων).

Στη συνέχεια παρουσιάζεται η μέθοδος ελέγχου της ισοδυναμίας της μέτρησης με την εφαρμογή της ΕΠΑ πολλαπλών ομάδων.

ΔΕΙΚΤΕΣ ΕΛΕΓΧΟΥ ΤΗΣ ΙΣΟΔΥΝΑΜΙΑΣ ΤΗΣ ΜΕΤΡΗΣΗΣ

Όπως προαναφέρθηκε, η κατάλληλη μέθοδος για την εξέταση των παραγοντικών μοντέλων πολλαπλών ομάδων που διαφέρουν στον αριθμό περιορισμών ισότητας σε διάφορες παραμέτρους τους (π.χ., παραγοντικές φορτίσεις, σφάλματα, κ.τ.λ.) ανά ομάδα είναι η ΕΠΑ πολλαπλών ομάδων. Προτού ο ερευνητής/τρια προχωρήσει σε σύγκριση συγκεκριμένων παραγοντικών μοντέλων πολλαπλών ομάδων μεταξύ τους (π.χ., σύγκριση του μοντέλου συμμορφικής ισοδυναμίας με το μοντέλο της μετρικής ισοδυναμίας), είναι σημαντικό το προηγούμενο ιεραρχικά μοντέλο να επιδείξει καλή προσαρμογή στα δεδομένα. Αυτό είναι πολύ σημαντικό και αποτελεί προϋπόθεση ώστε να εξεταστούν και επόμενα μοντέλα, τα οποία περι-

λαμβάνουν επιπρόσθετους περιορισμούς ισότητας. Για παράδειγμα, δεν μπορούν να εξεταστούν περαιτέρω μοντέλα με επιπρόσθετους περιορισμούς ισότητας εάν το μοντέλο της συμμορφικής ισοδυναμίας δεν αντανακλά την ίδια παραγοντική μορφή ανά ομάδα μέσω του ίδιου αριθμού παραγόντων οι οποίοι, ταυτοχρόνως, προσδιορίζονται από τα ίδια ακριβώς θέματα (Vandenberg & Lance, 2000).

Σύμφωνα με τους Cheung και Rensvold (2002), ο πιο συχνά χρησιμοποιούμενος δείκτης προσαρμογής στη βιβλιογραφία της *μοντελοποίησης δομικών εξισώσεων*³³ είναι ο δείκτης χ^2 . Μια μη στατιστικώς σημαντική τιμή του ($p > .05$) σημαίνει ότι δεν απορρίπτεται η μηδενική υπόθεση, δηλαδή, ότι ο *υποτιθέμενος πίνακας συνδιακυμάνσεων*³⁴ είναι πανομοιότυπος με τον *παρατηρούμενο πίνακα συνδιακυμάνσεων*³⁵. Επειδή, όμως, ο δείκτης αυτός επηρεάζεται σε μεγάλο βαθμό από το μέγεθος του δείγματος, έχουν προταθεί και άλλοι δείκτες οι οποίοι πρέπει να συνοδεύουν το δείκτη του χ^2 στην αξιολόγηση της προσαρμογής του μοντέλου. Ανάμεσα σε άλλους δείκτες, αυτοί που χρησιμοποιούνται συχνά είναι ο Συγκριτικός Δείκτης Προσαρμογής (Comparative Fit Index, CFI. Bentler, 1990), ο Μη Σταθμισμένος Δείκτης Προσαρμογής (Non-normed Fit Index, NNFI. Bentler & Bonett, 1980), η Ρίζα του Μέσου Τετραγωνικού Σφάλματος Προσέγγισης (Root Mean Square Error of Approximation, RMSEA. Steiger, 1990) και ο Δείκτης Gamma Hat (Steiger, 1989). Η κατανομή που ακολουθεί ο RMSEA είναι γνωστή, γεγονός που επιτρέπει τον υπολογισμό *διαστήματος εμπιστοσύνης*³⁶ (ΔE) του RMSEA το οποίο συνήθως ορίζεται στο 90% ΔE του RMSEA (Browne & Cudeck, 1993).

Επιπροσθέτως, οι Vandenberg και Lance (2000) προτείνουν το Δείκτη της Σχετικής Μη Κεντρικότητας (Relative Noncentrality Index, RNI. McDonald & Marsh, 1990) και τη Ρίζα του Τυποποιημένου Μέσου Τετραγωνικού Υπολοίπου Σφάλματος (Standardized Root Mean Square Residual, SRMR. Bentler, 1995). Για τους δείκτες NNFI και RNI προτείνεται η τιμή .90 ως το χαμηλότερο όριο και η τιμή .95 ή και μεγαλύτερη ως επιθυμητά όρια των δεικτών, ώστε να αντανακλούν καλή προσαρμογή του μοντέλου στα δεδομένα. Επίσης, για το δείκτη RMSEA η τιμή .08 πρέπει

³³ Structural equation modeling.

³⁴ Hypothesized covariance matrix.

³⁵ Observed covariance matrix.

³⁶ Confidence interval (CI).

να θεωρείται το υψηλότερο αποδεκτό όριο (Browne & Cudeck, 1993) ενώ για το δείκτη SRMR η τιμή .08 αντανακλά καλή προσαρμογή του μοντέλου με την τιμή .10 να θεωρείται το μέγιστο όριο (Vandenberg & Lance, 2000). Επίσης, για τον CFI η χαμηλότερη αποδεκτή τιμή είναι η τιμή του .90 (Cheung & Rensvold, 2002), ενώ η τιμή του .95 αντανακλά μία καλή προσαρμογή του μοντέλου στα δεδομένα (Hu & Bentler, 1999).

Όσον αφορά την εξέταση των τύπων της ισοδυναμίας της μέτρησης, ο βαθμός ισοδυναμίας με βάση τα είδη των περιορισμών ισότητας που ισχύουν ανάμεσα στα διάφορα μοντέλα εξετάζεται με τον Έλεγχο της Διαφοράς χ^2 ($\Delta\chi^2$) ή, αλλιώς, Λόγο Πιθανοφάνειας³⁷. Αυτός ο έλεγχος εξετάζει τη διαφορά στην τιμή του χ^2 που αντιστοιχεί σε κάθε ένα από τα δύο μοντέλα πολλαπλών ομάδων (δηλαδή αυτό με τους λιγότερους και αυτό με τους περισσότερους περιορισμούς ισότητας. Βλ. Cheung & Rensvold, 2002). Με αφορμή την έλλειψη γενικώς αποδεκτών κριτηρίων στην ΕΠΑ πολλαπλών ομάδων για τον καθορισμό του μεγέθους της μεταβολής στις τιμές των δεικτών προσαρμογής των μοντέλων που διαφέρουν στους περιορισμούς ισότητας, οι Cheung και Rensvold (2002) πρότειναν κρίσιμες τιμές μεταβολής των δεικτών προσαρμογής τέτοιων μοντέλων. Συγκεκριμένα, οι δείκτες προσαρμογής οι οποίοι πρέπει να εξετάζονται ως προς το εύρος της μεταβολής των τιμών τους ανάμεσα στα μοντέλα με τους λιγότερους και τους περισσότερους περιορισμούς ισότητας είναι ο CFI, ο Δείκτης Gamma Hat, και ο Δείκτης της Μη Κεντρικότητας του McDonald (McDonald's Noncentrality Index). Αυτοί οι δείκτες δεν επηρεάζονται από τη συνθετότητα του μοντέλου και το μέγεθος του δείγματος. Έτσι, μια τιμή αλλαγής μικρότερη ή ίση του -.01 για τον CFI, του -.001 για το Gamma Hat, και -.02 για το Δείκτη Μη Κεντρικότητας του McDonald σημαίνει ότι η μηδενική υπόθεση της ισοδυναμίας δεν πρέπει να απορριφθεί ή, με άλλα λόγια, πρέπει να γίνει αποδεκτή.

Συνολικά, κατά την εξέταση της καταλληλότητας των περιορισμών ισότητας σε ένα συγκεκριμένο μοντέλο πολλαπλών ομάδων (π.χ., της ισχυρής παραγοντικής ισοδυναμίας) θα πρέπει (α) να εξεταστούν οι δείκτες προσαρμογής που αντιστοιχούν στο μοντέλο για να εξακριβωθεί αρχικά εάν το μοντέλο έχει καλή προσαρμογή στα δεδομένα και (β) να εξεταστεί εάν το μοντέλο είναι χειρότερο ή ισοδύναμο με το προηγούμενο στη σειρά

³⁷ Chi-Square Difference ($\Delta\chi^2$) ή Likelihood Ratio (LR).

μοντέλο με την εφαρμογή του ελέγχου $\Delta\chi^2$. Πιο σημαντικό είναι, όμως, αυτός ο έλεγχος της διαφοράς των μοντέλων να συμπληρωθεί με τον έλεγχο της διαφοράς των τιμών είτε του CFI είτε του Gamma Hat είτε του Δείκτη Μη Κεντρικότητας του McDonald (Cheung & Rensvold, 2002) για να διασφαλισθεί ότι τα μοντέλα δε διαφέρουν ουσιαστικά μεταξύ τους. Επιπλέον, στην περίπτωση που τα αποτελέσματα των δύο πρώτων ελέγχων είναι ικανοποιητικά (δηλαδή δείξουν ότι τα δύο μοντέλα δε διαφέρουν μεταξύ τους), πρέπει να ακολουθήσει έλεγχος της στατιστικής σημαντικότητας κάθε ενός περιορισμού ισότητας του μοντέλου μέσω της εφαρμογής της δοκιμασίας Πολλαπλασιαστικής Lagrange (Lagrange Multiplier test, LM test. Βλ. Bentler, 1995). Το συγκεκριμένο στατιστικό κριτήριο παρέχεται από το πρόγραμμα EQS (Bentler, 1995) και χρησιμοποιείται προκειμένου να ελεγχθεί αν η αφαίρεση κάποιων συγκεκριμένων περιορισμών ισότητας αντιστοιχεί σε στατιστικώς σημαντική βελτίωση του μοντέλου, δηλαδή, μείωση της τιμής του χ^2 του μοντέλου πολλαπλών ομάδων. Ένα τέτοιο αποτέλεσμα θα σήμαινε ότι ο συγκεκριμένος περιορισμός ισότητας δεν ταιριάζει με τα δεδομένα και, άρα, η αντίστοιχη παράμετρος, που αρχικά θεωρήθηκε ισοδύναμη ανά ομάδα, ουσιαστικά δεν είναι.

ΑΠΟΥΣΙΑ ΙΣΟΔΥΝΑΜΙΑΣ ΤΗΣ ΜΕΤΡΗΣΗΣ ΣΤΗ ΔΙΑΠΟΛΙΤΙΣΜΙΚΗ ΣΥΓΚΡΙΣΗ ΚΑΙ ΑΠΕΙΛΕΣ ΕΓΚΥΡΟΤΗΤΑΣ

Η έγκυρη διαπολιτισμική σύγκριση προϋποθέτει την υποστήριξη τριών τύπων ισοδυναμίας: (α) της *ισοδυναμίας της εννοιολογικής κατασκευής*³⁸, (β) της *ισοδυναμίας της μονάδος μέτρησης*³⁹, και (γ) της *ισοδυναμίας της κλίμακας μέτρησης*⁴⁰ (van de Vijver & Leung, 1997). Οι τύποι αυτοί ισοδυναμίας που προτείνονται από τους van de Vijver και Leung (1997) στο πλαίσιο της διαπολιτισμικής σύγκρισης αφορούν τα επίπεδα του παράγοντα και του θέματος αλλά όχι της μεθόδου. Δηλαδή, ως στόχο έχουν την εξέταση της ισοδυναμίας είτε σε επίπεδο νοήματος της εννοιολογικής κατασκευής που εξετάζεται είτε της ισοδυναμίας σε επίπεδο θέματος, δηλαδή των ψυχομετρικών ιδιοτήτων των απαντήσεων που δίνονται στο ερωτηματολόγιο.

³⁸ Construct equivalence.

³⁹ Measurement unit equivalence.

⁴⁰ Scalar equivalence.

Ο κίνδυνος απουσίας εγκυρότητας στο εννοιολογικό επίπεδο, δηλαδή η *μονομέρεια της εννοιολογικής κατασκευής*⁴¹, είναι παρών όταν δεν αξιολογείται αν ο αριθμός ή/και η σύσταση των παραγόντων διαφέρουν ανά πληθυσμό. Ο κίνδυνος απουσίας εγκυρότητας σε επίπεδο θέματος, δηλαδή η *μονομέρεια του θέματος*⁴², αντανακλά προβλήματα στο επίπεδο του θέματος – κάτι που αναφέρεται και ως *διαφορική λειτουργία του θέματος*⁴³. Δηλαδή, ο συγκεκριμένος κίνδυνος σχετίζεται με τη διαφοροποίηση του περιεχομένου του θέματος ανάλογα με τον πληθυσμό στον οποίο ανήκουν τα άτομα (Byrne & Watkins, 2003). Έτσι συνολικά, η απουσία διαφορετικών τύπων ισοδυναμίας θέτει σε κίνδυνο την εγκυρότητα της διαπολιτισμικής σύγκρισης καθώς μειώνει το βαθμό της ισοδυναμίας στην ερμηνεία των θεμάτων ανά πληθυσμό (van de Vijver & Leung, 1997).

Στο πλαίσιο της ΕΠΑ πολλών διαφορετικών ομάδων, διαφορετικά είδη περιορισμών ισότητας (π.χ., στις παραγοντικές φορτίσεις, στα σφάλματα των θεμάτων, κ.ά.) αντανακλούν διαφορετικούς ελέγχους για την εξάλειψη συγκεκριμένων απειλών εγκυρότητας της διαπολιτισμικής σύγκρισης που προαναφέρθηκαν. Για παράδειγμα, η συμμορφική ισοδυναμία υποστηρίζεται όταν το παραγοντικό μοντέλο πολλαπλών ομάδων που περιέχει μηδενικές φορτίσεις των θεμάτων (δηλαδή, απουσία φορτίσεων) στους παράγοντες που δε θα έπρεπε να προσδιορίζουν θεωρητικώς, ταιριάζει στα δεδομένα (Steenkamp & Baumgartner, 1998). Αυτής της μορφής η ισοδυναμία αντικατοπτρίζει την ύπαρξη ισοδυναμίας εννοιολογικής κατασκευής. Η μετρική ισοδυναμία αντικατοπτρίζει την έννοια των ισοδύναμων διαστημάτων της κλίμακας μέτρησης ανά πληθυσμό (Rock, Werts, & Flaugher, 1978). Έτσι, η υποστήριξη της μετρικής ισοδυναμίας υποστηρίζει ισοδυναμία της μονάδος μέτρησης. Η ισχυρή παραγοντική ισοδυναμία, η οποία περιλαμβάνει ισοδυναμία και στο επίπεδο των σταθερών των θεμάτων αντικατοπτρίζει ισοδυναμία της κλίμακας μέτρησης (van de Vijver & Leung, 1997).

Στην περίπτωση της απουσίας ισοδυναμίας της μέτρησης σε κάποιο επίπεδο εξέτασης της (π.χ., μετρική ισοδυναμία, ισχυρή παραγοντική ισοδυναμία, κ.ά.), οι van de Vijver και Leung (1997) έχουν προτείνει την περαιτέρω εξέταση των αιτιών αυτού του φαινομένου με την εφαρμογή της *ανά-*

⁴¹ Construct bias.

⁴² Item bias.

⁴³ Differential item functioning (DIF).

λυσης διακύμανσης δύο παραγόντων⁴⁴, χρησιμοποιώντας τη συγκεκριμένη τιμή του μη ισοδύναμου θέματος⁴⁵ ως εξαρτημένη μεταβλητή και τους παράγοντες Κουλτούρα και *Επίπεδα των Τιμών της Υποκλίμακας*⁴⁶ (δηλαδή του παράγοντα του ερωτηματολογίου) ως ανεξάρτητες μεταβλητές. Όσον αφορά τον παράγοντα Επίπεδα των Τιμών της Υποκλίμακας, αναφερόμαστε σε έναν αυθαίρετο διαχωρισμό κάθε ενός από τα δείγματά μας σε έναν αριθμό υποομάδων, ίδιο ανά δείγμα, ανάλογα με τον αριθμό των απαντήσεων στην υποκλίμακα και με την πρόνοια να υπάρχουν τουλάχιστον 50 άτομα σε κάθε μια υποομάδα ανά δείγμα. Στόχος είναι να εξεταστεί εάν υπάρχουν κύριες επιδράσεις ή αλληλεπίδραση των παραγόντων Κουλτούρας και Επιπέδων των Τιμών της Υποκλίμακας στις τιμές των ατόμων στη συγκεκριμένη εξαρτημένη μεταβλητή. Εκτός από αυτή την προσέγγιση, που είναι και η πιο ισχυρή για την εξέταση του φαινομένου, οι Byrne και Watkins (2003) προτείνουν υποστηρικτικά και την περιγραφική εξέταση του αριθμού των ατόμων του δείγματος που επιλέγουν κάθε μία τιμή της κλίμακας απαντήσεων, όπως και της μορφής της κατανομής των τιμών (δηλ. της λοξότητας και της κύρτωσης) για συγκεκριμένα θέματα που αποδείχθηκαν μη ισοδύναμα κατά τους ελέγχους της ισοδυναμίας της μέτρησης.

Στη συνέχεια παρουσιάζεται παράδειγμα ελέγχου ισοδυναμίας της μέτρησης.

ΠΑΡΑΔΕΙΓΜΑ ΕΛΕΓΧΟΥ ΤΗΣ ΙΣΟΔΥΝΑΜΙΑΣ ΤΗΣ ΜΕΤΡΗΣΗΣ

Το παράδειγμα ελέγχου ισοδυναμίας της μέτρησης προέρχεται από το χώρο της ψυχολογικής ευεξίας (Ryan & Deci, 2001) και αφορά την Κλίμακα της Υποκειμενικής Ζωτικότητας (Subjective Vitality Scale. Ryan & Frederick, 1997). Η έννοια της “υποκειμενικής ζωτικότητας” (Ryan & Frederick, 1997) ανήκει στο θεωρητικό πλαίσιο της *θεωρίας του αυτο-καθορισμού*⁴⁷ (Ryan & Deci, 2002). Η έννοια αυτή αντανakλά την ψυχολογική εμπειρία κατά την οποία το άτομο βιώνει ενθουσιασμό και αισθάνεται γεμάτο ενέργεια και όρεξη για ζωή (Ryan & Frederick, 1997). Σύμφωνα με

⁴⁴ Two-way analysis of variance (two-way ANOVA).

⁴⁵ Item score.

⁴⁶ Subscale score levels.

⁴⁷ Self-determination theory.

τη θεωρία του αυτο-καθορισμού, όπως αυτή εφαρμόζεται στο χώρο της άσκησης ως συμπεριφοράς υγείας (Edmunds, Ntoumanis, & Duda, 2007), η υποκειμενική ζωτικότητα μελετάται ως δείκτης αποτελεσματικότητας διάφορων διαπροσωπικών τρόπων καθοδήγησης στο πλαίσιο ενός προγράμματος άσκησης με στόχο την ενίσχυση της συστηματικότητας της συμμετοχής των ατόμων στην άσκηση. Η Κλίμακα της Υποκειμενικής Ζωτικότητας χρησιμοποιείται για να αξιολογήσει τα επίπεδα της *καταστασιακής υποκειμενικής ζωτικότητας*⁴⁸ που αναφέρουν άντρες και γυναίκες ασκούμενοι, δηλαδή, το βαθμό στον οποίο νιώθουν το αίσθημα της ενεργητικότητας και της ζωντάνιας έπειτα από μία συνεδρία άσκησης. Παραδείγματα θεμάτων της κλίμακας είναι “Αυτή τη στιγμή έχω πολλή ενέργεια μέσα μου” και “Αυτή τη στιγμή νιώθω σε εγρήγορση και πολύ ζωντανή/ος”. Οι απαντήσεις δίνονται σε μια 7-βάθμια κλίμακα τύπου Likert από το 1 (καθόλου αληθές) έως το 7 (πολύ αληθές). Σχετική έρευνα έχει δείξει ότι αφαίρεση του δεύτερου από τα επτά θέματα, δηλαδή του *αντεστραμμένου θέματος*⁴⁹, βελτίωσε τη *μονοδιάστατη φύση*⁵⁰ της κλίμακας (Bostic, Rubio, & Hood, 2000). Οι αναλύσεις διεξήχθησαν με το στατιστικό πρόγραμμα μοντελοποίησης δομικών εξισώσεων EQS για Windows 6.1.

Πριν εξεταστεί ο βαθμός της ισοδυναμίας της μέτρησης, απαραίτητη προϋπόθεση αποτελεί η καλή προσαρμογή του μονοδιάστατου παραγοντικού μοντέλου σε κάθε έναν πληθυσμό ξεχωριστά. Οι δείκτες προσαρμογής για την Κλίμακα της Υποκειμενικής Ζωτικότητας ήταν για τους άντρες, $\chi^2(9, N = 165) = 18.91, p < .05, CFI = .986, RMSEA = .082, RMSEA\ 90\% \Delta E = .027 - .133$, και για τις γυναίκες, $\chi^2(9, N = 223) = 30.52, p < .05, CFI = .972, RMSEA = .104, RMSEA\ 90\% \Delta E = .065 - .145$. Παρά τις ικανοποιητικές τιμές του δείκτη CFI για τους άντρες και τις γυναίκες, οι τιμές του δείκτη RMSEA βρέθηκαν υψηλότερες από ό,τι θα έπρεπε. Αναφορικά με τις τιμές αυτές, οι Curran, Bollen, Chen, Paxton, και Kirby (2003) έχουν δείξει ότι σε δείγματα μικρότερα των 200 ατόμων και για παραγοντικά μοντέλα τριών παραγόντων με 3-4 θέματα ανά παράγοντα, ο δείκτης RMSEA τείνει να είναι σχετικά υψηλότερος από όσο θα έπρεπε. Έτσι, οι Curran et al. (2003) συμπεραίνουν ότι τα συνήθη επιθυμητά επίπεδα του δείκτη RMSEA θεωρούνται τόσο χρήσιμα σε δείγματα μέτρου μεγέθους και προτιμούν την

⁴⁸ State subjective vitality.

⁴⁹ Reverse item.

⁵⁰ Uni-dimensionality.

αποδοχή του μοντέλου, δεδομένου ότι οι υπόλοιποι δείκτες δείχνουν καλή προσαρμογή στα δεδομένα. Έτσι, με βάση το εύρημα των Curran et al. (2003) και τους περιορισμένους αριθμούς δείγματος ανά φύλο εξηγούνται και τα παρόντα αποτελέσματα αναφορικά με το δείκτη RMSEA.

Οι τύποι της ισοδυναμίας της μέτρησης που ελέγχθηκαν στην έρευνα αυτή, ακολουθώντας την περιγραφή του Gregorich (2006), ήταν η διαστασιακή ισοδυναμία, η συμμορφική ισοδυναμία, η μετρική ισοδυναμία, η ισχυρή παραγοντική ισοδυναμία, και η αυστηρή παραγοντική ισοδυναμία. Οι αντίστοιχες αναλύσεις διεξήχθησαν για να εξεταστεί η ισοδυναμία της μέτρησης των τιμών της κλίμακας ανάμεσα σε Έλληνες άντρες και γυναίκες ασκούμενους. Στην περίπτωση του μονοδιάστατου παραγοντικού μοντέλου η ισοδυναμία του αριθμού των παραγόντων και η ισοδυναμία της μορφής των παραγόντων είναι ταυτόσημη (Gregorich, 2006). Έτσι ο πρώτος τύπος ισοδυναμίας που εξετάστηκε ήταν η συμμορφική ισοδυναμία.

Ειδικότερα, υπολογίστηκαν τέσσερα παραγοντικά μοντέλα πολλαπλών δειγμάτων μέσω της ΕΠΑ πολλαπλών δειγμάτων (Πίνακας 2). Το Μοντέλο 1 εξέτασε την υπόθεση της συμμορφικής ισοδυναμίας. Στο μοντέλο αυτό δεν υπήρχε κανένας περιορισμός ισότητας σε καμία παράμετρο των έξι θεμάτων της κλίμακας. Το μοντέλο ήταν συμβατό με τα δεδομένα σε πολύ ικανοποιητικό βαθμό (Πίνακας 2).

Το Μοντέλο 2 εξέτασε την πλήρη μετρική ισοδυναμία. Συγκριτικά με το Μοντέλο 1 προστέθηκαν περιορισμοί ισότητας σε όλες τις παραγοντικές φορτίσεις εκτός της φόρτισης του Θέματος 1 η οποία σταθεροποιήθηκε στην τιμή 1 για να προσδιοριστεί η κλίμακα του λανθάνοντα παράγοντα, δηλαδή ο δείκτης αναφοράς⁵¹ (Steenkamp & Baumgartner, 1998). Σύμφωνα με τους Rensvold και Cheung (2001) το θέμα το οποίο επιλέγεται για να προσδιορίσει την κλίμακα του λανθάνοντα παράγοντα πρέπει να είναι μετρικά ισοδύναμο ανά πληθυσμό διότι διαφορετικά τα αποτελέσματα της ανάλυσης είναι ανακριβή και αυτό το φαινόμενο διαιωνίζεται και στις περαιτέρω δοκιμασίες διάφορων τύπων της ισοδυναμίας της μέτρησης. Οι δείκτες προσαρμογής του Μοντέλου 2 ήταν πολύ ικανοποιητικοί (βλ. Πίνακα 2). Η σύγκριση του Μοντέλου 2 με το Μοντέλο 1 έδειξε ότι τα δύο μοντέλα δε διέφεραν στατιστικώς σημαντικά. Αυτό φάνηκε και από το $\Delta\chi^2$ το οποίο ήταν μη στατιστικώς σημαντικό, και από το ΔCFI το οποίο οδήγησε στο ίδιο συμπέρασμα αφού η τιμή του ήταν μικρότερη του .01 (Cheung & Rensvold, 2002).

⁵¹ Reference indicator.

Πίνακας 2. Δείκτες καλής προσαρμογής για τα μοντέλα ισοδυναμίας της μέτρησης ανά φύλο

Παραγοντικά μοντέλα πολλαπλών ομάδων της Κλίμακας Υποκειμενικής Ζωτικότητας	χ^2	BE	$\Delta\chi^2$	ΔBE	CFI	ΔCFI	RMSEA	90% ΔE του RMSEA
Μοντέλο 1: Συμμορφική ισοδυναμία	49.44	18	--	--	.979	--	.067	.045-.090
Μοντέλο 2: Πλήρης μετρική ισοδυναμία	50.85	23	1.41	5	.981	.002	.056	.035-.077
Μοντέλο 3: Πλήρης ισχυρή ισοδυναμία	74.64	29	23.79*	6	.981	.000	.064	.046-.081
Μοντέλο 3 ^α : Μερική ισχυρή ισοδυναμία	59.77	28	8.92	5	.981	.000	.054	.035-.073
Μοντέλο 4: Μερική αυστηρή ισοδυναμία	63.79	33	4.02	5	.982	.001	.049	.031-.067

Σημείωση. BE = βαθμοί ελευθερίας. CFI = Συγκριτικός Δείκτης Προσαρμογής. RMSEA = Ρίζα του Μέσου Τετραγωνικού Σφάλματος Προσέγγισης. Το Μοντέλο 2 συγκρίνεται με το Μοντέλο 1. Τα Μοντέλα 3 και 3α συγκρίνονται με το Μοντέλο 2. Το Μοντέλο 4 συγκρίνεται με το Μοντέλο 3α. *σημαντικά διαφορετικό σε επίπεδο $p < .05$.

Στη συνέχεια πραγματοποιήθηκε έλεγχος με την εφαρμογή της δοκιμασίας Πολλαπλασιαστικής Lagrange για να διαπιστωθεί εάν υπήρξε κάποιος περιορισμός ισότητας των παραγοντικών φορτίσεων στο Μοντέλο 2 ο οποίος δεν ήταν συμβατός με τα δεδομένα. Η δοκιμασία Πολλαπλασιαστικής Lagrange παρέχει πληροφορίες σχετικά με το εάν η συνολική προσαρμογή του μοντέλου θα βελτιωνόταν στατιστικώς σημαντικά εάν αφαιρούνταν ένας συγκεκριμένος περιορισμός ισότητας από το μοντέλο πολλαπλών δειγμάτων. Αυτές οι πληροφορίες παρέχονται για κάθε έναν περιορισμό ισότητας ξεχωριστά, ανεξαρτήτως των υπόλοιπων περιορισμών ισότητας. Για κάθε έναν περιορισμό ισότητας, όταν η σχετική τιμή σημαντικότητας είναι μικρότερη του .05 ($p < .05$) σημαίνει ότι η διατήρηση του περιορισμού ισότητας στο μοντέλο χειροτερεύει σημαντικά τη συνολική προσαρμογή του μοντέλου στα δεδομένα και άρα ο περιορισμός ισότητας δεν είναι συμβατός με τα δεδομένα. Το συμπέρασμα είναι ότι η συγκεκριμένη παράμετρος του μοντέλου (π.χ., παραγοντική φόρτιση), η οποία θεωρήθηκε πιθανώς

ισοδύναμη ανά δείγμα, τελικά δεν είναι. Στην προκειμένη λοιπόν περίπτωση των δεδομένων της Κλίμακας Υποκειμενικής Ζωτικότητας, οι περιορισμοί ισότητας και των πέντε θεμάτων ήταν συμβατοί με τα δεδομένα (δηλαδή, το επίπεδο σημαντικότητας για κάθε έναν περιορισμό ήταν $p > .05$). Δεν υπήρξε περιορισμός ισότητας για το Θέμα 1 διότι το θέμα λειτούργησε ως δείκτης του παράγοντα. Έτσι, οι δείκτες προσαρμογής του Μοντέλου 2, η απουσία διαφοράς των Μοντέλων 2 και 1, και η συμβατότητα των περιορισμών ισότητας των παραγοντικών φορτίσεων με τα δεδομένα υποστήριξαν την πλήρη μετρική ισοδυναμία του μοντέλου.

Στο Μοντέλο 3 για να ελεγχθεί η ισχυρή παραγοντική ισοδυναμία διατηρήθηκαν όλοι οι συμβατοί με τα δεδομένα περιορισμοί ισότητας του Μοντέλου 2 (δηλαδή, στις παραγοντικές φορτίσεις) και προστέθηκαν περιορισμοί ισότητας στις σταθερές των θεμάτων που επέδειξαν ισοδυναμία των παραγοντικών φορτίσεων στο Μοντέλο 2, δηλαδή, σε όλα τα πέντε θέματα. Οι δείκτες προσαρμογής του Μοντέλου 3 ήταν επίσης πολύ ικανοποιητικοί (βλ. Πίνακα 2). Η διαφορά, ωστόσο, ($\Delta\chi^2$) του Μοντέλου 3 από το Μοντέλο 2 αποδείχθηκε στατιστικώς σημαντική ενώ το ΔCFI οδήγησε στο συμπέρασμα ότι τα μοντέλα δε διέφεραν. Όπως συζητήθηκε και προηγουμένως, επειδή η τιμή χ^2 επηρεάζεται σε μεγάλο βαθμό από το μέγεθος του δείγματος, το συμπέρασμα για τη διαφορά των μοντέλων πρέπει να βασίζεται στο δείκτη CFI (Cheung & Rensvold, 2002). Έτσι οδηγούμαστε στο συμπέρασμα ότι τα Μοντέλα 3 και 2 δε διαφέρουν. Στη συνέχεια, η δοκιμασία Πολλαπλασιαστής Lagrange έδειξε ότι ο περιορισμός ισότητας στο Θέμα 2 που ορίστηκε ως σταθερό ήταν στατιστικά σημαντικός. Αυτό θεωρείται ένδειξη ότι ο περιορισμός δεν είναι συμβατός με τα δεδομένα αλλά για να οδηγηθούν οι ερευνητές σε τέτοιο συμπέρασμα πρέπει να κρίνουν και πόσο μεγάλο είναι το μέγεθος της βελτίωσης του μοντέλου (δηλαδή, μείωση της τιμής χ^2) σε περίπτωση που ο περιορισμός ισότητας αφαιρεθεί. Δηλαδή, το συμπέρασμα σχετικά με το εάν ένας περιορισμός ισότητας είναι συμβατός με τα δεδομένα ή όχι επαφίεται στην υποκειμενική κρίση των ερευνητών, οι οποίοι πρέπει να συνεκτιμήσουν το επίπεδο σημαντικότητας του περιορισμού, το μέγεθος της βελτίωσης της τιμής χ^2 του μοντέλου καθώς και τους δείκτες προσαρμογής του μοντέλου.

Έτσι, στη συγκεκριμένη περίπτωση το Μοντέλο 3 με βάση τους δείκτες προσαρμογής έχει σχεδόν άριστη προσαρμογή στα δεδομένα και ταυτόχρονα με βάση το ΔCFI δε διαφέρει από το Μοντέλο 2. Όμως, στα πλαίσια της δοκιμασίας Πολλαπλασιαστής Lagrange φαίνεται ότι ένας από τους πε-

ριορισμούς ισότητας είναι στατιστικώς σημαντικός και ταυτόχρονα μεγάλος σε μέγεθος, $\chi^2 = 14.41$, $p < .05$, συγκρινόμενος με τη συνολική τιμή $\chi^2 = 51.86$ του Μοντέλου 3. Δηλαδή, αφαίρεση του συγκεκριμένου περιορισμού ισότητας αναμένεται να βελτιώσει σε μεγάλο βαθμό την τιμή χ^2 του συνολικού μοντέλου, δηλαδή, να οδηγήσει σε μείωση της τιμής χ^2 του μοντέλου κατά το ένα τέταρτο της συνολικής τιμής του. Αυτό σημαίνει ότι ο συγκεκριμένος περιορισμός ισότητας είναι σε μεγάλο βαθμό ασύμβατος με τα δεδομένα και, άρα, είναι ξεκάθαρο ότι, τουλάχιστον στα δεδομένα μας, το συγκεκριμένο θέμα δεν πρέπει να χρησιμοποιηθεί κατά τη σύγκριση των μέσων όρων της καταστασιακής υποκειμενικής ζωτικότητας ανάμεσα σε άντρες και γυναίκες ασκούμενους. Συμπεραίνουμε, λοιπόν, με βάση τους πολύ ικανοποιητικούς δείκτες προσαρμογής του Μοντέλου 3 και την απουσία διαφοράς με το Μοντέλο 2, ότι οι περιορισμοί των σταθερών των θεμάτων είναι σε γενικές γραμμές συμβατοί με τα δεδομένα και, άρα, ισοδύναμοι ανά ομάδα. Ωστόσο, η μεγάλη τιμή χ^2 του περιορισμού ισότητας του Θέματος 2 πρέπει να μας ανησυχήσει και να μην επαναπαυθούμε στη χρήση του. Σίγουρα κάτω από τις ίδιες συνθήκες μια πολύ μικρότερη τιμή χ^2 του συγκεκριμένου περιορισμού, παρά τη στατιστική σημαντικότητά του, δε θα ήταν ανησυχητική όσον αφορά τη συμβατότητά του με τα δεδομένα.

Έτσι, λοιπόν, το Μοντέλο 3 επαναυπολογίστηκε, αυτή τη φορά αφαιρώντας τον περιορισμό ισότητας από τη σταθερά του Θέματος 2 (Μοντέλο 3α: Πίνακας 2). Το νέο Μοντέλο 3α δε διέφερε από το Μοντέλο 2. Τώρα λοιπόν είναι πιο σίγουρο ότι δε θα χρησιμοποιηθεί το Θέμα 2 για τη σύγκριση των μέσων όρων της καταστασιακής υποκειμενικής ζωτικότητας, θέμα το οποίο φαίνεται να μη είναι κατάλληλο για μια τέτοια σύγκριση.

Τέλος, στο Μοντέλο 4 ελέγχθηκε η μερική αυστηρή παραγοντική ισοδυναμία, όπου διατηρήθηκαν όλοι οι συμβατοί με τα δεδομένα περιορισμοί ισότητας των παραγοντικών φορτίσεων και των σταθερών των θεμάτων από το Μοντέλο 3α εκτός του Θέματος 2. Έτσι, προστέθηκαν περαιτέρω περιορισμοί ισότητας και στα σφάλματα θέματος όλων των πέντε θεμάτων εκτός του Θέματος 2. Το Μοντέλο 4 έδειξε επίσης πολύ ικανοποιητική προσαρμογή στα δεδομένα (βλ. Πίνακα 2). Εφαρμογή της δοκιμασίας Πολλαπλασιαστικής Lagrange έδειξε ότι όλοι οι περιορισμοί ισότητας ήταν συμβατοί με τα δεδομένα, δηλαδή μη στατιστικώς σημαντικοί.

Συνολικά και με βάση την τιμή κριτήριο του -0.01 (Cheung & Rensvold, 2002) για το δείκτη ΔCFI κανένα μοντέλο δε διαφέρει από κάποιο άλλο (βλ. Πίνακα 2). Στο σημείο αυτό πρέπει να σημειωθεί ότι, σε περίπτωση

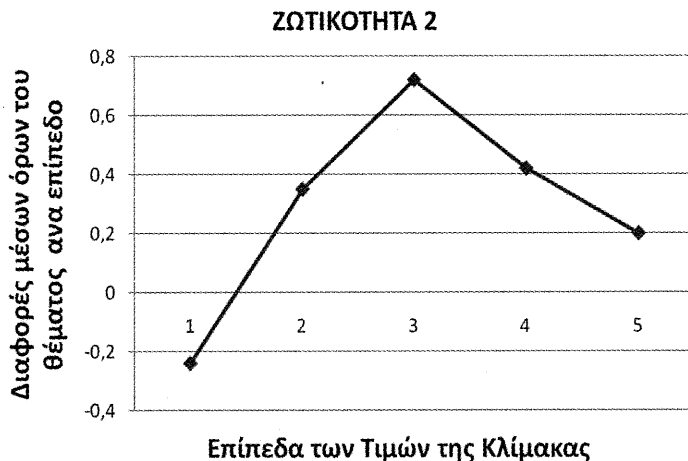
που τα δεδομένα ακολουθούν πολυμεταβλητή μη κανονική κατανομή, κατά τον έλεγχο $\Delta\chi^2$ των Satorra-Bentler πρέπει να χρησιμοποιηθεί και το *Satorra-Bentler διορθωμένο χ^2* (Satorra & Bentler, 2001)⁵².

Ανάλυση της απουσίας ισοδυναμίας ανά φύλο

Σύμφωνα με το παράδειγμα που παρουσιάζεται από τους Byrne και Watkins (2003) σχετικά με την αξιολόγηση της απουσίας ισοδυναμίας μέτρησης στο επίπεδο των θεμάτων, χρησιμοποιείται η ανάλυση διακύμανσης δύο παραγόντων, όπου οι τιμές του θέματος λειτουργούν ως η εξαρτημένη μεταβλητή ενώ το φύλο (ή η κουλτούρα, το έθνος, ή οποιαδήποτε άλλη ομαδοποίηση για σύγκριση) και τα επίπεδα των τιμών με βάση ολόκληρη την (υπο)κλίμακα του ερωτηματολογίου λειτουργούν ως οι ανεξάρτητες μεταβλητές. Αυτά τα επίπεδα των τιμών αναφέρονται στον αριθμό των ομάδων που συγκρίνονται μεταξύ τους και που δημιουργούνται από τους ερευνητές χρησιμοποιώντας ως μεταβλητή το άθροισμα των τιμών των θεμάτων της (υπο)κλίμακας (van de Vijver & Leung, 1997). Δηλαδή, σε μία μονοδιάστατη κλίμακα με έξι θέματα και 7-βάθμια κλίμακα απαντήσεων (από 1 έως 7), η ελάχιστη συνολική τιμή του αθροίσματος των θεμάτων θα είναι 6 και η μέγιστη συνολική τιμή θα είναι 42. Έτσι ο αριθμός των επιπέδων των τιμών της κλίμακας θα έπρεπε να είναι 36 ($42-6 = 36$). Επειδή, όμως, με τον τρόπο αυτό μειώνεται δραματικά ο αριθμός των ατόμων που εμπίπτει σε κάθε μία από τις 36 ομάδες ατόμων, οι ερευνητές θα πρέπει να αυξήσουν τεχνητά τον αριθμό των ατόμων σε κάθε μία από αυτές τις ομάδες μειώνοντας αντιστοίχως τον αριθμό των ομάδων που θα δημιουργήσουν. Ο αριθμός των ατόμων σε κάθε ομάδα θα πρέπει να είναι σχετικά ισοδύναμος ανά ομάδα και όχι μικρότερος του 50 (van de Vijver & Leung, 1997). Στην ανάλυση αυτή ενδιαφέρον έχουν τα αποτελέσματα της στατιστικής σημαντικότητας σχετικά με την επίδραση του φύλου (ή της κουλτούρας ή άλλης ομαδοποίησης), όπως και της αλληλεπίδρασης του φύλου (ή της κουλτούρας) με τα επίπεδα των τιμών, αλλά όχι η κύρια επίδραση των επιπέδων των τιμών αυτών καθαυτών (Byrne & Watkins, 2003).

Σύμφωνα με την πρακτική των Byrne και Watkins (2003), η ανάλυση για την απουσία ισοδυναμίας έγινε στο παρόν παράδειγμα στο Θέμα 2 που δεν

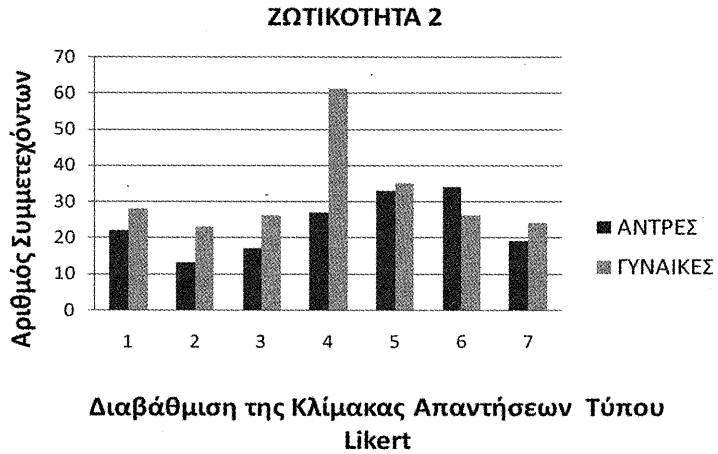
⁵² Satorra-Bentler scaled χ^2 .



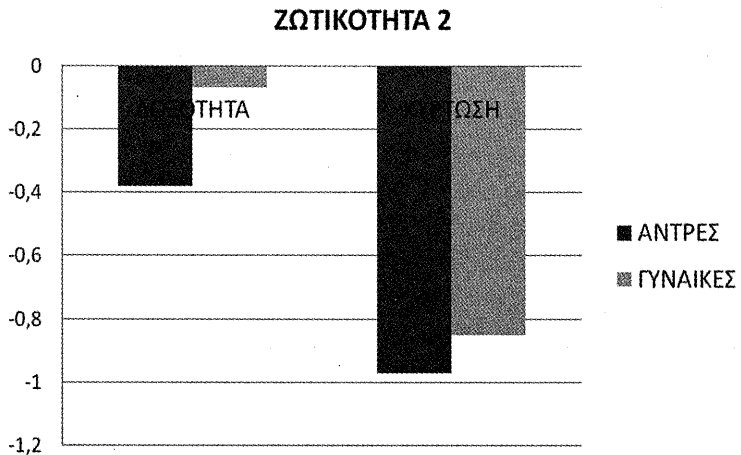
Γράφημα 1. Διαφορές μέσων όρων μεταξύ ανδρών-γυναικών στο μη ισοδύναμο Θέμα 2 ανά επίπεδο των τιμών της (υπο)κλίμακας.

παρουσίασε ισοδύναμη σταθερά (απουσία ισχυρής παραγοντικής ισοδυναμίας) στις αναλύσεις ισχυρής ισοδυναμίας ανά φύλο (Μοντέλο 3. Βλ. Πίνακα 2). Η μεταβλητή του φύλου είχε δύο επίπεδα ενώ το δείγμα χωρίστηκε σε πέντε επίπεδα τιμών, διότι ο χωρισμός των δειγμάτων σε περισσότερες από πέντε ομάδες κατέληγε στην ύπαρξη λιγότερων των 50 ατόμων ανά ομάδα (ασχέτως της 7-βάθμιας κλίμακας απαντήσεων). Περαιτέρω, για κάθε δείγμα (α) υπολογίσθηκαν οι μέσοι όροι του Θέματος 2 για κάθε ένα από τα πέντε επίπεδα των τιμών στα οποία διασπάσθηκε κάθε δείγμα αντρών και γυναικών ξεχωριστά, (β) ο μέσος όρος του Θέματος 2 της μιας ομάδος αφαιρέθηκε από τον αντίστοιχο μέσο όρο της άλλης ομάδος για κάθε ένα από τα πέντε επίπεδα των τιμών και (γ) αυτές οι διαφορές των μέσων όρων ανά επίπεδο τιμών (δηλαδή ανά ομάδα ατόμων που δημιουργήθηκε) παρουσιάσθηκαν σε σχετικό γράφημα (Γράφημα 1). Επίσης, υπολογίσθηκε η συχνότητα των ατόμων που επέλεξαν κάθε τιμή της 7-βάθμιας κλίμακας απαντήσεων τύπου Likert ανά πληθυσμιακή ομάδα (Γράφημα 2), όπως και οι τιμές της λοξότητας και της κύρτωσης για το μη ισοδύναμο θέμα ανά πληθυσμιακή ομάδα (Γράφημα 3).

Σχετικά με το Γράφημα 1, και σύμφωνα με τους Byrne και Watkins (2003), εάν η γραμμή του γραφήματος παραμένει σταθερά επάνω ή κάτω από τη γραμμή που αντιστοιχεί στο μηδέν, οι τιμές στο θέμα τείνουν να



Γράφημα 2. Σύγκριση των συχνοτήτων των ατόμων που επέλεξαν κάθε μία βαθμίδα της κλίμακας απαντήσεων του μη ισοδύναμου Θέματος 2 της Κλίμακας Υποκειμενικής Ζωτικότητας.



Γράφημα 3. Σύγκριση των τιμών της λοξότητας και της κύρτωσης για το μη ισοδύναμο Θέμα 2 της Κλίμακας Υποκειμενικής Ζωτικότητας ανάμεσα σε άντρες και γυναίκες Έλληνες ασκούμενους/ες.

είναι καθολικώς επηρεασμένες, δηλαδή σε *ομοιόμορφη μονομέρεια*⁵³ (Mellenbergh, 1982), και πιθανώς να ευθύνεται ο παράγοντας φύλο ανάλογα με τους πληθυσμούς που συγκρίνονται (στη συγκεκριμένη περίπτωση το φύλο). Όταν, όμως, η γραμμή δεν παραμένει συστηματικά και σταθερά επάνω ή κάτω της γραμμής του μηδενός και, ταυτοχρόνως, την τέμνει, οι απαντήσεις δεν είναι καθολικώς επηρεασμένες μόνον από το βασικό παράγοντα (π.χ., το φύλο) αλλά διαφαίνεται ταυτοχρόνως και η αλληλεπίδραση του φύλου με τα επίπεδα των τιμών της (υπο)κλίμακας, δηλαδή *μη ομοιόμορφη μονομέρεια*⁵⁴ (βλ. Byrne & Watkins, 2003). Σημασία έχει κατά πόσο η επίδραση του φύλου από μόνο του ή η αλληλεπίδραση του φύλου με τα επίπεδα των τιμών συνεισφέρουν περισσότερο στη μορφή της γραμμής του σχήματος. Όταν, όμως, η γραμμή τέμνει με έναν μη συστηματικό τρόπο τη γραμμή του μηδενός, τότε πιστοποιείται απουσία της επίδρασης είτε του φύλου είτε της αλληλεπίδρασης του φύλου με τα επίπεδα των τιμών (βλ. Byrne & Watkins, 2003).

Στο παρόν παράδειγμα και με βάση το Γράφημα 1, οι απαντήσεις στο Θέμα 2 επηρεάστηκαν σημαντικά μόνο από το φύλο. Αυτό συμφωνεί με την τιμή $F = 6.21, p < .05$, μερικό $\eta^2 = .01$ για την κύρια επίδραση του φύλου συγκριτικά με την τιμή $F = 1.92, p > .05$, για την αλληλεπίδραση του φύλου με τα επίπεδα της υποκλίμακας, η οποία δεν ήταν στατιστικώς σημαντική. Αξίζει, ωστόσο, να επισημανθεί ότι η διαφορά στην τιμή F δεν ήταν τόσο μεγάλη ανάμεσα στην κύρια επίδραση του φύλου και στην αλληλεπίδρασή του με τα επίπεδα της υποκλίμακας. Έτσι συμπεραίνουμε ότι η απουσία ισοδυναμίας του Θέματος 2 όσον αφορά τη σύγκριση ανάμεσα σε άντρες και γυναίκες μπορεί να μην είναι και τόσο ισχυρή, όπως υποστηρίζεται και από τα αποτελέσματα της ανάλυσης διακύμανσης δύο παραγόντων.

ΣΥΜΠΕΡΑΣΜΑ

Γενικά, ο έλεγχος των βαθμών της ισοδυναμίας της μέτρησης σε διάφορες παραμέτρους που απασχολούν τους ερευνητές στη συγκριτική ψυχολογική έρευνα είναι σημαντικός για τη διασφάλιση έγκυρων διομαδικών συγκρίσεων. Τέτοιοι έλεγχοι μπορούν να διασφαλίσουν σε μεγαλύτερο

⁵³ Uniform bias.

⁵⁴ Non-uniform bias.

βαθμό την εγκυρότητα των συγκρίσεων μεταξύ ομάδων απαλείφοντας την πιθανότητα το όργανο μέτρησης να λειτουργεί διαφορετικά κατά τη χρήση του από διαφορετικούς πληθυσμούς. Έτσι, ο ερευνητής/τρια μπορεί να προχωρήσει σε διομαδικές συγκρίσεις έχοντας διασφαλίσει ότι οι πιθανές διαφορές που θα προκύψουν από τους ελέγχους των διομαδικών συγκρίσεων δε θα οφείλονται στη διαφοροποιημένη συμπεριφορά του οργάνου μέτρησης κατά τη χρήση του από τους διαφορετικούς πληθυσμούς.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Baumgartner, H., & Steenkamp, J. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38, 143-156.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bostic, T. J., Rubio, D. M., & Hood, M. (2000). A validation of the Subjective Vitality Scale using structural equation modeling. *Social Indicators Research*, 52, 313-324.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Buss, A. R., & Royce, J. R. (1975). Detecting cross-cultural commonalities and differences: Intergroup factor analysis. *Psychological Bulletin*, 82, 128-136.
- Byrne, B. M. (1994). Testing for the factorial validity, replication, and invariance of a measurement instrument: A paradigmatic application based on the Maslach Burnout Inventory. *Multivariate Behavioral Research*, 29, 289-311.
- Byrne, B. M., Shavelson, R. J., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456-466.
- Byrne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, 34, 155-175.
- Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equation modeling. *Journal of Cross-Cultural Psychology*, 31, 187-212.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255.
- Curran, P. J., Bollen, K. A., Chen, F., Paxton, P., & Kirby, J. B. (2003). Finite sampling properties of the point estimates and confidence intervals of the RMSEA. *Sociological Methods and Research*, 32, 208-252.

- Dragow, F., & Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology, 70*, 662-680.
- Edmunds, J., Ntoumanis, N., & Duda, J. L. (2007). Adherence and well-being in overweight and obese patients referred to an exercise on prescription scheme: A self-determination theory perspective. *Psychology of Sport and Exercise, 8*, 722-740.
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care, 44*(Suppl. 3), 78-94.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*, 117-144.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.
- Irvine, S. H. (1969). Contributions of ability and attainment testing in Africa to a general theory of intellect. *Journal of Biosocial Science, 1*, 91-102.
- Janssens, M., Brett, J. M., & Smith, F. J. (1995). Confirmatory cross-cultural research: Testing the viability of a corporation-wide safety policy. *Academy of Management Journal, 38*, 364-382.
- Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*, 53-76.
- Little, T. D. (2000). On the comparability of constructs in cross-cultural research: A critique of Cheung & Rensvold. *Journal of Cross-Cultural Psychology, 31*, 213-219.
- Malpass, R. S. (1977). Theory and method in cross-cultural psychology. *American Psychologist, 32*, 1069-1079.
- Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher-order factor models and their invariance across groups. *Psychological Bulletin, 97*, 562-582.
- McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness-of-fit. *Psychological Bulletin, 107*, 247-255.
- Mellenbergh, G. J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics, 7*, 105-118.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543.
- Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care, 44*(Suppl. 3), 69-77.
- Millsap, R. E., & Everson, H. (1991). Confirmatory measurement model comparisons using latent means. *Multivariate Behavioral Research, 26*, 479-497.
- Millsap, R. E., & Hartog, S. B. (1988). Alpha, beta, and gamma changes in evaluation research: A structural equation approach. *Journal of Applied Psychology, 73*, 574-584.
- Mullen, M. (1995). Diagnosing measurement equivalence in cross-national research. *Journal of International Business Studies, 3*, 573-596.
- Pentz, M. A., & Chou, C. (1994). Measurement invariance in longitudinal clinical research assuming change from development and intervention. *Journal of Consulting and Clinical Psychology, 62*, 450-462.
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*, 552-566.

- Rensvold, R. B., & Cheung, G. W. (2001). Testing for metric invariance using structural equation models: Solving the standardization problem. In C. A. Schriesheim & L. L. Neider (Eds.), *Research in management: Vol. 1. Equivalence in measurement* (pp. 51-98). Greenwich, CT: Information Age.
- Riordan, C. M., & Vandenberg, R. J. (1994). A central question in cross-cultural research: Do employees of different cultures interpret work-related measures in an equivalent manner? *Journal of Management*, 20, 643-671.
- Rock, D. A., Werts, C. E., & Flaugher, R. L. (1978). The use of analysis of covariance structures for comparing the psychometric properties of multiple variables across populations. *Multivariate Behavioral Research*, 13, 403-418.
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52, 141-166.
- Ryan, R. M., & Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3-33). Rochester, NY: University of Rochester Press.
- Ryan, R. M., & Frederick, C. (1997). On energy, personality, and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality*, 65, 529-565.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.
- Singh, J. (1995). Measurement issues in cross-national research. *Journal of International Business Studies*, 26, 597-619.
- Steenkamp, J., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78-90.
- Steiger, J. H. (1989). *EzPATH: Causal modeling*. Evanston, IL: SYSTAT.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173-180.
- Suzuki, S., & Rancer, A. S. (1994). Argumentativeness and verbal aggressiveness: Testing for conceptual and measurement equivalence across cultures. *Communication Monographs*, 6, 256-279.
- Tayeb, M. (1994). Organizations and national culture: Methodology considered. *Organization Studies*, 15, 429-446.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-69.
- van de Vijver, F., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.
- Yoo, B. (2002). Cross-cultural comparisons: A cautionary note. *Psychology and Marketing*, 19, 357-368.

MEASUREMENT INVARIANCE IN COMPARATIVE PSYCHOLOGICAL RESEARCH

Symeon P. Vlachopoulos

Aristotle University of Thessaloniki, Greece

Abstract: Important questions that need to be addressed in comparative psychological research include whether participants from different groups interpret the same scale items in a conceptually similar manner and whether ethnic, gender, or other individual differences preclude responding to instruments in similar ways. The present article addresses the topic of measurement invariance within a confirmatory factor analysis framework and its application in psychological research. Specifically, (a) the reason is discussed why evaluation of measurement invariance across populations is critical in psychological research, (b) the process of measurement invariance evaluation is described, (c) theoretical interpretations are provided and practical applications of various degrees of measurement invariance are discussed, and (d) an example of measurement invariance evaluation in psychological research is provided.

Key words: Cross-cultural validity, Cross-group comparison, Multi-group confirmatory factor analysis.

Address: Symeon P. Vlachopoulos, Laboratory of Social Research on Physical Activity, Department of Physical Education and Sport Science at Serres, Aristotle University of Thessaloniki, Aghios Ioannis, 621 10 Serres, Greece. Phone: +30-2310-991045. Fax: +30-23210-64806. E-mail: vlachop@phed-sr.auth.gr