

ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΚΑΙ ΕΦΑΡΜΟΓΗ ΤΗΣ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ ΔΟΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ ΣΤΙΣ ΚΟΙΝΩΝΙΚΕΣ ΕΠΙΣΤΗΜΕΣ

Νικόλαος Τσιγγίλης

Πανεπιστήμιο Θεσσαλίας, Τρίκαλα

Περίληψη: Τις τελευταίες δύο δεκαετίες η μοντελοποίηση δομικών εξισώσεων τείνει να είναι η κυρίαρχη μέθοδος με την οποία εξετάζονται ερευνητικές υποθέσεις σε διάφορες επιστήμες, συμπεριλαμβανομένων και των κοινωνικών επιστημών. Η μεγάλη διάδοση και χρήση της στηρίζεται στο ότι επιτρέπει την εξέταση σύνθετων θεωρητικών μοντέλων στα οποία εκ των προτέρων έχει καθοριστεί ο τρόπος με τον οποίο μεταβλητές (μετρήσιμες ή λανθάνουσες) συνδέονται μεταξύ τους. Σκοπός της παρούσας εργασίας είναι να εισαγάγει τον ερευνητή των κοινωνικών επιστημών στις βασικές έννοιες και στις διαδικασίες διεξαγωγής της μοντελοποίησης δομικών εξισώσεων. Τέλος, δίνεται ένα παράδειγμα από το χώρο της εκπαιδευτικής ψυχολογίας.

Λέξεις κλειδιά: Ανάλυση λανθανουσών μεταβλητών, Επιβεβαιωτική παραγοντική ανάλυση.

Διεύθυνση: Νικόλαος Τσιγγίλης, Τμήμα Επιστήμης Φυσικής Αγωγής και Αθλητισμού, Πανεπιστήμιο Θεσσαλίας, Καριές, 421 00 Τρίκαλα. Τηλ.: 24310-47062. Fax.: 24310-47062. E-mail: tsigilis@uom.gr

ΕΙΣΑΓΩΓΗ

Τα τελευταία χρόνια όλο και περισσότεροι ερευνητές στο χώρο των κοινωνικών επιστημών χρησιμοποιούν τη *μοντελοποίηση δομικών εξισώσεων*¹. Πρόσφατη αναζήτηση στη βάση δεδομένων PsycINFO έδειξε ότι η μοντελοποίηση δομικών εξισώσεων τείνει να έχει την ίδια συχνότητα χρήσης με την ανάλυση διακύμανσης (Nachtigall, Kroehne, Funke, & Steyer, 2003). Η μεγάλη διάδοσή της έγκειται στο γεγονός ότι πρόκειται για μια ευέλικτη ανάλυση με αρκετές δυνατότητες η οποία επιτρέπει την εξέταση σύνθετων θεωριών. Αυτό επιτυγχάνεται με τη μοντελοποίηση σύνθετων αλληλοεπιδράσεων μεταξύ μετρήσιμων και λανθανουσών μεταβλητών, επιτυγχάνοντας έτσι μια καλύτερη προσέγγιση της ανθρώπινης συμπεριφοράς (Lei & Wu, 2007). Η μοντελοποίηση δομικών εξισώσεων είναι μία στατιστική τεχνική, η οποία ενσωματώνει και συνδυάζει “κάτω από την ίδια ομπρέλα” διαφορετικά είδη αναλύσεων, όπως πολλαπλή γραμμική παλινδρόμηση, παραγοντική ανάλυση, ανάλυση διακύμανσης, κ.ά. (Nachtigall et al., 2003). Σκοπός της παρούσας εργασίας είναι να εισαγάγει τον ερευνητή των κοινωνικών επιστημών στις έννοιες της μοντελοποίησης δομικών μοντέλων, στα πλεονεκτήματά της, καθώς και στα απαραίτητα βήματα που πρέπει να ακολουθηθούν προκειμένου να διεξαχθεί η ανάλυση αυτή. Για την καλύτερη εμπέδωση παρέχεται συγκεκριμένο παράδειγμα εφαρμογής της μοντελοποίησης δομικών εξισώσεων από το χώρο της εκπαιδευτικής ψυχολογίας.

Πλεονεκτήματα της μοντελοποίησης δομικών εξισώσεων

Υπάρχουν αρκετοί λόγοι για τους οποίους ο ερευνητής στην περιοχή των κοινωνικών επιστημών θα επέλεγε τη μοντελοποίηση δομικών εξισώσεων. Με την εφαρμογή της ανάλυσης αυτής μπορεί να μελετηθεί η ορθότητα διάφορων θεωριών. Αυτό συμβαίνει επειδή οι αιτιακές σχέσεις μεταξύ των μεταβλητών του μοντέλου καθορίζονται εκ των προτέρων σύμφωνα με την

¹ Structural equation modeling. Άλλες ονομασίες με τις οποίες αναφέρεται η μοντελοποίηση δομικών εξισώσεων στη διεθνή βιβλιογραφία είναι “latent variable analysis” (ανάλυση λανθανουσών μεταβλητών), “covariance structure analysis” (ανάλυση της δομής της συνδιακύμανσης), ή και ακόμα LISREL (από το όνομα ενός δημοφιλούς στατιστικού προγράμματος).

υπάρχουσα θεωρία. Στη συνέχεια, αφού οι σχέσεις εκφραστούν, με τη βοήθεια συνήθως γραμμικών εξισώσεων, πραγματοποιείται ταυτόχρονη ανάλυση του πλέγματος των αλληλεξαρτήσεων. Έτσι, υπάρχει μια επιβεβαιωτική παρά διερευνητική προσέγγιση στην εξέταση της θεωρίας, η οποία ανταποκρίνεται στη σύγχρονη ανάγκη για συστηματική και ολιστική αντιμετώπιση των ερευνητικών προβληματισμών.

Μια σημαντική διαφορά από τις υπάρχουσες πολυμεταβλητές αναλύσεις αφορά το σφάλμα της μέτρησης, που στη μοντελοποίηση δομικών εξισώσεων λαμβάνεται υπόψη. Το γεγονός αυτό επιτρέπει την ανάλυση των αλληλεπιδράσεων και αλληλεξαρτήσεων απαλλαγμένων από σφάλματα μέτρησης (Nachtigall et al., 2003). Επιπλέον, στο υπό εξέταση μοντέλο μπορεί να εισαχθούν τόσο μετρήσιμες όσο και λανθάνουσες μεταβλητές, δυνατότητα που δεν υπάρχει στις πολυμεταβλητές αναλύσεις. Συγχρόνως, η μοντελοποίηση δομικών εξισώσεων επιτρέπει την εξέταση διάφορων ειδών μοντέλων, όπως *μοντέλα διαδρομών*², *μοντέλα λανθάνουσών μεταβλητών*³, *πολυεπίπεδα μοντέλα*⁴, *διαχρονικά μοντέλα*⁵, και *μοντέλα ανάπτυξης*⁶. Επιπλέον, με τη χρήση της μοντελοποίησης δομικών εξισώσεων μπορεί να εξεταστεί η *ισοδυναμία του μοντέλου*⁷ σε διαφορετικές πληθυσμιακές ομάδες καθώς και ο *διαμεσολαβητικός*⁸ ή ο *ρυθμιστικός*⁹ ρόλος των μεταβλητών.

Τέλος, στη διάδοση της μοντελοποίησης δομικών εξισώσεων συντέλεσε και η ανάπτυξη φιλικών προς το χρήστη στατιστικών προγραμμάτων. Ανάμεσα στα πιο διαδεδομένα είναι το AMOS (Arbuckle, 2006), το EQS (Bentler, 2005), το LISREL (Jöreskog & Sörbom, 1996), και το MPLUS (Muthén & Muthén, 1998-2006). Αρκετά από τα παραπάνω προγράμματα διαθέτουν φοιτητική έκδοση (π.χ., AMOS, LISREL, MPLUS), η οποία προσφέρεται δωρεάν στο διαδίκτυο μαζί με το εγχειρίδιό τους. Σε άρθρο του ο Kline (1998a) παρουσιάζει και συγκρίνει τα πρώτα τρία από αυτά τα προγράμματα.

² Path models.

³ Latent variables models.

⁴ Multilevel models.

⁵ Longitudinal models.

⁶ Growth models.

⁷ Model equivalence.

⁸ Mediating role.

⁹ Moderating role.

Βασικές έννοιες και συμβολισμοί στη μοντελοποίηση δομικών εξισώσεων

Προτού προχωρήσουμε στον τρόπο διεξαγωγής της μοντελοποίησης δομικών εξισώσεων είναι απαραίτητο να παρουσιαστούν και να διευκρινιστούν ορισμένες βασικές έννοιες και συμβολισμοί που χρησιμοποιούνται στη μεθοδολογία αυτή.

Μετρήσιμες (ή παρατηρούμενες) μεταβλητές. Μετρήσιμες (ή παρατηρούμενες) μεταβλητές¹⁰ είναι αυτές που μπορούν να μετρηθούν άμεσα. Παραδείγματα άμεσα μετρήσιμων μεταβλητών είναι το ύψος και το βάρος ενός ατόμου, ο χρόνος αντίδρασης, τα διάφορα ερωτήματα στα ερωτηματολόγια αυτο-αναφορών. Στη διεθνή βιβλιογραφία αναφέρονται και ως *ενδεικτικές μεταβλητές*¹¹, *εμφανείς μεταβλητές*¹², ή *μεταβλητές αναφοράς*¹³. Οι μετρήσιμες μεταβλητές συνήθως χρησιμοποιούνται για να προσδιορίσουν μία λανθάνουσα μεταβλητή. Συνιστάται κάθε λανθάνουσα μεταβλητή να καθορίζεται από τέσσερις ή περισσότερες μετρήσιμες μεταβλητές· τρεις φαίνεται να είναι αποδεκτό, ενώ η χρήση δύο ή μιας μεταβλητής οδηγεί σε προβλήματα εκτίμησης των παραμέτρων του μοντέλου. Έχει καθιερωθεί οι μετρήσιμες μεταβλητές να παριστάνονται με τετράγωνο ή παραλληλόγραμμο. Έτσι στο Σχήμα 1 οι μεταβλητές από V1 έως V9 συμβολίζουν μετρήσιμες μεταβλητές.

Λανθάνουσες μεταβλητές. Οι λανθάνουσες μεταβλητές¹⁴ παριστάνουν έννοιες ή δομές οι οποίες δεν μπορούν να μετρηθούν ή να παρατηρηθούν άμεσα αλλά μόνο έμμεσα μέσω των εκφάνσεών τους. Παραδείγματα λανθάνουσών μεταβλητών είναι η φυσική κατάσταση (η οποία μετριέται μέσω της δύναμης, ταχύτητας, αντοχής και ευκαμψίας), η επαγγελματική ικανοποίηση (η οποία μετριέται μέσω της ικανοποίησης του εργαζομένου από τη σχέση του με τον άμεσο προϊστάμενο, τις αποδοχές του, τις δυνατότητες εξέλιξής του, τις εργασιακές συνθήκες κ.ά.), τα κίνητρα, η επαγγελματική εξουθένωση, η νοημοσύνη, το άγχος, η κατάθλιψη, η ποιότητα ζωής κ.ά. Άλλες ονομασίες με τις οποίες μπορεί να αναφερθούν οι λανθάνουσες μεταβλητές είναι *μη παρατηρούμενες μεταβλητές*¹⁵, *νοητικές κα-*

¹⁰ Measured (or observed) variables.

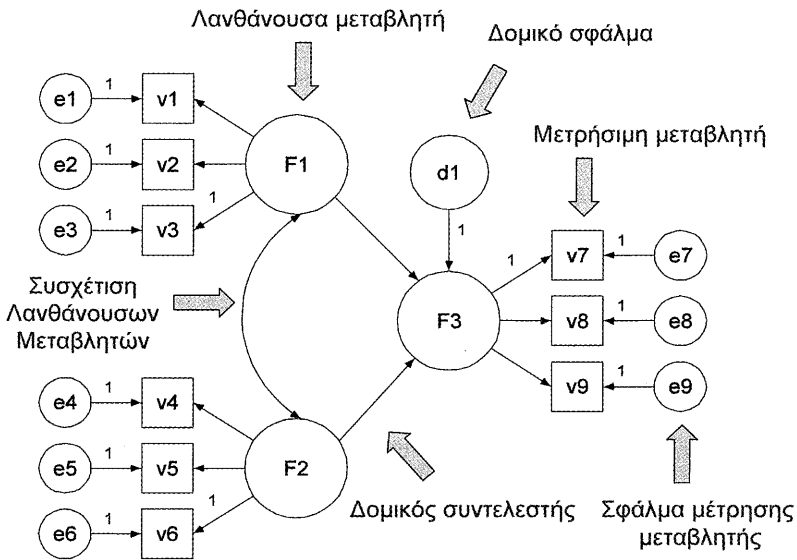
¹¹ Indicators.

¹² Manifest variables.

¹³ Reference variables.

¹⁴ Latent variables.

¹⁵ Unobserved variables.



Σχήμα 1. Συμβολισμοί στη μοντελοποίηση δομικών εξισώσεων.

τασκευές¹⁶, ή παράγοντες¹⁷. Οι λανθάνουσες μεταβλητές παριστάνονται με κύκλο ή έλλειψη. Έτσι, στο Σχήμα 1 οι παράγοντες F1, F2 και F3 συμβολίζουν λανθάνουσες μεταβλητές. Επιπλέον, τα σφάλματα μέτρησης¹⁸ ή τα υπόλοιπα σφάλματος¹⁹ τόσο των μετρήσιμων μεταβλητών (e1 έως e9) όσο και των λανθανουσών μεταβλητών (d1) παριστάνονται επίσης με κύκλο ή έλλειψη επειδή δε μετρούνται άμεσα.

Μία άλλη διάκριση των μεταβλητών στη μοντελοποίηση δομικών εξισώσεων αφορά τη αιτιακή σχέση μεταξύ των μεταβλητών. Ποια ή ποιες δηλαδή μεταβλητές προκαλούν, προβλέπουν ή ερμηνεύουν κάποια ή κάποιες άλλες. Έτσι, σ' ένα μοντέλο μπορεί να υπάρχουν εξωγενείς και ενδογενείς μεταβλητές.

Εξωγενείς μεταβλητές. Οι εξωγενείς μεταβλητές²⁰ συνήθως βρίσκονται

¹⁶ Constructs.

¹⁷ Factors.

¹⁸ Measurement errors.

¹⁹ Residuals.

²⁰ Exogenous variables.

στην αρχή του μοντέλου και δε δέχονται την αιτιακή επίδραση καμιάς άλλης μεταβλητής, αλλά αντίθετα επιδρούν και διαμορφώνουν άλλες μεταβλητές. Οι μεταβλητές αυτές θεωρείται ότι έχουν μετρηθεί με ακρίβεια και επομένως δεν έχουν σφάλμα μέτρησης. Μπορούν όμως να αλληλοσυνδέονται με άλλες εξωγενείς μεταβλητές. Η σύνδεση αυτή παριστάνεται με καμπύλη γραμμή, η οποία φέρει βέλη και στις δύο άκρες της (Σχήμα 1, η σύνδεση μεταξύ των F1 και F2) και δηλώνει ότι οι δύο αυτές μεταβλητές συνδέονται αλλά χωρίς να υπάρχει σχέση αιτίας-αιτιατού. Οι εξωγενείς μεταβλητές μπορεί να είναι είτε λανθάνουσες είτε μετρήσιμες. Όταν πρόκειται για παλινδρόμηση τότε αναφέρονται ως ανεξάρτητες μεταβλητές. Στο Σχήμα 1 οι μεταβλητές F1 και F2 θεωρούνται εξωγενείς μεταβλητές αφού προβλέπουν την F3 και επομένως δε φέρουν σφάλμα μέτρησης. Οι ευθείες γραμμές από τις F1 και F2 με τα βέλη να δείχνουν προς την F3 δηλώνουν ότι υπάρχει μία σχέση αιτίας-αιτιατού και πιο συγκεκριμένα ότι η F3 δέχεται την επίδραση των F1 και F2.

Ενδογενείς μεταβλητές. Σε αντίθεση με τις εξωγενείς οι ενδογενείς μεταβλητές²¹ δέχονται την αιτιακή επίδραση μεταβλητών είτε αυτές είναι ενδογενείς είτε διαμεσολαβητικές μεταβλητές²². Με τη σειρά τους οι ενδογενείς μεταβλητές μπορεί να επιδρούν σε άλλες εξαρτημένες ή διαμεσολαβητικές. Οι ενδογενείς μεταβλητές μπορεί να είναι μετρήσιμες ή λανθάνουσες και φέρουν σφάλμα μέτρησης, γεγονός που σημαίνει ότι δεν μπορεί να προβλεφθούν τέλεια. Όταν η ενδογενής μεταβλητή είναι λανθάνουσα, τότε το σφάλμα της αναφέρεται ως *σφάλμα λανθάνουσας μεταβλητής* ή *δομικό σφάλμα*²³. Η μεταβλητή F3 του Σχήματος 1 θεωρείται ενδογενής μεταβλητή αφού προβλέπεται από τις F1 και F2 και συνεπώς φέρει σφάλμα μέτρησης (d1).

Μοντέλα δομικών εξισώσεων

Στην μοντελοποίηση δομικών εξισώσεων εξετάζονται δύο είδη μοντέλων, το *μοντέλο μέτρησης*²⁴ και το *δομικό μοντέλο*²⁵.

²¹ Endogenous variables.

²² Mediating variables.

²³ Structural error or Disturbance.

²⁴ Measurement model.

²⁵ Structural model.

Μοντέλο μέτρησης. Το μοντέλο αυτό αναφέρεται στη σχέση των λανθανουσών μεταβλητών με τις αντίστοιχες μετρήσιμες μεταβλητές. Με πιο απλά λόγια σ' ένα μοντέλο μέτρησης εξετάζεται ο βαθμός στον οποίο οι μετρήσιμες μεταβλητές πράγματι είναι κατάλληλες για να αντιπροσωπεύσουν μια συγκεκριμένη λανθάνουσα μεταβλητή. Στην περίπτωση της επιβεβαιωτικής παραγοντικής ανάλυσης²⁶ η εξέταση της σχέσης αυτής αποτελεί τον κύριο στόχο της ανάλυσης. Το μοντέλο μέτρησης αποτελεί ένα τμήμα της μοντελοποίησης δομικών εξισώσεων και πρέπει να προηγείται της εξέτασης του δομικού μοντέλου, δηλαδή του μοντέλου που προβλέπει σχέσεις μεταξύ των μεταβλητών. Δεν έχει νόημα κάποιος να προχωρήσει στην εξέταση του δομικού μοντέλου εάν πρώτα δεν εξασφαλίσει την εγκυρότητα του μοντέλου μέτρησης (Garson, 2007). Στο Σχήμα 1 οι σχέσεις του κάθε λανθάνοντα παράγοντα (π.χ., του F1) με τις μετρήσιμες μεταβλητές, οι οποίες χρησιμοποιούνται για τον προσδιορισμόν (π.χ., οι V1, V2, V3), συνιστούν το μοντέλο μέτρησης.

Δομικό μοντέλο. Σε αντίθεση με το μοντέλο μέτρησης το δομικό μοντέλο εστιάζεται στη σχέση μεταξύ των λανθανουσών μεταβλητών (λανθανόντων παραγόντων). Στο δομικό μοντέλο μπορεί να υπάρχουν και μετρήσιμες μεταβλητές, οι οποίες όμως δε χρησιμοποιούνται για τον καθορισμό κάποιας λανθάνουσας μεταβλητής. Στο μοντέλο του Σχήματος 1 οι σχέσεις μεταξύ των τριών λανθανόντων παραγόντων (η συσχέτιση μεταξύ των F1 και F2 καθώς και οι διαδρομές από F1 και F2 προς F3) συνιστούν το δομικό μοντέλο.

Έχει προταθεί η μοντελοποίηση δομικών εξισώσεων να πραγματοποιείται σε δύο βήματα (Andreson & Gerbing, 1988). Στο πρώτο βήμα ελέγχεται το μοντέλο μέτρησης. Εάν το μοντέλο αυτό έχει αποδεκτή προσαρμογή στα δεδομένα τότε η ανάλυση συνεχίζεται με το δομικό μοντέλο. Η διαδικασία αυτή όμως δε φαίνεται να έχει ενσωματωθεί στην ερευνητική πρακτική.

ΒΑΣΙΚΑ ΒΗΜΑΤΑ ΤΗΣ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ ΔΟΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ

Κατά τη μοντελοποίηση δομικών εξισώσεων ο ερευνητής ακολουθεί συγκεκριμένα βήματα, τα οποία είναι: (α) ο καθορισμός του μοντέλου²⁷, (β) ο

²⁶ Confirmatory factor analysis (CFA).

²⁷ Model specification.

προσδιορισμός του μοντέλου²⁸, (γ) η εκτίμηση του μοντέλου²⁹, (δ) ο έλεγχος του μοντέλου³⁰, και (ε) η τροποποίηση του μοντέλου³¹. Τα βήματα αυτά παρουσιάζονται πιο αναλυτικά παρακάτω.

Καθορισμός του μοντέλου

Ο καθορισμός του μοντέλου αποτελεί το πρώτο και ίσως το πιο καθοριστικό βήμα της μοντελοποίησης δομικών εξισώσεων. Με βάση την υπάρχουσα θεωρία, τη δημοσιευμένη επιστημονική έρευνα καθώς και την προσωπική άποψη του ερευνητή δημιουργείται το προς εξέταση μοντέλο, δηλαδή το θεωρητικό μοντέλο. Αυτό σημαίνει ότι επιλέγονται προσεκτικά οι μεταβλητές που θα εισαχθούν στο μοντέλο (και φυσικά ποιες θα αποκλειστούν), καθορίζεται με σαφήνεια η θέση τους μέσα στο μοντέλο (π.χ., εάν θα είναι ενδογενείς ή εξωγενείς) καθώς και η μεταξύ τους σχέση. Σκοπός της διαδικασίας αυτής είναι να δημιουργηθεί ένα μοντέλο το οποίο να περιγράφει όσο το δυνατόν ακριβέστερα τα εμπειρικά δεδομένα που θα συλλεχθούν.

Αποτυχία επιβεβαίωσης του μοντέλου σημαίνει ότι αυτό μπορεί να είναι ένα *μη ορθά καθορισμένο μοντέλο*³². Αυτό μπορεί να συμβαίνει επειδή μία η περισσότερες σημαντικές μεταβλητές δεν έχουν συμπεριληφθεί στο μοντέλο (π.χ., προηγούμενη εμπειρία ή εκπαίδευση), ή απουσιάζει κάποια σημαντική παράμετρος (π.χ., η συσχέτιση μεταξύ δύο μεταβλητών Α και Β). Φυσικά στην περίπτωση που έχουν εισαχθεί όλες οι απαραίτητες μεταβλητές και έχουν καθοριστεί όλες οι παράμετροι αλλά το μοντέλο εξακολουθεί να μην περιγράφει ικανοποιητικά τα εμπειρικά δεδομένα, τότε γεννιούνται σοβαρές αμφιβολίες ως προς την ορθότητα της θεωρίας στην οποία στηρίζεται το μοντέλο.

Προσδιορισμός του μοντέλου

Βασικός σκοπός της μοντελοποίησης δομικών εξισώσεων είναι ο υπολογισμός των παραμέτρων του μοντέλου (π.χ., γραμμικών συντελεστών, συσχετίσεων). Για να εκτιμηθούν οι παράμετροι είναι απαραίτητο να μπορεί

²⁸ Model identification.

²⁹ Model estimation.

³⁰ Model testing.

³¹ Model modification.

³² Misspecified model.

να βρεθεί μία και μοναδική λύση για τις τιμές των παραμέτρων. Για να γίνει πιο κατανοητό ας υποθέσουμε ότι σύμφωνα με το προς εξέταση μοντέλο ισχύει ταυτόχρονα ότι η διακύμανση της μεταβλητής Y ισούται με 10 και η διακύμανση της $Y = a + \beta$. Στην περίπτωση αυτή υπάρχουν άπειροι συνδυασμοί αριθμών που έχουν άθροισμα 10. Για να μπορέσει να υπάρξει λύση θα πρέπει να τεθεί κάποιος περιορισμός³⁴. Εάν, για παράδειγμα, περιορίσουμε το a να ισούται με 1 τότε αναγκαστικά η τιμή του β είναι 9. Επομένως, για κάθε δυνατή παράμετρο του μοντέλου θα πρέπει να οριστεί εάν θα είναι *ελεύθερα εκτιμώμενη παράμετρος*³⁵ ή εάν η τιμή της θα έχει εκ των προτέρων καθοριστεί, δηλαδή θα είναι *σταθερή παράμετρος*³. Η τιμή της ελεύθερα εκτιμώμενης παραμέτρου δεν είναι γνωστή και επομένως θα πρέπει να υπολογιστεί. Αντιθέτως, η τιμή της σταθερής παραμέτρου έχει εκ των προτέρων καθοριστεί (συνήθως λαμβάνει τιμές 0 ή 1). Εάν για όλες τις παραμέτρους του μοντέλου μπορεί να βρεθεί μία και μοναδική λύση τότε το μοντέλο έχει προσδιοριστεί (*υπερ-προσδιορισμένο μοντέλο*³⁶). Στην περίπτωση που έστω και μία παράμετρος δεν μπορεί να προσδιοριστεί, τότε όλο το μοντέλο δεν έχει προσδιοριστεί επαρκώς (*υπο-προσδιορισμένο μοντέλο*³⁷).

Μία απαραίτητη αλλά όχι αναγκαία συνθήκη για τον προσδιορισμό ενός μοντέλου είναι ο αριθμός των ελεύθερα εκτιμώμενων παραμέτρων να είναι μικρότερος ή ίσος με τον αριθμό των στοιχείων του πίνακα διακύμανσης-συνδιακύμανσης (S) των εμπειρικών δεδομένων. Ο αριθμός των στοιχείων του πίνακα δίνεται από τον τύπο $p(p + 1)/2$, ενώ ο αριθμός των ελεύθερα εκτιμώμενων παραμέτρων που μπορεί να προσδιοριστούν δίνεται από τον τύπο $p(p + 3)/2$, όπου p ο αριθμός των μετρήσιμων μεταβλητών. Στην περίπτωση που υπάρχουν τρεις μετρήσιμες μεταβλητές ο αριθμός των στοιχείων του πίνακα διακύμανσης-συνδιακύμανσης είναι $3(3 + 1)/2 = 6$ (τρεις διακυμάνσεις των μεταβλητών και τρεις συνδιακυμάνσεις), ενώ των ελεύθερα εκτιμώμενων παραμέτρων $3(3 + 3)/2 = 9$. Επομένως, θα πρέπει να εισαχθούν τουλάχιστον τρεις περιορισμοί για να είναι δυνατόν οι παράμετροι του μοντέλου να προσδιοριστούν.

³³ Constraint.

³⁴ Free parameter.

³⁵ Fix parameter.

³⁶ Over-identified model

³⁷ Under-identified model

Στην πράξη ο ερευνητής επιλέγει μία μετρήσιμη μεταβλητή από κάθε λανθάνουσα μεταβλητή και καθορίζει την τιμή της μεταξύ τους σχέσης (φόρτιση) ως 1. Αυτό έχει ως αποτέλεσμα (α) να δοθεί μία κλίμακα μέτρησης³⁸ στη λανθάνουσα μεταβλητή που είναι η ίδια με την κλίμακα της μετρήσιμης μεταβλητής, και (β) να μπορεί να υπολογιστούν οι τιμές των υπόλοιπων παραμέτρων του μοντέλου, αποφεύγοντας πιθανά προβλήματα προσδιορισμού του μοντέλου. Συνήθως η μεταβλητή που επιλέγεται ως σταθερή είναι αυτή που σε προηγούμενη ανάλυση (π.χ., διερευνητική παραγοντική ανάλυση) είχε την υψηλότερη σχέση (φόρτιση) με τον παράγοντα, είναι δηλαδή η μετρούμενη μεταβλητή που εκπροσωπεί καλύτερα τη λανθάνουσα μεταβλητή.

Εναλλακτικά ο ερευνητής μπορεί να θέσει τη διακύμανση της λανθάνουσας μεταβλητής στη μονάδα. Αυτή η επιλογή έχει ως αποτέλεσμα να προκύψει τυποποιημένη λύση. Πρέπει, όμως, να σημειωθεί ότι η διακύμανση της λανθάνουσας μεταβλητής δεν μπορεί να περιοριστεί στη μονάδα στην περίπτωση που είναι ενδογενής μεταβλητή στο μοντέλο, επειδή η διακύμανση της ενδογενούς λανθάνουσας μεταβλητής δεν αποτελεί παράμετρο προς εκτίμηση.

Εκτίμηση του μοντέλου

Το επόμενο βήμα αφορά την επιλογή της μεθόδου για την εκτίμηση των παραμέτρων του μοντέλου. Ο ερευνητής θα πρέπει να επιλέξει την πιο κατάλληλη για τα δεδομένα του μέθοδο εκτίμησης³⁹. Ανάμεσα στις πιο συχνά χρησιμοποιούμενες είναι η μέθοδος της Μέγιστης Πιθανοφάνειας (Maximum Likelihood, ML), των Γενικευμένων Ελάχιστων Τετραγώνων (Generalized Least Squares, GLS), και η μέθοδος των Ελάχιστων Τετραγώνων με Συντελεστές Βαρύτητας (Weighted Least Squares, WLS). Κατά την εφαρμογή των μεθόδων εκτίμησης από τα εμπειρικά δεδομένα, δηλαδή από τον πίνακα διακύμανσης-συνδιακύμανσης δείγματος S , δημιουργείται ένας πίνακας διακύμανσης-συνδιακύμανσης του πληθυσμού Σ . Σκοπός της κάθε μεθόδου είναι να μειώσει τη διαφορά των στοιχείων μεταξύ των δύο πινάκων S και Σ . Το στατιστικό κριτήριο για τη σύγκριση των δύο πινάκων είναι το $T = (N - 1)F_{min}$, όπου N είναι το μέγεθος του δείγματος και F_{min}

³⁸ Metric.

³⁹ Estimation method

η τιμή της συνάρτησης ελαχιστοποίησης της διαφοράς μεταξύ των δύο πινάκων (Hu & Bentler, 1995). Πρέπει να τονιστεί ότι η τιμή της F_{min} εξαρτάται από τη μέθοδο εκτίμησης που θα χρησιμοποιηθεί. Σε μεγάλα μεγέθη δείγματος το T ακολουθεί (ασυμπτωτικά) την κατανομή χ^2 . Το T συχνά αναφέρεται και ως “δοκιμασία χ^2 ”. Μη στατιστικά σημαντικές τιμές χ^2 δηλώνουν ότι οι δύο πίνακες δε διαφέρουν και επομένως το μοντέλο ανταποκρίνεται στα εμπειρικά δεδομένα.

Οι διάφορες μέθοδοι εκτίμησης έχουν διαφορετικά χαρακτηριστικά. Η μέθοδος ML είναι η προεπιλεγμένη σε αρκετά γνωστά στατιστικά πακέτα (π.χ., AMOS, EQS). Η ML εκτιμά ορθά τις παραμέτρους του μοντέλου καθώς και τα τυπικά τους σφάλματα υπό την προϋπόθεση ότι οι μετρήσιμες μεταβλητές είναι συνεχείς (σε κλίμακα ίσων διαστημάτων⁴⁰ ή σε αναλογική κλίμακα⁴¹) και ακολουθούν την πολυμεταβλητή κανονική κατανομή (West, Finch, & Curran, 1995).

Αρκετές όμως είναι οι περιπτώσεις που στην ερευνητική πρακτική οι παραπάνω προϋποθέσεις δεν ικανοποιούνται. Αυτό μπορεί να συμβαίνει επειδή (α) η κατανομή των συνεχών μεταβλητών αποκλίνει από την κανονική, (β) στα δεδομένα υπάρχουν ακραίες τιμές⁴², και (γ) στο μοντέλο έχουν εισαχθεί μεταβλητές σε κλίμακα διάταξης⁴³ ή δίτιμες μεταβλητές⁴⁴. Στην περίπτωση αυτή η τιμή του χ^2 αυξάνεται με αποτέλεσμα την απόρριψη της μηδενικής υπόθεσης. Επιπλέον, υπάρχει μία τάση για μείωση των τιμών των τυπικών σφαλμάτων, που έχει ως αποτέλεσμα την ανάδειξη παραμέτρων του μοντέλου με χαμηλές τιμές ως στατιστικά σημαντικών.

Ο τρόπος αντιμετώπισης εξαρτάται από την αιτία και την έκταση της απόκλισης από την πολυμεταβλητή κανονική κατανομή. Η εξέταση για ακραίες τιμές θα πρέπει να αποτελεί ρουτίνα πριν την τελική ανάλυση των δεδομένων. Οι Tabachnick και Fidell (2001) περιγράφουν αναλυτικά τον τρόπο εντοπισμού τόσο των μονομεταβλητών όσο και των πολυμεταβλητών ακραίων τιμών καθώς και τις ενέργειες που ακολουθούνται στη συνέχεια. Στην περίπτωση που η κατανομή ορισμένων συνεχών μεταβλητών δεν ακολουθεί την κανονική κατανομή μπορεί να εφαρμοστεί κάποιο εί-

⁴⁰ Interval scale.

⁴¹ Ratio scale.

⁴² Outliers.

⁴³ Ordinal scale.

⁴⁴ Binary variables.

δος μετασχηματισμού⁴⁵. Γενικά, μεταβλητές με θετική λοξότητα⁴⁶ μετασχηματίζονται υψώνοντας τις τιμές τους σε δύναμη μικρότερη της μονάδος (π.χ., λογαριθμικός μετασχηματισμός), ενώ μεταβλητές με αρνητική λοξότητα μετασχηματίζονται υψώνοντας τις τιμές τους σε δύναμη μεγαλύτερη της μονάδος.

Εάν το πρόβλημα της απόκλισης από την πολυμεταβλητή κανονική κατανομή εξακολουθεί να υφίσταται τότε μπορεί να χρησιμοποιηθεί η μέθοδος GLS, που, αν και προϋποθέτει την ύπαρξη πολυμεταβλητής κανονικής κατανομής, δε λαμβάνει υπόψη κατά τον υπολογισμό των παραμέτρων και των τυπικών τους σφαλμάτων την *πολυμεταβλητή κύρτωση*⁴⁷ (Schumacker & Lomax, 2004). Η μέθοδος WLS στηρίζεται στη GLS και είναι *ανεξάρτητη από την κατανομή*⁴⁸ των μετρήσιμων μεταβλητών. Τα μειονεκτήματά της είναι ότι (α) απαιτεί περισσότερους υπολογισμούς και (β) χρειάζεται μεγαλύτερο δείγμα για να δώσει σταθερές εκτιμήσεις.

Ένας άλλος τρόπος αντιμετώπισης των *μη κανονικών μεταβλητών*⁴⁹ είναι η εφαρμογή της μεθόδου *Διορθωμένο χ^2* (Scaled χ^2 ή Satorra-Bentler scaled χ^2 . Satorra & Bentler, 1994). Με τη μέθοδο αυτή η τιμή του χ^2 προσαρμόζεται ανάλογα με το βαθμό απόκλισης από την πολυμεταβλητή κανονική κατανομή. Όσο η πολυμεταβλητή κύρτωση αυξάνεται τόσο η τιμή του χ^2 μειώνεται (Schumacker & Lomax, 2004). Επιπλέον, τα τυπικά σφάλματα που εκτιμώνται με τη μέθοδο αυτή θεωρούνται αρκετά σταθερά (*ανθεκτικά τυπικά σφάλματα*⁵⁰). Η δυνατότητα υπολογισμού του Satorra-Bentler διορθωμένου χ^2 καθώς και των σταθερών τυπικών σφαλμάτων παρέχονται από τα προγράμματα EQS, MPLUS, και LISREL.

Σύγκριση των δύο μεθόδων (Satorra-Bentler διορθωμένο χ^2 και WLS) έδειξε ότι η μέθοδος εκτίμησης WLS δίνει σταθερά αποτελέσματα όταν το μέγεθος του δείγματος είναι γύρω στα 1000 άτομα, το μοντέλο που εξετάζεται είναι σχετικά απλό, και οι αποκλίσεις από την πολυμεταβλητή κανονική κατανομή είναι μέτριες. Αντιθέτως, η μέθοδος Satorra-Bentler διορθωμένο χ^2 φαίνεται ότι δίνει καλές εκτιμήσεις του χ^2 με μικρότερου με-

⁴⁵ Transformation.

⁴⁶ Skewness.

⁴⁷ Multivariate kurtosis.

⁴⁸ Distribution free.

⁴⁹ Nonnormal variables.

⁵⁰ Robust standard errors.

γέθους δείγματα (π.χ., ≥ 200 άτομα), ενώ επηρεάζεται λιγότερο από υψηλές αποκλίσεις από την πολυμεταβλητή κανονική κατανομή. Επιπλέον, τα σταθερά τυπικά σφάλματα δίνουν σχετικά ακριβείς εκτιμήσεις.

Στην περίπτωση που η απόκλιση από την πολυμεταβλητή κανονικότητα οφείλεται στην ύπαρξη μεταβλητών σε κλίμακα διάταξης (με τέσσερις ή λιγότερες κατηγορίες) ή δίτιμων μεταβλητών μπορεί να εφαρμοστεί η μέθοδος *Μεθοδολογία Συνεχών/Κατηγορικών Μεταβλητών* (Continuous/Categorical Variable Methodology, CVM. Muthén, 1984). Η μέθοδος εκτίμησης CVM επιτρέπει την ανάλυση συνεχών, κατηγορικών, και δίτιμων μεταβλητών, καθώς και συνδυασμό αυτών και δίνει σταθερές και συνεπείς εκτιμήσεις των παραμέτρων του μοντέλου. Το πρόγραμμα MPLUS παρέχει τη δυνατότητα αυτή αφού δεν περιορίζεται από το είδος της κλίμακας στην οποία έχουν μετρηθεί οι μεταβλητές.

Έλεγχος του μοντέλου

Μετά την εφαρμογή κάποιας μεθόδου εκτίμησης και τον υπολογισμό των παραμέτρων του μοντέλου, ο ερευνητής θα πρέπει να αποφασίσει κατά πόσο το θεωρητικό του μοντέλου υποστηρίζεται από τα εμπειρικά δεδομένα. Για την απόφαση αυτή εξετάζονται (α) η προσαρμογή του μοντέλου στα δεδομένα, (β) η σημαντικότητα των παραμέτρων του μοντέλου, και (γ) το μέγεθος καθώς και το πρόσημο των παραμέτρων.

Προσαρμογή του μοντέλου στα δεδομένα. Η βασική δοκιμασία για την εξέταση της προσαρμογής του μοντέλου στα δεδομένα στηρίζεται στο στατιστικό T που ακολουθεί τη χ^2 κατανομή. Μη στατιστικώς σημαντικές τιμές του χ^2 δηλώνουν ότι οι δύο πίνακες S και Σ είναι παρόμοιοι και επομένως το μοντέλο περιγράφει ικανοποιητικά τα εμπειρικά δεδομένα. Η τιμή, όμως, του χ^2 έχει τα παρακάτω μειονεκτήματα (Byrne, 1994. Hu & Bentler, 1995. Kline, 1998b): (α) Το στατιστικό T ακολουθεί ασυμπτωτικώς την κατανομή χ^2 . Επομένως, σε μικρά δείγματα είναι πιθανό να αποκλίνει από τη συγκεκριμένη κατανομή και να μην οδηγεί σε ορθά συμπεράσματα ως προς την προσαρμογή του μοντέλου. (β) Ο ερευνητής μπορεί να οδηγηθεί σε εσφαλμένα συμπεράσματα ως προς την καταλληλότητα του μοντέλου όταν η προϋπόθεση της πολυμεταβλητής κανονικής κατανομής δεν ικανοποιείται. (γ) Επιπλέον, η τιμή του χ^2 αυξάνεται όσο αυξάνει το μέγεθος του δείγματος. Επομένως, σε μεγάλα δείγματα είναι πιθανόν μικρές διαφορές μεταξύ των δύο πινάκων να βρεθούν στατιστικά σημαντικές. (δ)

Τέλος, η εκτίμηση της προσαρμογής του μοντέλου με βάση το χ^2 οδηγεί σε μία απόφαση αποδοχής/απόρριψης χωρίς ο ερευνητής να μπορεί να έχει μία εικόνα για το βαθμό αποδοχής ή απόρριψης.

Σε μία προσπάθεια αντιμετώπισης των παραπάνω μειονεκτημάτων αναπτύχθηκαν διάφοροι δείκτες προσαρμογής⁵¹, οι οποίοι μπορούν να ταξινομηθούν σε τρεις κατηγορίες: (α) απόλυτοι δείκτες προσαρμογής⁵², (β) δείκτες σχετικής βελτίωσης του μοντέλου⁵³, και (γ) δείκτες φειδωλότητας του μοντέλου⁵⁴.

Απόλυτοι δείκτες. Οι δείκτες αυτοί εξετάζουν το βαθμό στον οποίο έχει αναπαραχθεί ο πίνακας διακύμανσης-συνδιακύμανσης του πληθυσμού Σ . Στην κατηγορία των απόλυτων δεικτών ανήκουν ο Δείκτης Καλής Προσαρμογής (Goodness of Fit Index, GFI), ο Προσαρμοσμένος Δείκτης Καλής Προσαρμογής (Adjusted Goodness of Fit Index, AGFI), ο δείκτης Ρίζα του Τυποποιημένου Μέσου Τετραγωνικού Υπολοίπου Σφάλματος (Standardized Root-Mean-square Residual, SRMR), καθώς και ο δείκτης Ρίζα του Μέσου Τετραγωνικού Σφάλματος Προσέγγισης (Root Mean Square Error of Approximation, RMSEA).

Η τιμή του δείκτη GFI δηλώνει το ποσοστό της διακύμανσης-συνδιακύμανσης του πίνακα S που προβλέπεται από τον πίνακα Σ . Σύμφωνα με τον Kline (1998b) ο GFI μπορεί να θεωρηθεί ανάλογος του δείκτη R^2 της ανάλυσης πολλαπλής παλινδρόμησης. Για την εκτίμηση του AGFI η τιμή του GFI προσαρμόζεται ανάλογα με το λόγο των βαθμών ελευθερίας του μοντέλου προς τον αριθμό των μεταβλητών. Οι τιμές των δύο αυτών δεικτών κυμαίνονται από 0-1, ενώ τιμές πάνω από .90 θεωρούνται ικανοποιητικές.

Η τιμή του δείκτη SRMR είναι η τετραγωνική ρίζα της τυποποιημένης μέσης τιμής της διαφοράς των στοιχείων των πινάκων S και Σ . Είναι κατανοητό ότι όσο πιο μικρή είναι η τιμή του SRMR τόσο καλύτερα έχει αναπαραχθεί ο πίνακας Σ . Τιμές του SRMR μικρότερες από .05 αποτελούν ένδειξη καλής προσαρμογής του μοντέλου στα δεδομένα.

Ο δείκτης RMSEA παρουσιάστηκε πρώτη φορά από τους Steiger και Lind (1980. Βλ. Curran, Bollen, Chen, Paxton, & Kirby, 2003). Στη συνέχεια έγινε ευρύτερα γνωστός από τους Browne and Cudeck (1993) και Steiger

⁵¹ Fit indices.

⁵² Absolute fit indices.

⁵³ Incremental fit indices.

⁵⁴ Parsimony fit indices.

(1990). Η τιμή του δείκτη RMSEA παριστάνει τη διαφορά των δύο πινάκων διακύμανσης-συνδιακύμανσης S και Σ ανά βαθμό ελευθερίας. Εάν ένα μοντέλο έχει καλή προσαρμογή στα δεδομένα, η τιμή αυτού του δείκτη αναμένεται να είναι μικρότερη του .05, ενώ τιμές μικρότερες του .08 είναι ενδεικτικές ικανοποιητικής προσαρμογής⁵⁵. Τα πλεονεκτήματα του δείκτη RMSEA είναι ότι (α) το μέγεθος του δείγματος έχει μικρή επίδραση στη τιμή του (Fan, Thompson, & Wang, 1999) και (β) η κατανομή που ακολουθεί ο RMSEA είναι γνωστή, γεγονός που επιτρέπει τον υπολογισμό διαστήματος εμπιστοσύνης (ΔE) που συνήθως ορίζεται στο 90% ΔE . Σε ένα μοντέλο με καλή προσαρμογή το κάτω όριο ενός 90% ΔE θα πρέπει να πλησιάζει στο μηδέν ενώ το άνω όριο να είναι μικρότερο του .08 (Garson, 2007).

Δείκτες σχετικής βελτίωσης του μοντέλου. Οι δείκτες αυτοί αξιολογούν τη σχετική βελτίωση του μοντέλου ως προς ένα βασικό μοντέλο⁵⁶. Το μοντέλο αυτό είναι συνήθως το μηδενικό μοντέλο⁵⁷ ή, αλλιώς, ανεξάρτητο μοντέλο⁵⁸. Από τους δείκτες της κατηγορίας αυτής θα αναφερθούν ο Tucker-Lewis Δείκτης (Tucker-Lewis Index, TLI) ή αλλιώς Μη Σταθμισμένος Δείκτης Προσαρμογής (Non-Normed Fit Index, NNFI), ο Σταθμισμένος Δείκτης Προσαρμογής (Normed Fit Index, NFI), και ο Συγκριτικός Δείκτης Προσαρμογής (Comparative Fit Index, CFI).

Αρχικά, ο TLI (ή NNFI) αναπτύχθηκε για να εφαρμοστεί στη διερευνητική παραγοντική ανάλυση. Στη συνέχεια όμως η χρήση του επεκτάθηκε και στη μοντελοποίηση δομικών εξισώσεων και χρησιμοποιείται για τη σύγκριση του προς εξέταση μοντέλου με το μηδενικό μοντέλο. Τιμές του TLI μεγαλύτερες από .90 θεωρούνται ικανοποιητικές. Πρόσφατη, όμως, έρευνα υποστήριξε την ανάγκη για πιο αυστηρές τιμές, οι οποίες κυμαίνονται γύρω στο .95 (Hu & Bentler, 1999). Οι τιμές του όμως μπορεί να βρίσκονται και εκτός διαστήματος 0-1.

Σε μια προσπάθεια να περιοριστεί το εύρος τιμών του TLI μέσα στο διάστημα 0-1 αναπτύχθηκε ο NFI. Ο δείκτης αυτός παριστάνει το ποσοστό

⁵⁵ Σύμφωνα με τους Hu και Bentler (1999) η τιμή κριτήριο του δείκτη RMSEA για καλή προσαρμογή ενός μοντέλου είναι μικρότερη ή ίση του .06.

⁵⁶ Baseline model

⁵⁷ Null model.

⁵⁸ Independent model. Μηδενικό μοντέλο ή ανεξάρτητο μοντέλο είναι το μοντέλο εκείνο στο οποίο οι μετρήσιμες μεταβλητές δε συνδέονται μεταξύ τους, δηλαδή οι μεταξύ τους συσχετίσεις είναι μηδέν. Αυτό συνεπάγεται ότι και η σχέση μεταξύ των λανθανουσών μεταβλητών είναι μηδέν. Το ανεξάρτητο μοντέλο είναι το πιο φειδωλό μοντέλο.

βελτίωσης του προς εξέταση μοντέλου ως προς το μηδενικό μοντέλο. Για παράδειγμα τιμή του NFI = .90 σημαίνει ότι η προσαρμογή του μοντέλου που θέτει ο ερευνητής είναι 90% καλύτερη του μηδενικού μοντέλου. Τιμές του NFI μεγαλύτερες του .90 δηλώνουν ικανοποιητική προσαρμογή του μοντέλου ενώ γύρω στο .95 καλή προσαρμογή (Hu & Bentler, 1999. Schumacker & Lomax, 2004). Το μειονέκτημα του δείκτη αυτού είναι το γεγονός ότι δε λαμβάνει υπόψη του την πολυπλοκότητα του μοντέλου. Έτσι, μοντέλα με μεγαλύτερο αριθμό παραμέτρων τείνουν να έχουν και καλύτερες NFI τιμές (Garson, 2007).

Ο δείκτης CFI παρουσιάστηκε από τον Bentler (1990) σε μία προσπάθειά του να περιορίσει το εύρος τιμών ενός προηγούμενου δείκτη (Bentler's Fit Index, BFI) μέσα στο διάστημα 0-1. Η εκτίμησή του CFI στηρίζεται στη μετατοπισμένη⁵⁹ χ^2 κατανομή [$\chi^2(df, \lambda)$], της οποίας η κατανομή χ^2 είναι μερική περίπτωση (όταν $\lambda = 0$). Αν και η ερμηνεία του είναι παρόμοια με του NFI, για τον υπολογισμό του CFI λαμβάνεται υπόψη ο αριθμός των παραμέτρων. Τιμές του CFI μεγαλύτερες από .90 θεωρούνται αποδεκτές, ενώ οι Hu και Bentler (1999) προτείνουν πιο υψηλές τιμές (γύρω στο .95) για την αποδοχή ικανοποιητικής προσαρμογής ενός μοντέλου στα δεδομένα.

Δείκτες φειδωλότητας του μοντέλου. Οι δείκτες φειδωλότητας εκτιμούν την προσαρμογή του μοντέλου, λαμβάνοντας υπόψη και τον αριθμό των παραμέτρων που πρέπει να υπολογιστούν. Από τους πρώτους δείκτες φειδωλότητας είναι ο δείκτης Σταθμισμένο χ^2 [Normed Chi-Square, NC = (χ^2/df)], οποίος παριστάνει το πηλίκο της τιμής χ^2 προς τους βαθμούς ελευθερίας. Αρκετοί ερευνητές προτείνουν τη χρήση του NC ως μέτρου προσαρμογής του μοντέλου (Byrne, Shavelson, & Muthén, 1989. Kline, 1998b). Γενικά, τιμές κοντά στη μονάδα θεωρείται ότι αντανακλούν ένα ορθό μοντέλο. Όμως δεν είναι ακόμα ξεκάθαρο ποια τιμή του NC αντανακλά μη ικανοποιητικό μοντέλο (Arbuckle, 1997). Για παράδειγμα, η Byrne (1989) εκτιμά ότι τιμές μεγαλύτερες από 2 δείχνουν μη ικανοποιητική προσαρμογή, ενώ ο Kline (1998b) αναφέρει ότι τιμές NC μικρότερες του 3 για μεγάλα δείγματα και μικρότερες του 2.5 για μικρά δείγματα μπορεί να θεωρούνται ικανοποιητικές

Ένας άλλος δείκτης φειδωλότητας που συχνά χρησιμοποιείται στη βιβλιογραφία είναι το Πληροφοριακό Κριτήριο Akaike (Akaike Information

⁵⁹ Noncentral.

Criterion, AIC). Υπάρχουν δύο τρόποι υπολογισμού του AIC. Ο πρώτος είναι $AIC = \chi^2 + 2q$, όπου q ο αριθμός των παραμέτρων στο μοντέλο (στο πρόγραμμα AMOS), και ο δεύτερος $AIC = \chi^2 - 2df$ (στο πρόγραμμα EQS), όπου df οι βαθμοί ελευθερίας. Ανεξάρτητα όμως από τον τρόπο υπολογισμού του, τιμές κοντά στο μηδέν δηλώνουν πιο φειδωλό μοντέλο (Schumacker & Lomax, 2004). Ο AIC είναι ιδιαίτερα χρήσιμος όταν ο ερευνητής επιθυμεί να συγκρίνει δύο ή περισσότερα μη εμφωλευμένα μοντέλα⁶⁰. Στην περίπτωση αυτή το μοντέλο με τη μικρότερη τιμή AIC θεωρείται ότι έχει την καλύτερη προσαρμογή.

Οι Browne και Cudeck (1989) παρουσίασαν το Δείκτη Αναμενόμενης Διασταυρούμενης Εγκυρότητας (Expected Cross-Validation Index, ECVI). Ο δείκτης αυτός χρησιμοποιείται για την επιλογή ανάμεσα σε μία ομάδα πιθανών μοντέλων. Μικρότερες τιμές του ECVI δείχνουν καλύτερη προσαρμογή του μοντέλου στα δεδομένα. Επιπλέον, η τιμή του ECVI αναμένεται να είναι μία καλή εκτίμηση του δείκτη προσαρμογής στην περίπτωση που το προτεινόμενο μοντέλο επανεξεταστεί χρησιμοποιώντας καινούργιο δείγμα. Τέλος, είναι δυνατόν να υπολογιστεί και διάστημα εμπιστοσύνης για την τιμή του (συνήθως 90%ΔΕ).

Σύγκριση των διάφορων δεικτών. Όσον αφορά τη δοκιμασία χ^2 , έχει αναφερθεί παραπάνω ότι εξαρτάται από τη μέθοδο εκτίμησης που θα χρησιμοποιηθεί για να υπολογιστεί το στατιστικό T . Έτσι, υπάρχει η τιμή χ^2 που στηρίζεται στη μέθοδο εκτίμησης ML ($ML \chi^2$), στη μέθοδο GLS ($GLS \chi^2$), στη μέθοδο ADF ($ADF \chi^2$), καθώς και στη μέθοδο Satorra-Bentler διορθωμένο χ^2 . Η τιμή χ^2 που βασίζεται στη ML μέθοδο εκτίμησης έχει βρεθεί ότι έχει την ίδια απόδοση με την $GLS \chi^2$ και Satorra-Bentler διορθωμένο χ^2 σε μεγάλο μέγεθος δείγματα. Η μέθοδος ADF έχει αρκετά καλά αποτελέσματα, τα οποία μάλιστα είναι ανεξάρτητα από την κατανομή των παρατηρούμενων μεταβλητών. Η ιδιότητα αυτή ισχύει μόνο στην περίπτωση που υπάρχουν ασυνήθιστα μεγάλα μεγέθη δειγμάτων (π.χ., 5000 άτομα). Σε μικρά μεγέθη όμως η ADF φαίνεται ότι δεν έχει καλή απόδοση και επομένως δε συνιστάται. Το $GLS \chi^2$ αποδίδει καλύτερα σε μικρά μεγέθη σε σχέση με το $ML \chi^2$. Τέλος, το Satorra-Bentler διορθωμένο χ^2 έχει βρεθεί ότι αποδίδει αρκετά καλά ακόμα και σε πολύ μικρά δείγματα και θα πρέπει να επιλέγεται σε ανάλογες περιπτώσεις (Hu & Bentler, 1995).

⁶⁰ Non nested models.

Η ύπαρξη πολλών δεικτών προσαρμογής του μοντέλου συχνά προβληματίζει τους ερευνητές για ποιον ή ποιους από αυτούς θα χρησιμοποιήσουν για να συναγάγουν συμπέρασμα σχετικά με το βαθμό που το μοντέλο τους περιγράφει τα εμπειρικά δεδομένα. Το πρόβλημα είναι ακόμα πιο οξύ όταν διαφορετικοί δείκτες οδηγούν σε διαφορετικά συμπεράσματα. Σε πρόσφατη έρευνά τους οι Fan et al. (1999) εξέτασαν την επίδραση του μεγέθους του δείγματος, της μεθόδου εκτίμησης καθώς και το βαθμό της απόκλισης από το ορθό μοντέλο σε δέκα δείκτες προσαρμογής που συχνά εμφανίζονται στη βιβλιογραφία. Τα αποτελέσματά τους έδειξαν ότι οι τιμές των δεικτών CFI, TLI και RMSEA επηρεάζονταν σε μικρότερο βαθμό σε σχέση με τους υπόλοιπους δείκτες από το μέγεθος του δείγματος και τη μέθοδο εκτίμησης. Επιπλέον, οι δείκτες CFI και TLI ήταν οι πιο ανθεκτικοί σε εσφαλμένως καθορισμένα μοντέλα. Η καλή επίδοση των CFI και TLI έχει επισημανθεί και σε προγενέστερη μελέτη (Hu & Bentler, 1995). Σε άλλη έρευνα οι Williams και Holahan (1994) μελέτησαν την απόδοση διάφορων δεικτών φειδωλότητας. Από τα αποτελέσματά τους φάνηκε ότι ο AIC ήταν ο καλύτερος.

Εξέταση των παραμέτρων του μοντέλου. Αφού ο ερευνητής επιλέξει το μοντέλο που αντιπροσωπεύει καλύτερα τα εμπειρικά του δεδομένα πρέπει να προχωρήσει στην ερμηνεία των παραμέτρων του μοντέλου. Αυτό σημαίνει ότι θα πρέπει να εξεταστεί η στατιστική τους σημαντικότητα, καθώς και το μέγεθος και το πρόσημό τους (θετικό ή αρνητικό). Αξίζει να αναφερθεί ότι το γεγονός ότι ένα μοντέλο έχει καλή προσαρμογή στα δεδομένα, δεν εγγυάται ότι όλοι οι παράμετροί του έχουν θεωρητική και πρακτική σημασία.

Στατιστική σημαντικότητα των παραμέτρων. Η εξέταση της σημαντικότητας των παραμέτρων ενός μοντέλου μπορεί να πραγματοποιηθεί με δύο τρόπους. Ο πρώτος τρόπος αφορά την εφαρμογή του κριτηρίου του Κρίσιμου Λόγου⁶¹ (CR), κατά την οποία η τιμή της παραμέτρου διαιρείται με το τυπικό της σφάλμα. Η τιμή που προκύπτει ακολουθεί την τυποποιημένη κανονική κατανομή, και στην περίπτωση που είναι μεγαλύτερη μίας τιμής κριτηρίου (π.χ., 1.96 για δίπλευρο έλεγχο, σε επίπεδο σημαντικότητας .05) τότε η παράμετρος θεωρείται στατιστικώς διαφορετική του μηδενός. Η τιμή όμως των τυπικών σφαλμάτων των παραμέτρων εξαρτάται από τον τρόπο που έχει προσδιοριστεί το μοντέλο. Έχει αναφερθεί ότι διαφορετικός

⁶¹ Critical Ratio (CR).

προσδιορισμός του μοντέλου μπορεί να οδηγήσει σε διαφορετικά συμπεράσματα ως προς τη σημαντικότητα των παραμέτρων (Gonzalez & Griffin, 2001). Για το λόγο αυτό οι Gonzalez και Griffin (2001) προτείνουν ο έλεγχος της σημαντικότητας των παραμέτρων να πραγματοποιείται με τη δοκιμασία Διαφορά χ^2 ($\Delta\chi^2$) ή, αλλιώς, Λόγο Πιθανοφάνειας⁶² και η οποία δεν επηρεάζεται από τον τρόπο προσδιορισμού του μοντέλου.

Πιο αναλυτικά, η τιμή χ^2 του αρχικού μοντέλου αφαιρείται από την τιμή χ^2 του νέου μοντέλου, στο οποίο η συγκεκριμένη παράμετρος έχει οριστεί να είναι μηδέν. Η $\Delta\chi^2$ ακολουθεί την χ^2 κατανομή με βαθμούς ελευθερίας τη διαφορά των βαθμών ελευθερίας των δύο μοντέλων. Εάν η $\Delta\chi^2$ δεν είναι στατιστικά σημαντική, σημαίνει ότι η ύπαρξη της παραμέτρου αυτής δε συνεισφέρει στην καλύτερη προσαρμογή του μοντέλου. Στο πρόγραμμα AMOS έκδοση 7.0 η παραπάνω δοκιμασία μπορεί να πραγματοποιηθεί με την εντολή “specification search”.

Μέγεθος και πρόσημο των παραμέτρων. Οπτική εξέταση του μεγέθους και του πρόσημου των παραμέτρων αποκαλύπτει εάν αυτές λαμβάνουν επιτρεπτές τιμές και εάν βρίσκονται σε συμφωνία με τους ισχυρισμούς της υπάρχουσας θεωρίας. Για παράδειγμα, εάν αναμένεται ότι όσο μεγαλύτερη είναι η επαγγελματική εξουθένωση τόσο μικρότερη θα είναι η επαγγελματική ικανοποίηση των εργαζομένων τότε θα πρέπει να βρεθεί μία αρνητική συσχέτιση μεταξύ των δύο αυτών εννοιών. Από την άλλη μεριά, μία συσχέτιση δεν μπορεί να είναι μεγαλύτερη της μονάδας ή η διακύμανση να έχει αρνητικό πρόσημο. Οι περιπτώσεις αυτές αναφέρονται στη βιβλιογραφία ως *Heywood case* (Dillon, Kumar, & Munali, 1987). Οι αιτίες μπορεί να είναι διάφορες, όπως η ύπαρξη ακραίων τιμών, μη επαρκές μέγεθος δείγματος, πολυσυγγραμμικότητα, και μικρός αριθμός μετρήσιμων μεταβλητών ανά λανθάνοντα παράγοντα (συνήθως λιγότεροι από τρεις).

Τροποποίηση του μοντέλου

Στην περίπτωση που το αρχικό μοντέλο δεν έχει καλή προσαρμογή στα δεδομένα ο ερευνητής έχει δύο επιλογές. Η πρώτη του επιλογή είναι να ακολουθηθεί μία αυστηρή τακτική και να απορρίψει την ορθότητα του μοντέλου. Η δεύτερη επιλογή αφορά τη διερεύνηση των σημείων του μοντέλου

⁶² Chi-Square Difference ($\Delta\chi^2$) ή Likelihood Ratio (LR).

που ήταν υπεύθυνα για τη μειωμένη προσαρμογή. Αυτά συνήθως είναι παράμετροι οι οποίοι δεν έχουν επαρκώς καθοριστεί, δηλαδή *σφάλματα καθορισμού*⁶³. Για παράδειγμα, μια συσχέτιση ή μία φόρτιση που έχει παραλειφθεί ή πλεονάζει. Μετά τον εντοπισμό των παραμέτρων αυτών το μοντέλο τροποποιείται και το νέο τροποποιημένο μοντέλο εξετάζεται εκ νέου ως προς την προσαρμογή του στα δεδομένα.

Ένας τρόπος για τον έλεγχο των σημείων του μοντέλου που δεν έχουν επαρκώς καθοριστεί είναι η οπτική εξέταση του πίνακα υπολοίπων σφάλματος⁶⁴ και ιδιαίτερα των τυποποιημένων υπολοίπων σφάλματος. Σύμφωνα με τον Arbuckle (1997), εάν το μοντέλο είναι ορθό, τότε η αναμενόμενη τιμή της *συνδιακύμανσης των τυποποιημένων υπολοίπων σφάλματος*⁶⁵ θα πρόκειται σε απόλυτη τιμή να είναι μικρότερη του 2. Υψηλές τιμές τυποποιημένων υπολοίπων σφάλματος μεταξύ πολλών και διαφορετικών μεταβλητών δηλώνουν σφάλμα γενικά στον καθορισμό του μοντέλου. Όταν οι υψηλές τιμές αφορούν μία μεταβλητή αυτό σημαίνει ότι η συγκεκριμένη συνδιακύμανση της μεταβλητής αυτής δεν έχει εξηγηθεί καλά από το μοντέλο. Επομένως, το μοντέλο θα πρέπει να επανεξεταστεί και να βρεθούν τρόποι έτσι ώστε η συγκεκριμένη συνδιακύμανση να εξηγηθεί καλύτερα. Αυτό συνήθως γίνεται ορίζοντας κάποιες παραμέτρους από σταθερές σε ελεύθερα εκτιμώμενες.

Τα υπάρχοντα στατιστικά προγράμματα έχουν ενσωματώσει διαδικασίες αναζήτησης βελτίωσης του μοντέλου. Για παράδειγμα, στο πρόγραμμα AMOS και LISREL υπάρχει η δυνατότητα υπολογισμού του Δείκτη Τροποποίησης⁶⁶ για κάθε σταθερή παράμετρο του μοντέλου. Η τιμή του δείκτη αυτού παρουσιάζει την εκτιμώμενη μείωση της τιμής του χ^2 του αρχικού μοντέλου στην περίπτωση που κάποια συγκεκριμένη σταθερή παράμετρος εκτιμηθεί ελεύθερα. Υψηλές τιμές του Δείκτη Τροποποίησης για κάποια παράμετρο φανερώνουν ότι μπορεί να επιτευχθεί καλύτερη προσαρμογή στα δεδομένα όταν η σταθερή αυτή παράμετρος εκτιμηθεί ελεύθερα.

Άλλες μέθοδοι εκτίμησης της τροποποίησης του μοντέλου είναι η Αναμενόμενη Μεταβολή Παραμέτρου⁶⁷, ο Πολλαπλασιαστής Lagrange⁶⁸, και το

⁶³ Specification errors.

⁶⁴ Residual matrix.

⁶⁵ Standardized covariance of residuals.

⁶⁶ Modification Index.

⁶⁷ Expected Parameter Change (EPC).

⁶⁸ Lagrange Multiplier.

Wald τεστ. Η πρώτη μέθοδος υποστηρίζεται από τα προγράμματα AMOS, EQS, και LISREL και δείχνει την εκτιμώμενη μεταβολή στο μέγεθος και στο πρόσημο για κάθε σταθερή παράμετρο εάν αυτή εκτιμηθεί ελεύθερα από τα δεδομένα του δείγματος. Με την εφαρμογή της μεθόδου Αναμενόμενης Μεταβολής Παραμέτρου ο ερευνητής μπορεί να διαπιστώσει εάν το μέγεθος και το πρόσημο της επιπρόσθετης παραμέτρου που θα εκτιμηθεί είναι σε συμφωνία με τη θεωρία. Οι δοκιμασίες Πολλαπλασιαστικής Lagrange και Wald τεστ παρέχονται από το πρόγραμμα EQS. Το στατιστικό Πολλαπλασιαστικής Lagrange δείχνει την εκτιμώμενη μεταβολή του χ^2 στην περίπτωση που στο μοντέλο εισαχθεί μία νέα παράμετρος. Αντίθετα, το στατιστικό Wald τεστ φανερώνει ποιες παράμετροι του μοντέλου είναι περιττές και συνεπώς θα πρέπει να αφαιρεθούν. Επίσης, επειδή το στατιστικό Wald τεστ εξετάζει εάν η τιμή μιας ομάδας παραμέτρων μπορεί ταυτόχρονα να θεωρηθεί μηδέν χωρίς ουσιαστική μείωση της προσαρμογής του μοντέλου, μπορεί να λειτουργήσει και ως η πολυμεταβλητή δοκιμασία για την εξέταση της σημαντικότητας των παραμέτρων ενός μοντέλου (Byrne, 1994, p. 149).

Πρέπει να σημειωθεί ότι η τροποποίηση του αρχικού μοντέλου φαίνεται να είναι μια κοινή πρακτική στη μοντελοποίηση δομικών εξισώσεων (MacCallum, 1986). Η τροποποίηση του αρχικού μοντέλου πρέπει προσεκτικά να αποφασιστεί και θεωρείται επιτυχημένη όταν (α) οι τροποποιήσεις είναι ελάχιστες (Schumacker & Lomax, 2004), (β) βελτιώνουν ουσιαστικά την προσαρμογή του μοντέλου (MacCallum, 1986), (γ) οι τιμές των επιπρόσθετων παραμέτρων είναι σημαντικές και ουσιαστικές (Byrne, 2006. Byrne et al., 1989), και (δ) παρέχεται τεκμηριωμένη αιτιολόγηση για την εισαγωγή τους (Byrne, 2006. Byrne et al., 1989. MacCallum, 1986).

Σε κάθε τροποποίηση του αρχικού μοντέλου υπάρχει ο κίνδυνος η καλύτερη προσαρμογή των δεδομένων να οφείλεται σε *τυχαίους λόγους*⁶⁹, οι οποίοι συνδέονται με χαρακτηριστικά του συγκεκριμένου δείγματος που χρησιμοποιήθηκε (MacCallum, Roznowski, & Necowitz, 1992). Για να μπορέσει ο ερευνητής να είναι περισσότερο σίγουρος ότι η τροποποίηση του μοντέλου του πράγματι δεν ήταν τυχαία θα πρέπει να επιβεβαιώσει τις τροποποιήσεις του σε ένα άλλο ανεξάρτητο από το αρχικό δείγμα, δηλαδή να ελέγξει τη *διασταυρούμενη εγκυρότητα*⁷⁰ του μοντέλου. Εάν αυτό δεν

⁶⁹ Capitalization on chance.

⁷⁰ Cross-validation.

είναι δυνατόν τότε μπορεί να χρησιμοποιηθεί ο δείκτης ECVI. Παράδειγμα τεκμηριωμένης τροποποίησης του αρχικού μοντέλου και επιβεβαίωσης της τροποποίησής του σε άλλο δείγμα μπορεί να βρεθεί στο άρθρο των Evaggelínou, Tsigilis, και Papa (2002) και Tsigilis και Gregoriadis (2008).

Αναφορά των αποτελεσμάτων μοντελοποίησης δομικών εξισώσεων

Μετά από τη διεξαγωγή της μοντελοποίησης δομικών εξισώσεων ο ερευνητής συνήθως επιθυμεί να κάνει ευρύτερα γνωστά τα αποτελέσματα της μελέτης του. Στο κείμενο των αποτελεσμάτων θα πρέπει να αναφέρονται τα εξής: (α) Η μέθοδος που χρησιμοποιήθηκε για την εκτίμηση των παραμέτρων του μοντέλου. (β) Ο βαθμός στον οποίο ικανοποιούνται οι προϋποθέσεις για την εφαρμογή της συγκεκριμένης μεθόδου. Για παράδειγμα, εάν εφαρμόστηκε η μέθοδος της μέγιστης πιθανοφάνειας (ML) θα πρέπει να τεκμηριωθεί ότι οι μεταβλητές είναι συνεχείς και ότι ακολουθούν την πολυμεταβλητή κανονική κατανομή. (γ) Τυχόν εναλλακτικά μοντέλα που εξετάστηκαν. (δ) Η τιμή του χ^2 και οι βαθμοί ελευθερίας και τουλάχιστον ένας δείκτης από κάθε ομάδα δεικτών προσαρμογής (απόλυτοι δείκτες, δείκτες σχετικής βελτίωσης και δείκτες φειδωλότητας). (ε) Το μέγεθος, το πρόσημο και η σημαντικότητα των παραμέτρων μαζί με τα τυπικά τους σφάλματα.

ΠΑΡΑΔΕΙΓΜΑ ΕΠΙΒΕΒΑΙΩΤΙΚΗΣ ΠΑΡΑΓΟΝΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

Στο τμήμα αυτό της εργασίας παρουσιάζεται η εφαρμογή της μοντελοποίησης δομικών εξισώσεων για την εξέταση της εσωτερικής δομής ενός οργάνου μέτρησης της αυτο-αποτελεσματικότητας των εκπαιδευτικών. Η εφαρμογή αυτή συναντάται με το όνομα επιβεβαιωτική παραγοντική ανάλυση και εξετάζει τη σχέση των λανθανουσών μεταβλητών με τις αντίστοιχες μετρήσιμες μεταβλητές. Ελέγχει δηλαδή το βαθμό στον οποίο οι μετρήσιμες μεταβλητές είναι πράγματι κατάλληλες για να αντιπροσωπεύσουν μια συγκεκριμένη λανθάνουσα μεταβλητή.

Τα δεδομένα προέρχονται από τη διδακτορική διατριβή του Τσιγγίλη (2005), στην οποία μεταξύ άλλων εξετάστηκε η δομική εγκυρότητα ενός οργάνου μέτρησης της αυτο-αποτελεσματικότητας εκπαιδευτικών. Ως αυτο-αποτελεσματικότητα του εκπαιδευτικού μπορεί να οριστεί η πίστη ή η πεποίθησή του ότι μπορεί να επηρεάσει τη μάθηση των μαθητών ακόμα και

εκείνων που δεν έχουν υψηλά κίνητρα (Guskey & Passaro, 1994). Σειρά ερευνών δείχνουν ότι η αποτελεσματικότητα του εκπαιδευτικού είναι ένας σημαντικός παράγοντας, ο οποίος επηρεάζει πολλές πτυχές της εκπαιδευτικής διαδικασίας. Πιο συγκεκριμένα, η αποτελεσματικότητα των εκπαιδευτικών έχει βρεθεί ότι συνδέεται με τα κίνητρα και την επίτευξη των μαθητών καθώς και με την αυτο-εκτίμησή τους, με τη διαχείριση της τάξης και τη διατήρηση της πειθαρχίας, την παρατεταμένη απουσία των καθηγητών και την επαγγελματική τους εξουθένωση και το άγχος (Brouwers & Tomic, 2000. Soodak & Podell, 1996. Tschannen-Moran & Woolfolk-Hoy, 2001).

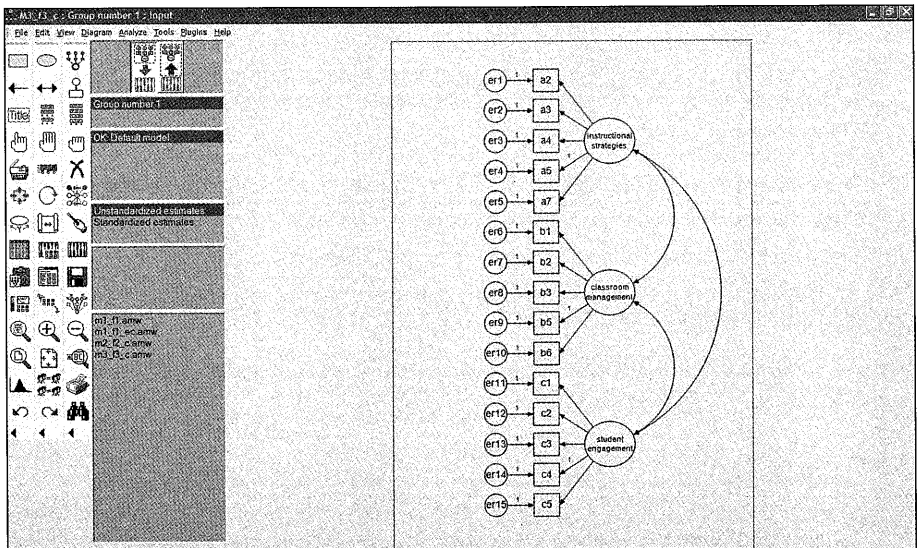
Η αυτο-αποτελεσματικότητα 284 καθηγητών φυσικής αγωγής δημόσιων σχολείων πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης μελετήθηκε με μια συντομευμένη μορφή της Κλίμακας Αίσθησης Αποτελεσματικότητας των Εκπαιδευτικών (Teachers' Sense of Efficacy Scale, TSES. Tschannen-Moran & Woolfolk-Hoy, 2001)⁷¹. Η συντομευμένη TSES αξιολογεί τρεις διαστάσεις της αυτο-αποτελεσματικότητας του εκπαιδευτικού: (α) την αποτελεσματικότητά τους σε στρατηγικές διδασκαλίας, η οποία αναφέρεται στην αντίληψη των εκπαιδευτικών σχετικά με την ικανότητά τους να εφαρμόζουν μεθόδους και στρατηγικές διδασκαλίας στην τάξη καθώς και να προσαρμόζουν το μάθημά τους στο επίπεδο των μαθητών (πέντε θέματα, π.χ., “Σε ποιο βαθμό μπορείς να εφαρμόξεις εναλλακτικές μεθόδους διδασκαλίας;”), (β) την αποτελεσματικότητά τους στη διαχείριση της τάξης, η οποία αναφέρεται στην αντίληψη των εκπαιδευτικών σχετικά με την ικανότητά τους να μπορούν να πειθαρχούν την τάξη και να επιβάλλουν κανόνες συμπεριφοράς (πέντε θέματα, π.χ., “Σε ποιο βαθμό μπορείς να ελέγχεις συμπεριφορές που διασπούν τη συνοχή της τάξης;”) και (γ) την αποτελεσματικότητά τους ως προς την ενασχόληση των μαθητών, η οποία αναφέρεται στην αντίληψη των εκπαιδευτικών σχετικά με την ικανότητά τους να κινητοποιούν και να εμπλέκουν τους μαθητές σε διαδικασίες μάθησης (πέντε θέματα, π.χ., “Σε ποιο βαθμό μπορείς να βελτιώσεις την ικανότητα κατανόησης ενός μαθητή που δεν τα πάει καθόλου καλά;”). Οι απαντήσεις δίνονταν σε 9-βάθμια κλίμακα τύπου Likert κλιμακούμενη από το 1 (καθόλου) έως το 9 (σε πολύ μεγάλο βαθμό).

⁷¹ Αρχικά η κλίμακα ονομάστηκε Ohio State Teacher Efficacy Scale (Tschannen-Moran & Woolfolk-Hoy, 2001). Στη συνέχεια, όμως, εμφανίστηκε με την ονομασία Teachers' Sense of Efficacy Scale (http://mxtsch.people.wm.edu/research_tools.php).

Παρακάτω περιγράφονται αναλυτικά τα βήματα που ακολουθήθηκαν για την εξέταση της δομής της συντομευμένης TSES. Για την ανάλυση των δεδομένων χρησιμοποιήθηκε το πρόγραμμα AMOS 7.0 (Arbuckle, 2006) και EQS 6.1 (Bentler, 2005).

Καθορισμός του μοντέλου

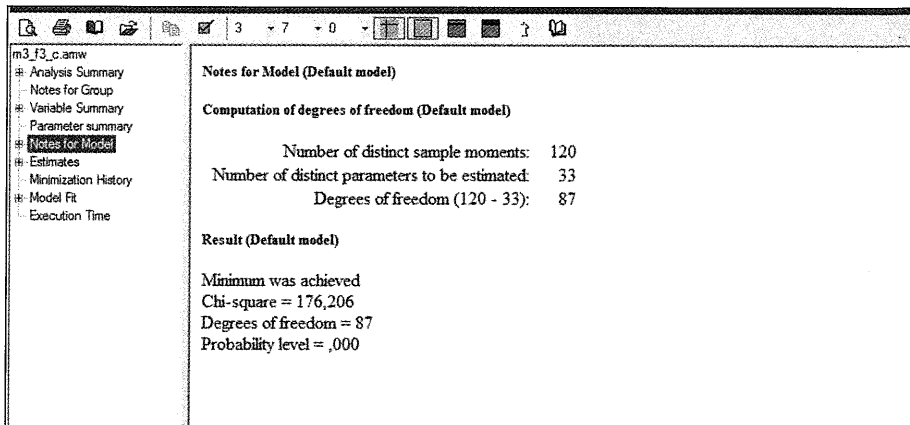
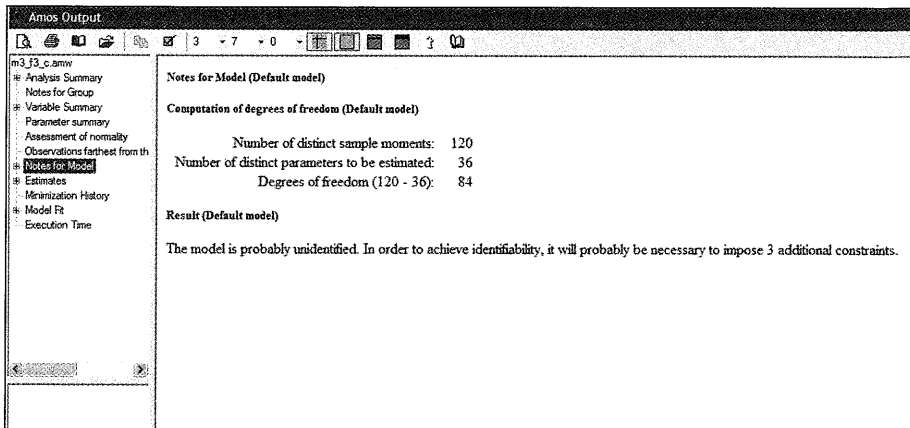
Το μοντέλο που τέθηκε προς εξέταση στηρίχθηκε στην έρευνα των Tschannen-Moran και Woolfolk-Hoy (2001). Με βάση τα αποτελέσματα της έρευνας αυτής τα 15 θέματα της συντομευμένης TSES θεωρήθηκε ότι είναι εκφράσεις τριών λανθανόντων παραγόντων, δηλαδή της Αποτελεσματικότητας των Στρατηγικών Διδασκαλίας, της Αποτελεσματικότητας στη Διαχείριση της Τάξης, και της Αποτελεσματικότητας ως προς την Ενασχόληση των Μαθητών. Όπως φαίνεται και από το Σχήμα 2, οι λανθάνοντες παράγοντες σχετίζονταν μεταξύ τους, ενώ το κάθε ερώτημα συνδεόταν μόνο με τον παράγοντα τον οποίο αξιολογεί, και τα σφάλματα των παρατηρούμενων μεταβλητών ήταν ανεξάρτητα μεταξύ τους.



Σχήμα 2. Καθορισμός του μοντέλου της δομής της συντομευμένης TSES.

Προσδιορισμός του μοντέλου

Αρχική προσπάθεια εκτίμησης του μοντέλου έδειξε ότι οι παράμετροι του μοντέλου δεν ήταν δυνατόν να εκτιμηθούν επειδή δεν υπήρχαν οι αναγκαίοι περιορισμοί. Όπως φαίνεται και από το Σχήμα 3, το πρόγραμμα AMOS ενημερώνει το χρήστη ότι θα πρέπει να εισαχθούν, στην περίπτωση του συγκεκριμένου μοντέλου, τρεις περιορισμοί. Ήταν απαραίτητο, δηλαδή, σε ορισμένες παραμέτρους να δοθεί μία αρχική τιμή. Έτσι λοιπόν οι φορτίσεις των μετρήσιμων μεταβλητών #a5, #b5 και #c4 στους αντίστοι-



Σχήμα 3. Αποτελέσματα εκτίμησης του μοντέλου της συντομευμένης TSES πριν και μετά την εισαγωγή τριών περιορισμών.

χους παράγοντες τέθηκαν στη μονάδα. Οι μεταβλητές αυτές επιλέχθηκαν επειδή σε προκαταρκτικές αναλύσεις είχαν εμφανίσει την υψηλότερη φόρτιση σε σχέση με τις άλλες μεταβλητές του παράγοντα στον οποίο αναφέρονται. Ο περιορισμός των τριών αυτών μεταβλητών στη μονάδα είχε ως αποτέλεσμα: (α) το μοντέλο να μπορεί να προσδιοριστεί με βαθμούς ελευθερίας $120 - 33 = 87$, όπου 120 ο αριθμός των στοιχείων του πίνακα διακύμανσης-συνδιακύμανσης και 33 ο αριθμός των ελεύθερα εκτιμώμενων παραμέτρων (12 φορτίσεις μετρήσιμων μεταβλητών, 15 διακυμάνσεις μετρήσιμων μεταβλητών, 3 συσχετίσεις λανθανόντων παραγόντων και 3 διακυμάνσεις λανθανόντων παραγόντων), και (β) οι λανθάνοντες παράγοντες να λάβουν την κλίμακα μέτρησης των μετρήσιμων αυτών μεταβλητών.

Εκτίμηση του μοντέλου

Στο Σχήμα 4 παρουσιάζονται τα αποτελέσματα εξέτασης της κατανομής των μετρήσιμων μεταβλητών. Τιμές του δείκτη CR μεγαλύτερες του 1.96 είναι ένδειξη απόκλισης από τις τιμές λοξότητας και κύρτωσης της κανονικής κατανομής σε επίπεδο σημαντικότητας .05. Από τα αποτελέσματα γίνεται φανερό ότι η πλειονότητα των μετρήσιμων μεταβλητών αποκλίνουν από την κανονική κατανομή. Επιπλέον, ο δείκτης πολυμεταβλητής κύρτωσης Mardia's (Mardia, 1970) ήταν στατιστικώς σημαντικός (Mardia's coefficient = 56.58, $p < .001$), γεγονός που φανερώνει ότι η προϋπόθεση της πολυμεταβλητής κανονικότητας δεν ισχύει. Για το λόγο αυτό αποφασίστηκε ως μέθοδος εκτίμησης να χρησιμοποιηθεί το Satorra-Bentler διορθωμένο χ^2 που υποστηρίζεται από το πρόγραμμα EQS.

Έλεγχος του μοντέλου

Για τον έλεγχο της προσαρμογής του μοντέλου στα δεδομένα εκτός από το χ^2 εξετάστηκαν οι δείκτες χ^2/df , CFI, SRMR, και RMSEA. Οι τιμές των δεικτών αυτών (που είναι αντιπροσωπευτικοί και των τριών κατηγοριών, δηλαδή απόλυτοι δείκτες, δείκτες σχετικής βελτίωσης, και δείκτες φειδωλότητας) παρουσιάζονται στον Πίνακα 1. Η τιμή του χ^2 ήταν στατιστικώς σημαντική, γεγονός που δείχνει μη ικανοποιητική προσαρμογή. Οι τιμές όμως των υπόλοιπων δεικτών προσαρμογής ήταν εξαιρετικά ικανοποιητικές. Πιο αναλυτικά, η τιμή του χ^2/df ήταν κάτω από 2 (Byrne, 1989. Kline, 1998b), ενώ η τιμή του CFI ήταν μεγαλύτερη του .90 και μάλιστα ικανοποιεί το πιο αυστηρό κριτήριο του .95 που προτείνουν οι Hu και Bentler

Amos Output

m3_f3_c.amw

- # Analysis Summary
- Notes for Group
- # Variable Summary
- Parameter summary
- Assessment of normality**
- Observations farthest from th
- # Notes for Model
- Estimates
- Minimization History
- # Model Fit
- Execution Time

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
b6	1,000	9,000	-,565	-3,889	,416	1,431
c5	2,000	9,000	-,537	-3,696	-,089	-,307
a7	2,000	9,000	-,732	-5,035	,743	2,555
c1	1,000	9,000	-,879	-6,045	1,697	5,836
c2	1,000	9,000	-,592	-4,072	,726	2,499
c3	1,000	9,000	-,816	-5,614	,770	2,648
c4	1,000	9,000	-,329	-2,266	-,247	-,848
b1	3,000	9,000	-,860	-5,920	,890	3,063
b2	1,000	9,000	-,869	-5,979	1,562	5,374
b3	2,000	9,000	-,754	-5,190	,566	1,946
b5	2,000	9,000	-,760	-5,229	,177	,610
a2	1,000	9,000	-1,250	-8,602	2,519	8,664
a3	1,000	9,000	-1,208	-8,314	3,129	10,764
a4	3,000	9,000	-,701	-4,822	,126	,434
a5	3,000	9,000	-,766	-5,270	,633	2,177
Multivariate					56,584	21,113

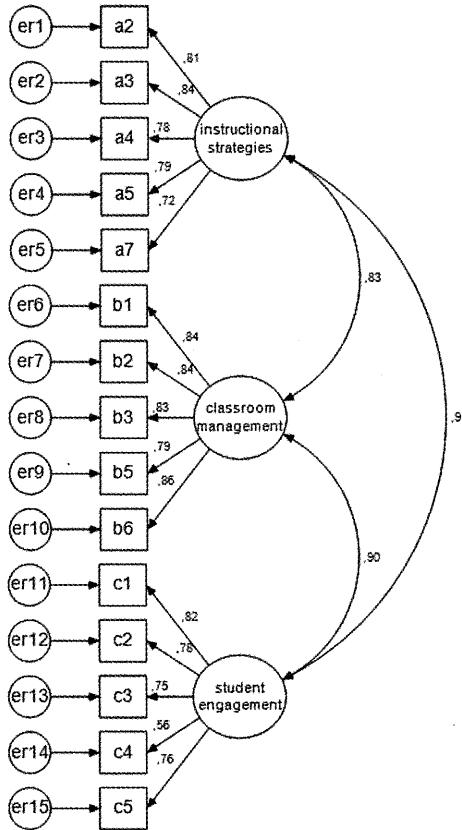
Σχήμα 4. Αποτελέσματα εξέτασης της κατανομής των μετρήσιμων μεταβλητών της συντομευμένης μορφής της TSES.

Πίνακας 1. Αποτελέσματα επιβεβαιωτικής παραγοντικής ανάλυσης στις απαντήσεις της συντομευμένης TSES

Μοντέλο	S-B χ^2	df	χ^2/df	CFI	SRMR	RMSEA	90%ΔΕ RMSEA
M1	269.4*	90	2.99	.867	.048	.084	.072 - .095
M3	145.1*	87	1.67	.957	.034	.049	.034 - .062

Σημείωση: M1 = μοντέλο ενός λανθάνοντα παράγοντα. M3 = μοντέλο τριών λανθάνοντων παραγόντων. S-B χ^2 : Satorra-Bentler scaled χ^2 . ΔΕ: Διάστημα εμπιστοσύνης. * $p < .01$.

(1999). Τέλος, οι τιμές τόσο του SRMR και όσο και του RMSEA ήταν μικρότερες του .05 και .06, αντιστοίχως (Hu & Bentler, 1999). Τα παραπάνω αποτελέσματα συνηγορούν στο ότι το μοντέλο των τριών συσχετιζόμενων παραγόντων περιγράφει αρκετά καλά τα εμπειρικά δεδομένα και επομένως θα πρέπει να γίνει αποδεκτό.



Σχήμα 5. Φορτίσεις των θεμάτων της συντομευμένης TSES στους παράγοντες και συσχετίσεις μεταξύ των τριών λανθανόντων παραγόντων.

Στο Σχήμα 5 παρουσιάζονται οι φορτίσεις των θεμάτων της συντομευμένης TSES σε κάθε παράγοντα καθώς και οι συσχετίσεις μεταξύ των τριών λανθανόντων παραγόντων, όπως προέκυψαν από την επιβεβαιωτική παραγοντική ανάλυση. Επίσης, στο Σχήμα 6 παρουσιάζονται τα αποτελέσματα της εκτίμησης των παραμέτρων του εξεταζόμενου μοντέλου. Στον πίνακα του σχήματος με τίτλο “Regression weights” υπάρχουν οι εκτιμήσεις των μη-τυποποιημένων γραμμικών συντελεστών (Estimate), τα τυπικά τους σφάλματα (SE) και ο δείκτης CR, που παριστάνει το λόγο του μη-τυποποιημένου γραμμικού συντελεστή προς το τυπικό του σφάλμα. Τι-

Αξίζει να σημειωθεί ότι οι συσχετίσεις μεταξύ των παραγόντων ήταν αρκετά υψηλές. Έτσι, κρίθηκε απαραίτητο να εξεταστεί η *διακρίνουσα εγκυρότητα*⁷² της συντομευμένης TSES, δηλαδή ο βαθμός στον οποίο οι τρεις διαστάσεις της πράγματι αποτελούν τρεις διακριτούς λανθάνοντες παράγοντες ή εάν στην ουσία πρόκειται για ένα μονοδιάστατο εργαλείο.

Δύο κριτήρια χρησιμοποιήθηκαν για την εξέταση της διακριτότητας των τριών λανθάνοντων παραγόντων (Anderson & Gerbing, 1988, Evagelinou et al., 2002). Και τα δύο κριτήρια φαίνεται να στηρίζουν τη διακριτότητα των διαστάσεων της συντομευμένης μορφής κλίμακας TSES. Το πρώτο κριτήριο αφορούσε τη σύγκριση του μοντέλου των τριών παραγόντων (M_3) με το μοντέλο ενός παράγοντα (M_1) (μονοδιάστατο μοντέλο). Επειδή το μοντέλο M_1 είναι εμφωλευμένο⁷³ στο M_3 τα δύο μοντέλα μπορεί να συγκριθούν. Η σύγκριση βασίζεται στη διαφορά των τιμών των δύο χ^2 , η οποία ακολουθεί την κατανομή χ^2 με βαθμούς ελευθερίας τη διαφορά των βαθμών ελευθερίας των δύο μοντέλων. Μη στατιστικώς σημαντικό $\Delta\chi^2$ δηλώνει παρόμοια προσαρμογή στα δεδομένα. Από τη σύγκριση⁷⁴ φάνηκε ότι η διαφορά των δύο μοντέλων ήταν στατιστικά σημαντική, γεγονός που δηλώνει ότι το μοντέλο με τη μικρότερη χ^2 τιμή, δηλαδή το M_3 , έχει σαφώς καλύτερη προσαρμογή στα δεδομένα σε σχέση με το μονοδιάστατο μοντέλο, διαφορά του Satorra-Bentler διορθωμένου $\chi^2 = 49.59$, $df = 3$, $p < .001$.

Σύμφωνα με το δεύτερο κριτήριο το τυπικό σφάλμα της συσχέτισης χρησιμοποιείται για την κατασκευή ενός 95%ΔΕ γύρω από τον κάθε συντελεστή συσχέτισης. Η διακριτότητα των παραγόντων υποστηρίζεται όταν το 95%ΔΕ του συντελεστή συσχέτισης δεν περιλαμβάνει τη μονάδα. Τα ΔΕ όπως υπολογίστηκαν ήταν .753 - .909 για το συντελεστή συσχέτισης $r = .831$ (τυπικό σφάλμα .039) μεταξύ της Αποτελεσματικότητας των Στρατηγικών

⁷² Discriminant validity.

⁷³ Ένα μοντέλο, M_A θεωρείται εμφωλευμένο σε ένα άλλο μοντέλο M_B όταν το πρώτο μοντέλο (M_A) μπορεί να προκύψει από την τοποθέτηση ή άρση περιορισμών στο δεύτερο (M_B). Στο συγκεκριμένο παράδειγμα το μοντέλο M_1 θεωρείται εμφωλευμένο στο M_3 επειδή το M_1 μπορεί να προκύψει από το M_3 εάν η σχέση μεταξύ των τριών λανθάνουσών μεταβλητών τοποθετηθεί στη μονάδα, γεγονός που σημαίνει ότι οι τρεις λανθάνουσες μεταβλητές είναι στην ουσία μία.

⁷⁴ Επειδή ως μέθοδος εκτίμησης χρησιμοποιήθηκε το Satorra-Bentler διορθωμένο χ^2 η σύγκριση των δύο εμφωλευμένων μοντέλων πραγματοποιήθηκε με το πρόγραμμα SBDIFF.EXE, το οποίο προσφέρεται δωρεάν στην εξής ιστοσελίδα: <http://www.abdn.ac.uk/~psy086/dept/psychom.htm>. Περισσότερες πληροφορίες για τη μέθοδο σύγκρισης παρέχονται στο άρθρο των Satorra και Bentler (2001).

Διδασκαλίας και της Αποτελεσματικότητας στη Διαχείριση της Τάξης, .857 - .961 για το συντελεστή συσχέτισης $r = .909$ (τυπικό σφάλμα .026) μεταξύ της Αποτελεσματικότητας των Στρατηγικών Διδασκαλίας και της Αποτελεσματικότητας ως προς την Ενασχόληση των Μαθητών, και .844 - .960 για το συντελεστή συσχέτισης $r = .902$ (τυπικό σφάλμα .029) μεταξύ της Αποτελεσματικότητας στη Διαχείριση της Τάξης και της Αποτελεσματικότητας ως προς την Ενασχόληση των Μαθητών. Σε κανένα από τα διαστήματα εμπιστοσύνης δε συμπεριλαμβάνεται η μονάδα, επομένως και το δεύτερο κριτήριο φαίνεται να υποστηρίζει τη διακριτότητα των λανθανόντων παραγόντων.

ΠΑΡΑΡΤΗΜΑ

Ο ερευνητής των κοινωνικών επιστημών που ενδιαφέρεται για περισσότερες πληροφορίες πάνω στη μοντελοποίηση δομικών εξισώσεων εκτός από τα συναφή εγχειρίδια και βιβλία μπορεί να αξιοποιήσει και τις δυνατότητες που προσφέρει το διαδίκτυο. Για το λόγο αυτό δίνονται συγκεκριμένες ιστοσελίδες. Στις ιστοσελίδες αυτές υπάρχουν βίντεο για τη χρήση στατιστικών προγραμμάτων που ειδικεύονται στην μοντελοποίηση δομικών εξισώσεων, λεπτομερή κείμενα καθώς και ταξινόμηση της σχετικής βιβλιογραφίας.

1. <http://www.upa.pdx.edu/IOA/newsom/semrefs.htm>

Μια αρκετά χρήσιμη ιστοσελίδα όπου υπάρχει το σύνολο σχεδόν των βιβλιογραφικών αναφορών, ταξινομημένων ανάλογα με το θέμα που πραγματεύονται (π.χ., εισαγωγικά κείμενα, ιστορία της μοντελοποίησης δομικών εξισώσεων, θέματα μεγέθους δείγματος κ.ά.).

2. <http://amosdevelopment.com/video/index.htm>

Πέντε βίντεο σχετικά με εφαρμογές του προγράμματος AMOS.

3. <http://www2.chass.ncsu.edu/garson/pa765/structur.htm>

Λεπτομερής παρουσίαση της μοντελοποίησης δομικών εξισώσεων.

4. <http://www.utexas.edu/its-archive/rc/tutorials/stat/amos/#overview%20of%20SEM>

Λεπτομερής παρουσίαση της μοντελοποίησης δομικών εξισώσεων και εφαρμογής της με το πρόγραμμα AMOS.

5. <http://www.ats.ucla.edu/stat/seminars/>

Διάφορα βίντεο για την μοντελοποίηση δομικών εξισώσεων με το πρόγραμμα Mplus.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*, 411-423.
- Arbuckle, J. L. (1997). *Amos users' guide version 3.6*. Chicago, IL: Small Waters.
- Arbuckle, J. L. (2006). *Amos 7.0 users' guide*. Chicago, IL: SPSS Inc.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246.
- Bentler, P. M. (2005). *EQS 6 structural equations program manual*. Encino, CA: Multivariate Software.
- 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). Beverly Hills, CA: Sage.
- Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, *24*, 445-455.
- Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teacher and Teaching Education*, *16*, 239-253.
- Byrne, B. M. (1989). *A primer of LISREL: Basic application and programming for confirmatory factor analytic models*. New York: Springer.
- Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications and programming*. Thousand Oaks, CA: Sage.
- Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications and programming* (2nd ed.). Mahwah, NJ: Erlbaum.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456-466.
- 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.
- Dillon, W. R., Kumar, A., & Munali, N. (1987). Offending estimates in covariance structure analysis: Comments on the causes of and solutions to Heywood cases. *Psychological Bulletin*, *101*, 126-135.
- Evaggelinou, Ch., Tsigilis, N., & Papa, A. (2002). Construct validity of the gross motor development test: A cross-validation approach. *Adapted Physical Activity Quarterly*, *19*, 482-494.
- Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling*, *6*, 56-83.
- Garson, D. G. (2007). *Structural equation modelling*. Retrieved May 21, 2008, from the World Wide Web <http://www2.chass.ncsu.edu/garson/pa765/structur.htm>
- Gonzalez, R., & Griffin, D. (2001). Testing parameters in structural equation modeling: Every "one" matters. *Psychological Methods*, *6*, 258-269.
- Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, *31*, 627-643.
- Hu, L., & Bentler, P. M. (1995). *Evaluation of model fit*. In R. H. Hoyle (Ed.), *Structural equation modelling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: Sage.

- 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.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International.
- Kline, R. B. (1998a). Software programs for structural equation modeling: AMOS, EQS and LISREL. *Journal of Psychoeducational Assessment*, 16, 343-364.
- Kline, R. B. (1998b). *Principles and practice of structural equation modeling*. New York: Guilford.
- Lei, P-W., & Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. *Educational Measurement: Issues and Practice*, 26, 33-43.
- MacCallum, R. C. (1986). Specification searches in covariance structure modelling. *Psychological Bulletin*, 100, 107-120.
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modification in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111, 490-504.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.
- Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49, 115-132.
- Muthén, L. K., & Muthén, B. O. (1998-2006). *Mplus User's guide* (4th ed). Los Angeles, CA: Muthén & Muthén.
- Nachtigall, C., Kroehne, U., Funke, F., & Steyer, R. (2003). (Why) Should we use SEM; Pros and cons of structural equation modeling. *Methods of Psychological Reports Online*, 8, 1-22.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399-419). Thousand Oaks, CA: Sage.
- Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Erlbaum.
- Soodak, L. C., & Podell, D. M. (1996). Teacher efficacy: Toward the understanding of a multi-faceted construct. *Teacher and Teaching Education*, 12, 401-411.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173-180.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Understanding multivariate statistics* (4th ed). Neadham Heights, MA: Allyn & Bacon.
- Tschannen-Moran, M., & Woolfolk-Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Τσιγγίλης, Ν. (2005). Το μοντέλο των χαρακτηριστικών της εργασίας και ο ρόλος της αυτο-αποτελεσματικότητας σε καθηγητές φυσικής αγωγής. Αδημοσίευστη διδακτορική διατριβή. Τμήμα Επιστήμης Φυσικής Αγωγής και Αθλητισμού, Τρίκαλα, Πανεπιστήμιο Θεσσαλίας.
- Tsigilis, N., & Gregoriadis, A. (2008). Measuring teacher-child relationships in Greek kindergarten setting: A validity study of the short form of Student-Teacher Relationship Scale. *Early Education and Development*, 19, 816-835.

- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Thousand Oaks, CA: Sage.
- Williams, L. J., & Holahan, P. J. (1994). Parsimony-based fit indices of multiple indicator models: Do they work? *Structural Equation Modeling, 1*, 161-189.

BASIC CONCEPTS AND APPLICATION OF STRUCTURAL EQUATION MODELING IN THE FIELD OF SOCIAL SCIENCES

Nikolaos Tsigilis

University of Thessaly, Trikala, Greece

Abstract: During the last two decades structural equation modeling (SEM) has become the dominant method for testing hypotheses in various disciplines, including social sciences. Its increased popularity is based on its ability to examine complex theoretical models, in which relationships among various variables (either measured or latent) have been determined a priori. The purpose of this paper is to introduce social sciences researchers to the basic concepts underlying structural equation modeling and to describe the procedures for conducting it. Finally, an example from the field of educational psychology is provided.

Key words: Confirmatory factor analysis, Latent variable analysis.

Address: Nikolaos Tsigilis, Department of Physical Education and Sport Sciences, University of Thessaly, Karies, 421 00 Trikala, Greece. Tel.: +30-24310-47062. E-mail: tsigilis@uom.gr