

Η αναγκαιότητα σταθμισμένων ψυχομετρικών εργαλείων στην κλινική πράξη. Εισαγωγή στο πρώτο μέρος του Ειδικού τεύχους με τίτλο «Ψυχομετρική αξιολόγηση γνωστικών λειτουργιών: εφαρμογές στη νευροψυχολογία και την εκπαίδευση»

Η σημασία της κατασκευής και συστηματικής χρήσης σταθμισμένων ψυχομετρικών εργαλείων στην κλινική πράξη είναι αυτονόητη, επαρκώς τεκμηριωμένη και απόλυτα καθιερωμένη διεθνώς, αν και υπολείπεται σημαντικά στη χώρα μας. Με εξαίρεση το MMPI, το WISC-III και, πιθανώς, το Αθηνά Τεστ οι ειδικοί στο χώρο της κλινικής και σχολικής ψυχολογίας και της νευροψυχολογίας δεν είχαν στη διάθεσή τους, μέχρι πρότινος, κατάλληλα σταθμισμένα ψυχομετρικά εργαλεία-απολύτως απαραίτητα, αν και όχι απολύτως επαρκή, όπως θα δούμε παρακάτω, για την έγκυρη και αξιόπιστη εκτίμηση της έκτασης και βαρύτητας γνωσιακών, γλωσσικών, συναισθηματικών και ακαδημαϊκών προβλημάτων. Η έλλειψη εθνικής πολιτικής στο χώρο της εφαρμοσμένης ψυχολογίας και η απουσία χρηματοδότησης εν μέρει μόνο δικαιολογούν την απροθυμία των ειδικών στο χώρο της ψυχομετρίας να αναλάβουν μεγάλης κλίμακας μελέτες στάθμισης τέτοιων εργαλείων. Πόσοι άραγε από τους συμμετέχοντες στο παρόν ειδικό τεύχος, αλλά και τους αναγνώστες αυτού του κειμένου δεν έχουν γίνει αποδέκτες κατηγορηματικών απόψεων για τα ψυχομετρικά εργαλεία που κυμαίνονται από απλώς «άχρηστα» έως και «επικίνδυνα» για την κοινωνική ισότητα και δικαιοσύνη; Πολλοί αυτόκλητοι «ειδικοί» αμφισβητούν, προβάλλοντας πλήθος ιδεολογικών επιχειρημάτων, τόσο την πρόθεση όσο και τη χρησιμότητα των ψυχομετρικών εργαλείων. Καθώς όμως η «άπρακτη» ιδεολογία, ακόμη και όταν φθάνει τα όρια της ιδεοληψίας, είναι πάντοτε πιο ηχηρή και φλύαρη από την πρακτικά θεμελιωμένη άποψη, οι φωνές αυτές τυγχάνουν ευρείας διάδοσης τόσο στο χώρο της επαγγελματικής εκπαίδευσης (ψυχολόγων και εκπαιδευτικών), όσο και στην κοινή γνώμη, ανάμεσα δηλαδή στους ανθρώπους που έχουν τα μέγιστα να ωφεληθούν από την εφαρμογή έγκυρων και αξιόπιστων εργαλείων αξιολόγησης. Η σύντομη αυτή εισαγωγή δεν θα προσπαθήσει να αντικρούσει αυτές τις χλιοειπωμένες απόψεις, υιοθετώντας τη λογική ότι μόνο η δεοντολογικά σωστή και έλλογη χρήση ψυχομετρικών εργαλείων, που πληρούν τις βασικές προϋποθέσεις αξιοπιστίας και εγκυρότητας, θα μπορέσει σε βάθος χρόνου να αποκαταστήσει το επιστημονικό κύρος και τη θέση τους στην κλινική πράξη στη χώρα μας.

Τα ψυχομετρικά εργαλεία, όπως και οι πειραματικές διαδικασίες, αποσκοπούν στην αποτύπωση ψυχολογικών φαινομένων, δηλαδή παρατηρήσιμων πράξεων (συμπεριφορών), όπως η ανάκληση ενός παρελθόντος επεισοδίου και η κατονομασία ενός καταλόγου γραπτών λέξεων που συνθέτουν μια ψυχολογική λειτουργία, όπως η μνήμη και η ανάγνωση, αντίστοιχα. Βασική, λοιπόν, παραδοχή της ψυχομετρίας, όπως συμβαίνει σε όλες τις επιστήμες, είναι ότι τα ψυχολογικά φαινόμενα μπορούν να μετρηθούν με υψηλό βαθμό πιστότητας, καθώς η εκδήλωσή τους δεν είναι τυχαία, και μπορεί να επαναληφθεί με συστηματικό τρόπο στο πλαίσιο ενός έργου. Ακολουθώντας την παράδοση της Γνωσιακής Ψυχολογίας, η ψυχομετρία αποδέχεται ότι κάθε ψυχολογική λειτουργία ρυθμίζεται άμεσα από ένα μηχανισμό, ένα σύνολο δηλαδή από αλληλένδετες διεργασίες, αν και στις περισσότερες κλινικές εφαρμογές δεν απαιτείται λεπτομερής διερεύνηση των επιδόσεων του εξεταζόμενου σε σχέση με επιμέρους συνιστώσες διεργασίες.

Σταθμισμένες δοκιμασίες χρησιμοποιούνται για την ανίχνευση, διάγνωση και ενδελεχή διερεύνηση του είδους και της έκτασης των δυσκολιών και ελλειμμάτων στην κλινική, νευροψυχολογική και εκπαιδευτική/σχολική πρακτική. Συμβατικά διακρίνονται σε ανιχνευτικά και διαγνωστικά εργαλεία καθώς και μέσα ποιοτικής και ποσοτικής εκτίμησης επιμέρους δυσκολιών και ειδικών προβλημάτων που χρήζουν θεραπευτικής αντιμετώπισης. Απώτερος σκοπός των ανιχνευτικών εργαλείων είναι ο εντοπισμός ατόμων που χρήζουν περαιτέρω αξιολόγησης ώστε να διαπιστωθεί με μεγαλύτερη βεβαιότητα αν πληρούν τα κριτήρια μίας ή περισσότερων διαγνωστικών κατηγοριών και αν έχουν άμεση ανάγκη παρέμβασης. Στόχος της τελευταίας είναι η βελτίωση ελλειμματικών λειτουργιών (π.χ. στην περίπτωση ενός ασθενούς με αφασία), η καθυστέρηση της επιδείνωσης συγκεκριμένων γνωσιακών ικανοτήτων (π.χ. σε έναν ασθενή με Ήπια Γνωσιακή Έκπτωση) ή η πρόληψη της εκδήλωσης σοβαρότερων ακαδημαϊκών και ψυχοκοινωνικών προβλημάτων (πρώιμη εκπαιδευτική ανίχνευση). Οι ανιχνευτικές δοκιμασίες και κλίμακες περιέχουν μικρό αριθμό υποκλιμάκων, ερωτήσεων ή προβλημάτων καθιστώντας δυνατή την εφαρμογή τους σε μεγάλο αριθμό ατόμων από λιγότερο ειδικευμένους εξεταστές (π.χ., εκπαιδευτικούς), καθώς και σε βεβαρυσμένους ασθενείς (π.χ., στην οξεία ή υποξεία φάση μιας νόσου). Εκτός από τη συντομία τους, οι δοκιμασίες αυτές στοχεύουν στον εντοπισμό όσο το δυνατόν περισσότερων ατόμων με πραγματικό πρόβλημα, διατηρώντας ταυτόχρονα ένα σχετικά υψηλότερο αλλά αποδεκτό ποσοστό «Ψευδώς θετικών» αποτελεσμάτων.

Απώτερος σκοπός των διαγνωστικών εργαλείων είναι, αντίθετα, ο προσδιορισμός της διαγνωστικής κατηγορίας στην οποία πιθανώς μπορεί να ενταχθεί η κλινική εικόνα του εξεταζόμενου με παράλληλο αποκλεισμό άλλων εναλλακτικών διαγνώσεων. Η διαφορική διάγνωση είναι απαραίτητη καθότι ένα σύμπτωμα ή σημείο είναι δυνατόν να εμφανίζεται σε περισσότερες από μία νευρολογικές ή ψυχιατρικές νόσους. Πολλά διαγνωστικά εργαλεία έχουν σχεδιαστεί για να αποδίδουν διχοτομικού τύπου πληροφορίες, δηλαδή εάν μπορεί να τεθεί ή όχι μια συγκεκριμένη διάγνωση. Διαγνωστικά εργαλεία που έχουν σχεδιαστεί με την παραπάνω λογική έχουν ένα σημαντικό μειονέκτημα, καθώς τελικά καταλήγουν να παράσχουν στον κλινικό λιγότερες πληροφορίες από αυτές που συλλέγησαν κατά τη διαδικασία της αξιολόγησης. Αντίθετα, ο σχεδιασμός της κατάλληλης παρέμβασης απαιτεί μεγαλύτερο εύρος και πλούτο πληροφοριών σχετικά τόσο με τις ιδιαίτερες αδυναμίες όσο και με τα «δυνατά» σημεία του ατόμου στα οποία θα μπορούσε να στηριχθεί η παρέμβαση.

Οποιοσδήποτε και αν είναι πάντως ο απώτερος σκοπός για τον οποίον έχει σχεδιαστεί ένα ψυχομετρικό εργαλείο, η ψυχοδιαγνωστική διαδικασία επιχειρεί να ανιχνεύσει την ύπαρξη πιθανών ελλειμμάτων σε μία ή περισσότερες γνωσιακές ή γλωσσικές λειτουργίες ή ακαδημαϊκές δεξιότητες. Τις περισσότερες φορές ένα έλλειμμα αναφέρεται στην ανεπαρκή έκφραση μίας ή περισσότερων επιμέρους υπολειτουργιών μιας βασικής λειτουργίας, λόγου χάρη στη δευτερογενή μνήμη λεκτικών επεισοδίων. Η παρουσία ενός συγκεκριμένου ελλείμματος τεκμαίρεται από χαμηλή επίδοση του εξεταζόμενου σε ένα ή περισσότερα τεστ, η εκτέλεση των οποίων θεωρητικά απαιτεί την εμπλοκή της εν λόγω λειτουργίας. Απαραίτητη προϋπόθεση για την ανίχνευση ενός λειτουργικού ελλείμματος είναι η σύγκριση των επιδόσεων του εξεταζόμενου με αντίστοιχες τυπικές τιμές. Ιδανικά οι τιμές αυτές θα προέρχονταν από τον ίδιο τον εξεταζόμενο πριν από την εμφάνιση των πρώτων σημείων της νόσου για την οποία επισκέπτεται τον ειδικό (αυτό βέβαια δεν ισχύει για αναπτυξιακές διαταραχές). Ελλείψει τέτοιων «βασικών» δεδομένων, ο κλινικός πρέπει να συμβιβαστεί με τη σύγκριση της επίδοσης του εξεταζόμενου με το μέσο όρο της επίδοσης μιας ομάδας ατόμων που θεωρείται αντιπροσωπευτική του πληθυσμού από τον οποίο εκείνος προέρχεται. Σε αυτό το σημείο ας μας επιτραπεί μια σύντομη παρέκβαση για να αναφερθούμε εν παρόδω σε μερικές από τις πιο σημαντικές διαδικασίες στην πορεία συλλογής και επεξεργασίας των δεδομένων για τη στάθμιση ενός ψυχομετρικού εργαλείου¹. Από την επάρκεια αυτής της διαδικασίας θα εξαρτηθεί σε μεγάλο βαθμό η δια-

γνωστική αξία ενός τεστ, η δυνατότητα δηλαδή ενός τεστ να εντοπίζει την ύπαρξη πραγματικών ελλειμμάτων.

Ενώ, λοιπόν, ο βαθμός αντιπροσωπευτικότητας της «ομάδας σύγκρισης» του εξεταζόμενου σε σχέση με τον πληθυσμό είναι θέμα κατάλληλης δειγματοληψίας της έρευνας στάθμισης του ψυχομετρικού εργαλείου, τα ιδιαίτερα χαρακτηριστικά της ομάδας εξαρτώνται σε μεγάλο βαθμό από τη φύση της λειτουργίας που εξετάζεται, καθώς και από το πεδίο εφαρμογής του εργαλείου. Για παράδειγμα, ένα τεστ δευτερογενούς λεκτικής μνήμης που στοχεύει στην ανίχνευση προδρομικής αμνησίας μπορεί να χρησιμοποιηθεί στην αξιολόγηση πληθώρας νευρολογικών διαταραχών με μεγάλο εύρος ηλικίας έναρξης: από την εφηβεία (όπως η σκλήρυνση κατά πλάκας) μέχρι την τρίτη ηλικία (όπως οι διάφορες νευροεκφυλιστικές νόσοι που προκαλούν άνοια). Αφού προσδιοριστεί το ηλικιακό εύρος των ατόμων για τους οποίους προορίζεται, το επόμενο βήμα στο σχεδιασμό της έρευνας στάθμισης είναι η διερεύνηση (με βάση προηγούμενες έρευνες αλλά και καλά στοχευμένα πιλοτικά δεδομένα) εκείνων των χαρακτηριστικών του ατόμου τα οποία επηρεάζουν την επίδοσή του στο τεστ. Δύο είναι τα δημογραφικά χαρακτηριστικά που σχετίζονται με την επίδοση στις περισσότερες ψυχομετρικές δοκιμασίες: η ηλικία και το μορφωτικό επίπεδο. Ο βαθμός συσχέτισης καθεμιάς από τις παραπάνω μεταβλητές με την επίδοση σε μια συγκεκριμένη δοκιμασία εξαρτάται, όπως αναφέρθηκε και προηγουμένως, από τη φύση της εξεταζόμενης λειτουργίας (ή υπολειτουργίας), καθώς και από το ηλικιακό εύρος εφαρμογής της. Έτσι η επίδραση της ηλικίας μπορεί να είναι αμελητέα για μια δοκιμασία που σταθμίζεται σε πληθυσμό ηλικίας 20 έως 35 ετών. Αντίθετα στο παράδειγμα που αναφέρθηκε παραπάνω η επίδραση της ηλικίας αναμένεται να είναι σημαντική εντός του ηλικιακού εύρους των 17 έως 80 ετών. Στη συνέχεια, και κυρίως για λόγους οικονομίας ο ερευνητής καλείται να προσδιορίσει τις ηλικιακές υποομάδες, αντιπροσωπευτικό δείγμα των οποίων θα πρέπει να αναζητήσει. Εφόσον, παραδείγματος χάριν, διαπιστώσει ότι η επίδραση της ηλικίας είναι μηδαμινή μεταξύ ατόμων 17 έως 35 ετών αλλά αισθητά μεγαλύτερη εντός του εύρους των 17 έως 40 ετών, τότε θα μπορούσε να καθορίσει ως νεότερη ομάδα του δείγματος στάθμισης την ομάδα των ατόμων ηλικίας 17 έως 35 ετών. Μεταξύ ενηλίκων, η επίδραση της ηλικίας είναι συνήθως μεγαλύτερη για δοκιμασίες που εξετάζουν, εν μέρει ή κατά βάση, ταχύτητα επεξεργασίας ερεθισμάτων και μάθηση καινούργιων πληροφοριών ή επεισοδίων, ο δε βαθμός επίδρασης της ηλικίας αυξάνεται με την ίδια την ηλικία. Σημαντικό επόμενο βήμα είναι να προσδιορίσει το βαθμό της σχέσης μεταξύ μορφωτικού επιπέδου και επίδοσης στην υπό στάθμιση δοκιμασία. Σημειώτεον ότι η μεταβλητή «μορφωτικό επίπεδο» (όπως υποδηλώνεται από τα έτη επίσημης εκπαίδευσης του ατόμου) δεν χρησιμοποιείται παρά ως ένας εξαιρετικά αδρός δείκτης των γνώσεων και εμπειριών του ατόμου καθώς και του γενικού νοητικού του επιπέδου στο βαθμό που η νοημοσύνη σχετίζεται με τη σχολική επιτυχία και την πιθανότητα προόδου στις βαθμίδες της επίσημης εκπαίδευσης. Η επίδραση του μορφωτικού επιπέδου είναι συνήθως μεγαλύτερη για δοκιμασίες που απαιτούν υψηλό επίπεδο παγιωμένης γνώσης (π.χ. δοκιμασίες λεξιλογίου), δεν είναι όμως καθόλου αμελητέα, ακόμη και σε δοκιμασίες οι οποίες εμμέσως απαιτούν γλωσσική επεξεργασία, όπως οι δοκιμασίες λεκτικής μνήμης.

Το επόμενο στάδιο της διαδικασίας στάθμισης είναι η εξέταση της κατανομής των ατομικών βαθμολογιών ώστε να διαπιστωθεί αν είναι κανονική ή όχι. Αυτό θα εξαρτηθεί εν μέρει από την αρχική επιλογή των επιμέρους προβλημάτων ή ερωτήσεων ενός τεστ και εν μέρει από τη φύση των δεδο-

1. Σημαντικά βήματα της διαδικασίας στάθμισης, όπως η εξέταση της εσωτερικής συνέπειας των προβλημάτων, ερωτήσεων ή προτάσεων ενός εργαλείου, ο προσδιορισμός του βαθμού επαναληψιμότητας των αποτελεσμάτων ενός τεστ, η εσωτερική δομή μιας συστοιχίας δοκιμασιών που συναποτελούν ένα σύνθετο ψυχομετρικό εργαλείο κ.ά., δεν αναφέρονται εδώ λόγω έλλειψης χώρου.

μένων που αυτό παρέχει. Έτσι, τα περισσότερα ανιχνευτικά εργαλεία περιέχουν μεγάλο αριθμό «εύκολων» για το γενικό πληθυσμό ερωτήσεων και επομένως η κατανομή των ατομικών βαθμολογιών σε αυτά κλίνει προς τα δεξιά, σημειώνεται δηλαδή μεγαλύτερη συχνότητα υψηλών βαθμολογιών από τη συχνότητα των χαμηλών βαθμολογιών. Χαρακτηριστικό παράδειγμα τέτοιου εργαλείου είναι η Δοκιμασία Αφασίας της Βοστώνης, η στάθμιση της οποίας περιγράφεται σε ένα από τα άρθρα του παρόντος τεύχους. Επίσης, δοκιμασίες στις οποίες η επίδοση μετριέται με το χρόνο αντίδρασης (π.χ. η Δοκιμασία Διαδρομών) συχνά εμφανίζουν μη κανονική κατανομή. Μια λύση σε αυτό το πρόβλημα είναι ο μετασχηματισμός των δεδομένων από τις «ωμές» τιμές σε τιμές με άλλες ιδιότητες (π.χ. υψωμένες στη «ν» δύναμη) ή η απαλοιφή των ακραίων τιμών (π.χ. η χρήση του μέσου όρου που περιλαμβάνει το 50% των δεδομένων μόνο, δηλ. του *trimmed mean*). Όλοι αυτοί οι μετασχηματισμοί έχουν σαν αποτέλεσμα την τελική παραγωγή κατανομών που εμφανίζουν ιδιότητες παρόμοιες με αυτές της κανονικής κατανομής. Όταν αντίθετα η κατανομή των ατομικών βαθμολογιών του δείγματος στάθμισης μπορεί να θεωρηθεί ως κανονική, η επίδοση του εξεταζόμενου εκφράζεται συχνότερα με τον αριθμό των τυπικών αποκλίσεων πάνω ή κάτω από το μέσο όρο της κατανομής των τυπικών βαθμολογιών, δηλαδή με τη μορφή τυπικού βαθμού. Ο προσδιορισμός κριτηρίων, ώστε να χαρακτηριστεί μια τυπική βαθμολογία ως ελλειμματική, βασίζεται σε μεγάλο βαθμό σε κλινικές έρευνες που στοχεύουν στην εκτίμηση της σχέσης μεταξύ της βαθμολογίας σε ένα τεστ και της βαρύτητας λειτουργικών ελλειμμάτων του ασθενούς είτε συγχρονικά είτε διαχρονικά. Λόγου χάρη, τυπικός βαθμός $z = -1,0$ σε μια δοκιμασία μνήμης μπορεί να θεωρηθεί ενδεικτικός μνημονικού ελλείμματος μόνο εφόσον κλινικές έρευνες έχουν καταδείξει ότι άτομα με παρόμοια επιδόσεις (α) εμφανίζουν ταυτόχρονα σημαντικά προβλήματα καθημερινής λειτουργικότητας σε τομείς της καθημερινότητας που απαιτούν μνήμη, ή/και (β) είναι πολύ πιθανό να εμφανίσουν ταχύτερη επιδείνωση των μνημονικών λειτουργιών τους από το γενικό πληθυσμό στο άμεσο μέλλον, παρουσιάζοντας ταυτόχρονα δυσκολίες στην καθημερινή τους λειτουργικότητα. Στις περισσότερες δοκιμασίες, πάντως, επίδοση χαμηλότερη από 1,5 τυπικές αποκλίσεις κάτω του μέσου όρου της κατανομής των επιδόσεων της κατάλληλης ομάδας σύγκρισης του γενικού πληθυσμού θεωρούνται ενδεικτικές κλινικά σημαντικού ελλείμματος. Η επίδοση αυτή αντιστοιχεί στο κατώτερο 7ο εκατοστημόριο της αντίστοιχης κατανομής. Σημειώτεον πάντως ότι για κάθε τεστ που είτε κατασκευάζεται εκ νέου είτε προσαρμόζεται από ξενόγλωσση έκδοση και σταθμίζεται στην Ελλάδα, τα όρια κλινικής σημαντικότητας ατομικών επιδόσεων καλό είναι να προσδιορίζονται εκ νέου.

Η βαθμολογία σε ένα τεστ δεν είναι παρά ένα αφηρημένο στοιχείο στην προσπάθεια εκτίμησης των πραγματικών ικανοτήτων και αδυναμιών του ατόμου. Ακόμη και αν έχουν τηρηθεί όλες οι προϋποθέσεις μιας καλής στάθμισης που αναφέρθηκαν παραπάνω (καθώς και αυτές που παραλήφθηκαν) η βαθμολογία σε ένα και μόνο τεστ παρέχει ελάχιστες πληροφορίες στην ψυχοδιαγνωστική διαδικασία. Μερικά από τα προβλήματα που θα πρέπει να λαμβάνει υπόψη του ο κλινικός κατά τη χρήση των τεστ είναι τα εξής:

1. Η εγκυρότητα περιεχομένου ενός τεστ δεν είναι ποτέ τέλεια. Ένα και μόνο τεστ είναι εκ των πραγμάτων αδύνατο να καλύψει επαρκώς μια ψυχολογική λειτουργία ή υπολειτουργία. Επομένως η αξιολόγηση θα πρέπει να βασιστεί στην επίδοση σε περισσότερα από ένα τεστ, η εκτέλεση των οποίων απαιτεί την υπό εξέταση λειτουργία. Καθώς όμως είναι πρακτικά αδύνατο να συσταθεί και να χορηγηθεί μια πλήρως αντιπροσωπευτική συστοιχία όλων των διαστάσεων μιας ψυχολογικής λειτουργίας, είναι αναμενόμενο ότι κάποιες από τις επιμέρους αυτές διαστάσεις δεν θα αξιολογηθούν επαρκώς. Ένα σχετικό ζήτημα αφορά την οικολογική εγκυρότητα ενός εργαλείου, το βαθμό δηλαδή στον οποίο βοηθά στην εκτίμηση-πρόβλεψη της λειτουργικότητας του εξεταζόμενου στο «φυσικό» του περιβάλλον. Συνήθως ο περιορισμός αυτός της αμιγώς ψυχομετρικής αξιολόγησης δεν μειώνει την εγκυρότητα της

κλινικής διάγνωσης. Μπορεί όμως να μειώσει την ακρίβεια της εκτίμησης όλων εκείνων των δυσκολιών του ατόμου που χρήζουν παρεμβατικής αντιμετώπισης αποτυγχάνοντας να «καθοδηγήσουν» με ακρίβεια το σχεδιασμό της παρέμβασης. Ποιοτικές παρατηρήσεις του κλινικού κατά τη χορήγηση των ψυχομετρικών εργαλείων μπορούν να παράσχουν χρήσιμες πληροφορίες προς αυτή την κατεύθυνση.

2. Η οικολογική εγκυρότητα ενός τεστ, το οποίο αποτελεί μετάφραση/προσαρμογή ξενόγλωσσας έκδοσης, εξαρτάται σε μεγάλο βαθμό από την καταλληλότητά του για το πολιτισμικό περιβάλλον και εκπαιδευτικό υπόβαθρο των ατόμων στα οποία απευθύνεται. Αν και η εξαρχής κατασκευή ενός εργαλείου είναι προτιμότερη, εκτενής και πολυσταδιακή πιλοτική διερεύνηση της καταλληλότητας των ερωτήσεων/προβλημάτων και ερεθισμάτων ενός τεστ είναι απαραίτητη για την εξασφάλιση της οικολογικής εγκυρότητας της προσαρμοσμένης στα Ελληνικά εκδόχης του. Ενδεικτικά αυτής της στρατηγικής είναι τα τεστ μνήμης και γλωσσικής ικανότητας που παρουσιάζονται στο παρόν τεύχος από τους Οικονόμου και Παπαθανασίου και τους συνεργάτες τους.

3. Ο τυπικός βαθμός του εξεταζόμενου, ο οποίος προέρχεται από τη σύγκριση της επίδοσής του με το μέσο όρο της καταλληλότερης υποομάδας του δείγματος στάθμισης δεν αντιστοιχεί με απόλυτη ακρίβεια με τη διαφορά μεταξύ της τωρινής και της προνοσηρής επίδοσής του. Αυτό γιατί όπως αναφέρθηκε και προηγουμένως η επιλογή της υποομάδας σύγκρισης βασίζεται σε εξαιρετικά αδρούς δείκτες των στοιχείων που χαρακτηρίζουν τον εξεταζόμενο (δηλαδή την ηλικία και το μορφωτικό επίπεδο). Άλλα εγγενή και κοινωνικά χαρακτηριστικά δεν είναι πρακτικά δυνατό να αξιολογηθούν σε αυτή τη διαδικασία. Δυστυχώς, τα δεδομένα στάθμισης σπανίως λαμβάνουν υπόψη τους το μορφωτικό επίπεδο και το πολιτισμικό περιβάλλον από το οποίο προέρχεται ο εξεταζόμενος, αν και τελευταία, τόσο στη χώρα μας όσο και διεθνώς, σημειώνονται αξιόλογες προσπάθειες προς αυτή την κατεύθυνση (Lucas, Ivnik, Smith et al., 2005. Marcopoulos, McLain, & Giuliano, 1997. Ostrosky-Solis, Ardila, Roselli et al., 1998. Welch, Doineau, Johnson, & King, 1996).

Συμπερασματικά, η χρήση ψυχομετρικών δοκιμασιών θεωρείται πλέον απολύτως απαραίτητη στο πλαίσιο της διαγνωστικής διαδικασίας, ιδιαίτερα για την εκτίμηση γλωσσικών, γνωσιακών και ακαδημαϊκών προβλημάτων, και λιγότερο για την εκτίμηση συναισθηματικών και γενικότερα εσωτερικευμένων κλινικών καταστάσεων. Με τη χρήση τους περιορίζεται σημαντικά ο κίνδυνος υποκειμενικών κρίσεων που είναι αναπόφευκτος στη διάρκεια μιας πρόσωπο-με-πρόσωπο αλληλεπίδρασης εξεταζόμενου και εξεταστή. Παρέχουν επίσης τη δυνατότητα τυποποιημένης εξέτασης, με τα ίδια ερεθίσματα και διαδικασίες χορήγησης και βαθμολόγησης, επιτρέποντας τη σύγκριση των επιδόσεων του εξεταζόμενου τόσο με τις επιδόσεις άλλων ατόμων του γενικού πληθυσμού, όσο και με τις επιδόσεις του ίδιου του ατόμου διαχρονικά ή μετά από ειδική παρέμβαση. Παρ' όλες τις πρακτικές αδυναμίες τους, οι οποίες σε μεγάλο βαθμό μπορούν να ξεπεραστούν εφαρμόζοντας πιστά τις διαδικασίες τις οποίες έχει υποδείξει η μακρόχρονη εμπειρία στο χώρο της Ψυχομετρίας, η χρήση των ψυχομετρικών εργαλείων θα πρέπει να αποτελεί αναπόσπαστο κομμάτι της κλινικής αξιολόγησης, καθώς η προσεκτική και έλλογη εφαρμογή τους τελικά αποβαίνει προς όφελος του ίδιου του ατόμου που αναζητεί τη βοήθεια του ειδικού.

Η ιδέα για τη σύνθεση ενός ειδικού τεύχους με θέμα τη στάθμιση ψυχομετρικών εργαλείων για την αξιολόγηση γνωσιακών, γλωσσικών και ακαδημαϊκών δεξιοτήτων στα Ελληνικά προέκυψε με βάση το σκεπτικό που αναπτύχθηκε παραπάνω. Το ενδιαφέρον από συναδέλφους για συμβολή σε αυτή την προσπάθεια ήταν μεγαλύτερο από το αναμενόμενο, με αποτέλεσμα τα άρθρα που έγιναν τελικώς αποδεκτά μετά από τυφλή κρίση να δημοσιευτούν σε δύο διαδοχικά τεύχη του περιοδικού *Ψυχολογία*. Το παρόν τεύχος περιλαμβάνει εργασίες που αφορούν την κατασκευή και στάθμιση ψυχομετρικών εργαλείων αξιολόγησης ακαδημαϊκών δεξιοτήτων παιδιών και εφήβων, ενώ άρθρα που περιγράφουν τη στάθμιση και εφαρμογή εργαλείων για την αξιολόγηση γνωσιακών και γλωσσικών δε-

ξιοτήτων σε ενήλικες δημοσιεύονται στο επόμενο τεύχος του περιοδικού. Παρακάτω παρουσιάζονται αναλυτικά οι συγκεκριμένες εργασίες σε αλφαβητική σειρά.

Ο Ι. Δημάκος πραγματοεύθηκε την ψυχομετρική αξιολόγηση μιας κλίμακας γραπτής γλωσσικής έκφρασης. Να σημειωθεί πως δεν υπήρχε ανάλογη κλίμακα στην ελληνική πραγματικότητα για τη μέτρηση του εν λόγω χαρακτηριστικού. Χρησιμοποιώντας αξιολογήσεις εγκυρότητας προσώπου και συγκλίνουσας εγκυρότητας από δύο διαφορετικές πηγές πληροφοριών (μαθητή και εκπαιδευτικό), ο συγγραφέας συμπέρανε ότι η προτεινόμενη κλίμακα θα μπορούσε να αποτελέσει αφητηρία για τη μέτρηση του συγκεκριμένου χαρακτηριστικού στον ελληνικό πληθυσμό αφού και οι δείκτες εσωτερικής συνοχής ήταν ικανοποιητικοί.

Επίσης, οι Παντελιάδου, Αντωνίου και Σιδερίδης παρουσίασαν τη δημιουργία μιας κλίμακας για την αξιολόγηση χαρακτηριστικών μαθησιακών δυσκολιών από εκπαιδευτικούς. Δεδομένου ότι οι εκπαιδευτικοί αποτελούν τα τελευταία χρόνια αρκετά έγκυρες πηγές πληροφόρησης εκπαιδευτικών χαρακτηριστικών των μαθητών, η κλίμακα αυτή βασίζεται σε αξιολογήσεις μόνο από αυτούς. Η κλίμακα αυτή σταθμιστήκε στον ελληνικό χώρο σε ένα μεγάλο δείγμα εκπαιδευτικών και των αντιστοιχών μαθητών τους από διαφορετικά μέρη της χώρας. Η αξιοπιστία αξιολογήθηκε με επαναληπτικές μετρήσεις και η εσωτερική συνοχή της κλίμακας με τη μέθοδο άλφα του Cronbach. Επίσης, διαφορετικά είδη εγκυρότητας αξιολογήθηκαν με το μη παραμετρικό μοντέλο Rasch. Ενδεικτικά, η παραγωγική εγκυρότητα και η σωστή κλιμάκωση της κλειδας απαντήσεων επιβεβαιώθηκαν με αυτή τη μέθοδο. Οι ερευνητές προχώρησαν στη δημιουργία προβλεπτικών εξισώσεων, έτσι ώστε να μπορούν να μετασχηματιστούν οι «ωμές» απαντήσεις των μαθητών (raw scores) σε πιθανότητες να διαγνωσθούν ως παιδιά με μαθησιακές δυσκολίες.

Οι Πρωτόπαπας και Σκαλούμπκας αξιολόγησαν την αναγνωστική ευχέρεια με σκοπό τον εντοπισμό των μαθησιακών δυσκολιών. Χρησιμοποιώντας μια σειρά από αναλύσεις παραγόντων, ανέδειξαν την ύπαρξη συστοιχιών που όριζαν εκ νέου το χαρακτηριστικό της ευχέρειας διευρύνοντας την εννοιολογική εγκυρότητα του χαρακτηριστικού. Τα αποτελέσματα αυτά ανέδειξαν το γεγονός ότι οι εννοιολογικές δομές χαρακτηριστικών δεν θα πρέπει να θεωρούνται δεδομένες, αλλά ότι μπορούν να διευρυνθούν και να οριστούν από παράγοντες που αποτελούν συνοδά χαρακτηριστικά (π.χ. χαρακτηριστικά ευφυΐας). Έτσι οι ερευνητές εξήγαγαν συμπεράσματα με βάση διαφορετικά κριτήρια αναγνωστικής ευχέρειας.

Οι Σιδερίδης, Μουζάκη, Πρωτόπαπας και Σίμος παρουσίασαν τις ψυχομετρικές ιδιότητες μιας ορθογραφικής δοκιμασίας για μαθητές του δημοτικού. Το ενδιαφέρον των ερευνητών εστιάστηκε σε ζητήματα αξιοπιστίας και εγκυρότητας οι οποίες αξιολογήθηκαν ως επαρκείς. Για παράδειγμα η ανάλυση των σφαλμάτων (principal components of residuals) με το μοντέλο Rasch επιβεβαίωσε τη μονοδιάστατη δομή της ορθογραφικής δοκιμασίας. Στη συνέχεια οι ερευνητές ασχολήθηκαν με την πιθανή προκατάληψη κατά γλωσσική ομάδα για συγκεκριμένες λέξεις, δεδομένης της υπόθεσης ότι πιθανώς οι γηγενείς μαθητές έχουν καλύτερη ορθογραφική επίδοση σε κάποιες λέξεις. Τα αποτελέσματα από το μοντέλο Rasch συνέκλιναν στο συμπέρασμα ότι υπήρχαν πολύ χαμηλά επίπεδα προκατάληψης σε συγκεκριμένες λέξεις με ιδιαίτερα φωνολογικά χαρακτηριστικά.

Βιβλιογραφία

- Lucas, J. A., Ivnik, R. J., Smith, G. E., Ferman, T. J., Willis, F. B., Petersen, R. C., & Graff-Radford, N. R. (2005). Mayo's older African Americans normative studies: norms for Boston Naming Test, controlled oral word association, category fluency, animal naming, Token Test, WRAT-3 reading, trail making test, Stroop test, and Judgment of Line Orientation. *The Clinical Neuropsychologist, 19*, 243-269.
- Marcopulos, B. A., McLain, C. A., & Giuliano, A. J. (1997). Cognitive impairment or inadequate norms? A study of healthy, rural, older adults with limited education. *The Clinical Neuropsychologist, 11*, 111-131.
- Ostrosky-Solis, F., Ardila, A., Roselli, M., Lope-Arango, G., & Uriel-Mendoza, V. (1998). Neuropsychological test performance in illiterate subjects. *Archives of Clinical Neuropsychology, 13*, 645-660.
- Welch, L. W., Doineau, D., Johnson, S., & King, D. (1996). Educational and gender normative data for the Boston Naming Test in a group of older adults. *Brain and Language, 53*, 260-266.