

Psychology: the Journal of the Hellenic Psychological Society

Vol 15, No 4 (2008)



Measuring visuospatial ability: development and psychometric properties of the Spatial Perception Test

Ιωάννης Τσαούσης

doi: [10.12681/psy_hps.23846](https://doi.org/10.12681/psy_hps.23846)

Copyright © 2020, Ιωάννης Τσαούσης



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0](https://creativecommons.org/licenses/by-sa/4.0/).

To cite this article:

Τσαούσης Ι. (2020). Measuring visuospatial ability: development and psychometric properties of the Spatial Perception Test. *Psychology: The Journal of the Hellenic Psychological Society*, 15(4), 411–431.
https://doi.org/10.12681/psy_hps.23846

Μετρώντας τη χωρο-αντιληπτική ικανότητα: Η ανάπτυξη και τα ψυχομετρικά χαρακτηριστικά του Τεστ Αντίληψης Χώρου (TAX)

ΙΩΑΝΝΗΣ ΤΣΑΟΥΣΗΣ¹

ΠΕΡΙΛΗΨΗ

Η εργασία παρουσιάζει την ανάπτυξη, καθώς και την ψυχομετρική αξιολόγηση ενός νέου εργαλείου μέτρησης της χωρο-αντιληπτικής ικανότητας, το οποίο βασίζεται στο θεωρητικό μοντέλο που πρότείνει ο Lohman (1979). Το Τεστ Αντίληψης Χώρου (TAX) αποτελείται από 36 στοιχεία τα οποία μετρούν τρεις διαφορετικές χωρο-αντιληπτικές δεξιότητες: (α) νοητικές περιστροφές, (β) χωρικός προσανατολισμός και (γ) σχηματισμός νοερων εικονων. Τέσσερις διαφορετικές ομάδες συμμετεχόντων (N=694) πήραν μέρος στα διαφορετικά στάδια ανάπτυξης του TAX. Τα στοιχεία της τελικής έκδοσης του τεστ επιλέχθηκαν με βάση τη Θεωρία του Υπολογισμού των Ικανοτήτων (Item Response Theory), και πιο συγκεκριμένα με το μοντέλο μιας παραμέτρου ή με το μοντέλο Rasch. Τα αποτελέσματα έδειξαν ότι τα στοιχεία που επιλέχθηκαν για την τελική έκδοση του τεστ ικανοποιούσαν τις προϋποθέσεις του μοντέλου. Για την αξιολόγηση της αξιοπιστίας του TAX, χρησιμοποιήθηκαν τόσο ο δείκτης Kuder-Richardson (0.85) όσο και η αξιοπιστία επαναληπτικής χρήσης (0.84). Ως προς τη συγχρονική του εγκυρότητα, το TAX παρουσίασε μέτριες συσχετίσεις με δύο τεστ που μετρούν το ίδιο χαρακτηριστικό από τις συστοιχίες Γενικών Ικανοτήτων (από 0.52 έως 0.65) και Διαφορικής Ικανότητας (από 0.36 έως 0.56). Ως προς τη συγκλίνουσα εγκυρότητά του, το TAX παρουσίασε μέτρια συσχέτιση με τις προοδευτικές μήτρες του Raven (από 0.20 ως 0.53), ενώ στα πλαίσια της αποκλίνουσας εγκυρότητάς του, παρουσίασε πολύ χαμηλή (από 0.16 έως 0.27) ως καθόλου (από 0.07 έως 0.17) συσχέτιση με δύο τεστ που μετρούσαν αφαιρετικό συλλογισμό και μη λεκτικό συλλογισμό, αντίστοιχα. Συνοψίζοντας, τα αποτελέσματα της συγκεκριμένης έρευνας τεκμηριώνουν το TAX ως ένα εγκυρο και αξιοπιστο ψυχομετρικο εργαλειο, το οποίο μπορεί να χρησιμοποιηθεί για τη μέτρηση της χωρο-αντιληπτικής ικανότητας.

Λέξεις-κλειδιά: Χωρο-αντιληπτική ικανότητα, Θεωρία

1. Εισαγωγή

Η χωρο-αντιληπτική ικανότητα (spatial ability) είναι μια από τις πιο σημαντικές δεξιότητες που έχει αναπτύξει ο άνθρωπος στην προσπάθειά

του να επιτύχει την καλύτερη προσαρμογή του στο περιβάλλον στο οποίο ζει και αναπτύσσεται (Bishop, 1980). Είναι μια δεξιότητα που εμπλέκεται σε μια σειρά από πολύπλοκες διεργασίες, όπως η μάθηση, η εκπαίδευση, η άσκηση εργα-

1. Διεύθυνση: Ιωάννης Τσαούσης, Πανεπιστήμιο Κρήτης, Τμήμα Ψυχολογίας, Πανεπιστημιούπολη Γάλλου, Ρέθυμνο 74100. Τηλ.: 28310-77519. e-mail: tsaousis@psy.soc.uoc.gr.

σίας, ακόμη και το παιχνίδι (Rafi et al., 2005). Μερικά χαρακτηριστικά παραδείγματα στα οποία διαπιστώνουμε έμπρακτα την εμπλοκή της χωρο-αντιληπτικής ικανότητας στην καθημερινότητά μας είναι τα εξής: όταν προσπαθούμε να βρούμε το δρόμο της επιστροφής προς το σπίτι μας ξεκινώντας από ένα συγκεκριμένο σημείο της πόλης· όταν θέλουμε να τοποθετήσουμε τα καινούργια μας έπιπλα σε ένα άδειο δωμάτιο· όταν θέλουμε να ετοιμάσουμε τη βαλίτσα μας για ταξίδι· όταν προσπαθούμε να πιάσουμε ένα αντικείμενο στον αέρα· όταν επιχειρούμε να στείλουμε την μπάλα στα δίχτυα σε έναν ποδοσφαιρικό αγώνα κ.λπ. (Hegarty & Waller, 2005).

Η χωρο-αντιληπτική ικανότητα είχε αρχίσει να απασχολεί τους ερευνητές ψυχολόγους από την εποχή του Galton (1880), ο οποίος για πρώτη φορά προσπάθησε να μελετήσει με συστηματικό τρόπο τις διάφορες δεξιότητες της ανθρώπινης συμπεριφοράς. Αργότερα, όταν η μεθοδολογία της έρευνας, αλλά και η στατιστική, κυριάρχησαν στον τρόπο μελέτης των ατομικών διαφορών, ερευνητές όπως ο Spearman (1927) και ο Thurstone (1938) προσπάθησαν να τεκμηριώσουν τη σπουδαιότητα της χωρο-αντιληπτικής ικανότητας με πιο σαφή κριτήρια και να την εντάξουν ως αυτόνομη δεξιότητα στο δομικό μοντέλο της ανθρώπινης νοημοσύνης. Στην αρχή, η χωρο-αντιληπτική ικανότητα είχε θεωρηθεί ως μια δευτερεύουσας σημασίας δεξιότητα, η οποία κυρίως αφορούσε άτομα με αυτό που λέμε «πρακτικό» μυαλό (Bishop, 1980). Σήμερα όμως, θεωρείται μια πρωτογενής δεξιότητα (higher order), η οποία μάλιστα παίζει σημαντικό ρόλο τόσο στον επιστημονικό τρόπο σκέψης (π.χ. μαθηματικά), όσο και σε άλλους τομείς της ανθρώπινης δραστηριότητας, όπως στην ικανότητα του ατόμου να χειρίζεται με αποτελεσματικό τρόπο τη γλώσσα και κατ' επέκταση τη σκέψη (Bishop, 1980).

Πολλοί ερευνητές έχουν ορίσει τη χωρο-αντιληπτική ικανότητα ως κομβικό σημείο σε μια σειρά από επιστήμες και επιστημονικές δραστηριότητες, όπως τα μαθηματικά (Gallagher, 1989), το μηχανικό σχέδιο και η γραφιστική (Gimmestad

& Sorby, 1996. Ferguson, 1977), οι σπουδές στις θετικές επιστήμες (Pribyl & Bonder, 1987. Lord, 1990), οι εμπορικές σπουδές (Bennett, Seashore & Wesman, 1974), η φυσική αγωγή (Meeker, 1991), η δημιουργικότητα (Lohman, 1993). Επιπρόσθετα, έχει αποδειχθεί ότι κυριαρχεί σε επαγγελματικούς χώρους όπου η οπτικό-χωρο-αντιληπτική ικανότητα αποτελεί σημαντικό στοιχείο της άσκησης του επαγγέλματος, όπως στους μηχανικούς, στους αρχιτέκτονες, στους πιλότους, στους τεχνικούς ηλεκτρονικών υπολογιστών, στους χειριστές μηχανών κ.ά. (Lohman, 1993).

Ορισμός και εννοιολογικό περιεχόμενο

Κατά καιρούς, έχουν διατυπωθεί αρκετοί διαφορετικοί ορισμοί για την έννοια της χωρο-αντιληπτικής ικανότητας και έχουν χρησιμοποιηθεί διαφορετικά ονόματα για να περιγραφεί το ίδιο εννοιολογικό κατασκεύασμα, όπως χωρο-αντιληπτική διεργασία, χωρο-αντιληπτική νοημοσύνη, χωρο-αντιληπτικός συλλογισμός και χωρο-αντιληπτική αίσθηση (Lohman, 1988). Ο Bortoline (1998) ορίζει τη χωρο-αντιληπτική διεργασία ως τη νοητική διεργασία που χρησιμοποιούμε για να αντιλαμβανόμαστε, να αποθηκεύουμε, να ανακαλούμε, να δημιουργούμε, να επεξεργαζόμαστε και να μεταδίδουμε χωρικές εικόνες. Ο Gardner (1983) ορίζει τη χωρο-αντιληπτική νοημοσύνη ως μια από τις επτά βασικές και ανεξάρτητες μορφές νοημοσύνης και θεωρεί ότι στηρίζεται στην ικανότητα του ατόμου να αντιλαμβάνεται και να μεταβιβάζει «το όλο» ενός αντικειμένου βασιζόμενος σε μια μορφολογική οργάνωση ("gestalt") της πραγματικότητας, η οποία διαφοροποιείται από τη λογικο-μαθηματική οργάνωση της πραγματικότητας. Οι Clements και Battista (1992) θεωρούν ότι ο χωρο-αντιληπτικός συλλογισμός περιλαμβάνει νοητικές δεξιότητες με τη βοήθεια των οποίων δημιουργούνται και ελέγχονται νοητικές αναπαραστάσεις χωρικών αντικειμένων, σχέσεων και μετασχηματισμών. Επίσης, οι Linn και Petersen (1985) ορίζουν το χωρο-αντιληπτικό συλλογισμό ως μια δεξιότητα αναφορικά με την

αναπαράσταση, το μετασχηματισμό, την αναπαραγωγή και την ανάκληση συμβολικών, μη λεκτικών πληροφοριών. Τέλος, το National Council of Teachers of Mathematics (1989) των ΗΠΑ, ορίζει τη χωρο-αντιληπτική αίσθηση ως μια διαισθητική ικανότητα που βοηθά το άτομο να κατανοήσει το περιβάλλον που τον περικλείει, καθώς και τα αντικείμενα που περιλαμβάνονται σε αυτό. Επιπρόσθετα, ορίζει ως βασικά συστατικά στοιχεία αυτής της ικανότητας την αίσθηση της κατεύθυνσης, τον προσανατολισμό και την αντίληψη της προοπτικής των αντικειμένων στο χώρο: την αναφορά των σχημάτων και του μεγέθους εικόνων και αντικειμένων: καθώς και τον τρόπο με τον οποίο η αλλαγή στο σχήμα επηρεάζει την αλλαγή στο μέγεθος.

Ένα σημαντικό ζητούμενο των ερευνών που πραγματοποιούνται τα τελευταία χρόνια αναφορικά με τη χωρο-αντιληπτική ικανότητα, σχετίζεται με τη δομή της συγκεκριμένης έννοιας. Πιο συγκεκριμένα, το ερώτημα που φαίνεται ότι απασχολεί τους ειδικούς είναι το κατά πόσο η συγκεκριμένη δεξιότητα είναι μονοδιάστατη ή πολυδιάστατη. Επιπρόσθετα, ενδιαφέρον παρουσιάζει τόσο ο αριθμός όσο και το εννοιολογικό περιεχόμενο των βασικών παραγόντων που φαίνεται να προκύπτουν από την έρευνα.

Ο Lohman (1979) από μια μετα-αναλυτική μελέτη διαπίστωσε την ύπαρξη τριών ανεξάρτητων διαστάσεων και διατύπωσε το αντίστοιχο θεωρητικό μοντέλο. Σύμφωνα με αυτό το μοντέλο, η χωρο-αντιληπτική ικανότητα περιλαμβάνει τρεις βασικές διαστάσεις: τη διάσταση του Σχηματισμού Νοερής Εικόνας (ΣΝΕ) (Visualization), τη διάσταση του Χωρικού Προσανατολισμού (ΧΠ) (Spatial Orientation) και τη διάσταση των Χωρικών Σχέσεων (ΧΣ) (Spatial Relations). Ο ΣΝΕ αναφέρεται στην ικανότητα του ατόμου να χειρίζεται νοερά κινήσεις αντικειμένων και χωρικές φόρμες. Χαρακτηριστικά παραδείγματα ερεθισμάτων που χρησιμοποιούνται για να μετρήσουν αυτή τη διάσταση είναι οι αναδιπλώσεις αντικειμένων (surface development), όπου τα άτομα καλούνται να διπλώσουν ή να ξεδιπλώσουν νοερώς διάφορα γεωμετρικά σχήματα. Ο ΧΠ αναφέρεται στην

ικανότητα του ατόμου να φαντάζεται την εμφάνιση ενός τρισδιάστατου αντικειμένου υπό διαφορετικές οπτικές γωνίες (διαφορετικό προσανατολισμό). Ένα παράδειγμα ερεθίσματος που χρησιμοποιείται για τη μέτρηση αυτής της διάστασης είναι η παρουσίαση ενός τρισδιάστατου αντικειμένου ως κάτοψη, και διαφορετικές παρουσιάσεις του ίδιου αντικειμένου από διαφορετικές οπτικές γωνίες, όπου τα άτομα καλούνται να αναγνωρίσουν το αντικείμενο που είδαν σε κάτοψη. Τέλος, οι ΧΣ περιλαμβάνουν τις νοητικές περιστροφές αντικειμένων (mental rotation), όπου τα άτομα καλούνται να περιστρέψουν νοερά διάφορα αντικείμενα, δυσδιάστατα ή τρισδιάστατα σχήματα.

Παρά το γεγονός ότι το μοντέλο του Lohman, αρχικά θεωρήθηκε ένα από τα πιο αξιόπιστα θεωρητικά μοντέλα στο χώρο (Carroll, 1993. McGee, 1979), μελλοντικές έρευνες που χρησιμοποίησαν την παραγοντική ανάλυση ως μέθοδο διερεύνησης της δομής της χωρο-αντιληπτικής ικανότητας διαπίστωσαν την ύπαρξη δύο μόνο βασικών διαστάσεων: τον ΣΝΕ και τον ΧΠ, καθώς θεωρούν ότι η διάσταση ΧΣ αποτελεί συνθετικό στοιχείο της διάστασης ΧΠ (Carroll, 1993. Colom et al., 2003. Eliot & Smith, 1983. Linn & Petersen, 1985). Ακόμη και σήμερα όμως, δεν έχει δοθεί μια πειστική απάντηση για το κατά πόσο οι δύο αυτές διαστάσεις είναι ανεξάρτητες μεταξύ τους ή όχι με αποτέλεσμα το θεωρητικό μοντέλο του Lohman να είναι ακόμη επίκαιρο και να χρησιμοποιείται ως θεωρητικό πλαίσιο για τη μέτρηση της χωρο-αντιληπτικής ικανότητας.

Η μέτρηση της χωρο-αντιληπτικής ικανότητας

Σήμερα υπάρχουν πολλά και διαφορετικά ψυχομετρικά εργαλεία που μετρούν τη χωρο-αντιληπτική ικανότητα και τα οποία μπορούν να ταξινομηθούν σε τέσσερις γενικές κατηγορίες: τεστ επίδοσης (performance tests), τεστ χαρτιού-και-μολυβιού (paper-and-pencil tests), λεκτικά τεστ (verbal tests) και δυναμικά τεστ (dynamic tests). Τα τεστ επίδοσης είναι τα παλαιότε-

ρα και περιλαμβάνουν συνήθως δραστηριότητες όπως οι αναδιπλώσεις αντικειμένων, ο χειρισμός κύβων, ο χειρισμός αντικειμένων σε φόρμες. Όλες αυτές οι δραστηριότητες χρησιμοποιούνται σήμερα και για τη μέτρηση της νοημοσύνης, γι' αυτό και τις συναντάμε στα πιο γνωστά τεστ γενικής νοημοσύνης για τη μέτρηση της μη λεκτικής νοημοσύνης (e.g., Wechsler, 1955). Ένα άλλο είδος τεστ επίδοσης αναφέρεται στη μέτρηση της ικανότητας του ατόμου να αντιλαμβάνεται το χώρο σε πραγματικές συνθήκες, όπως να βρίσκει το συντομότερο δρόμο μεταξύ διαφορετικών περιοχών, να μπορεί να βρίσκει τον προσανατολισμό του στο χώρο κ.λπ. Πάντως, τέτοιου είδους δραστηριότητες έχουν βρεθεί ότι στην καλύτερη περίπτωση παρουσιάζουν μέτρια συσχέτιση με άλλα τεστ που μετρούν τη χωρο-αντιληπτική ικανότητα (Allen, 1982. Lorenz & Neisser, 1986), γι' αυτό και χρησιμοποιούνται σπάνια.

Τα τεστ χαρτιού-και-μολυβιού είναι τα πιο πολυχρησιμοποιημένα και βασίζονται στην κλασική μορφή των ψυχολογικών τεστ, όπου σχήματα ή προβλήματα παρουσιάζονται σε ένα βιβλιάρκι και το άτομο χρησιμοποιώντας ένα στυλό ή ένα μολύβι σημειώνει τη σωστή απάντηση. Οι Eliot και Smith (1983) παρουσιάζουν, δίνοντας και παραδείγματα, 392 διαφορετικά τεστ αντίληψης χώρου που εντάσσονται σε αυτή την κατηγορία.

Από την άλλη, τα λεκτικά τεστ, παρά το γεγονός ότι παρουσιάζουν υψηλές συσχετίσεις με άλλα τεστ αντίληψης χώρου αλλά και με άλλα εξωτερικά κριτήρια (Ackerman & Kanfer, 1993), δεν έχουν λάβει της προσοχής που ίσως τους αξίζει. Σε αυτού του είδους τα τεστ το άτομο ακούει το πρόβλημα, το οποίο απαιτεί συνήθως τη δημιουργία μιας νοητικής εικόνας, και στη συνέχεια καλείται να απαντήσει σε μια σειρά από ερωτήσεις. Για παράδειγμα: «Φαντάσου ότι κατευθύνεσαι βόρεια στη συνέχεια στρίψε δεξιά, έπειτα προχώρα ευθεία και στη συνέχεια στρίψε πάλι δεξιά. Προς ποια κατεύθυνση κοιτάς;». Σε αυτού του είδους τα τεστ το άτομο καλείται να χρησιμοποιήσει τις χωρο-αντιληπτικές του

ικανότητες με έναν τρόπο που είναι πιο κοντά στην καθημερινή χρήση τους σε σχέση με τα τεστ χαρτιού-και-μολυβιού.

Τέλος, ενώ στα περισσότερα τεστ αντίληψης χώρου τα ερεθίσματα που παρουσιάζονται είναι στατικά, τα δυναμικά τεστ ακολουθούν μια διαφορετική λογική. Ο Gibson (1947) και αργότερα οι Seibert και Snow (1965) ανέπτυξαν μια σειρά από τεστ που χρησιμοποιούσαν κινούμενες εικόνες ως ερεθίσματα και οι οποίες παρουσιάζονταν συνήθως στην οθόνη ενός υπολογιστή. Πιο πρόσφατα, οι Pellegrino και Hunt (1989) δημιούργησαν μια σειρά από δυναμικά τεστ, τα οποία μετρούν την ικανότητα του ατόμου να υπολογίζει πραγματική κίνηση, να προβλέπει τροχιές κινούμενων αντικειμένων και να υπολογίζει χρονικές αφίξεις δύο ή περισσότερων αντικειμένων. Όλες οι παραπάνω διεργασίες έχουν αποδειχθεί ότι μετρούν διαφορετικές δεξιότητες από αυτές που μετρούν τα τεστ χαρτιού-και-μολυβιού (Colom et al., 2003).

Εκτός όμως από τις διαφορές στα ερεθίσματα, τα τεστ αντίληψης χώρου διαφέρουν επίσης και ως προς τον τύπο των απαντήσεων που ζητούν. Για παράδειγμα, υπάρχουν τεστ που ζητούν τη σωστή απάντηση ανάμεσα σε εναλλακτικές επιλογές, υπάρχουν τεστ που ζητούν από τα άτομα να κατασκευάσουν ή να αναπαραγάγουν από μνήμης ένα γεωμετρικό σχήμα και, τέλος, υπάρχουν τεστ που ζητούν από τα άτομα μια λεκτική δήλωση. Σύμφωνα με τον Lohman (1988), για τα τεστ χαρτιού-και-μολυβιού είναι προτιμότερο να χρησιμοποιούνται ερεθίσματα στα οποία ζητάμε από τα άτομα να κάνουν κάτι (κατασκευή ή αναπαραγωγή ενός σχήματος ή ενός αντικειμένου), αφού τέτοιου είδους ερεθίσματα μπορεί να είναι πιο δύσκολα στη βαθμολόγησή τους, μας δίνουν όμως περισσότερες πληροφορίες τόσο για τη χωρο-αντιληπτική ικανότητα του ατόμου όσο και για άλλα στοιχεία, όπως για παράδειγμα, τη στρατηγική που έχει ακολουθήσει το άτομο στην προσπάθειά του να επιλύσει το πρόβλημα, καθώς μας επιτρέπουν την ανάλυση των λαθών που κάνει (Kyllonen, 1984).

Μια άλλη διαφορά εντοπίζεται ανάμεσα στα τεστ που βασίζονται στο χρόνο και στα τεστ που βασίζονται στα διαφορετικά επίπεδα δυσκολίας των ερεθισμάτων. Τα τεστ που βασίζονται στο χρόνο τείνουν να μετρούν πιο συγκεκριμένα χαρακτηριστικά της χωρο-αντιληπτικής ικανότητας σε σχέση με τα τεστ που δεν έχουν χρονικό περιορισμό αφού η λογική τους είναι διαφορετική. Έτσι, σε αυτού του είδους τα τεστ είναι πιο εύκολο να μελετηθούν σε βάθος περισσότερα χαρακτηριστικά της έννοιας.

Σήμερα, υπάρχουν αρκετά τεστ που μετρούν τη χωρο-αντιληπτική ικανότητα, όμως τα περισσότερα από αυτά χρησιμοποιούν στοιχεία από μία μόνο διάσταση, όπως το Τεστ Αντίληψης Χώρου της συστοιχίας Τεστ Γενικής Νοημοσύνης (General Ability Test-Smith & Whetton, 1988), όπου χρησιμοποιούνται μόνο αναδιπλώσεις ή το Τεστ Νοητικών Περιστροφών (Mental Rotation Test – Shepard & Metzler, 1971), στο οποίο χρησιμοποιούνται μόνο ερεθίσματα για τα οποία απαιτείται νοητική περιστροφή. Αυτό έχει ως αποτέλεσμα να γεινόνται πάντοτε αντιρρήσεις για το κατά πόσο τα διάφορα τεστ μετρούν όλες τις πιθανές εκφάνσεις της έννοιας της χωρο-αντιληπτικής ικανότητας.

Κάτω από αυτό το πρίσμα, στόχος της συγκεκριμένης έρευνας είναι η κατασκευή ενός τεστ αντίληψης χώρου, το οποίο να μετρά και τις τρεις βασικές διαστάσεις της χωρο-αντιληπτικής ικανότητας: (α) το *Σχηματισμό Νοερής Εικόνας*, (β) το *Χωρικό Προσανατολισμό* και (γ) τις *Χωρικές Σχέσεις*. Ο λόγος για τον οποίο αποφασίσαμε να χρησιμοποιήσουμε το μοντέλο των τριών παραγόντων του Lohman (1979) αντί του μοντέλου των δύο παραγόντων, όπως προτείνουν νεότερες έρευνες, είναι γιατί θέλουμε να διερευνήσουμε κατά πόσο ο χωρικός προσανατολισμός και οι χωρικές σχέσεις είναι ανεξάρτητες μεταξύ τους ή συνθέτουν μια κοινή διάσταση, καθώς υπάρχουν έρευνες που διαπιστώνουν ότι υπάρχουν διαφορές ανάμεσα τους (Smith & Whetton, 1988). Η βασική διαφορά ανάμεσα στη διάσταση ΧΣ και στη διάσταση ΧΠ είναι ότι στα ερεθίσματα που χρησιμοποιούνται για τη μέτρηση του ΧΠ, ο προ-

σανατολισμός του σώματος του ατόμου που συμπληρώνει ένα τέτοιο τεστ παίζει σημαντικό ρόλο στην επίλυση του προβλήματος: κάτι που δεν συμβαίνει με τα ερεθίσματα που χρησιμοποιούνται για τη μέτρηση της διάστασης ΧΣ (π.χ. στις νοητικές περιστροφές). Και αυτός είναι ένας σημαντικός λόγος για να χρησιμοποιηθούν και τα δύο είδη ερεθισμάτων για τη δημιουργία ενός τεστ αντίληψης χώρου.

2. Μέθοδος

Συμμετέχοντες

Τέσσερα διαφορετικά συμπτωματικά δείγματα χρησιμοποιήθηκαν στη συγκεκριμένη έρευνα. Το πρώτο δείγμα χρησιμοποιήθηκε για την έρευνα πιλότο, στο πλαίσιο της οποίας επιλέχθηκαν τα τελικά στοιχεία του τεστ. Συμμετείχαν 349 άτομα, τα οποία δοκιμάστηκαν ως υποψήφιοι στη διαδικασία επιλογής προσωπικού σε μια ιδιωτική τράπεζα. Από αυτά, 118 (33,80%) ήταν άνδρες και 223 (63,90%) ήταν γυναίκες· οκτώ άτομα (2,30%) δεν δήλωσαν το φύλο τους. Ο μέσος όρος ηλικίας τους ήταν 27,1 έτη (TA=10,24). Το δεύτερο δείγμα αποτελούσαν 110 άτομα τα οποία συμμετείχαν σε ένα σεμινάριο συναισθηματικής νοημοσύνης. Από αυτά, τα 35 (31,8%) ήταν άνδρες και 67 (60,9%) ήταν γυναίκες· οκτώ άτομα (7,3%) δεν δήλωσαν το φύλο τους. Ο μέσος όρος ηλικίας τους ήταν 26,39 έτη (TA=9,66). Το τρίτο δείγμα περιλάμβανε 120 προπτυχιακούς και μεταπτυχιακούς φοιτητές ψυχολογίας, οι οποίοι συμμετείχαν σε ένα μάθημα ψυχομετρίας, για το οποίο υπήρχε υποχρεωτική η συμπλήρωση ενός αριθμού ψυχομετρικών εργαλείων. Από αυτούς, οι 35 (29,2%) ήταν άνδρες και οι 85 (70,8%) ήταν γυναίκες, ενώ ο μέσος όρος ηλικίας τους ήταν 29,30 έτη (TA=11,19). Το τέταρτο δείγμα περιλάμβανε 115 άτομα, τα οποία συμπλήρωσαν το τεστ σε μια ημερίδα ψυχολογικής αξιολόγησης. Από αυτά, 48 (41,70%) ήταν άνδρες και 67 (58,30%) ήταν γυναίκες. Ο μέσος όρος ηλικίας του συγκεκριμένου δείγματος ήταν 27,40 έτη (TA=9,64).

Εργαλεία

Το Τεστ Αντίληψης Χώρου (TAX). Η αρχική έκδοση του TAX αποτελούνταν από τρία μέρη και περιλάμβανε 60 στοιχεία τα οποία προέρχονταν και από τις τρεις βασικές διαστάσεις της χωρο-αντιληπτικής ικανότητας (Σχηματισμός Νοερής Εικόνας, Χωρικός Προσανατολισμός και Χωρικές Σχέσεις), σύμφωνα με το θεωρητικό μοντέλο του Lohman (1979). Για κάθε μέρος δημιουργήθηκαν 20 στοιχεία-προβλήματα με ένα γεωμετρικό σχήμα ως στέλεχος και τέσσερις (για τον SNE και τον ΧΠ) ή πέντε (για τις ΧΣ) εναλλακτικές επιλογές ως απαντήσεις, από τις οποίες το άτομο θα έπρεπε να επιλέξει τη μία ως σωστή. Η τελική έκδοση του τεστ περιλαμβάνει 36 στοιχεία και είναι επίσης χωρισμένο σε τρία μέρη (12 στοιχεία από κάθε διάσταση), καθένα από τα οποία έχει δικές του οδηγίες και παραδείγματα. Παραδείγματα αντιπροσωπευτικών στοιχείων από κάθε διάσταση βρίσκονται στο Παράρτημα.

Η Συστοιχία Διαφορικής Ικανότητας - Το Τεστ Χωρικών Σχέσεων (The Differential Aptitude Test - Space Relations - DAT; Hyde & Trickey, 1995). Το συγκεκριμένο εργαλείο έχει δημιουργηθεί για να μετρά τη χωρο-αντιληπτική ικανότητα. Περιλαμβάνει 50 μη διπλωμένα επίπεδα σχέδια και ζητά από τα άτομα που το συμπληρώνουν να φανταστούν το τρισδιάστατο σχήμα που θα προκύψει από το δίπλωμα του συγκεκριμένου σχεδίου. Για κάθε στοιχείο υπάρχουν τέσσερις απαντητικές επιλογές (παρουσιάζουν διαφορετικές εκδοχές ενός τρισδιάστατου σχήματος) από τις οποίες μόνο η μία είναι σωστή. Ο δείκτης εσωτερικής συνοχής Kuder-Richardson για το συγκεκριμένο δείγμα ήταν 0,70.

Η Συστοιχία Γενικών Ικανοτήτων- Το Τεστ Χωρο-αντιληπτικής Ικανότητας (The General Ability Test - Spatial Ability - GAT; Smith & Whetton, 1988). Αυτό είναι ένα εργαλείο που σχεδιάστηκε για να μετρά τη χωρο-αντιληπτική ικανότητα. Περιλαμβάνει 20 μητρικά μη διπλωμένα επίπεδα σχήματα, και 80 τρισδιάστατα σχήματα (τέσσερα για καθένα από τα μη διπλωμένα σχήματα) ως απαντητικές επιλογές. Τα άτομα που καλούνται να συμπληρώ-

σουν το τεστ προσπαθούν να εντοπίσουν κατά πόσο το καθένα από τα τρισδιάστατα σχήματα που τους παρουσιάζονται μπορεί να προκύψει από την αναδίπλωση του μητρικού μη διπλωμένου σχήματος. Για το συγκεκριμένο δείγμα, ο δείκτης εσωτερικής συνοχής Kuder-Richardson ήταν 0,89.

Η Συστοιχία Διαφορικών Ικανοτήτων - Το Τεστ Αφαιρετικού Συλλογισμού (The Differential Aptitude Test - Abstract Reasoning - DAT; Hyde & Trickey, 1995). Το συγκεκριμένο εργαλείο έχει κατασκευαστεί για να μετρά αφαιρετικό συλλογισμό. Καθένα από τα 40 στοιχεία που περιλαμβάνονται στο τεστ αποτελεί μια γεωμετρική σειρά στην οποία τα στοιχεία που τη συνθέτουν αλλάζουν σύμφωνα με ένα συγκεκριμένο κανόνα. Τα άτομα που καλούνται να συμπληρώσουν το τεστ θα πρέπει να εντοπίσουν τον κανόνα σύμφωνα με τον οποίο αλλάζουν τα στοιχεία της σειράς και να εντοπίσουν το στοιχείο που θα πρέπει να ακολουθήσει. Ο δείκτης εσωτερικής συνοχής Kuder-Richardson για το συγκεκριμένο δείγμα ήταν 0,78.

Η Συστοιχία Γενικών Ικανοτήτων - Το Μη Λεκτικό Τεστ (The General Ability Test - Non Verbal -GAT; Smith & Whetton, 1988). Το συγκεκριμένο εργαλείο έχει κατασκευαστεί για να μετρά τη μη λεκτική ικανότητα. Περιλαμβάνει 36 στοιχεία τα οποία είναι χωρισμένα σε δύο διαφορετικές θεματικές ενότητες: στην πρώτη ενότητα περιλαμβάνονται στοιχεία που απαιτούν την κωδικοποίηση σχημάτων, ενώ στη δεύτερη ενότητα περιλαμβάνονται στοιχεία που αναφέρονται σε αναλογίες, ακολουθίες σειρών και μοτίβα σχημάτων. Για το συγκεκριμένο δείγμα, ο δείκτης εσωτερικής συνοχής Kuder-Richardson ήταν 0,87.

Το Τεστ Προοδευτικών Μητρών του Raven - Βασική Έκδοση (The Raven Progressive Matrices - Standard Version. Raven, Court & Raven, 1979). Το συγκεκριμένο εργαλείο έχει κατασκευαστεί για να μετρά τη ρέουσα νοημοσύνη (fluid intelligence). Περιλαμβάνει πέντε σετ από 12 μήτρες ακολουθιών σχημάτων, οι οποίες είναι τοποθετημένες με προοδευτικά αυξανόμενη δυσκολία. Τα άτομα καλούνται να μελετήσουν τις ακολουθίες των σχημάτων που περιλαμβάνει η

κάθε μήτρα και να εντοπίσουν το σχήμα που θα πρέπει να τοποθετηθεί στην ακολουθία ώστε η μήτρα να είναι ολοκληρωμένη. Ο δείκτης εσωτερικής συνοχής Kuder-Richardson για το συγκεκριμένο δείγμα ήταν 0,94.

Διαδικασία

Από όλους τους συμμετέχοντες ζητήθηκε να συμπληρώσουν το TAX. Η πρώτη ομάδα συμπλήρωσε επιπρόσθετα και κάποια άλλα γνωστικά τεστ, τα οποία χρησιμοποιήθηκαν ως κριτήρια επιλογής για τη θέση για την οποία είχαν κάνει αίτηση. Σε όλα τα άτομα δηλώθηκε από την αρχή, ότι τα αποτελέσματα από το TAX δεν θα λαμβάνονταν υπόψη στη διαδικασία επιλογής. Η δεύτερη ομάδα συμπλήρωσε το TAX δύο φορές με χρονική απόσταση ανάμεσα στις δύο χορηγήσεις τις τέσσερις εβδομάδες, καθώς εξηγήθηκε στους συμμετέχοντες ότι συμμετείχαν στη διαδικασία αξιολόγησης της αξιοπιστίας του τεστ. Η τρίτη ομάδα συμπλήρωσε επιπρόσθετα το Τεστ Χωρικών Σχέσεων της συστοιχίας των Διαφορικών Ικανοτήτων και το Τεστ Χωρο-αντιληπτικής Ικανότητας της συστοιχίας των Γενικών Ικανοτήτων, καθώς τους εξηγήθηκε ότι συμμετείχαν στη διαδικασία αξιολόγησης της συγχρονικής εγκυρότητας του τεστ. Τέλος, η τέταρτη ομάδα συμπλήρωσε επιπρόσθετα το Τεστ Αφαιρετικού Σύλλογισμού της συστοιχίας των Διαφορικών Ικανοτήτων, το Μη Λεκτικό Τεστ της συστοιχίας των Γενικών Ικανοτήτων, καθώς και το τεστ Προοδευτικών Μητρών του Raven, καθώς συμμετείχαν στη διαδικασία αξιολόγησης της εγκυρότητας εννοιολογικής κατασκευής. Όλα τα άτομα συμμετείχαν εθελοντικά και όσοι το επιθυμούσαν έλαβαν τα αποτελέσματα ταχυδρομικώς, αφού πρώτα συμπλήρωσαν τη διεύθυνσή τους στο φύλο απαντήσεων.

3. Αποτελέσματα

Μέθοδος ανάλυσης

Η επιλογή των στοιχείων της τελικής έκδοσης του TAX πραγματοποιήθηκε με τη χρήση της Θεω-

ρίας Υπολογισμού των Ικανοτήτων (Item Response Theory-IRT). Η Θεωρία Υπολογισμού των Ικανοτήτων (ΘΥΙ) βασίζεται στη θεωρία των λανθανόντων χαρακτηριστικών (Birnbaum, 1968) και αναφέρεται σε ένα σύνολο μαθηματικών μοντέλων, τα οποία περιγράφουν την πιθανότητα ενός ατόμου με μια συγκεκριμένη ικανότητα να απαντήσει σωστά ένα στοιχείο με ένα δεδομένο βαθμό δυσκολίας. Μας δίνει δηλαδή πληροφορίες, σε επίπεδο πιθανοτήτων, για τη σχέση που υπάρχει ανάμεσα στην απάντηση που δίνει ένα άτομο σε ένα στοιχείο και της πραγματικής ικανότητας που κατέχει (Hambleton, Swaminathan & Rogers, 1991).

Στο πλαίσιο της κατασκευής ενός ψυχομετρικού εργαλείου, η συγκεκριμένη μέθοδος παρουσιάζει σημαντικά πλεονεκτήματα σε σχέση με την Κλασική Θεωρία των Τεστ - ΚΘΤ (Classical Test Theory-CTT), μέθοδος με την οποία έχουν κατασκευαστεί τα περισσότερα ψυχομετρικά εργαλεία τις προηγούμενες δεκαετίες. Το μεγαλύτερο μειονέκτημα της ΚΘΤ ως μεθόδου κατασκευής εργαλείων είναι ότι τα τεστ που προκύπτουν από τη χρήση της είναι πιθανόν να παραγάγουν διαφορετικά αποτελέσματα ανάλογα με το δείγμα των ατόμων που τα συμπληρώνουν ή ανάλογα με το δείγμα των στοιχείων που θα επιλεγούν να τα συνθέσουν. Για παράδειγμα, εάν άτομα με διαφορετικό επίπεδο ικανότητας σε μια έννοια συμπληρώσουν διαφορετικής δυσκολίας τεστ είναι πολύ πιθανό να παρουσιάσουν διαφορετικές επιδόσεις. Τα τεστ που προκύπτουν όμως από τη χρήση της ΘΥΙ δεν αντιμετωπίζουν τέτοιου είδους προβλήματα, αφού τα χαρακτηριστικά των στοιχείων που επιλέγονται δεν εξαρτώνται από το δείγμα των ατόμων που χρησιμοποιούνται για την κατασκευή του τεστ. Επιπρόσθετα, ο υπολογισμός των παραμέτρων μπορεί να πραγματοποιηθεί σε επίπεδο στοιχείου (item level), κάτι που επιτρέπει τη χρήση αυτής της ιδιότητας στη σύγκριση διαφορετικών ομάδων ως προς ένα συγκεκριμένο στοιχείο. Τέλος, στη ΘΥΙ δεν είναι υποχρεωτικό να εξασφαλίζεται η ισότητα των σφαλμάτων μέτρησης, μια προϋπόθεση που είναι σημαντική στην ΚΘΤ (Hambleton, Swaminathan & Rogers, 1991).

Στη ΘΥΙ, η σχέση ανάμεσα στην απόκριση ενός

ατόμου σε ένα στοιχείο και της έννοιας που μετρά το τεστ εκφράζεται μέσω της Καμπύλης των Χαρακτηριστικών της Ερώτησης – ΚΧΕ (Item Characteristic Curve – ICC). Η ΚΧΕ είναι μια γραφική παράσταση που περιγράφει με κάθε λεπτομέρεια την πιθανότητα που έχει το άτομο να απαντήσει με ένα συγκεκριμένο τρόπο σε ένα στοιχείο (π.χ. σωστά) όσο αυξάνεται το επίπεδο της ιδιότητας που μετρά το τεστ. Οι ΚΧΕ που μπορούμε να υπολογίσουμε είναι τριών ειδών: μιας παραμέτρου (βάσει του υπολογισμού μόνο του βαθμού δυσκολίας), δύο παραμέτρων (βάσει του υπολογισμού του βαθμού δυσκολίας και του βαθμού διαφοροποίησης) και τριών παραμέτρων (βάσει του υπολογισμού του βαθμού δυσκολίας, του βαθμού διαφοροποίησης και της εικασίας). Στη συγκεκριμένη έρευνα, χρησιμοποιήσαμε την ΚΧΕ με βάση τον υπολογισμό της μίας παραμέτρου, ή του μοντέλου Rasch (1980), όπως είναι ευρύτερα γνωστό.

Η μέθοδος αξιολογεί την επίδοση αλλά και τη συμπεριφορά των ατόμων ή των μερών του τεστ με βάση τις προσδοκίες του μοντέλου (Smith & Smith, 2004). Το μοντέλο Rasch υποθέτει ότι η πιθανότητα (P) που έχει ένα άτομο να απαντήσει σε ένα στοιχείο σωστά, βασίζεται αποκλειστικά (α) στη θέση (θ), που κατέχει ένα άτομο σε ένα λανθάνον χαρακτηριστικό θ (δηλαδή την ικανότητά του σε αυτό το χαρακτηριστικό στη δική μας περίπτωση σε καθεμιά από τις τρεις διαστάσεις της αντίληψης χώρου), και (β) στη θέση d της ΚΧΕ στο συγκεκριμένο λανθάνον χαρακτηριστικό (δηλαδή, στη δυσκολία του στοιχείου). Η ΚΧΕ υποτίθεται ότι ακολουθεί μια λογιστική εξίσωση, η οποία βρίσκεται γύρω από το σημείο d , που σημαίνει ότι η πιθανότητα που έχει ένα άτομο να απαντήσει σωστά ένα στοιχείο υπολογίζεται από την παρακάτω εξίσωση:

$$P(\theta, A) = \frac{1}{1 + e^{-\theta_i - d_i}}$$

Όπου: d είναι η παράμετρος δυσκολίας
 θ είναι το επίπεδο της ικανότητας που μετράμε, και
 e είναι η βάση του συστήματος των φυσικών λογαρίθμων

Προκαταρκτική ανάλυση

Προτού ξεκινήσουμε τη διαδικασία επιλογής των στοιχείων που θα συνθέσουν την τελική έκδοση του TAX μέσω της ΘΥΙ, θα πρέπει να ελέγξουμε δύο βασικές προϋποθέσεις που είναι αναγκαίο να εξασφαλίζονται ώστε να μπορεί να εφαρμοστεί η συγκεκριμένη μέθοδος: α) το μονοδιάστατο της κάθε έννοιας που μελετάμε, και β) την ανεξαρτησία των στοιχείων σε κάθε διάσταση (Swaminathan, Hambleton & Rogers, 2007). Για να ελέγξουμε το κατά πόσο κάθε διάσταση ήταν μονοπαραγοντική, προχωρήσαμε στην ανάλυση κυρίων συνιστωσών για διχοτομικά δεδομένα. Σύμφωνα με τον Anderson (1994), για να τεκμηριωθεί η ύπαρξη μονοδιάστατης έννοιας θα πρέπει το πηλίκο της πρώτης από τη δεύτερη ιδιοτιμή να είναι τουλάχιστον 3:1 ή 4:1. Για τη διάσταση ΣΝΕ το αντίστοιχο πηλίκο ήταν 6,76 (11,5:1,70), για τη διάσταση ΧΠ ήταν 9,88 (15,8:1,60) και για τη διάσταση ΧΣ ήταν 4,95 (9,9:2,0). Αναφορικά με το ποσοστό της διακύμανσης που ερμηνεύει ο πρώτος παράγοντας κάθε διάστασης, είχαμε 64,9% για τον ΣΝΕ, 79,6% για τον ΧΠ και 58,2% για τις ΧΣ. Τα ποσοστά αυτά κρίνονται ιδιαίτερα ικανοποιητικά και τεκμηριώνουν τη μονοδιάστατη έννοια σε κάθε διάσταση, καθώς, σύμφωνα με τον Linacre (2004), ο πρώτος παράγοντας υποστηρίζεται όταν εξηγεί περισσότερο από 60% της συνολικής διακύμανσης.

Αναφορικά με τη δεύτερη προϋπόθεση, και για να αποδείξουμε την ανεξαρτησία των στοιχείων σε κάθε διάσταση, αρχικά χωρίσαμε το δείγμα σε 3 ομάδες ανάλογα με την επίδοσή τους (χαμηλή, μέτρια και υψηλή επίδοση). Στη συνέχεια, για κάθε ομάδα χωριστά αξιολογήθηκαν και οι 189 συσχετίσεις μεταξύ των 20 στοιχείων και ελέγχθηκε ο βαθμός συσχέτισής τους. Από τα αποτελέσματα διαπιστώθηκε ότι σε καμιά από τις τρεις ομάδες δεν εντοπίστηκαν περισσότερες συσχετίσεις ($r > 0,30$) από αυτές που αναμέναμε λόγω τυχαίων παραγόντων ($189 * 0,05 = 9,45$). Πιο συγκεκριμένα, για τη διάσταση ΣΝΕ και για την ομάδα με χαμηλές επιδόσεις ο αριθμός των σημαντικών συσχετίσεων ήταν 10, για την ομάδα με

μέτριες επιδόσεις ήταν 3 και για την ομάδα με υψηλές επιδόσεις ήταν 4. Για τη διάσταση ΧΠ, οι αντίστοιχες τιμές ήταν 7, 2 και 6, ενώ για τη διάσταση ΧΣ, οι αντίστοιχες τιμές ήταν 3, 7 και 9. Μετά την εξασφάλιση των δύο βασικών προϋποθέσεων για την ορθή χρήση της ΘΥΙ, προχωρήσαμε στην επιλογή των πιο κατάλληλων στοιχείων για καθεμιά από τις 3 διαστάσεις, ώστε να δημιουργήσουμε την τελική έκδοση του TAX.

Τέλος, μια επιπρόσθετη προκαταρκτική ανάλυση αφορούσε την αξιολόγηση της αξιοπιστίας των συμμετεχόντων. Χρησιμοποιώντας τους σταθμισμένους δείκτες τιμών *Z-infit/outfit*, διερευνήσαμε την ύπαρξη συμμετεχόντων οι οποίοι συνεισέφεραν σημαντικά στην αύξηση του σφάλματος μέτρησης (*standard error of measurement*). Με τη βοήθεια των συγκεκριμένων αυτών στατιστικών δεικτών εντοπίσαμε άτομα τα οποία για διάφορους λόγους (π.χ. κούραση, αδιαφορία κ.λπ.) δεν ανταποκρίθηκαν σωστά στη συμπλήρωση των τεστ, και τα αφαιρέσαμε από τη συνέχεια των αναλύσεων ώστε να μην αυξήσουν το σφάλμα της μέτρησης (Linacre, 2003).

Επιλογή των στοιχείων της τελικής έκδοσης

Για να αποφασίσουμε κατά πόσο ένα στοιχείο είναι κατάλληλο ή όχι σύμφωνα με τη ΘΥΙ, χρησιμοποιήσαμε το στατιστικό κριτήριο χ^2 για καλή προσαρμογή (*chi-square fit statistic*). Τα περισσότερα στατιστικά πακέτα χρησιμοποιούν δύο συγκεκριμένους δείκτες που ονομάζονται *Infit* και *Outfit*, και οι οποίοι εκφράζουν τις αποκλίσεις των στοιχείων αλλά και ατόμων από τις προσδοκίες του μοντέλου Rasch. Στη συγκεκριμένη ανάλυση, η προσδοκία του μοντέλου, εφόσον τα εμπειρικά δεδομένα ταιριάζουν στο μοντέλο Guttman (με βάση το οποίο αξιολογούνται τα δεδομένα), είναι ότι η τιμή του δείκτη ισούται με τη μονάδα (*expected value* = 1). Αυτό σημαίνει ότι η μονάδα εκφράζει την απόλυτη ταύτιση των εμπειρικών δεδομένων με το μαθηματικό μοντέλο. Σύμφωνα με τον Linacre (1999), όταν οι δείκτες *infit/outfit* είναι ίσοι με τη μονάδα, τότε το

συγκεκριμένο στοιχείο έχει τα ιδανικά ψυχομετρικά χαρακτηριστικά. Όσο απομακρύνονται οι δείκτες από τη μονάδα τόσο χαμηλότερης ποιότητας ψυχομετρικά χαρακτηριστικά έχουν τα στοιχεία. Εάν κάποιο στοιχείο παρουσιάζει τιμή > 1,5 σε κάποιον από τους δύο ή και τους δύο δείκτες, το στοιχείο αυτό δεν είναι κατάλληλο (*misfit*) και απορρίπτεται (Bond & Fox, 2007). Στον Πίνακα 1 που ακολουθεί παρουσιάζονται οι δείκτες *Infit* και *Outfit* για όλα τα στοιχεία της αρχικής έκδοσης του TAX ανά διάσταση.

Όπως παρατηρούμε από τα δεδομένα που παρουσιάζονται στον Πίνακα 1, σε κάθε διάσταση όλα τα στοιχεία που επιλέχθηκαν βρίσκονταν σε αρκετά κοντινή απόσταση από την ιδανική τιμή 1. Πιο συγκεκριμένα, για τις Νοητικές Περιτροφές ο δείκτης *infit* κυμαινόταν από 0,93 έως 1,06 (MO = 1,03), ενώ ο δείκτης *outfit* από 0,88 έως 1,19 (MO = 1,02). Για το Χωρικό Προσανατολισμό, ο δείκτης *infit* κυμαινόταν από 0,81 έως 1,11 (MO = 0,95), ενώ ο δείκτης *outfit* από 0,75 έως 1,05 (MO = 0,90). Τέλος, για το Σχηματισμό Νοερών Εικόνων, ο δείκτης *infit* κυμαινόταν από 0,89 έως 1,03 (MO = 0,95), ενώ ο δείκτης *outfit* από 0,78 έως 0,97 (MO = 0,89).

Σε αυτό το σημείο είναι σημαντικό να παρουσιάσουμε τη στρατηγική που ακολουθήσαμε στο πλαίσιο της επιλογής των πιο κατάλληλων στοιχείων για κάθε διάσταση. Πιο συγκεκριμένα, βασιζόμενοι στους δείκτες δυσκολίας (*p*) αλλά και στους δείκτες *infit* και *outfit* (που είναι μια άλλη μορφή του δείκτη διαφοροποίησης *d*), προσαθήσαμε να συμπεριλάβουμε στοιχεία που να καλύπτουν όλο το φάσμα του χαρακτηριστικού που μετρά η κάθε διάσταση. Έτσι, επιλέξαμε περίπου 20% εύκολα στοιχεία, 60% μέτριας δυσκολίας στοιχεία και 20% δύσκολα στοιχεία. Αξίζει να σημειωθεί ότι μερικά στοιχεία, παρά το γεγονός ότι παρουσίαζαν χαμηλότερους δείκτες *infit* και *outfit* από άλλα στοιχεία (μέσα όμως στα επιτρεπτά όρια), επιλέχθηκαν επειδή ακριβώς είχαν βαθμό δυσκολίας ο οποίος δεν καλυπτόταν από κάποιο άλλο στοιχείο. Επιπρόσθετα, σε περίπτωση που κάποια από τα στοιχεία είχαν τον ίδιο βαθμό δυσκολίας, το κριτήριο το οποίο χρη-

Πίνακας 1
Δείκτες Infit και Outfit για όλα τα στοιχεία της αρχικής έκδοσης του TAX

Αριθμός Στοιχείου	Infit	ZSTD	Outfit	ZSTD
Νοητικές Περιστροφές				
8	1,06	0,9	0,93	-0,4
4	1,04	0,6	1,00	0,1
7	0,97	-0,5	0,88	-1,1
10	0,90	-1,70	0,84	-1,5
6	0,94	-1,1	0,96	-0,4
3	1,00	0,1	1,20	2,0
13	1,00	0,00	0,99	-0,1
11	0,97	-0,05	0,97	-0,3
15	0,79	-4,2	0,70	-3,8
5	1,11	2,1	1,16	1,9
14	0,84	-3,4	0,79	-2,9
2	1,17	3,2	1,24	2,8
12	0,93	-1,3	0,89	-1,4
1	1,22	4,0	1,35	4,0
16	0,84	-3,3	0,76	-3,3
9	0,99	-0,3	0,99	-0,1
18	1,02	0,4	1,01	0,2
17	1,06	1,1	1,11	1,1
20	1,12	1,9	1,08	0,7
19	1,03	0,4	1,19	1,3
Χωρικός Προσανατολισμός				
22	0,91	-1,2	0,98	0,0
33	0,82	-2,6	0,57	-2,4
23	0,83	-2,5	0,69	-1,7
24	0,98	-0,2	1,02	0,1
35	0,95	-0,7	0,91	-0,5
27	0,95	-0,9	0,86	-0,9
25	0,82	-3,3	0,74	-2,1
26	0,97	-0,5	0,95	-0,4
34	1,11	1,9	1,05	0,5
28	0,94	-1,0	0,88	-1,1
38	0,98	-0,3	0,92	-0,7
31	0,79	-4,0	0,69	-3,5
37	0,85	-2,8	0,75	-2,7
36	0,99	-0,1	0,94	-0,6
29	1,44	6,3	1,86	6,7
30	0,73	-4,8	0,66	-3,7
21	1,78	9,2	2,94	9,9
32	0,81	-2,9	0,79	-1,7
39	1,24	2,9	1,82	4,1
40	0,90	-0,9	0,72	-1,0

Πίνακας 1
Δείκτες Infit και Outfit για όλα τα στοιχεία της αρχικής έκδοσης του TAX (συνέχεια)

Αριθμός Στοιχείου	Infit	ZSTD	Outfit	ZSTD
Σχηματισμός Νοερών Εικόνων				
50	0,89	-1.9	0,83	-1.8
46	1,08	1.7	1.12	1.5
41	1,02	0.4	1.05	0.7
43	1,02	0.5	0,97	-0.4
48	0,94	-1.4	0,91	-1.3
42	1,07	1.6	1.09	1.4
49	0,94	-1.3	0,91	-1.4
44	1,14	2.7	1.23	2.8
51	1,03	0.5	1,02	0.2
47	0,99	-0.1	0,95	-0.05
58	0.90	-1.7	0.84	-1.8
54	0,99	-0.1	0,97	-0.3
60	0,89	-1.5	0,78	-1.8
59	0,89	-1.5	0,79	-1.7
55	0,92	-1.0	0,95	-0.3
56	1,05	0.6	1.14	1.0
53	0,92	-0.9	0,84	-1.1
57	1,03	0.3	0,99	0.0
45	1,29	2.5	2,06	4.2
52	0,99	0.0	0,80	-0.7

ZSTD: Standardized Residuals

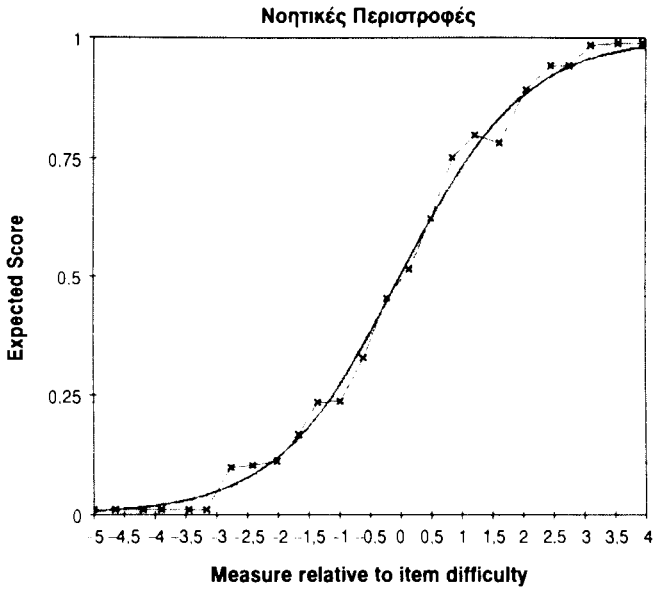
Σημείωση: Τα στοιχεία που είναι σημειωμένα με έντονους χαρακτήρες είναι αυτά που επιλέχθηκαν για την τελική έκδοση του TAX

σιμοποιήθηκε για να επιλέξουμε κάποιο ή κάποια από αυτά, ήταν οι δείκτες infit/outfit. Πάντοτε επιλέγονταν τα στοιχεία με τις καλύτερες τιμές infit/outfit (πιο κοντά στην αναμενόμενη τιμή 1).

Στα Γραφήματα 1, 2 και 3 που ακολουθούν, παρουσιάζονται οι συνολικές ΚΧΕ για καθεμιά από τις τρεις ξεχωριστές διαστάσεις, αφού λόγω του περιορισμένου χώρου δεν είναι δυνατόν να παρουσιάσουμε τις γραφικές παραστάσεις των ΚΧΕ για καθένα από τα στοιχεία που τελικά αποτέλεσαν την τελική έκδοση του TAX. Κύριο χαρακτηριστικό και στις τρεις περιπτώσεις είναι η μορφή που παρουσιάζουν οι ΚΧΕ (σχήμα s), γεγονός που υποδηλώνει την πολύ καλή λειτουργία του κάθε στοιχείου από αυτά που συμπεριλήφθηκαν στην τελική έκδοση του τεστ. Αναλυτικό-

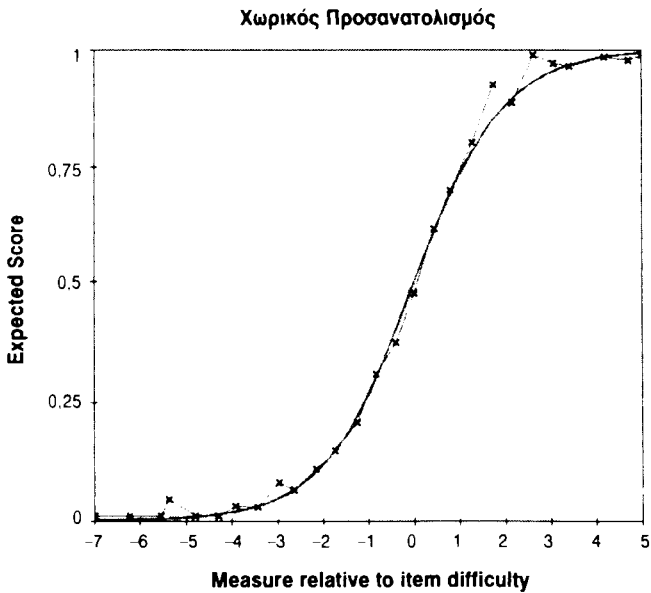
τερα, η σχηματική μορφή s που παρουσιάζει καθεμιά από τις τρεις ΚΧΕ, προτείνει ότι ένα άτομο μεσαίας ικανότητας (μηδέν στον οριζόντιο άξονα) έχει πιθανότητα περίπου 50% να απαντήσει σωστά στα στοιχεία που συνθέτουν τη συγκεκριμένη διάσταση (βλέπε κάθετο άξονα). Επιπρόσθετα, στα αρχικά σε σειρά στοιχεία της κάθε διάστασης η προσδοκία είναι ότι ένα άτομο μεσαίας νοητικής ικανότητας θα απαντήσει σωστά στα περισσότερα από αυτά, ενώ στα τελευταία σε σειρά στοιχεία η προσδοκία είναι ότι θα απαντήσει λανθασμένα στα περισσότερα από αυτά.

Τέλος, στον Πίνακα 2 που ακολουθεί, παρουσιάζονται οι μέσοι όροι, οι τυπικές αποκλίσεις καθώς και οι ενδοσυνάψεις μεταξύ των ξεχωριστών κλιμάκων του TAX.



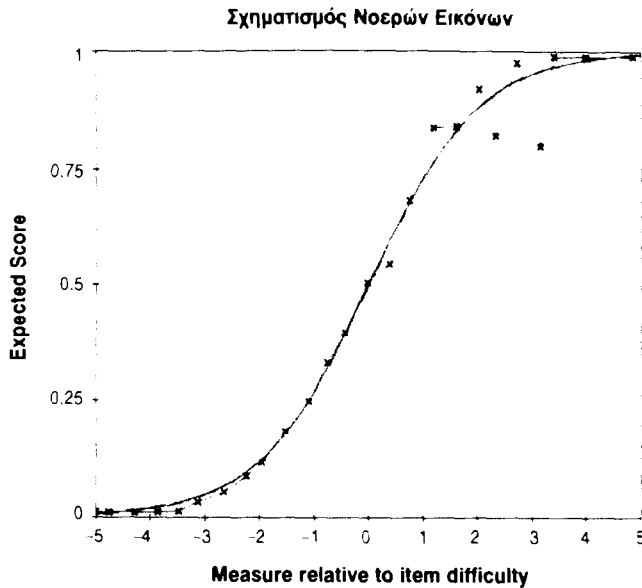
Σχήμα 1

Η καμπύλη των χαρακτηριστικών της ερώτησης για την κλίμακα Νοητικές Περιστροφές



Σχήμα 2

Η καμπύλη των χαρακτηριστικών της ερώτησης για την κλίμακα Χωρικός Προσανατολισμός



Σχήμα 3

Η καμπύλη των χαρακτηριστικών της ερώτησης για την κλίμακα Σχηματισμός Νοερών Εικόνων

Πίνακας 2

Μέσοι όροι, τυπικές αποκλίσεις και ενδοσυνάψεις μεταξύ των υποκλιμάκων του TAX (N=349)

Κλίμακες	ΜΟ	ΤΑ	ΝΠ	ΧΠ	ΣΝΕ	ΣΕ
Νοητικές Περιστροφές	6.42	2.87	1	0.50**	0.35**	0.78**
Χωρικός Προσανατολισμός	6.68	3.14		1	0.36**	0.83**
Σχηματισμός Νοερών Εικόνων	4.08	2.56			1	0.63**
Συνολικό Τεστ	28.84	10.60				1

ΝΠ=Νοητικές Περιστροφές, ΧΠ=Χωρικός Προσανατολισμος, ΣΝΕ=Σχηματισμός Νοερών Εικόνων, ΣΕ=Συνολική Επίδοση

** p<0.01.

Αξιοπιστία του TAX

Για την εκτίμηση της αξιοπιστίας του TAX χρησιμοποιήθηκε ο δείκτης Kuder-Richardson (Kuder & Richardson, 1937), ο οποίος είναι ένας δείκτης υπολογισμού της εσωτερικής συνοχής

του τεστ, κατάλληλος για τεστ που χρησιμοποιούν διχοτομική κλίμακα (π.χ. Σωστό - Λάθος). Επιπρόσθετα, χρησιμοποιήθηκε, ο δείκτης επαναληπτικής χορήγησης (test - retest), ο οποίος αξιολογεί τη σταθερότητα των αποτελεσμάτων στο χρόνο (το διάστημα που μεσολάβησε ανά-

Πίνακας 3
Δείκτες Αξιοπιστίας Kuder-Richardson και επαναληπτικής χορήγησης για το TAX.

Κλίμακες	Δείκτης Αξιοπιστίας Kuder-Richardson (N=349)	Δείκτης Αξιοπιστίας Επαναληπτικής Χορήγησης (N=108)	Δείκτης kappa (N=108)
Νοητικές Περιστροφές	0,72	0,77	0,69
Χωρικός Προσανατολισμός	0,80	0,88	0,81
Σχηματισμός Νοερών Εικόνων	0,69	0,71	0,61
Συνολικό Τεστ	0,85	0,84	0,71

μεσα στις δύο χορηγήσεις ήταν 4 εβδομάδες), ενώ χρησιμοποιήθηκε και ο δείκτης kappa που αξιολογεί το βαθμό συμφωνίας μεταξύ δύο σετ δεδομένων (π.χ. 1η χορήγηση - 2η χορήγηση), ως ένας εναλλακτικός τρόπος αξιολόγησης της σταθερότητας τόσο των υποκλίμακων όσο και του συνολικού τεστ. Τα αποτελέσματα από τους δύο πρώτους δείκτες έδειξαν ότι τόσο οι υποκλίμακες όσο και το συνολικό τεστ παρουσίαζαν δείκτες υψηλότερους από το ελάχιστο αποδεκτό όριο του 0,70 (Nunnally & Berstein, 1994). Τα αποτελέσματα από το δείκτη kappa έδειξαν επίσης αποδεκτές τιμές τόσο για τις υποκλίμακες όσο και για το συνολικό τεστ, καθώς οι τιμές του ξεπερνούσαν το ελάχιστο αποδεκτό όριο του 0,60, αφού ο δείκτης kappa αξιολογείται με διαφορετικό τρόπο από τους υπόλοιπους δείκτες αξιοπιστίας (Landis & Koch, 1977). Στον Πίνακα 3 που ακολουθεί παρουσιάζονται τα αποτελέσματα από την αξιολόγηση της αξιοπιστίας.

Εγκυρότητα του TAX

Αναφορικά με την αξιολόγηση του κατά πόσο το TAX μετρά πράγματι χωρο-αντιληπτική ικανότητα, πραγματοποιήθηκαν τρεις διαφορετικές εκτιμήσεις: συγχρονικής εγκυρότητας, συγκλίνουσας εγκυρότητας και αποκλίνουσας εγκυρότητας. Πιο συγκεκριμένα, αναφορικά με τη συγχρονική εγκυρότητα, το TAX χορηγήθηκε μαζί με δύο ήδη γνωστά τεστ που μετρούν χωρο-αντιλη-

πτική ικανότητα: το Τεστ Χωρικών Σχέσεων της συστοιχίας Διαφορικών Ικανοτήτων και το Τεστ Χωρο-αντιληπτικής Ικανότητας της συστοιχίας Γενικών Ικανοτήτων. Και με τα δύο αυτά τεστ το TAX έδειξε μέτρια προς υψηλή συσχέτιση (από 0,36 έως 0,65), γεγονός που αποδεικνύει ότι αφενός μετρά την ίδια έννοια που μετρούν και τα δύο αυτά αναγνωρισμένα τεστ, αφετέρου μετρά και επιπρόσθετα χαρακτηριστικά της έννοιας που εκείνα δεν μετρούν.

Για την αξιολόγηση της συγκλίνουσας εγκυρότητας, το TAX χορηγήθηκε μαζί με το Τεστ Προοδευτικών Μητρώων του Raven, ένα τεστ που μετρά γενική νοημοσύνη και το οποίο, σύμφωνα με τη θεωρία, θα πρέπει να εμφανίζει συσχέτιση με ένα τεστ Αντίληψης Χώρου, αφού πολλές από τις δεξιότητες που χρησιμοποιούνται για να επιλυθούν τα προβλήματα σε ένα τέτοιο τεστ είναι κοινές με τις δεξιότητες που χρησιμοποιούνται για να επιλυθούν προβλήματα γενικής νοημοσύνης (Raven, Court & Ravan, 1979). Τα αποτελέσματα έδειξαν ότι υπήρχε μέτρια συσχέτιση (από 0,20 έως 0,53), επιβεβαιώνοντας την αρχική υπόθεση. Τέλος, αναφορικά με την αποκλίνουσα εγκυρότητα, το TAX χορηγήθηκε μαζί με ένα μη λεκτικό τεστ της συστοιχίας Γενικών Ικανοτήτων και ένα τεστ Αφαιρετικού Συλλογισμού της συστοιχίας Διαφορικών Ικανοτήτων, δύο έννοιες οι οποίες, ενώ χρησιμοποιούν παρόμοιου τύπου ερεθίσματα ως στοιχεία για να μετρηθούν (γεωμετρικά σχήματα), εντούτοις σύμφωνα με τη θε-

Πίνακας 4
Αποτελέσματα Συγχρονικής, Συγκλίνουσας και Αποκλίνουσας Εγκυρότητας για το TAX

Κλίμακες	Σύγχρονική Εγκυρότητα		Συγκλίνουσα Εγκυρότητα		Αποκλίνουσα Εγκυρότητα
	ΤΧΣ -DAT (N= 120)	ΤΧΑΙ-GAT (N= 120)	Raven (N=115)	ΤΑΣ-DAT (N= 115)	ΜΛΤ-GAT (N= 115)
Νοητικές Περιστροφές	0.52**	0.50**	0.38**	0.20*	0.17
Χωρικός Προσανατολισμός	0.56**	0.48**	0.54**	0.26**	0.11
Σχηματισμός Νοερών Εικόνων	0.48**	0.36**	0.20*	0.16	0.07
Συνολικό Τεστ	0.65**	0.56**	0.53**	0.27**	0.15

ΤΧΣ: Τεστ Χωρικών Σχέσεων. ΤΧΑΙ: Τεστ Χωρο-Αντιληπτικής Ικανότητας. ΤΑΣ: Τεστ Αφαιρετικού Συλλογισμού. ΜΛ: Μη Λεκτικό Τεστ. DAT: Differential Aptitude Test. GAT: General Ability Test.

* $p < 0.05$, ** $p < 0.01$

ωρία διαφέρουν από την αντίστοιχη της χωρο-αντιληπτικής ικανότητας, καθώς οι στρατηγικές τεχνικές που χρησιμοποιούν τα άτομα για να επιλύσουν τα προβλήματα που περιλαμβάνονται σε αυτά τα τεστ είναι διαφορετικές. Και πάλι, η έρευνα έδειξε ότι με το μεν μη λεκτικό τεστ της συστοιχίας Γενικών Ικανοτήτων δεν υπάρχει καμιά συσχέτιση (από 0.07 έως 0.17), ενώ με το τεστ Αφαιρετικού Συλλογισμού της συστοιχίας Διαφορικής Ικανότητας υπάρχει πολύ χαμηλή συσχέτιση (< 0.30), επιβεβαιώνοντας με αυτό τον τρόπο πλήρως την αρχική υπόθεση. Στον Πίνακα 4 που ακολουθεί παρουσιάζονται τα αποτελέσματα από τις έρευνες εγκυρότητας.

4. Συζήτηση

Βασικός στόχος της συγκεκριμένης έρευνας ήταν η κατασκευή ενός τεστ μέτρησης της χωρο-αντιληπτικής ικανότητας, το οποίο να αντανάκλα ένα σύγχρονο θεωρητικό μοντέλο και να παρουσιάζει ισχυρά ψυχομετρικά χαρακτηριστικά (αξιοπιστία και εγκυρότητα). Το TAX βασίζεται στο θεωρητικό μοντέλο του Lohman (1979), σύμφωνα με το οποίο η χωρο-αντιληπτική ικανότητα περιλαμβάνει τρεις ξεχωριστές και αυτόνο-

μες διαστάσεις: (α) το Σχηματισμό Νοερής Εικόνας, (β) το Χωρικό Προσανατολισμό και (γ) τις Χωρικές Σχέσεις. Στην αρχική του έκδοση το TAX περιλάμβανε 60 στοιχεία και από τις τρεις βασικές διαστάσεις (20 για κάθε διάσταση). Η δομή του κάθε στοιχείου περιλάμβανε ένα γεωμετρικό σχήμα ως στέλεχος και μια σειρά από εναλλακτικές επιλογές ως απαντήσεις, από τις οποίες το άτομο θα έπρεπε να επιλέξει μία ως σωστή.

Για την κατασκευή του TAX χρησιμοποιήθηκε η μέθοδος της Θεωρίας του Υπολογισμού των Ικανοτήτων, μέθοδος η οποία παρουσιάζει σημαντικά πλεονεκτήματα σε σχέση με την Κλασική Θεωρία, η οποία παραδοσιακά χρησιμοποιείται για την κατασκευή ψυχομετρικών εργαλείων. Το μεγαλύτερο πλεονέκτημα της συγκεκριμένης μεθόδου είναι ότι τα τεστ που προκύπτουν από τη χρήση της δεν βασίζονται στις ιδιότητες των στοιχείων που συνθέτουν το κάθε τεστ ούτε και στις ιδιότητες των ατόμων που συμμετέχουν στη φάση της κατασκευής τους, αλλά μόνο στο επίπεδο κατοχής του χαρακτηριστικού που μετρά το τεστ. Αυτό έχει ως αποτέλεσμα μια σειρά από πλεονεκτήματα σε σχέση με τα τεστ που έχουν προκύψει από την Κλασική Θεωρία, όπως ότι μπορούμε να συγκρίνουμε άτομα διαφορετικού επιπέδου ακόμη και αν και έχουν συ-

μπληρώσει διαφορετικά τεστ, ή να κάνουμε χρήση της μεθόδου αναπροσαρμοσμένης αξιολόγησης με ηλεκτρονικό υπολογιστή (computerized adaptive testing).

Στο πλαίσιο της πιλοτικής φάσης της κατασκευής του TAX, από τα 60 αρχικά στοιχεία επιλέχθηκαν τα καλύτερα 12 για κάθε διάσταση προνοώντας να επιλεγθούν στοιχεία που θα καλύπτουν όλο το φάσμα της ικανότητας που μετράμε. Έτσι, επιλέχθηκαν περίπου 20% εύκολα στοιχεία, 60% μέτριας δυσκολίας στοιχεία και 20% δύσκολα στοιχεία, φροντίζοντας καθένα από αυτά να έχει υψηλό βαθμό διαφοροποίησης, ο οποίος αξιολογήθηκε με τη βοήθεια των δεικτών *infit* και *outfit*. Η τελική έκδοση του TAX περιλάμβανε 3 ξεχωριστά μέρη (ένα για κάθε διάσταση). Κάθε μέρος περιλάμβανε 12 στοιχεία δίνοντας ένα σύνολο 36 στοιχείων.

Για την τεκμηρίωση των ψυχομετρικών χαρακτηριστικών του TAX πραγματοποιήθηκαν τόσο έρευνες για την αξιολόγηση της αξιοπιστίας του όσο και έρευνες για την αξιολόγηση της εγκυρότητάς του. Αναφορικά με την αξιολόγηση της αξιοπιστίας του TAX χρησιμοποιήσαμε το δείκτη εσωτερικής συνοχής Kuder-Richardson (Kuder & Richardson, 1937), το δείκτη επαναληπτικών μετρήσεων, καθώς και το δείκτη kappa. Από τα αποτελέσματα διαπιστώσαμε ότι και οι τρεις δείκτες αξιοπιστίας έδειξαν ότι το TAX είναι ένα αξιόπιστο τεστ. Πιο συγκεκριμένα, ο δείκτης εσωτερικής συνοχής τόσο για τις υποκλίμακες όσο και το συνολικό τεστ κυμάνθηκε από 0,69 έως 0,85, τιμές που υποδηλώνουν ότι τα στοιχεία που συνθέτουν τις διαφορετικές υποκλίμακες αλλά και ολόκληρο το τεστ μετρούν το ίδιο εννοιολογικό κατασκεύασμα. Επιπρόσθετα, και ο δείκτης επαναληπτικών μετρήσεων κυμάνθηκε μέσα στα αποδεκτά όρια ($>0,70$), αφού τόσο το συνολικό τεστ όσο και οι διαφορετικές υποκλίμακες κυμάνθηκαν από 0,71 έως 0,88, υποδηλώνοντας ότι το τεστ μετρά με ακρίβεια τις έννοιες που υποστηρίζει ότι μετρά. Τέλος, αναφορικά με το δείκτη kappa, ο βαθμός συμφωνίας ανάμεσα στην πρώτη και τη δεύτερη χορήγηση κυμάνθηκε από 0,61 έως 0,81, τιμές οι

οποίες σύμφωνα με τους Landis και Koch (1977) κρίνονται ως ικανοποιητικές.

Για την εκτίμηση της εγκυρότητας του TAX πραγματοποιήθηκαν έρευνες αξιολόγησης τόσο της συγχρονικής εγκυρότητας όσο και της εγκυρότητας εννοιολογικής κατασκευής (συγκλίνουσα και αποκλίνουσα). Πιο συγκεκριμένα, αναφορικά με την αξιολόγηση της συγχρονικής εγκυρότητας, τα αποτελέσματα έδειξαν ότι το TAX συσχετίζεται σε μέτριο επίπεδο με δύο τεστ που αποδεδειγμένα μετρούν χωρο-αντιληπτική ικανότητα: το τεστ Χωρικών Σχέσεων της συστοιχίας Διαφορικών Ικανοτήτων και το τεστ Χωρο-αντιληπτικής Ικανότητας της συστοιχίας Γενικών Ικανοτήτων (0,65 και 0,56, αντίστοιχα). Αυτοί οι δείκτες συσχέτισης κρίνονται πολύ ικανοποιητικοί εάν σκεφτούμε ότι τα δύο συγκεκριμένα τεστ μετρούν τη χωρο-αντιληπτική ικανότητα χρησιμοποιώντας στοιχεία από μία μόνο διάσταση (αναδιπλώσεις). Στο ίδιο επίπεδο κυμαίνονται και οι δείκτες συσχέτισης ανάμεσα στις διαφορετικές διαστάσεις του TAX με τα δύο αυτά τεστ (από 0,36 έως 0,56), υποδηλώνοντας ότι και οι υποκλίμακες του TAX μετρούν έννοιες σχετικές με τη χωρο-αντιληπτική ικανότητα.

Όσον αφορά τη συγκλίνουσα εγκυρότητα, τα αποτελέσματα έδειξαν μια μέτρια συσχέτιση με το τεστ γενικής νοημοσύνης Προοδευτικές Μητρες του Raven (0,53), επιβεβαιώνοντας και δεδομένα από άλλες έρευνες που έχουν δείξει ότι υπάρχει συσχέτιση ανάμεσα τη γενική νοητική ικανότητα και την αντίληψη χώρου (Hyde & Trickey, 1995. Smith & Whetton, 1988). Αναφορικά με την αποκλίνουσα εγκυρότητα, η χαμηλή συσχέτιση ($r > 0,30$) του TAX και των υποκλιμάκων του με το τεστ Αφαιρετικού Συλλογισμού της συστοιχίας Διαφορικών Ικανοτήτων καθώς και η μη συσχέτισή τους με το Μη Λεκτικό τεστ της συστοιχίας Γενικών Ικανοτήτων, υποδηλώνουν ότι το TAX μετρά κάτι διαφορετικό από τις παραπάνω έννοιες, παρά το γεγονός ότι τα στοιχεία που χρησιμοποιούνται για να μετρηθούν αυτές οι έννοιες είναι παρόμοια με τα στοιχεία του TAX (γεωμετρικά σχήματα). Αυτό θα μπορούσε να οδηγήσει κάποιον να ισχυριστεί ότι υπάρχει

εννοιολογική συγγένεια ανάμεσα στην έννοια που μετρά το TAX και στις έννοιες που μετρούν τα συγκεκριμένα εργαλεία.

Ένας επιπλέον στόχος αυτής της έρευνας ήταν να μελετηθεί και το κατά πόσο ο Χωρικός Προσανατολισμός (ΧΠ) και οι Χωρικές Σχέσεις (ΧΣ) είναι ανεξάρτητες μεταξύ τους ή συνθέτουν μία κοινή διάσταση, όπως υποστηρίζουν κάποιοι ερευνητές (Carroll, 1993. Colom et al., 2003). Από τη μελέτη των ενδοσυναφειών μεταξύ των υποκλιμάκων του TAX διαπιστώνουμε ότι ναι μεν υπάρχει κάποια συσχέτιση ανάμεσα στις δύο αυτές διαστάσεις ($r=0,50$), όμως όπως αποκαλύπτει ο δείκτης προσδιορισμού (r^2) το κοινό ποσοστό της διακύμανσης που ερμηνεύουν οι δύο αυτές έννοιες είναι μόλις 25%. Κατά συνέπεια μπορούμε να υποστηρίξουμε ότι αυτές οι δύο έννοιες είναι σχετικά ανεξάρτητες μεταξύ τους και ορθά συμπεριληφθήκαν ως ξεχωριστές διαστάσεις στο νέο αυτό εργαλείο, καθώς φαίνεται ότι μετρούν διαφορετικές έννοιες.

Ένα από τα σημαντικά μειονεκτήματα της συγκεκριμένης έρευνας ήταν η χρήση ευκαιριακού δείγματος. Συνεπώς, είναι αναγκαία η χορήγηση της τελικής έκδοσης του TAX σε ένα αντιπροσωπευτικό δείγμα του πληθυσμού με στόχο τόσο την επιβεβαίωση της παραγοντικής του δομής, όσο και τη δημιουργία ενός συστήματος νορμών, ώστε να είναι δυνατή η αξιολόγηση των επιδόσεων που προκύπτουν από τη χορήγησή του. Νόρμες θα πρέπει να κατασκευαστούν τόσο για το γενικό πληθυσμό, όσο και για επιμέρους πιο συγκεκριμένες ομάδες του πληθυσμού, όπως, φοιτητές, μαθητές, εργαζόμενοι κ.λπ. Ένα άλλο μειονέκτημα της συγκεκριμένης έρευνας είναι η έλλειψη δεδομένων τα οποία να τεκμηριώνουν την προβλεπτική εγκυρότητα του TAX. Μελλοντικές έρευνες θα πρέπει να επικεντρωθούν σε αυτή την κατεύθυνση, ώστε να ενισχυθούν ακόμη περισσότερο τα ελπιδοφόρα εμπειρικά δεδομένα που βρέθηκαν σε αυτή την έρευνα.

Ως προτάσεις για μελλοντικά ερευνητικά βήματα, θα είχε ενδιαφέρον να μελετηθεί η χρησιμότητα του συγκεκριμένου εργαλείου σε μια σειρά από διαφορετικά πλαίσια και διαδικασίες,

όπως η επιλογή ατόμων σε εργασιακούς χώρους αλλά και εκπαιδευτικά ιδρύματα όπου η αντίληψη χώρου παίζει σημαίνοντα ρόλο (π.χ. αρχιτέκτονες, μηχανολόγοι, ναυπηγοί, πολιτικοί μηχανικοί κ.λπ.). η διάγνωση ατόμων με νευρολογικές διαταραχές ή εγκεφαλικές βλάβες κ.λπ. Μια επιπρόσθετη έρευνα θα μπορούσε να αφορά τη μελέτη της επίδρασης των δημογραφικών χαρακτηριστικών, όπως του φύλου και της ηλικίας στη χωρο-αντιληπτική ικανότητα καθώς και της επίδρασης της εξοικείωσης στη χωρο-αντιληπτική ικανότητα. Ενδιαφέρον θα είχε επίσης η μελέτη της επίδρασης της εκτεταμένης ενασχόλησης με τρισδιάστατα ηλεκτρονικά παιχνίδια στην ανάπτυξη της αντίληψης χώρου.

Συνοψίζοντας, η συγκεκριμένη έρευνα παρουσίασε την κατασκευή ενός εργαλείου μέτρησης της χωρο-αντιληπτικής ικανότητας χρησιμοποιώντας τη θεωρία του Υπολογισμού των Ικανοτήτων ως μέθοδο επιλογής των στοιχείων που συνθέτουν το τεστ. Η αξιολόγηση των ψυχομετρικών χαρακτηριστικών του νέου τεστ τεκμηρίωσε το TAX ως ένα αξιόπιστο και έγκυρο τεστ, το οποίο μπορεί να χρησιμοποιηθεί για τη μέτρηση της αντίληψης του χώρου στον ελληνικό πληθυσμό.

Βιβλιογραφία

- Ackerman, P. L. & Kanfer, R. (1993). Integrating laboratory and field study for improving selection: Development of a battery for predicting air traffic controller success. *Journal of Applied Psychology*, 78, 413-432.
- Allen, G. L. (1982). *Assessment of visuospatial abilities using complex cognitive tasks*. Norfolk, VA: Department of Psychology, Old Dominion University.
- Anderson, R. O. (1994). *The theoretical development and empirical evaluation of a logistic regression, paired-comparisons procedure for assessing unidimensionality in the Rasch model*. Doctoral dissertation, University of Minnesota, Minneapolis, MN.

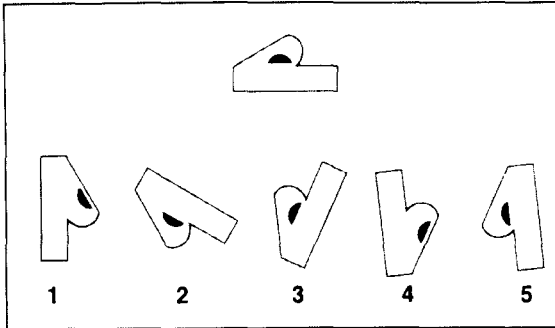
- Bennett, G. K., Seashore, H. G. & Wesman, A. G. (1974). *Manual for the Differential Aptitude Test* (5th ed.). New York: The Psychological Corporation.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M. R. Novick. *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Bishop, A. J. (1980). Spatial abilities and mathematics: Education review. *Educational Studies in Mathematics*, 11, 257-269.
- Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch model* (2nd ed.). London: Lawrence Erlbaum.
- Bortoline, G. R. (1998). Visual science: an emerging discipline. *Journal of Geometry and Graphics*, 2, 181-187.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Clements, D. H. & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.). *Handbook of research on mathematics, teaching and learning*. New York: Macmillan.
- Colom, R., Contreras, J. M., Chun Shih, P. & Santacreu, J. (2003). The assessment of spatial ability with a single computerized test. *European Journal of Psychological Assessment*, 19, 92-100.
- Eliot, J. C. & Smith, I. M. (1983). *An international directory of spatial tests*. Windsor: NFER-Nelson.
- Ferguson, E. S. (1977). The mind's eye: non verbal thought in technology. *Science*, 827-836.
- Gallagher, S. A. (1989). Predictors of SAT mathematics scores in a causal model of mathematics achievement. *Journal of Research in Mathematics Education*, 15, 361-377.
- Galton, F. (1880). Statistics of mental imagery. *Mind*, 5, 300-318.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligence*. New York: Basic.
- Gibson, J. J. (Ed.) (1947). *Motion picture testing and research. Army air forces aviation psychology program*, (Report No. 7). Washington, DC: GPO.
- Gimmestad, B. B. & Sorby, S. A. (1996). Making connections: spatial skills and engineering drawing. *Mathematics Teachers*, 89, 348-353.
- Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.
- Hegarty, M. & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.). *The Cambridge handbook of visuospatial thinking* (pp. 121-169). Cambridge: Cambridge University Press.
- Hyde, G. & Trickey, G. (1995). *Differential Aptitude Tests (DAT) for Guidance: Dat for Guidance. The Manual*. New York: The Psychological Corporation.
- Kuder, G. F. & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2, 151-160.
- Kyllonen, P. C. (1984). Information processing analysis of spatial ability. (Doctoral dissertation, Stanford University) *Dissertation Abstracts International*, 45, 819A.
- Landis J. R., Koch G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33, 159-174.
- Linacre, J. M. (1999). *A user's guide and manual to winsteps, 5.0*. Chicago: MESA press.
- Linacre, J. M. (2003). Size vs. significance: Standardized chi-square fit statistic. *Rasch Measurement Transactions*, 17, 918.
- Linn, M. C. & Petersen, A. C. (1985). Emergence and characterisation of sex differences in spatial ability: a meta-analysis. *Child Development*, 56, 1479-1498.
- Lohman, D. F. (1979). Spatial ability: A review and reanalysis of the correlational literature (Tech. Rep. No. 8), Stanford, CA: Stanford University, Aptitude Research project, School of Education. (NTIS NO. AD-A075 972).
- Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 181-248). Hillsdale, NJ: Erlbaum.

- Lohman, D. F. (1993). Effects of practice and training on the acquisition and transfer of spatial skills: Two speed-accuracy studies. (Final Report Grant AFOSR-91-9367). Iowa City, IA: Lindquist Center for Measurement.
- Lord, T. R. (1990). Enhancing learning in the life sciences through spatial perception. *Innovative Higher Education*, 15, 5-16.
- Lorenz, C. A. & Neisser, U. (1986). *Ecological and psychometric dimensions of spatial ability*. Atlanta, GA: Department of Psychology, Emory University, Report #10.
- McGee, M. G. (1979). Human spatial abilities: psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, 86, 889-918.
- Meeker, M. (1991). How do arithmetics and maths differ? *Gifted Child Today*, 14, 6-7.
- National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.
- Nunnally, J. & Berstein, I. (1994). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.
- Pellegrino, J. W. & Hunt, E. B. (1989). Computer-controlled assessment of static on dynamic spatial reasoning. In R. F. Dillon & J. W. Pellegrino (Eds), *Testing: Theoretical and applied perspectives* (pp. 174-198). New York: Praeger.
- Pribyl, J. R. & Bonder, G. M. (1987). Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of Research in Science Teaching*, 24, 229-240.
- Rafi, A., Anuar, K., Samad, A., Hayati, M. & Mahadzir, M. (2005). Improving spatial ability using a web-based virtual environment. *Automation in Construction*, 14, 707-715.
- Raven, J. C., Court, J. H. & Raven, J. (1979). *Manual for Raven's Progressive matrices and Vocabulary Scales*. London: Lewis.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.
- Seibert, W. F. & Snow, R. E. (1965). *Studies in cine-psychometry I: Preliminary factor analysis of visual cognition and memory*. Lafayette, IN: Audio Visual Center, Purdue University.
- Shepard, R. & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science*, 171, 701-703.
- Smith, E. V. & Smith, R. M. (2004). *Introduction to Rasch measurement: Theory, models and applications*. Mapple Grove, MN: JAM Press.
- Smith, P. & Whetton, C. (1988). *General Ability Tests: User's guide*. Windsor: NFER-Nelson.
- Spearman, C. E. (1927). *The abilities of man: their nature and measurement*. London: Macmillan.
- Swaminathan, H., Hambleton, R. K. & Rogers, H. J. (2007). Assessing the fit of item response theory. *Handbook of Statistics*, 26, 683-718.
- Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs*, 1.
- Wechsler, D. (1955). *Wechsler adult intelligence scale*. New York: Psychological Corporation.

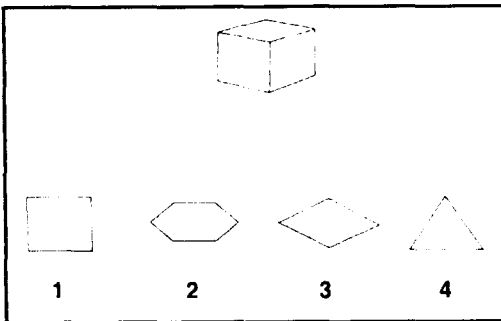
Παράρτημα

Παραδείγματα αντιπροσωπευτικών στοιχείων από κάθε διάσταση του TAX

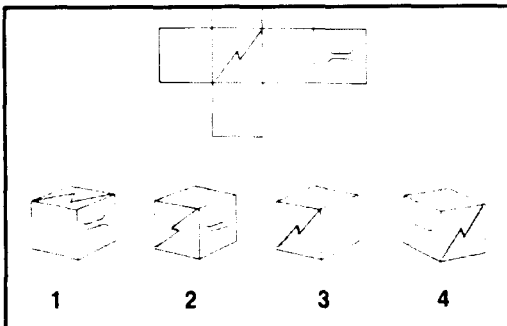
Νοητικές Περιστροφές



Χωρικός Προσανατολισμός



Σχηματισμός Νοερών Εικόνων



Measuring visuospatial ability: development and psychometric properties of the Spatial Perception Test

IOANNIS TSAOUSIS¹

ABSTRACT

This paper describes the development and psychometric evaluation of a novel tool for the assessment of visuospatial ability, based on Lohman's model (1979).

The Spatial Perception Test (SPT) is comprised of 36 items measuring three distinct visuospatial skills: (a) mental rotation, (b) spatial orientation, and (c) construction of mental images. Four groups of participants contributed to the development of the test (N=694). Final item selection was based on Item Response Theory (Rasch model) demonstrating adequate model fit. Internal consistency (Kuder-Richardson index = 0.85) and test-retest reliability indices were very good ($r=0.84$). Concurrent validity estimates were obtained through comparison with The Differential Aptitude Test-Space Relations and The General Ability Test-Spatial Ability demonstrating moderate correlations ($r=0.52 - 0.65$ and $r=0.36 - 0.56$, respectively). Both convergent and divergent validity estimates were satisfactory, as indicated by moderate correlations with Raven Progressive Matrices scores ($r=0.20 - 0.53$), and very low correlations with tests that measure abstract and non-verbal reasoning ($r=0.16 - 0.27$ and $r=0.07 - 0.17$, respectively). Our findings support the reliability and validity of SPT as a measure of visuo-spatial ability.

Key words: Visuospatial ability, Theory

1. Address: Ioannis Tsaousis, University of Crete, Department of Psychology, Gallos Campus, Rethymno 74100. Tel. 28310-77519. e-mail: tsaousis@psy.soc.uoc.gr.