



Αναζητήσεις στη Φυσική Αγωγή & τον Αθλητισμό
Τόμος 11 (2), 79 – 98
Δημοσιεύτηκε: Σεπτέμβριος 2013



Inquiries in Sport & Physical Education
Volume 11 (2), 79 - 98
Released: September 2013

www.pe.uth.gr/emag

ISSN 1790-3041

Κατασκευή Κλίμακας Διαβαθμισμένων Κριτηρίων (rubric) για την Αξιολόγηση του Ελληνικού Παραδοσιακού Χορού. Έλεγχος Εγκυρότητας και Αξιοπιστίας

Αθηνά Πίτση¹, Νικόλαος Διγγελίδης² & Φίλιππος Φιλίππου¹

¹ΤΕΦΑΑ, Δημοκρίτειο Πανεπιστήμιο Θράκης

²ΤΕΦΑΑ, Πανεπιστήμιο Θεσσαλίας

Περίληψη

Σκοπός της εργασίας ήταν η κατασκευή και ο έλεγχος εγκυρότητας και αξιοπιστίας αναλυτικής ρούμπρικας ειδικής για την αξιολόγηση της απόδοσης ελληνικών παραδοσιακών χορών. Η κατασκευή της βασίστηκε στα δύο θεμελιώδη χαρακτηριστικά της: α) τον κατάλογο των κριτηρίων, ή "αυτό που μετράει" και β) τις διαβαθμίσεις της ποιότητας απόδοσης κάθε κριτηρίου. Δύο ειδικοί του Ελληνικού παραδοσιακού χορού συμφώνησαν μεταξύ τους σχετικά με την εννοιολογική καταλληλότητα κάθε κριτηρίου, την επάρκεια του αριθμού τους και την ακρίβεια της περιγραφής κάθε επιπέδου απόδοσης, θεμελιώνοντας την εγκυρότητα περιεχομένου της ρούμπρικας. Έτσι, για την κατασκευή της επιλέχθηκαν οκτώ κριτήρια: η αναγνώριση του Χορού, η ανταπόκριση στο ρυθμικό πρότυπο, ο συγχρονισμός, η ακολουθία βημάτων, η κατεύθυνση-σχήμα και η λαβή χορού, η θέση και η στάση του σώματος, τα ποιοτικά στοιχεία κίνησης και η εκφραστικότητα. Επίσης, για κάθε κριτήριο ορίστηκαν πέντε επίπεδα απόδοσης: μη αποδεκτό, οριακά αποδεκτό, μέτρια αποδεκτό, πολύ καλό και άριστο. Ο έλεγχος της συγκλίνουσας εγκυρότητας πραγματοποιήθηκε συσχετίζοντας την αξιολόγηση μέσω ρούμπρικας με την υποκειμενική και ολιστική αξιολόγηση ενός ειδικού. Οι τιμές του συντελεστή συσχέτισης r κυμάνθηκαν από 0.845 $p < 0.01$ έως 0.904 $p < 0.01$. Για τον έλεγχο της αξιοπιστίας της ρούμπρικας αξιολογήθηκε η ικανότητα δεκατεσσάρων φοιτητών, χορευτικά αρχάριων, ηλικίας 18-20 ετών, να αναγνωρίσουν τρεις ελληνικούς χορούς από τη μελωδία τους, να ανταποκριθούν στο ρυθμικό τους πρότυπο με κρούσεις παλαμών και να αποδώσουν την κινητική τους φόρμα. Η απόλυτη συμφωνία μεταξύ των κριτών στη συνολική βαθμολογία κάθε χορού, κυμάνθηκε από 79% έως 93%. Αντίστοιχα, οι τιμές της σχετικής συμφωνίας ήταν στο 100%, του δείκτη ICC single από 0.908 έως 0.956, του ICC average από 0.940 έως 0.983, του Spearman's r από 0.921 έως 0.940 και του Cronbach's α από 0.950 έως 0.982. Επιπλέον, η απόλυτη συμφωνία μεταξύ δύο διαφορετικών βαθμολογιών του ίδιου κριτή, κυμάνθηκε από 79% έως 100% για τους τρεις χορούς. Αντίστοιχα, οι τιμές της σχετικής συμφωνίας ήταν στο 100%, του δείκτη ICC single από 0.964 έως 0.994, του ICC average από 0.982 έως 0.997, του Spearman's r από 0.956 έως 0.994 και του Cronbach's α από 0.981 έως 0.998. Συμπεραίνεται ότι η παραπάνω ρούμπρικα, αποτελεί ένα έγκυρο και αξιόπιστο στη χρήση του εργαλείο, με το οποίο είναι δυνατόν αντικειμενικά, δίκαια και με συνέπεια να τεκμηριωθεί το επίπεδο χορευτικών δεξιοτήτων των φοιτητών.

Λέξεις κλειδιά: *εγκυρότητα, αξιοπιστία, ρούμπρικα, ελληνικός χορός*

Construction Scale Classified Criteria (rubric) for the Evaluation of Greek Traditional Dance. Validity and Reliability

Athina Pitsi¹, Nikolaos Diggelidis², & Philippos Philippou¹

¹Department of Physical Education and Sports Sciences, Democritus University of Thrace, Komotini, Hellas

²Department of Physical Education and Sports Sciences, University of Thessaly, Trikala, Hellas

Abstract

The purpose of the study was to develop and examine the validity and reliability of an analytic rubric for the evaluation Greek traditional dance performance. The development of the rubric was based on the fundamentals of the list of criteria or "what it measures" and the performance levels of each criterion. Two experts Greek traditional dance agreed among themselves on the conceptual relevance of each criterion for the adequacy of their numbers and the accuracy of the description of each performance level, establishing the content validity of rubrics. So for the construction eight criteria were selected, dance identification, response to the rhythm of the dance, synchronization, step sequence, direction-scheme and type of hand holding of the dance, body stance and position, quality characteristics of the movement, and expressiveness. Also, five performance levels were defined for each criterion, not acceptable, marginally acceptable, moderate acceptable, very good, and excellent. From the grading of each criterion a total score for the dance was calculated. Convergent validity was examined correlating assessment through rubrics with the subjective and holistic evaluation of an expert. The values of the correlation coefficient r ranged from 0.845 $p < 0.01$ to 0.904 $p < 0.01$. In order to check rubric reliability fourteen students, beginners to the dance experience (age 18-20 yrs) were examined in their ability to identify three traditional dances from their music, to respond to the rhythm of the dance with hand clapping and perform the kinetic form of each dance. Absolute agreement between raters on the total score of the dances ranged from 79% to 93%, adjacent agreement was 100%, single ICC from 0.908 to 0.956, average ICC from 0.94 έως 0.983, Spearman's r from 0.921 to 0.940 and Cronbach's α from 0.95 to 0.982. In addition, the absolute agreement between two grading of the same rater ranged from 79% to 100%, adjacent agreement was 100%, single ICC from 0.964 to 0.994, average ICC from 0.982 to 0.997, Spearman's r from 0.956 έως 0.994 and Cronbach's α from 0.981 to 0.998. It is concluded that the above rubric is a valid and reliable to use tool, which can objectively, fairly and consistently documented the level of dancing skills of the students.

Key words: *validity, reliability, rubric, Greek traditional dance*

Εισαγωγή

Η αξιολόγηση, αναπόσπαστο κομμάτι της εκπαιδευτικής διαδικασίας, χρησιμοποιείται ως μέσο βελτίωσης της παρεχόμενης εκπαίδευσης. Στον τομέα της διδακτικής χρησιμοποιείται για να δηλώσει τη διαδικασία μέτρησης της μάθησης, τόσο ποιοτικά όσο και ποσοτικά, ως αποτέλεσμα της διδασκαλίας και της χρήσης των πιο πρόσφορων διδακτικών μέσων. Ωθεί τους εκπαιδευτικούς στο να εξετάζουν και να εφαρμόζουν διάφορες τεχνικές με στόχο την ικανοποίηση των μαθησιακών αναγκών των εκπαιδευομένων. Η αξιολόγηση υψηλής ποιότητας και η διαδικασία προσδιορισμού της απόδοσης εξαρτώνται από την ακριβή και αξιόπιστη μέτρηση των βασικών παραγόντων απόδοσης. Η εκτέλεση π.χ. μιας κινητικής δραστηριότητας ή το τελικό αποτέλεσμα μιας εργασίας ή μιας μαθησιακής ικανότητας, δύσκολα αξιολογούνται αν κάποιος δεν εφαρμόσει την ανάλογη τεχνική, με άλλα λόγια, αν δεν έχει το κατάλληλο εργαλείο. Στις μέρες μας όλο και περισσότερο εξελιγμένα συστήματα μετρήσεων απαιτούνται για την αξιολόγηση ή την εκτίμηση των συστημάτων σκέψης, διαδικασίας απόκτησης της γνώσης, καθώς και ανάπτυξης στάσεων-συμπεριφορών (Wiggins & McTighe, 2005).

Ο χορός διεθνώς έχει ενταχθεί στα πλαίσια της εκπαιδευτικής διαδικασίας, διαδραματίζοντας πολλούς ρόλους, λόγω της κοινωνικής, συναισθηματικής, σωματικής και γνωστικής του επίδρασης στο άτομο. Αναφέρεται ότι είναι μια εξαιρετικά πολύπλοκη εμπειρία, που καθιστά δύσκολη την προσπάθεια να μετρηθεί η εκπαιδευτική του αξία (Hagoood, 2001). Υποστηρίζεται επίσης ότι αποτελεί περισσότερο ένα πεδίο έρευνας παρά μια κοινωνική δραστηριότητα και βρίσκεται μέσα στα πλαίσια της φυσικής αγωγής και των τεχνών, με αποτέλεσμα να περιπλέκεται περισσότερο το έργο της αξιολόγησης (Oreck, 2007). Ωστόσο, η αναγκαιότητα της αξιολόγησης του τρόπου με τον οποίο αποδίδεται οποιοδήποτε είδος χορού, προκύπτει από το γεγονός αφενός της υποχρέωσης των δασκάλων να παράσχουν αναφορά για την πρόοδο των μαθητών τους είτε στους ίδιους είτε στους γονείς τους είτε στην εκπαιδευτική μονάδα που ανήκουν και αφετέρου μέσω αυτής, να αποτιμήσουν την παρεχόμενη

εκπαίδευση με στόχο τη βελτίωση.

Στην Ελλάδα ο παραδοσιακός χορός αποτελεί αντικείμενο φυσικής αγωγής, τόσο της πρωτοβάθμιας όσο και της δευτεροβάθμιας εκπαίδευσης και είναι ενταγμένος στα Αναλυτικά Προγράμματα Σπουδών των Τμημάτων Φυσικής Αγωγής και Αθλητισμού και των Παιδαγωγικών Σχολών. Κρίνεται πρωταρχικής σημασίας η κατασκευή και η χρήση έγκυρων και αξιόπιστων οργάνων για την αξιολόγηση της χορευτικής απόδοσης των μαθητών/σπουδαστών, με στόχο τη βελτίωση της εκτέλεσης του χορού και της γενικότερης χορευτικής προόδου τους. Βάσει των αναλυτικών προγραμμάτων φυσικής Αγωγής, Α/βάθμιας και Β/βάθμιας εκπαίδευσης, η αξιολόγηση των μαθητών στο αντικείμενο του παραδοσιακού χορού πρέπει να εστιάζεται στα στοιχεία τα οποία έχουν τεθεί ως στόχοι, να λειτουργεί με τρόπο ώστε να εξασφαλίζεται η αύξηση της παρακίνησης των μαθητών και η ανατροφοδότησή τους και να χρησιμοποιείται για τη διαπίστωση της επίτευξης των στόχων του μαθήματος και τη βαθμολόγηση των μαθητών, προτείνοντας ως εργαλεία διάφορα κινητικά, γνωστικά και συναισθηματικά τεστ ομαδικού επιπέδου, αλλά και φύλλα κριτηρίων με τα οποία αξιολογείται κυρίως ο ψυχοκινητικός, ο γνωστικός και ο συναισθηματικός τομέας (Φυσική Αγωγή στη Β/θμια εκπαίδευση, 2006).

Ωστόσο, η κατασκευή και η χρήση εργαλείων για την αξιολόγηση και τη βαθμολόγηση του μαθητή βρίσκεται σε ερευνητικό στάδιο. Ειδικότερα, κατασκευάστηκε όργανο αξιολόγησης της χορευτικής επίδοσης αποτελούμενο από λίστα κριτηρίων, τα οποία βασίζονται στη θεωρία ανάλυσης της κίνησης του Laban. Το συγκεκριμένο όργανο μπορεί να λειτουργήσει ως έγκυρη και αξιόπιστη μέθοδος αξιολόγησης της χορευτικής επίδοσης, αν χρησιμοποιηθεί από ειδικά εκπαιδευμένους κριτές (Dania, Koutsouba, Hatziharistos & Tyronola, 2010).

Τα κοινά χαρακτηριστικά στοιχεία κάθε χορευτικής δομής, αποτέλεσμα καρτελών δομικής ανάλυσης πέντε χορών, με βάση την αναλυτική μορφολογική μέθοδο των Martin και Pessovar (1961) και την ταξινομική χορολογική μέθοδο του Dance Study Group of the International Folk Music Council (1974), αποτέλεσαν τη βάση για την κατασκευή οργάνου αξιολόγησης παιδιών πρώτης σχολικής ηλικίας. Το συγκεκριμένο όργανο αξιολογεί χωριστά κάθε χορευτικό κύτταρο (χορευτικό βήμα) σε τρεις παράγοντες αξιολόγησης (χρονικά στοιχεία, χωρικά στοιχεία και τύπος της κίνησης). Το άθροισμα των αξιολογήσεων όλων των κυττάρων σε όλους τους παράγοντες, αποτέλεσε το συνολικό σκορ του κάθε παιδιού σε κάθε χορευτική φράση καθενός από τους πέντε χορούς. Οι συγγραφείς επισημαίνουν ότι τα σκορ των μαθητών είναι ανάλογα με τη σχέση μουσικού μέτρου και κίνησης, καθώς και με το στοιχείο της ρυθμικής αγωγής (Παναγοπούλου, Σερμπέζης, Αγγελούσης, 2009). Ωστόσο η αξιολόγηση της χορευτικής απόδοσης ανά κύτταρο, όπως προτείνεται με το παραπάνω όργανο, καθιστά τη διαδικασία πολύ χρονοβόρα.

Στην εκπαίδευση γενικά, δίνεται αυξανόμενη έμφαση στην αυθεντική αξιολόγηση με βάση την απόδοση. Μέσα σ' αυτά τα πλαίσια πολλά προγράμματα χορού χρησιμοποιούν πολλαπλές φόρμες αξιολόγησης της προόδου των μαθητών, όπως η άμεση παρατήρησή τους μέσα στην τάξη ή στη σκηνή, η αυτοαξιολόγηση μέσω συνέντευξης, συζήτησης και χαρτοφυλακίων, καθώς και η ετεροαξιολόγηση από συμμαθητή και τα γραπτά τεστ γνώσεων, αξιολογώντας μ' αυτόν τον τρόπο τεχνικές και φυσικές δεξιότητες, γνώσεις και χρήση λεξιλογίου διαφόρων ειδών χορού, αλλά και συμπεριφορές, όπως προσπάθεια και συνεργασία (Oreck, 2007).

Τα τελευταία χρόνια, κυρίως διεθνώς, για την αξιολόγηση και τη βαθμολόγηση των μαθητών σε οποιοδήποτε είδος χορού, προτείνονται επιπλέον κάποια εργαλεία όπως οι λίστες ελέγχου και οι κλίμακες διαβάθμισης (Kassing & Jay, 2003). Επιπλέον, σε εκπαιδευτικά συστήματα που βασίζονται σε στάνταρτς, χρησιμοποιούνται ευρέως ρούμπρικες (rubrics) οι οποίες αποτελούνται από κριτήρια βάσει των οποίων πραγματοποιείται η αξιολόγηση (π.χ. British Columbia Ministry of Education, 1998, New York City Department of Education, 2005, New Zealand Ministry of Education, 2003, Office for Standards in Education, 2002, Perpich Center for Arts Education, 2005, Saskatchewan Education, 1991, The Standards Site, 2005, κ.α.). Ο Schmid, (2003) υποστηρίζει ότι η ρούμπρικα εξασφαλίζει έναν μηχανισμό συστηματικής και αυθεντικής αξιολόγησης (δηλαδή αξιολόγησης σε πραγματικές συνθήκες) πολύπλοκων χορευτικών εκτελέσεων με έναν δίκαιο και αμερόληπτο τρόπο. Υποστηρίζεται επίσης ότι είναι ένα σύστημα ποσοτικής αξιολόγησης πολύπλοκων χορευτικών εκτελέσεων, καθώς ο τελικός βαθμός είναι η ποσότητα που αντιπροσωπεύει το σύνολο των ατομικών χορευτικών δεξιοτήτων, πάντα με σεβασμό στις αξιώσεις κάθε χορευτικής εκτέλεσης. Το πλήθος των πληροφοριών που παρέχει, επιτρέπει στο δάσκαλο χορού το σχεδιασμό μελλοντικών μαθησιακών εμπειριών για τη μεγιστοποίηση της χορευτικής απόδοσης, συμβάλλοντας αποτελεσματικά τόσο στη διαμορφωτική όσο και στην αθροιστική αξιολόγηση. Επίσης, η λεπτομερής αξιολόγηση της χορευτικής προόδου ενός ατόμου μέσω της ρούμπρικας, εξασφαλίζει μια πιο έγκυρη και αξιόπιστη αξιολόγηση (Warburton, 2002).

Ο όρος ρούμπρικα προέρχεται από τον Αγγλικό όρο rubric, που στα Ελληνικά θα μπορούσε να αποδοθεί ως

«κλίμακα διαβαθμισμένων κριτηρίων» και είναι μία τεχνική, ένα πολύ διαδεδομένο εργαλείο αξιολόγησης, που χρησιμοποιείται ευρέως τα τελευταία χρόνια τόσο στην αυτοαξιολόγηση όσο και στην ετεροαξιολόγηση στο χώρο της εκπαίδευσης, στοχεύοντας στην αύξηση της απόδοσης των μαθητών, στη βελτίωση της διδασκαλίας και στην αξιολόγηση διαφόρων προγραμμάτων (Διγγελίδης, 2009, Reddy & Andrade, 2010). Ορίζεται επίσης ως ένα έγγραφο περιγραφής διαφορετικών επιπέδων ποιότητας απόδοσης (π.χ. από το άριστο έως το μη αποδεκτό) ενός αντικειμένου προς μάθηση (Andrade, 2000). Με βασικό χαρακτηριστικό της τον καθορισμό σημαντικών κριτηρίων απόδοσης, η ρούμπρικα ταξινομεί και οργανώνει τις παρατηρήσεις απόδοσης, πάντα σε σχέση με διαφορετικά επίπεδα δεξιοτήτων, συμπεριφορών και ποιότητας έργου.

Έχει σχεδιαστεί για να βοηθήσει τους εκπαιδευτικούς να καταλήγουν σε παρόμοια συμπεράσματα όσον αφορά στην πρόοδο των μαθητών και να μετρούν την ικανότητά τους να χρησιμοποιούν και να εφαρμόζουν πραγματική, εννοιολογική, διαδικαστική και μεταγνωστική γνώση (Angelo, 2000). Η Bargainnier (2003), ισχυρίζεται ότι η συγκεκριμένη κλίμακα θα μπορούσε να χρησιμοποιηθεί στο χώρο της εκπαίδευσης, διότι η αξιολόγηση γίνεται πιο αντικειμενική και ουσιαστική. Χαρακτηριστικά αναφέρεται σε κάποια παραδείγματα, όπως αν ένας εκπαιδευτικός θέλει να αξιολογήσει την ικανότητα του μαθητή να χρησιμοποιήσει μια μαθηματική φόρμουλα για την επίλυση ενός προβλήματος ή την ικανότητά του να εκτελέσει με ποιότητα μια χορευτική δραστηριότητα ή τη δυνατότητά του να υιοθετήσει μια συγκεκριμένη στάση-συμπεριφορά. Η ρούμπρικα αποτελεί μια κλίμακα σύμφωνα με την οποία μια εργασία ή δεξιότητα αναλύεται στα θεμελιώδη-πρωτογενή χαρακτηριστικά της. Η Barbara Walvoord έχει τελειοποιήσει την παραπάνω μέθοδο αξιολόγησης που βασίζεται στη δύναμη της διαδικασίας βαθμολόγησης, αντιμετωπίζοντας ταυτόχρονα και τις αδυναμίες αυτής της διαδικασίας (Benander, 2000). Κάθε ένα από τα στοιχεία της κλίμακας αξιολογείται ή βαθμολογείται χωριστά, καθιστώντας δυνατό τον εντοπισμό τόσο των δυνατών σημείων όσο και των αδυναμιών του μαθητή. Η πληροφορία αυτή είναι σημαντική για την καθοδήγηση και τη βελτίωση τόσο του μαθητή όσο και της τάξης, αλλά επιπλέον, μπορεί να αποτελέσει σημαντικό οδηγό για τη βελτίωση, όχι μόνο της διδασκαλίας, αλλά και των προγραμμάτων σπουδών (Vars, 2002).

Τα βασικά δομικά στοιχεία της ρούμπρικας είναι 1) τα κριτήρια αξιολόγησης, τα θεμελιώδη-πρωτογενή ποιοτικά στοιχεία που θα αξιολογηθούν και 2) τα διαφορετικά επίπεδα διαβάθμισης της ποιότητας κάθε κριτηρίου (Andrade, 2000, Κουλουμπαρίτση & Ματσαγγούρας, 2004).

Οι ρούμπρικες κατηγοριοποιούνται σε ολιστικές και αναλυτικές (Moskal, 2000). Η ολιστική αξιολογεί τη συνολική απόδοση του μαθητή, δεν κρίνει τα συστατικά ξεχωριστά, δεν πληροφορεί τους μαθητές σχετικά με τα δυνατά και αδύνατα σημεία τους αναφορικά με την μάθηση, δεν θεωρείται τόσο χρήσιμη όσο μια αναλυτική, είναι όμως γρήγορη στη χρήση της καθώς παρέχει στον αξιολογητή τη δυνατότητα αξιολόγησης του τι ακριβώς διαδραματίζεται τη στιγμή της αξιολόγησης (Bargainnier, 2003). Στην αναλυτική, βαθμολογείται χωριστά καθένα από τα επίπεδα, ενώ οι βαθμολογίες κάθε επιπέδου μπορούν να αθροιστούν για να δώσουν τον τελικό βαθμό με αποτέλεσμα να χρειαστεί περισσότερος χρόνος για τη βαθμολόγηση. Βοηθά το μαθητή και τον εκπαιδευτικό στον προσδιορισμό των ισχυρών τομέων μάθησης όπως και αυτών που χρήζουν βελτίωσης αυξάνοντας επιπλέον τη συνέπεια στη βαθμολογία μεταξύ δύο βαθμολογητών (Zimmaro, 2007).

Τα βήματα που πρέπει να ακολουθηθούν για την κατασκευή μιας ρούμπρικας είναι α) ο σαφής καθορισμός του θέματος που θα αξιολογήσει ο ενδιαφερόμενος όπως και το τελικό αποτέλεσμα στο οποίο αναμένει να φτάσουν οι μαθητές, β) ο ακριβής καθορισμός των κριτηρίων για τα οποία υπάρχει ενδιαφέρον για αξιολόγηση (π.χ. για την αξιολόγηση του τρόπου γραφής ως βασικά κριτήρια αξιολόγησης μπορεί να καθοριστούν η συνοχή, το περιεχόμενο και η οργάνωση), γ) ο σαφής και ακριβής ορισμός κάθε κριτηρίου (π.χ. τι σημαίνει συνοχή; Τι σημαίνει οργάνωση;), δ) ο προσδιορισμός των διαφορετικών επιπέδων απόδοσης κάθε κριτηρίου, δηλαδή η περιγραφή π.χ. του υψηλού, του μεσαίου και του χαμηλού επιπέδου, ε) η ανάπτυξη μιας κλίμακας βαθμολόγησης με τόσα επίπεδα όσα και τα επίπεδα απόδοσης κάθε κριτηρίου, καθορίζοντας με σαφήνεια τη διαφορά μεταξύ τους. Το υψηλότερο αντανακλά την άριστη ποιότητα μιας εργασίας ή πιο σωστά εκείνο το επίπεδο το οποίο όλοι οι μαθητές προσπαθούν να επιτύχουν, ενώ το χαμηλότερο αντανακλά την πιο χαμηλή ποιότητα ή το επίπεδο το οποίο όλοι προσπαθούν να αποφύγουν (Zimmaro 2007). Κάθε επίπεδο αποτελεί φυσιολογική συνέχεια του προηγούμενου. Η Andrade (2000) προτείνει έναν τρόπο, μια τεχνική διαμόρφωσης-κατασκευής των επιπέδων ποιότητας κάθε κριτηρίου, βασισμένη στις απαντήσεις «ναι», «ναι, αλλά...», «όχι, αλλά...» και «όχι», που αντανακλούν το πώς ανταποκρίνεται ο μαθητής σε καθένα από τα επίπεδα απόδοσης κάθε κριτηρίου. Από αρκετούς επίσης ερευνητές (Κουλουμπαρίτση & Ματσαγγούρας, 2004, Kutlu, Yildirim, & Bilican, 2010), προτείνεται ως πολύ ωφέλιμη η συνεργασία μεταξύ εκπαιδευτικών για την επιλογή των κριτηρίων στην κατασκευή μιας ρού-

μπρικας, διότι οι από κοινού σκέψεις και αναζητήσεις συμβάλλουν στον ορθότερο εντοπισμό των αρχικών κριτηρίων.

Ανεξάρτητα από την τελική μορφή μιας ρούμπρικας, εφόσον χρησιμοποιείται ως όργανο αξιολόγησης, θα πρέπει να πληροί προϋποθέσεις εγκυρότητας (π.χ. όταν ένα εργαλείο μέτρησης είναι πράγματι έγκυρο, «αντανάκλα» την έννοια (τη μεταβλητή) που προτίθεται ότι μετράει) και αξιοπιστίας (π.χ. η συνεπής αξιολόγηση μετά από επαναλαμβανόμενες μετρήσεις είτε από το ίδιο είτε από διαφορετικά άτομα), στοιχεία που θεωρούνται σημαντικά για την εφαρμογή της. Για την αξιολόγηση της αξιοπιστίας μιας ρούμπρικας μπορεί να χρησιμοποιηθεί η *inter-rater agreement*, ο βαθμός στον οποίο υπάρχει απόλυτη συμφωνία στη βαθμολόγηση μεταξύ των κριτών, η *intra-rater agreement*, ο βαθμός στον οποίο ο ίδιος κριτής αποδίδει την ίδια βαθμολογία σε μία εργασία, όσες φορές τη βαθμολογήσει (Moskal & Laydens, 2000). Επίσης, η *inter-rater reliability*, που αναφέρεται στη συσχέτιση των βαθμολογιών διαφορετικών κριτών (Stellmack, 2009). Υποστηρίζεται ότι η *inter-rater agreement* είναι πολύ σημαντική για τους εκπαιδευτικούς, διότι χαρακτηρίζεται ως πιο αξιόπιστη πηγή ανατροφοδότησης αναφορικά με την απόδοση των μαθητών και αποτελεί βάση σχεδιασμού για επαγγελματική ανάπτυξη (Graham, Milanowski & Miller, 2012).

Σκοπός της παρούσας εργασίας ήταν η κατασκευή μιας αναλυτικής κλίμακας διαβαθμισμένων κριτηρίων (ρούμπρικας), ειδικής για την αξιολόγηση της απόδοσης ελληνικών παραδοσιακών χορών, με βάση τα θεμελιώδη συστατικά και τα βασικά βήματα κατασκευής, όπως αυτά προτείνονται από τη διεθνή βιβλιογραφία, καθώς και ο έλεγχος εγκυρότητας αξιοπιστίας της. Γίνεται μια περιγραφή του τρόπου ανάπτυξης του συγκεκριμένου εργαλείου και παρουσιάζονται τα δεδομένα σχετικά με τη συμφωνία μεταξύ των κριτών αλλά και μεταξύ επαναλαμβανόμενων αξιολογήσεων του ίδιου κριτή.

Μέθοδος

Συμμετέχοντες

Στην έρευνα συμμετείχαν, τυχαία επιλεγμένοι, δεκατέσσερις φοιτητές ηλικίας 19-20 ετών, αρχάριοι ως προς τη χορευτική τους εμπειρία, οι οποίοι διδάχθηκαν τρεις ελληνικούς χορούς (Χασαπιά, Ζωναράδικο, Έντεκα). Για τους χορούς τους ζητήθηκε να τους αναγνωρίσουν από τη μελωδία τους, να ανταποκριθούν στο ρυθμικό τους πρότυπο με κρούσεις των παλαμών και να αποδώσουν την κινητική τους φόρμα. Μετά τη διδασκαλία κάθε χορού, οι φοιτητές βιντεοσκοπήθηκαν χωριστά ακολουθώντας όλη την παραπάνω διαδικασία και ο καθένας αξιολογήθηκε από τρεις αξιολογητές (Α,Β,Γ) μέσω των βιντεοσκοπημένων εκτελέσεών του.

Κατασκευή ρούμπρικας

Στην προσπάθεια αξιολόγησης των προπτυχιακών φοιτητών ελληνικού παραδοσιακού χορού του Τμήματος Επιστήμης Φυσικής Αγωγής και Αθλητισμού του Δημοκρίτειου Πανεπιστημίου Θράκης, αναφορικά με την απόδοση ελληνικών χορών, αποφασίστηκε η κατασκευή ειδικής αναλυτικής ρούμπρικας. Αρχικά, από τους ερευνητές συντάχθηκε μια λίστα ποιοτικών παρατηρήσιμων κριτηρίων τα οποία αποτελούν θεμελιώδη-πρωταρχικά στοιχεία του χορού γενικά και του ελληνικού ειδικότερα και επιπλέον συνιστούν στοιχεία απόδοσης. Τα προαναφερθέντα παρατηρήσιμα κριτήρια βασίστηκαν στο σύστημα ανάλυσης και καταγραφής της κίνησης, *Labanotation* και *Effort-Shape Description*, αντίστοιχα (Hutchinson Guest, 2005), ενώ τα αντίστοιχα κριτήρια που αφορούσαν στους τρεις Ελληνικούς χορούς που αξιολογήθηκαν στην παρούσα μελέτη, καθορίστηκαν με βάση το σύστημα δομικής-μορφολογικής ανάλυσης των Martin και Pesovár (1961) και του International Folk Music Council (I.F.M.C., 1974). Τα κριτήρια αυτά ομαδοποιήθηκαν σε κατηγορίες, σχετικές με το χορευτικό ύφος, την ικανότητα του ατόμου να αναγνωρίζει και να αναπαράγει το χορό, να ακολουθεί το ρυθμό, να ελέγχει το σώμα του, να χρησιμοποιεί το χώρο, να επιδεικνύει ποιοτικά-δυναμικά στοιχεία της κίνησης, να μεταβιβάζει την κίνηση σε ολόκληρο το σώμα (Kassing & Jay, 2003, Τυροβολά, 1998, Warburton, 2002, Ζωγράφου, 2003). Με βάση τις κατηγορίες αυτές δημιουργήθηκε ένα αρχικό σχέδιο της ρούμπρικας η οποία δόθηκε σε ομάδα ειδικών του Ελληνικού παραδοσιακού χορού προκειμένου να αποφανθούν για την καταλληλότητα των κριτηρίων στη δομή της και τελικά απαρτίστηκε από οκτώ συγκεκριμένα και σαφή κριτήρια (Πίνακας 1.).

Πίνακας 1. Περιγραφή των οκτώ κριτηρίων της ρούμπρικας

Κριτήρια	Αξιολογείται...
Αναγνώριση Χορού	Η ικανότητα του φοιτητή να αναγνωρίσει και να ονοματίσει το χορό από το άκουσμα της μελωδίας του
Ανταπόκριση στο Ρυθμικό Πρότυπο	Η ικανότητα να ανταποκριθεί με κρούσεις παλαμών, στο ρυθμικό πρότυπο του χορού
Συγχρονισμός	Η ικανότητα να ακολουθεί το ρυθμό της μουσικής κατά την εκτέλεση του χορού
Ακολουθία Βημάτων Χορού	Πόσο καλά γνωρίζει τη σειρά των διδαχθέντων βημάτων του χορού
Κατεύθυνση-Σχήμα, Λαβή Χορού	Η ικανότητα χρήσης του χώρου (κατεύθυνση, επίπεδα) κατά την εκτέλεση του χορού ώστε να διατηρείται το ιδιαίτερο σχήμα του. Η επίδειξη της διδαχθείσας λαβής του χορού
Θέση και Στάση του Σώματος	Η επίδειξη της χαρακτηριστικής θέσης και στάσης του σώματος κατά την εκτέλεση του χορού
Ποιοτικά Στοιχεία Κίνησης	Η ικανότητα να εκτελεί το χορό τονίζοντας τα ιδιαίτερα ποιοτικά-δυναμικά στοιχεία του (το βάρος, ο χρόνος, ο χώρος, η ροή της κίνησης)
Εκφραστικότητα	Η ικανότητα να μεταφέρει την κίνηση και σε άλλα μέρη του σώματος (κεφάλι, ώμοι, κορμός, χέρια)

Βάσει της θεωρίας του ιδανικού αριθμού των επιπέδων (τρία έως πέντε) των κριτηρίων (Stevens & Levi, 2005), ορίστηκαν τελικά πέντε επίπεδα διαβάθμισης της απόδοσης κάθε κριτηρίου που περιγράφονται ως εξής: «1-Μη Αποδεκτό», «2-Οριακά Αποδεκτό», «3-Μέτρια Αποδεκτό», «4-Πολύ Καλό» και «5-Άριστο». Στον Πίνακα 2, περιγράφεται ένα από τα οκτώ κριτήρια μαζί με τα πέντε διαφορετικά επίπεδα ποιότητας απόδοσης του κριτηρίου.

Πίνακας 2. Παράδειγμα των επιπέδων απόδοσης ενός κριτηρίου

Κριτήριο	Μη αποδεκτό 1	Οριακά αποδεκτό 2	Μέτρια αποδεκτό 3	Πολύ καλό 4	Άριστο 5
Ακολουθία Βημάτων Χορού	Δε εκτελεί καθόλου τα βήματα του χορού	Δε εκτελεί ολόκληρη τη σειρά των βημάτων, ανακαλεί στη μνήμη του κάποιο μέρος της	Δεν εκτελεί από την αρχή του χορού τη σειρά των βημάτων, αλλά σταδιακά την ανακαλεί στη μνήμη του σχεδόν ολόκληρη	Εκτελεί από την αρχή τη σειρά των βημάτων, ανακαλώντας τη αμέσως στη μνήμη του, αλλά κατά διαστήματα σημειώνει κάποια λάθη	Εκτελεί ανακαλώντας αμέσως στη μνήμη του τη σειρά των βημάτων του χορού χωρίς καθόλου λάθη

Εγκυρότητα Περιεχομένου

Η εγκυρότητα περιεχομένου αναφέρεται στο βαθμό που ένα εργαλείο μέτρησης καλύπτει εννοιολογικά το εύρος της μεταβλητής που μετράει. Στην παρούσα μελέτη οι τύποι που επιλέχθηκαν προκειμένου να εκτιμηθεί η εγκυρότητα της ρούμπρικας, ήταν: α) η εγκυρότητα περιεχομένου (content validity) και γ) η συγκλίνουσα εγκυρότητα, ως τύπος της δομικής εγκυρότητας, σε μια προσπάθεια χρήσης ενός πιο ολοκληρωμένου πλαισίου θεμε-

λίωσης της εγκυρότητας (Jonsson & Svingby, 2007).

Για την καθιέρωση της εγκυρότητας περιεχομένου, τα ποιοτικά στοιχεία-κριτήρια της μεταβλητής που μετράται αποκτώνται συχνά μέσω ομάδων εμπειρογνομόνων αλλά και μέσω της πληθυσμιακής ομάδας στην οποία θα απευθυνθεί το όργανο αξιολόγησης. Τα ποιοτικά αυτά δεδομένα μετατρέπονται σε κατηγορίες ή θεματικές ενότητες και στη συνέχεια ποσοτικοποιούνται σε αριθμητικές τιμές ή κατατάξεις. Αυτή η διαδικασία συνήθως επαναλαμβάνεται καθώς η παροχή ανατροφοδότησης από τους εμπειρογνώμονες και από την υπάρχουσα βιβλιογραφία οδηγεί σε συναίνεση μεταξύ τους (Newman, Lim & Pineda, 2013). Είναι μια διαδικασία που μπορεί να εκτιμηθεί ποιοτικά, ποσοτικά ή με τη χρήση ενός συνδυασμού των δύο μεθόδων ώστε να προσδιοριστεί ο βαθμός συναίνεσης μεταξύ των εμπειρογνομόνων σχετικά με το εν λόγω όργανο αξιολόγησης (Haynes, Richard, & Kubany, 1995).

Η Newman και οι συνεργάτες (2013) υποστηρίζουν επίσης ότι οι εμπειρογνώμονες-ειδικοί θα πρέπει να αναφέρουν αν τα επιλεγμένα κριτήρια εκτιμούν την προς αξιολόγηση έννοια και η συμφωνία μεταξύ τους θα πρέπει να φτάνει το 80% τουλάχιστον, και επιπλέον να αξιολογούν την επάρκεια των επιλεγμένων κριτηρίων σε ποσοστό από 0% έως 100% και να προβαίνουν σε περαιτέρω προτάσεις προκειμένου να εξασφαλιστεί ένας επαρκής αριθμός κριτηρίων. Ο μέσος όρος εκτίμησης των εμπειρογνομόνων δηλώνει το τελικό ποσοστό επάρκειας των κριτηρίων. Ωστόσο, από τους παραπάνω ερευνητές αναφέρεται επίσης ότι είναι ασυνήθιστο να βρεθεί ένα σύνολο κριτηρίων που να μετρά το 100% του περιεχομένου της έννοιας που αξιολογείται, γι' αυτό κάθε ερευνητής θα πρέπει να στοχεύει σε ένα ποσοστό επάρκειας κάθε κριτηρίου, της τάξης του 80 %.

Με την παρούσα ρούμπρικα στόχος των ερευνητών είναι να αξιολογηθεί το επίπεδο απόδοσης ελληνικών χορών από φοιτητές. Συνεπώς, το πρώτο μέλημα για την κατασκευή της ήταν να προσδιοριστεί τι ακριβώς σημαίνει η έννοια απόδοση του ελληνικού χορού, να καθοριστεί η διαδικασία μέσω της οποίας θα εργαζόταν οι φοιτητές όπως και το αποτέλεσμα στο οποίο αναμενόταν ότι θα φτάσουν (Zimmaro, 2007).

Για την απόδοση του ελληνικού χορού, η οποία χωρίς ένα σύστημα βαθμολόγησης-ταξινόμησης είναι μια υποκειμενική ποιοτική αξιολόγηση, καθορίστηκαν από τους ερευνητές συγκεκριμένα επιδιωκόμενα αποτελέσματα τα οποία λειτούργησαν ως κριτήρια. Τα επιδιωκόμενα αποτελέσματα, βασισμένα κυρίως στη θεωρία του Laban, ήταν τόσο κινητικού (ρυθμοκινητική ανταπόκριση, εκτέλεση της χορογραφίας, χρήση του χώρου και του χρόνου, ποιοτική εκτέλεση της κίνησης) (Slettum, 2001, Wuarburton, 2002), όσο και γνωστικού περιεχομένου (αναγνώριση κάθε χορού και απόδοση του ονόματός του) (Kassing & Jay, 2003, NAEP, 1998).

Ταυτόχρονα, ένας σημαντικός αριθμός φοιτητών οι οποίοι συγκροτούσαν την πληθυσμιακή ομάδα στην οποία θα απευθυνόταν η ρούμπρικα, ρωτήθηκε για την καταλληλότητα των συγκεκριμένων επιδιωκόμενων αποτελεσμάτων ως κριτήρια απόδοσης-γνώσης ελληνικών χορών, ενώ παράλληλα, σύμφωνα με τους Ουζούνη και Νακάκη (2011), τους ζητήθηκε να κάνουν και τις δικές τους προτάσεις.

Το επόμενο βήμα ήταν η κατασκευή μιας αναλυτικής ρούμπρικας, ειδικής για την απόδοση του ελληνικού χορού που αποτελείτο από τα προαναφερθέντα παρατηρήσιμα κριτήρια και από διαφορετικά επίπεδα απόδοσης του καθενός. Σε αυτή τη χρονική στιγμή το αρχικό σχέδιο της ρούμπρικας δόθηκε στην ομάδα των δύο ειδικών του χορού προκειμένου να θεμελιωθεί η εγκυρότητα περιεχομένου του οργάνου. Οι προϋποθέσεις για να θεωρηθεί κάποιος ειδικός του χορού ήταν η κατοχή ειδικότητας ελληνικού παραδοσιακού χορού από πανεπιστημιακό τμήμα της χώρας με διδακτική και χορευτική εμπειρία μεγαλύτερη των δέκα ετών. Σημειώνεται ότι οι περισσότεροι ήταν διδακτικό επιστημονικό προσωπικό στην κατεύθυνση χορός σε πανεπιστημιακό τμήμα. Ειδικότερα ζητήθηκε από τους ειδικούς να αναφέρουν κατά την άποψή τους το ποσοστό (0%-100%) της εννοιολογικής καταλληλότητας κάθε κριτηρίου στην έννοια «απόδοση του ελληνικού χορού», αλλά και της επάρκειας του αριθμού των επιλεγμένων από τους ερευνητές κριτηρίων. Ζητήθηκε επίσης από τους ειδικούς να αξιολογήσουν με τον ίδιο τρόπο (ποσοστιαία βάση 0%-100%), την ακρίβεια της περιγραφής κάθε επιπέδου απόδοσης για τον σαφέστερο διαχωρισμό μεταξύ τους (Newman et al. 2013, Pineda, 2012).

Η ρούμπρικα αναθεωρήθηκε βάσει των σχολίων των ειδικών και τελικά αποφασίστηκε ομόφωνα τα κριτήρια να είναι συνολικά οκτώ, ενώ τα επίπεδα κάθε κριτηρίου συνολικά πέντε, για να υπάρξει δίκαιη διαφοροποίηση τόσο των φοιτητών με τη μη αποδεκτή απόδοση από τους φοιτητές με το κατώτατο αποδεκτό επίπεδο, όσο και των φοιτητών με άριστο επίπεδο απόδοσης (κανένα απολύτως λάθος) από εκείνους με πολύ καλή απόδοση (ελάχιστα λάθη). Το πέμπτο επίπεδο ορίστηκε για τους φοιτητές με τη μέτρια απόδοση.

Στο σημείο αυτό δομής της ρούμπρικας, κατόπιν υποδείξεων και προτάσεων από τους ειδικούς, δόθηκε ιδιαίτερη έμφαση από τους ερευνητές στον εντοπισμό και τη χρήση του καταλληλότερου λεξιλογίου, ώστε να μη χρησιμοποιηθεί συγκριτική γλώσσα, αλλά λέξεις μοναδικές για την περιγραφή κάθε επιπέδου και τη μεταξύ τους

διάκριση (Moni, Beswick & Moni, 2005, Zimmaro, 2007). Οι Reddy και Andrade (2010) υποστηρίζουν ότι η χρήση κατάλληλου λεξιλογίου είναι από τις πιο προκλητικές πτυχές δομής μιας ρούμπρικας, ενώ παράλληλα συνιστά στοιχείο εγκυρότητας.

Δομική εγκυρότητα

Η εγκυρότητα εννοιολογικής κατασκευής ή δομική εγκυρότητα ενσωματώνει διαφορετικές μορφές εγκυρότητας και απαντά στο ερώτημα, α) κατά πόσο το εργαλείο μέτρησης μετρά την έννοια του ενδιαφέροντος μας και β) κατά πόσο το εργαλείο μας είναι συναφές με παρόμοια εργαλεία ή παραλλαγές του ίδιου εργαλείου. Για την καθιέρωση δομικής εγκυρότητας της παρούσας ρούμπρικας χρησιμοποιήθηκε ο τύπος της συγκλίνουσας εγκυρότητας. Σύμφωνα με την συγκεκριμένη διαδικασία, η βαθμολογία των φοιτητών με τη χρήση της ρούμπρικας, τόσο από τον αξιολογητή Α όσο και από τον αξιολογητή Β, συσχετίστηκε με τη βαθμολογία κατόπιν υποκειμενικής και ολιστικής αξιολόγησης από έναν επίσης ειδικό του Ελληνικού χορού. Πιο συγκεκριμένα, ένας ανεξάρτητος αξιολογητής Γ, ειδικός του ελληνικού χορού, αξιολόγησε τους ίδιους φοιτητές δίνοντας μια συνολική βαθμολογία για τον καθένα σύμφωνα με ό,τι ο ίδιος θεωρούσε και εφάρμοζε στην αξιολόγηση του χορού. Η υποκειμενική αξιολόγηση από ειδικούς, είναι μια πρακτική που συνηθίζεται στην αξιολόγηση του τρόπου απόδοσης του ελληνικού χορού, εξαιτίας και της μέχρι τώρα απουσίας ειδικών οργάνων για το σκοπό αυτό. Στην παρούσα μελέτη ο αξιολογητής Γ ενημερώθηκε από τους ερευνητές απλώς για τις βασικές επιδιώξεις και τα επιδιωκόμενα αποτελέσματα της μαθησιακής διαδικασίας.

Έλεγχος συμφωνίας μεταξύ κριτών και μεταξύ δύο διαφορετικών βαθμολογιών του ίδιου κριτή

Για τον έλεγχο της αξιοπιστίας της ρούμπρικας, αξιολογήθηκε από δύο κριτές (Jonsson & Svingby, 2007, Thaler, Kazemi & Huscher, 2009), ειδικοί του ελληνικού παραδοσιακού χορού, η ικανότητα δεκατεσσάρων φοιτητών να αποδώσουν τρεις ελληνικούς χορούς, βάσει των οκτώ κριτηρίων που προαναφέρθηκαν. Οι φοιτητές βιντεοσκοπήθηκαν κατά την εκτέλεση των χορών και αξιολογήθηκαν μέσω των βιντεοσκοπημένων εκτελέσεών τους. Ο κάθε κριτής (Α και Β) αξιολόγησε και βαθμολόγησε και τους δεκατέσσερις φοιτητές και στους τρεις χορούς, ενώ μετά το πέρας δύο εβδομάδων οι ίδιοι φοιτητές επαναξιολογήθηκαν. Υπολογίστηκε η «απόλυτη» συμφωνία μεταξύ των κριτών (absolute agreement), όπου οι βαθμολογίες τους ήταν όμοιες και η «σχετική» συμφωνία (adjacent agreement), όπου οι βαθμολογίες των δύο κριτών (interrater agreement) διέφεραν κατά ένα επίπεδο απόδοσης. Για την απόλυτη συμφωνία υπολογίστηκε ο ακριβής αριθμός των βαθμολογιών των δύο κριτών που ήταν όμοιες μεταξύ τους και διαιρέθηκε με τον συνολικό αριθμό των βαθμολογιών. Για τη σχετική συμφωνία υπολογίστηκε και ο αριθμός των βαθμολογιών που διέφεραν κατά ένα επίπεδο και διαιρέθηκε με τον συνολικό αριθμό των βαθμολογιών. Οι παραπάνω υπολογισμοί χρησιμοποιήθηκαν και για τον έλεγχο συμφωνίας μεταξύ δύο βαθμολογιών του ίδιου κριτή (intrarater agreement) που πραγματοποιήθηκαν σε διαφορετικές χρονικές στιγμές. Επιπλέον, χρησιμοποιήθηκαν ο δείκτης ICC (Graham, Milanowski & Miller, 2012), ο δείκτης Cronbach's α , που χρησιμοποιήθηκε σε αρκετές μελέτες (Jonsson & Svingby, 2007) για την εκτίμηση της συνέπειας μεταξύ δύο διαφορετικών βαθμολογιών ενός κριτή (intrarater reliability) και ο δείκτης Spearman's r που έχει χρησιμοποιηθεί για την εκτίμηση της συνέπειας μεταξύ δύο κριτών (interrater reliability) (Thaler et al. 2009). Για τον έλεγχο της αξιοπιστίας (interrater reliability) προτείνεται και ο δείκτης Kendall's W (Olson, Schieve, Ruit & Vari, 2003), ο οποίος δεν χρησιμοποιήθηκε στον έλεγχο της συγκεκριμένης ρούμπρικας.

Για κάθε χορό δόθηκαν συνολικά 224 {2X(14 φοιτητές X 8 κριτήρια)} βαθμολογίες και από τους δύο κριτές, άρα υπήρχαν και 224 πιθανές περιπτώσεις συμφωνίας μεταξύ τους. Ο αριθμός 40 (8 κριτήρια X 5, ο μέγιστος βαθμός για κάθε ένα κριτήριο) ήταν ο υψηλότερος βαθμός που θα μπορούσε να συγκεντρώσει ένας φοιτητής σε κάθε χορό.

Αποτελέσματα

Δομική εγκυρότητα

Η τιμή του στατιστικού δείκτη Pearson Correlation μετά τη συσχέτιση των βαθμολογιών διαμορφώθηκε σε υψηλά επίπεδα όπως εμφανίζονται στον Πίνακα 3.

Πίνακας 3. Συσχετίσεις της υποκειμενικής και της αξιολόγησης με τη χρήση της ρούμπρικας

Υποκειμενική Αξιολόγηση Γ	
Αξιολογητής Α (ρούμπρικα)	.904**
Αξιολογητής Β (ρούμπρικα)	.845**
Μέση τιμή Αξιολογητών Α & Β	.906**

** $p < 0.01$ *Συμφωνία μεταξύ δύο κριτών (interrater agreement)*

Σύμφωνα με τους υπολογισμούς που πραγματοποιήθηκαν, στις 224 βαθμολογίες που σημειώθηκαν από τους δύο κριτές για κάθε χορό, υπήρξε 68% απόλυτη συμφωνία (152/224 περιπτώσεις) και 95% σχετική συμφωνία (212/224 περιπτώσεις) για το χορό Χασαπιά, 67% απόλυτη συμφωνία (150/224 περιπτώσεις) 91% σχετική συμφωνία (204/224 περιπτώσεις) για το Ζωναράδικο, ενώ για το χορό Έντεκα 64% απόλυτη συμφωνία (144/224 περιπτώσεις) και 91% σχετική συμφωνία (204/224 περιπτώσεις).

Για κάθε ένα από τα οκτώ κριτήρια της ρούμπρικας που αξιολογήθηκαν, οι τιμές της απόλυτης συμφωνίας κυμάνθηκαν από 36%-100%, της σχετικής από 86%-100%, του δείκτη ICC από 0.458-0.929, του δείκτη Spearman's r από 0.485-0.932 και του δείκτη Cronbach's α από 0.647-0.950 για το χορό Χασαπιά (Πίνακας 4). Για τον Ζωναράδικο (Πίνακας 5), αντίστοιχα από 43%-93% (απόλυτη συμφωνία), από 79%-100% (σχετική συμφωνία), από 0.552-0.960 (ICC), από 0.589-0.994 (Spearman's r), και από 0.732-0.960 (Cronbach's α), ενώ για το χορό Έντεκα (Πίνακας 6), από 36%-86% (απόλυτη συμφωνία), από 79%-100% (σχετική συμφωνία), από 0.776-0.978 (ICC), από 0.774-0.995 (Spearman's r) και από 0.872-0.977 (Cronbach's α).

Συμφωνία μεταξύ διαφορετικών βαθμολογιών του ίδιου κριτή (intrarater agreement)

Αναφορικά με τη συμφωνία των βαθμολογιών που δόθηκαν στους ίδιους φοιτητές σε δύο διαφορετικές χρονικές στιγμές, στο σύνολο των περιπτώσεων παρατηρήθηκε απόλυτη συμφωνία 82% (184/224), σχετική συμφωνία 100% (224/224) για το χορό Χασαπιά. Για τον Ζωναράδικο 88% (198/224) απόλυτη, 100% (224/224) σχετική και για το χορό Έντεκα 77% (172/224) απόλυτη και 95% (212/224) σχετική συμφωνία.

Αναλυτικότερα οι τιμές της απόλυτης συμφωνίας, της σχετικής συμφωνίας και όλων των στατιστικών δεικτών για τη βαθμολογία κάθε κριτηρίου σε κάθε χορό, που χρησιμοποιήθηκαν για τον έλεγχο της συμφωνίας των βαθμολογιών του ίδιου κριτή, διαμορφώθηκαν ως εξής: χορός Χασαπιά (Πίνακας 7), απόλυτη συμφωνία 64%-100%, σχετική συμφωνία 100%-100%, ICC 0.458-0.970 $p < 0,001$, Spearman's r 0.531-0.978, Cohen's k 0.440-0.796 και Cronbach's α 0.647-0.968. Ζωναράδικος χορός (Πίνακας 8), απόλυτη συμφωνία 80%-100%, σχετική συμφωνία 100%-100%, ICC 0.914-0.986, Spearman's r 0.771-0.994, Cohen's k 0.492-0.861 και Cronbach's α 0.959-1. Χορός Έντεκα (Πίνακας 9), απόλυτη συμφωνία 63%-100%, σχετική συμφωνία 88%-100%, ICC 0.869-1, Spearman's r 0.899-1 και Cronbach's α 0.946-1.

Πίνακας 4. Απόλυτη, σχετική συμφωνία και στατιστικοί δείκτες της συμφωνίας μεταξύ των δύο κριτών (inter-rater agreement) στο χορό Χασαπιά

ΧΟΡΟΣ ΧΑΣΑΠΙΑ							
Στατιστικοί δείκτες							
Κριτήρια	Απόλυτη συμφωνία	Σχετική συμφωνία	ICC Single	ICC Average	Spearman's r	Cohen's K	Cronbach's α
Αναγνώριση Χορού	100%	100%	1	1			1
Ανταπόκριση στο Ρυθμικό Πρότυπο	86%	100%	0.458*	0.629*	0.531*	0.440*	0.647*
Συγχρονισμός	50%	93%	0.746**	0.855**	0.810**		0.880**
Ακολουθία βημάτων χορού	57%	93%	0.742**	0.852**	0.687*		0.852**
Κατεύθυνση-Σχήμα, Λαβή Χορού	86%	100%	0.852**	0.920**	0.734*		0.915**
Θέση και Στάση του Σώματος	64%	86%	0.540*	0.701*	0.485		0.699*
Ποιότητα Κίνησης	36%	93%	0.703**	0.826**	0.855**		0.916**
Εκφραστικότητα	64%	100%	0.868**	0.929**	0.932**		0.950**
Συνολική Βαθμολογία	85%	100%	0.908**	0.952**	0.940**		0.956**

*p<0.05

**p<0.001

Πίνακας 5. Απόλυτη, σχετική συμφωνία και στατιστικοί δείκτες της συμφωνίας μεταξύ των δύο κριτών (interrater agreement) στο χορό Ζωναράδικο

ΧΟΡΟΣ ΖΩΝΑΡΑΔΙΚΟΣ							
Κριτήρια	Στατιστικοί δείκτες						
	Απόλυτη συμφωνία	Σχετική συμφωνία	ICC Single	ICC Average	Spearman's r	Cohen's k	Cronbach's α
Αναγνώριση Χορού	93%	100%	0.923**	0.960**	0.994**	0.803**	0.960**
Ανταπόκριση στο Ρυθμικό Πρότυπο	71%	86%	0.719**	0.837**	0.836**		0.861**
Συγχρονισμός	79%	100%	0.868**	0.929**	0.894**		0.939**
Ακολουθία βημάτων χορού	64%	93%	0.786**	0.880**	0.760**		0.888**
Κατεύθυνση-Σχήμα, Λαβή Χορού	43%	79%	0.559*	0.717*	0.653*		0.781*
Θέση και Στάση του Σώματος	43%	86%	0.552*	0.711*	0.589*		0.732*
Ποιότητα Κίνησης	71%	93%	0.793**	0.884**	0.784**		0.905**
Εκφραστικότητα	71%	93%	0.718**	0.836**	0.733**		0.864**
Συνολική Βαθμολογία	79%	100%	0.887**	0.940**	0.921**		0.950**

*p<0.05

**p<0.001

Πίνακας 6. Απόλυτη, σχετική συμφωνία και στατιστικοί δείκτες της συμφωνίας μεταξύ των δύο κριτών (interrater agreement) στο χορό Έντεκα

ΧΟΡΟΣ ΕΝΤΕΚΑ							
Στατιστικοί δείκτες							
Κριτήρια	Απόλυτη συμφωνία	Σχετική συμφωνία	ICC Single	ICC Average	Spearman's r	Cohen's k	Cronbach's α
Αναγνώριση Χορού	86%	100%	0.957**	0.978**	0.995**	-	0.977**
Ανταπόκριση στο Ρυθμικό Πρότυπο	79%	100%	0.935**	0.966**	0.940**	-	0.965**
Συγχρονισμός	64%	86%	0.878**	0.935**	0.893**	0.542**	0.935**
Ακολουθία βημάτων χορού	79%	86%	0.826**	0.905**	0.915**	0.712**	0.918**
Κατεύθυνση-Σχήμα, Λαβή Χορού	71%	79%	0.827**	0.905**	0.962**	0.616**	0.924**
Θέση και Στάση του Σώματος	71%	93%	0.883**	0.938**	0.820**	0.639**	0.934**
Ποιότητα Κίνησης	36%	100%	0.855**	0.922**	0.944**		0.968**
Εκφραστικότητα	64%	86%	0.776**	0.874**	0.774**		0.872**
Συνολική Βαθμολογία	93%	100%	0.967**	0.983**	0.980**		0.982**

*p=0.008

**p<0.001

Πίνακας 7. Απόλυτη, σχετική συμφωνία και στατιστικοί δείκτες της συμφωνίας μεταξύ δύο βαθμολογιών του ίδιου κριτή (intrarater agreement) στο χορό Χασαπιά

ΧΟΡΟΣ ΧΑΣΑΠΙΑ							
Στατιστικοί δείκτες							
Κριτήρια	Απόλυτη συμφωνία	Σχετική συμφωνία	ICC Single	ICC Average	Spearman's r	Cohen's k	Cronbach's α
Αναγνώριση Χορού	100%	100%					
Ανταπόκριση στο Ρυθμικό Πρότυπο	86%	100%	0.458*	0.629*	0.531*	0.440*	0.647*
Συγχρονισμός	86%	100%	0.930**	0.964**	0.978**		0.961**
Ακολουθία βημάτων χορού	64%	100%	0.838**	0.912**	0.884**	0.449*	0.937**
Κατεύθυνση-Σχήμα, Λαβή Χορού	79%	100%	0.828**	0.906**	0.848**		0.919**
Θέση και Στάση του Σώματος	79%	100%	0.812**	0.896**	0.696*	0.611*	0.891**
Ποιότητα Κίνησης	86%	100%	0.941**	0.970**	0.961**	0.796**	0.968**
Εκφραστικότητα	79%	100%	0.900**	0.948**	0.905**	0.708**	0.945**
Συνολική Βαθμολογία	79%	100%	0.964**	0.982**	0.956**		0.981**

* p<0.05

**p<0.001

Πίνακας 8. Απόλυτη, σχετική συμφωνία και στατιστικοί δείκτες της συμφωνίας μεταξύ δύο βαθμολογιών του ίδιου κριτή (intrarater agreement) στο χορό Ζωναράδικο

ΧΟΡΟΣ ΖΩΝΑΡΑΔΙΚΟΣ							
Στατιστικοί δείκτες							
Κριτήρια	Απόλυτη συμφωνία	Σχετική συμφωνία	ICC Single	ICC Average	Spearman's r	Cohen's k	Cronbach's α
Αναγνώριση Χορού	90%	100%	0.914**	0.955**	0.994**		0.955**
Ανταπόκριση στο Ρυθμικό Πρότυπο	100%	100%	1	1	0.913**		1
Συγχρονισμός	90%	100%	0.947**	0.973**	0.994**		0.973**
Ακολουθία βημάτων χορού	80%	100%	0.928**	0.963**	0.826*		0.959**
Κατεύθυνση-Σχήμα, Λαβή Χορού	80%	100%	0.972**	0.986**	0.771*	0.492*	0.986**
Θέση και Στάση του Σώματος	80%	100%	0.952**	0.976**	0.942**		0.976**
Ποιότητα Κίνησης	90%	100%	0.951**	0.975**	0.951**	0.861**	0.975**
Εκφραστικότητα	90%	100%	0.951**	0.975**	0.920**	0.730**	0.975**
Συνολική Βαθμολογία	90%	100%	0.994**	0.997**	0.988**		0.998**

* p<0.05

**p<0.001

Πίνακας 9. Απόλυτη, σχετική συμφωνία και στατιστικοί δείκτες της συμφωνίας μεταξύ δύο βαθμολογιών του ίδιου κριτή (intrarater agreement) στο χορό Έντεκα

ΧΟΡΟΣ ΕΝΤΕΚΑ							
Στατιστικοί δείκτες							
Κριτήρια	Απόλυτη συμφωνία	Σχετική συμφωνία	ICC Single	ICC Average	Spearman's r	Cohen's k	Cronbach's α
Αναγνώριση Χορού	100%	100%	1	1	1	1**	1**
Ανταπόκριση στο Ρυθμικό Πρότυπο	88%	88%	0.911**	0.954**	0.915**		0.954**
Συγχρονισμός	75%	100%	0.960**	0.980**	0.899*		0.982**
Ακολουθία βημάτων χορού	75%	100%	0.953**	0.976**	0.917**	0.667**	0.973**
Κατεύθυνση-Σχήμα, Λαβή Χορού	63%	88%	0.869**	0.930**	0.914**	0.467*	0.946** ⁵¹
Θέση και Στάση του Σώματος	75%	88%	0.894**	0.944**	0.986**	0.628*	0.950**
Ποιότητα Κίνησης	75%	100%	0.957**	0.978**	0.950**		0.981**
Εκφραστικότητα	63%	100%	0.921**	0.959**	0.956**		0.955**
Συνολική Βαθμολογία	100%	100%	0.972**	0.986**	0.994**		0.987**

*p<0.05

**p<0.001

Συζήτηση

Σκοπός της παρούσας εργασίας ήταν να κατασκευαστεί και να ελεγχθεί η εγκυρότητα και η αξιοπιστία ενός εργαλείου αυθεντικής αξιολόγησης. Για τον παραπάνω λόγο αναπτύχθηκε και δοκιμάστηκε μια ρούμπρικα για την αξιολόγηση της απόδοσης του ελληνικού χορού. Το συγκεκριμένο εργαλείο δομήθηκε από κριτήρια-κατεύθυνσεις τα οποία συμπεριέλαβαν στοιχεία σχετικά με το τι κινήσεις κάνει ο χορευτής χρησιμοποιώντας το σώμα του, πόσο δυναμικά τις αποδίδει, πώς χρησιμοποιεί το χώρο και το χρόνο για την απόδοση των κινήσεων. Ένας συνήθης δείκτης αξιοπιστίας εργαλείου αξιολόγησης είναι η απόλυτη (absolute agreement) και η σχετική συμφωνία (adjacent agreement) μεταξύ των κριτών (Tinsley & Weiss 2000).

Στην παρούσα μελέτη, επειδή είναι καλό να χρησιμοποιείται πάνω από ένας τρόπος μέτρησης, δεδομένου ότι μία μόνο μέθοδος δεν είναι πάντα κατάλληλη κάτω από όλες τις περιστάσεις, η ρούμπρικα ελέγχθηκε μέσα από τη διερεύνηση της συγκλίνουσας, της εγκυρότητας περιεχομένου, της απόλυτης και σχετικής συμφωνίας των μετρήσεων, του δείκτη ICC - ως αρκετά χρήσιμος στην περίπτωση ρούμπρικας με 5 ή περισσότερα επίπεδα απόδοσης - και των δεικτών Spearman's r, Cronbach's α και Cohen's k. Ο δείκτης Cohen's k προτείνεται για τον έλεγχο της συμφωνίας μεταξύ των κριτών (Oakleaf, 2009, Stellmack, Konheim-Kalkstein, Manora, Maseya & Schmitz, 2009) επειδή είναι πιο αυστηρός, καθώς υπολογίζει και τη συμφωνία που έχει προκύψει από τύχη. Στη

συγκεκριμένη εργασία χρησιμοποιήθηκε όπου ήταν εφικτό, επειδή για τον υπολογισμό του απαιτείται η απόδοση όλων των πιθανών βαθμολογιών απ' όλους τους κριτές.

Αναφορικά με την εγκυρότητα περιεχομένου, αυτή διασφαλίστηκε μέσω της κρίσης και της ομόφωνης (Newman et al. 2013, Pineda, 2012) τελικής απόφασης των ειδικών του χορού σχετικά με την καταλληλότητα και την επάρκεια των κριτηρίων, επειδή, με βάση και τη βιβλιογραφία (Kimberlin & Winterstein, 2008), δεν υπάρχει κάποιος στατιστικός δείκτης που να καθορίζει αν ένα όργανο καλύπτει επαρκώς το φάσμα όλων των χαρακτηριστικών της υπό διερεύνηση έννοιας, αυτή εξαρτάται από την γνώμη των ειδικών. Σύμφωνα με την Stellmack και τους συνεργάτες (2009), η θεμελίωση της εγκυρότητας μπορεί να ελεγχθεί διερευνώντας κατά πόσο ευθυγραμμίζεται-συσχετίζεται η βαθμολογία-σκορς που λαμβάνεται με τη χρήση ρούμπρικας με την αντίστοιχη που αποδίδει, χωρίς τη χρήση της, ένας ειδικός του χώρου-περιεχομένου μέσα στον οποίο εντάσσεται η υπό διερεύνηση έννοια. Στην παρούσα μελέτη πραγματοποιήθηκε συσχέτιση της βαθμολογίας που έλαβαν οι φοιτητές χρησιμοποιώντας τη ρούμπρικα με τη βαθμολογία που τους αποδόθηκε από έναν εμπειρογνώμονα-ειδικό του χορού, χωρίς να χρησιμοποιηθεί η ρούμπρικα, με τρόπο δηλαδή υποκειμενικό. Οι τιμές συσχέτισης, αρκετά υψηλές και θετικές (από 0.845 έως 0.906), κρίνονται αποδεκτές για τη θεμελίωση εγκυρότητας που στην παρούσα μελέτη χαρακτηρίζεται ως συγκλίνουσα (τύπος δομικής εγκυρότητας), διότι τόσο η ρούμπρικα όσο και η υποκειμενική αξιολόγηση του ειδικού-εμπειρογνώμονα εκτιμούν εννοιολογικά την ίδια μεταβλητή (απόδοση του ελληνικού χορού) επιφέροντας τα ίδια αποτελέσματα. Ταυτόχρονα εκφράζεται και ο βαθμός που η συγκεκριμένη ρούμπρικα αποτυπώνει το θεωρητικό πλαίσιο ανάλυσης και καταγραφής της κίνησης Laban στο οποίο βασίστηκε η κατασκευή της, συμφωνώντας και με τους Ουζούνη και Νακάκη (2011) και τους Kimberlin και Winterstein (2008), αναφορικά με την σχέση ενός εργαλείου μέτρησης και του θεωρητικού πλαισίου στο οποίο βασίστηκε.

Εξετάζοντας τη συμφωνία των κριτών στη συνολική βαθμολογία που έλαβε κάθε φοιτητής από καθέναν από τους κριτές, διαπιστώνεται ότι η απόλυτη συμφωνία μεταξύ τους βρίσκεται στο 85% και η σχετική στο 100% για το χορό Χασαπιά. Υψηλές εμφανίζονται και οι αντίστοιχες τιμές των κριτηρίων, με την απόλυτη και σχετική συμφωνία να αγγίζουν το 100%. Εξαιρέση αποτελεί το κριτήριο «ποιοτικά στοιχεία κίνησης» για το οποίο ο δείκτης ICC (0,826 $p < 0,001$) δείχνει πολύ υψηλή συμφωνία, ενώ οι δείκτες Spearman's r (0,855 $p < 0,001$) και Cronbach's α (0,916 $p < 0,001$) πολύ υψηλή συσχέτιση. Επίσης, οι τιμές 50% και 57% χαρακτηρίζουν μέτρια την απόλυτη συμφωνία μεταξύ των κριτών για τα κριτήρια «συγχρονισμός» και «ακολουθία βημάτων». Για τα συγκεκριμένα κριτήρια επίσης οι τιμές των παραπάνω δεικτών (ICC, Spearman's r και Cronbach's α) αντανάκλουν υψηλή συμφωνία (Πίνακας 3.). Πρέπει να τονιστεί ότι η βαθμολογία των δύο κριτών και στα τρία κριτήρια που αναφέρθηκαν παραπάνω, διέφερε μόνο κατά ένα επίπεδο, εκτός από μία περίπτωση, με αποτέλεσμα η σχετική συμφωνία και στα τρία να αγγίζει το 93%. Το αξιολογικό αυτό εύρημα είναι αρκετά ενθαρρυντικό για την αξιοπιστία στη χρήση της συγκεκριμένης ρούμπρικας, διότι, αν σε μία ρούμπρικα υπάρχουν περισσότερα από 4 ή 5 επίπεδα, η «σχετική» συμφωνία (adjacent agreement) θεωρείται ίσως πιο ρεαλιστική μέτρηση (Graham et al. 2012). Θα πρέπει επίσης να αναφερθεί ότι στο κριτήριο «ανταπόκριση στο ρυθμικό πρότυπο» ο δείκτης ICC είναι χαμηλός (0,458, $p < 0,001$), ενώ η απόλυτη και η σχετική συμφωνία είναι υψηλές (86%/100%). Αυτό συμβαίνει γενικά όταν δεν υπάρχει μεγάλη διακύμανση τιμών (στο παρόν κριτήριο σχεδόν όλοι οι φοιτητές έχουν βαθμολογηθεί με 5 και από τους δύο κριτές), αποτελεί ένα από τα μειονεκτήματα του δείκτη και μπορεί να παραπλανήσει κατά την εκτίμηση (Graham et al. 2012).

Η αντίστοιχη απόλυτη συμφωνία των κριτών στη συνολική βαθμολογία του Ζωναράδικου χορού ήταν 79% και η σχετική 100%, ενώ οι τιμές που σημειώθηκαν για κάθε κριτήριο χωριστά, κυμαίνονται πάνω από 71% και μέχρι 93% η απόλυτη, και σχεδόν 100% η σχετική συμφωνία. Οι υπόλοιποι στατιστικοί δείκτες σε αρκετά υψηλές τιμές, στοιχία τα οποία καθιστούν τη συγκεκριμένη ρούμπρικα αξιόπιστη και για τον συγκεκριμένο χορό. Εξαιρέση αποτελεί η τιμή 43% για τα κριτήρια «κατεύθυνση-σχήμα-λαβή» και «θέση και στάση του σώματος», για τα οποία όμως οι τιμές της σχετικής συμφωνίας φτάνουν το 79% και 93% αντίστοιχα (Πίνακας 4.). Επιπλέον, επειδή η βαθμολογία των δύο κριτών διαφέρει κατά ένα μόνο επίπεδο, υπολογίζοντας το μέσο όρο των βαθμολογιών τους στα συγκεκριμένα κριτήρια, η διαφορά ανέρχεται στο 0,5. Η διαφορά αυτή είναι πολύ μικρή ώστε να μην παραποιεί την πληροφόρηση που παίρνει ο φοιτητής αναφορικά με την απόδοσή του στα κριτήρια αυτά. Το στοιχείο αυτό ενισχύει τον εκπαιδευτικό ρόλο της ρούμπρικας.

Τέλος τα αντίστοιχα στοιχεία για το χορό Έντεκα (Πίνακας 5.) οδηγούν στα ίδια συμπεράσματα σχετικά με την αξιοπιστία, καθόσον οι περισσότερες τιμές που προέκυψαν, βρίσκονται σε αποδεκτά επίπεδα. Εξαιρείται η τιμή 36% που αντανάκλα την απόλυτη συμφωνία για το κριτήριο «ποιοτικά στοιχεία κίνησης» που όμως μετα-

τρέπεται στο 100% για τη σχετική συμφωνία και όπως προαναφέρθηκε, θεωρείται πιο ρεαλιστικός τρόπος μέτρησης σε ρούμπρικα με 5 ή περισσότερα επίπεδα απόδοσης.

Αξιολογώντας τις τιμές της απόλυτης συμφωνίας τόσο για τη συνολική βαθμολογία όσο και για κάθε κριτήριο χωριστά, σε όλους τους χορούς, φαίνεται ότι οι περισσότερες υπερβαίνουν το 75%, το οποίο θεωρείται ως αποδεκτό της απόλυτης συμφωνίας μεταξύ των κριτών, όταν μία ρούμπρικα, όπως και της παρούσας μελέτης, αποτελείται από 5 έως 7 επίπεδα απόδοσης (Graham et al. 2012). Αρκετές υπερβαίνουν επίσης και το 70%, τιμή που από άλλους ερευνητές (Jonsson & Svingby, 2007, Reddy & Andrade, 2010) θεωρείται ως αποδεκτό όριο απόλυτης συμφωνίας μεταξύ κριτών, αφού η απόλυτη συμφωνία τονίζουν, εξαρτάται σε πολύ μεγάλο βαθμό από τον αριθμό των επιπέδων. Όσο λιγότερα τα επίπεδα, τόσο περισσότερες οι πιθανότητες για απόλυτη συμφωνία. Αρκετές πάντως μελέτες στη βιβλιογραφία αναφέρουν ένα εύρος τιμών απόλυτης συμφωνίας μεταξύ των κριτών, το οποίο κυμαίνεται από 55%-75% (Garcia-Ros, 2011, Jonsson & Svingby, 2007). Όσον αφορά στις αντίστοιχες τιμές των υπόλοιπων δεικτών, σχεδόν όλες βρίσκονται μέσα στα αποδεκτά όρια και πάρα πολλές έχουν πάρει τιμές πάνω από 0,90, πληρώντας σε μεγάλο βαθμό τις προϋποθέσεις αξιοπιστίας. Θα πρέπει να αναφερθεί ότι, όπου ήταν εφικτό και υπολογίστηκε ο δείκτης Cohen's k , πήρε αρκετές τιμές πάνω από το όριο των 0,70, που θεωρείται πολύ υψηλή ένδειξη αξιοπιστίας και κάποιες, μέσα στο εύρος 0,40-0,70, που αποτελούν δείκτες καλής και επαρκούς αξιοπιστίας. Επιπλέον συμφωνούν με τις τιμές του συντελεστή που κυμάνθηκαν από 0,40-0,82 σε μελέτη για την αξιολόγηση οργάνου που κατασκευάστηκε για την καταγραφή κινήσεων ασθενών και θεραπειών σε συνεδρίες ψυχοθεραπείας (Davis, 1987).

Ανάλογα εμφανίζονται και τα αποτελέσματα αναφορικά με τη συμφωνία στη συνολική βαθμολογία δύο διαφορετικών αξιολογήσεων που έγιναν από τον ίδιο κριτή (intra-class agreement), σε διαφορετικές χρονικές στιγμές. Παρατηρείται απόλυτη συμφωνία που φτάνει στο 79%, 90% και 100% για τους χορούς Χασαπιά, Ζωνάραδικο και Έντεκα αντίστοιχα, και σχετική να αγγίζει το 100% και στους τρεις χορούς. Αναλύοντας χωριστά τα κριτήρια κάθε χορού, διαπιστώνεται ότι σχεδόν σε όλα η απόλυτη συμφωνία βρίσκεται πάνω από τα αποδεκτά όρια και η σχετική, εκτός από δύο κριτήρια στο χορό Έντεκα με τιμές 88%, αγγίζει το 100%. Οι υπόλοιποι στατιστικοί δείκτες κυμαίνονται από 0,90 και πάνω, εκτός από κάποιες τιμές που βρίσκονται μεταξύ 0,80 και 0,85. Τα παραπάνω αποτελέσματα ενισχύουν περαιτέρω την αξιοπιστία στη χρήση της συγκεκριμένης ρούμπρικας. Τονίζεται ότι και στην περίπτωση της επαναξιολόγησης παρατηρήθηκε χαμηλή τιμή του δείκτη ICC (0,629 $p < 0,05$) στο κριτήριο «ρυθμικό πρότυπο» του χορού Χασαπιά, λόγω της έλλειψης διακύμανσης των τιμών, όπως προαναφέρθηκε, ενώ όπου υπολογίστηκε ο δείκτης Cohen's k πήρε αρκετές τιμές από 0,70-1 και κάποιες από 0,40-0,70, τιμές τόσο υψηλές όσο και επαρκούς αξιοπιστίας αντίστοιχα (Davis, 1987).

Στην παρούσα μελέτη έγινε μια προσπάθεια θεμελίωσης της εγκυρότητας και αξιοπιστίας μιας ειδικής αναλυτικής ρούμπρικας, η οποία κατασκευάστηκε αποτελούμενη από οκτώ θεμελιώδη κριτήρια με πέντε επίπεδα απόδοσης το καθένα. Στόχος της η δίκαιη και συνεπής αξιολόγηση της μάθησης του ελληνικού χορού τόσο συνολικά όσο και κατά κριτήριο, αλλά και η συμβολή της στη σύνδεση της αξιολόγησης με τη διδασκαλία του χορού καθώς, με τον εντοπισμό των δυνατοτήτων και των αδυναμιών των φοιτητών, καθίστανται πιο ξεκάθαροι οι στόχοι της. Αυτό που πρέπει να τονιστεί είναι ότι η αξιολόγηση μέσω παρατηρούμενων κριτηρίων απαιτεί την ανάπτυξη ικανότητας παρατήρησης κάθε κριτηρίου χωριστά κι όχι της γενικής εικόνας ή εντύπωσης που προκαλεί ο εξεταζόμενος στον αξιολογητή. Ερμηνεύοντας τα αποτελέσματα της παρούσας μελέτης διαπιστώνεται ότι η ρούμπρικα, αποτελεί ένα αξιόπιστο στη χρήση του εργαλείο, καθόσον φαίνεται ότι καλύπτει εννοιολογικά το εύρος της μεταβλητής που μετράει, ενώ οδηγεί σε παρόμοιες βαθμολογίες, ανεξάρτητα από ποιον πραγματοποιείται η αξιολόγηση.

Σημασία για τη Φυσική Αγωγή

Η συνεπής χρήση μιας καλά σχεδιασμένης κλίμακας διαβαθμισμένων κριτηρίων (ρούμπρικας) βελτιώνει σημαντικά τη **διευκόλυνση** της μάθησης, παρέχοντας τόσο στους μαθητές όσο και στους εκπαιδευτικούς σαφήνεια για τον από κοινού στόχο. Οι μαθητές αξιολογούνται ατομικά, συγκρινόμενοι μόνο με την προηγούμενη προσπάθειά τους και όχι με εκείνη των συμμαθητών τους και μπορούν να επικυρώσουν καλύτερα τη δική τους πρόοδο. Οι εκπαιδευτές μπορούν δίκαια και με συνέπεια να τεκμηριώσουν το επίπεδο δεξιοτήτων και ανάπτυξης των μαθητών. Η χρήση ρούμπρικας σε διάφορα εκπαιδευτικά προγράμματα μπορεί να εξασφαλίσει, επίσης, τη συνεπή μέτρηση της ποιότητας των επιδόσεων των μαθητών που έχουν διαφορετικούς εκπαιδευτές.

Σημασία για την Ποιότητα Ζωής

Ένα υγιές εκπαιδευτικό σύστημα πρέπει να παρέχει τη δυνατότητα βελτίωσής του, ώστε να εγγυάται τη δυνατότητα των παιδιών να παρακολουθούν την εξέλιξη της γνώσης. Η χρήση της ρούμπρικας, ως εργαλείο αυθεντικής αξιολόγησης, μιας διαδικασίας που εμπλέκει και τους μαθητές στο μηχανισμό αξιολόγησης, τους εφοδιάζει με πρακτικές που αναδεικνύουν το δημιουργικό τρόπο σκέψης και δράσης, εφόδιο απαραίτητο καθ' όλη τη διάρκεια της ζωής.

Βιβλιογραφία

- Andrade, H. G. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, 57(5), 13-18.
- Angelo, T. (2002). Fostering critical thinking in our courses: Practical, research-based strategies to improve learning. The Pennsylvania State University Teaching and Learning Colloquy VIII. May 8, 2002.
- Bargainnier, S. (2003). *Fundamentals of Rubrics*. Kenesiology. The Pennsylvania State University, Faculty Development Series.
- Benander, R., Denton, J., Page, D. & Skinner, C. (2000). Primary Trait Analysis: Anchoring Assessment in the Classroom. *The Journal of General Education*, 49 (4), 279-302.
- British Columbia Ministry of Education. (1998). *Appendix D: Assessment and evaluation*. Ανακτήθηκε στις 20 Αυγούστου του 2012 από την ηλεκτρονική διεύθυνση http://www.bced.gov.bc.ca/irp/dan11_12/apdint.htm.
- Dania, A., Koutsouba, M., Hatziharistos, D. & Tyrovola, V. (2010). Establishing observer reliability in the context of dance performance assesement. VII congreso de la asociación internacional de escuelas superiores de educación física (AIESEP), A Coruña, 26-29 de Octubre.
- Davis, M. (1987). Steps to achieving observer agreement: The LIMS Reliability Project. *Movement Studies: Observer Agreement*, 2, 7-19.
- Διγγελίδης, Ν. (2009). Αξιολόγηση, Βαθμολόγηση και Μάθηση στην Τριτοβάθμια Εκπαίδευση. ΤΕΦΑΑ, Πανεπιστήμιο Θεσσαλίας. Ανακτήθηκε στις 10 Νοεμβρίου του 2009 από την ηλεκτρονική διεύθυνση http://www.uth.gr/static/miscdocs/Axiologisi_Mathimatwn.pdf.
- Garsia-Ros, R. (2011). Analysis and validation of a rubric to assess oral presentation skills in university contexts. *Electronic Journal of Research in Educational Psychology*, 9(3), 1043-1062.
- Graham, M., Milanowski, A. & Miller, J. (2012). *Measuring and Promoting Inter-Rater Agreement of Teacher and Principal Performance Ratings*. Center for Educator Compensation Reform. Ανακτήθηκε στις 12 Ιουλίου του 2012 από την ηλεκτρονική διεύθυνση http://cecr.ed.gov/pdfs/Inter_Rater.pdf
- Hagood, T. K. (2001). Dance to read or dance to dance? *Arts Education Policy Review*, 102(5), 21-26.
- Haynes, S. N., Richard, D. C. S. & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7, 238-247.
- International Folk Music Council - IFMC. (1974). Study group of folk dance terminology, foundations for the analysis of the structure and form of folk dance. A syllabus, *Yearbook of the I.F.M.C.*, 6, 115-135.
- Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130-144.
- Kassing, G., & Jay, D. M. (2003). *Dance Teaching Methods and Curriculum Design*. Unitade States: Human Ki-

netics.

- Kimberlin, C.L. & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacists*, 65(1), 2276-2284.
- Kutlu, O., Yildirim, O. & Bilican, S. (2010). The comparison of the views of teachers with positive and negative attitudes towards rubrics. *Procedia Social and Behavioral Sciences*, 9, 1566-1573.
- Κουλουμπαρίτση, Χ., Α. & Ματσαγγούρας, Γ., Η. (2004). «Φάκελος εργασιών του μαθητή (Portfolio): Η αυθεντική αξιολόγηση στη διαθεματική διδασκαλία» στο Π. Αγγελίδης και Γ. Μαυροειδής (επιμ.), *Εκπαιδευτικές Καινοτομίες για το Σχολείο του Μέλλοντος*. Αθήνα: Τυπωθείτω.
- Martin, G. & Pesovár, E. (1961). A Structural analysis of the Hungarian folk dance. *Acta Ethnographica*, 10, 1-40.
- Moni, R.W., E. Beswick, & K.B. Moni, (2005). Using student feedback to construct an assessment rubric for a concept map in physiology. *Advances in Physiology Education*, 29, 197-203.
- Moskal, B., M. (2000). Scoring rubrics: what, when and how? *Practical Assessment, Research & Evaluation*, 7(3). Ανακτήθηκε στις 4 Μαΐου του 2010 από την ηλεκτρονική διεύθυνση <http://PAREonline.net/getvn.asp?v=7&n=3>.
- Moskal, B. M. & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment Research & Evaluation*, 7(10). Ανακτήθηκε στις 11 Ιουλίου του 2012 από την ηλεκτρονική διεύθυνση <http://PAREonline.net/getvn.asp?v=7&n=10>.
- New York City Department of Education (2005). *Blueprint for teaching and learning in the arts: Dance grades pre K-12*. New York: Author.
- New Zealand Ministry of Education (2003). Ανακτήθηκε στις 20 Αυγούστου του 2012 από την ηλεκτρονική διεύθυνση www.nzqa.govt.nz/ncea/assessment/exemplar/index.html.
- Newman, I., Lim, J. & Pineda, F. (2013). Content Validity Using a Mixed Methods Approach: Its Application and Development Through the Use of a Table of Specifications Methodology. *Journal of Mixed Methods Research*, XX(X), 1-18.
- Oakleaf, M., (2009). Using Rubrics to Assess Information Literacy: An Examination of Methodology and Inter-rater Reliability. *Journal of the American Society for Information Science and Technology*, 60(5), 969-983.
- Office for Standards in Education (2002). *Inspecting dance 11-16 with guidance on self-evaluation*. Ανακτήθηκε στις 20 Αυγούστου του 2012 από την ηλεκτρονική διεύθυνση www.ofsted.gov.uk.
- Olson, L., Schieve, A., D., Ruit, G., K. & Vari, C. R., (2003). Measuring Inter-rater Reliability of the Sequenced Performance Inventory and Reflective Assessment of Learning (SPIRAL). *Academic Medicine*, 78, 844-850.
- Oreck, B. (2007). *To see and to share: Evaluating the dance experience in education*. In L. Bresler (Ed.), *International Handbook of Research in Arts Education* (pp. 341-356).
- Ουζούνη, Χ., & Νακάκης Κ. (2011). Η Αξιοπιστία και η Εγκυρότητα των Εργαλείων Μέτρησης σε Ποσοτικές Μελέτες. *ΝΟΣΗΛΕΥΤΙΚΗ*, 50(2), 231-239.
- Παναγοπούλου Β., Σερμπέζης Β. & Αγγελούσης Ν. (2009). Αξιολόγηση της χορευτικής απόδοσης παιδιών πρώτης σχολικής ηλικίας. Στο 17^ο Συνέδριο Φυσικής Αγωγής και Αθλητισμού, ΤΕΦΑΑ, Δημοκρίτειο Πανεπιστήμιο Θράκης. Ανακτήθηκε στις 10 Αυγούστου του 2010 από την ηλεκτρονική διεύθυνση <http://www-new.phyed.duth.gr/undergraduate/images/files/congress/2009/abstract/Dance.pdf>.
- Παπαϊωάννου, Α., & Λαπαρίδης, Κ., Μιχαλοπούλου, Μ., Πυλιανίδης, Θ., Ζέτου, Ε., Γουλιμάρης, Δ., Καρυπίδης, Α., Μυλώσης, Δ., Χριστοδουλίδης, Τ., Κούλη, Ο., Καλογιάννης, Π. & Κωνσταντινίδου, Ξ. (2006). *Φυσική Αγωγή Β' Γυμνασίου, Βιβλίο Εκπαιδευτικού*. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων.
- Perpich Center for Arts Education (2005). *Research, assessment and curriculum*. Ανακτήθηκε στις 20 Αυγούστου του 2012 από την ηλεκτρονική διεύθυνση www.pcae.k12.mn.us/rac/pubs/facs/dance.pdf.
- Pineda, M., F. (2012). Developing a Process to Create and Validate an Instrument Assessing Student Attainment of Competencies at an Intercultural University in Mexico. FIU Electronic Theses and Dissertations. Florida International University.
- Reddy, Y., M. & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35 (4), 435-448.
- Saskatchewan Education (1991). *Student evaluation: A teacher handbook*. Ανακτήθηκε στις 20 Αυγούστου του 2012 από την ηλεκτρονική διεύθυνση <http://www.sasked.gov.sk.ca/docs/policy/studeval/index.html>.
- Schmid, W., D. (2003). Authentic Assessment in the Arts. Empowering Students and Teachers. *Journal of Dance Education*, 3(2), 65-73.
- Slettum, B., S., Fox, C., Looney, M., A. & Jay, D., M. (2001). Validity and Reliability of a Folk-Dance Performance Checklist for Children. *Measurement in Physical Education and Exercise Science*, 5(1), 35-55.

- Stellmack, A., M., Konheim-Kalkstein, L., Y., Manora, E., J., Maseya, R., A. & Schmitz, P., J. A. (2009). An Assessment of Reliability and Validity of a Rubric for Grading APA Style Introductions. *Teaching of Psychology*, 36 (2), 102 - 107.
- Stevens, D., D. & Levi, A. (2005). *Introduction to rubrics: An assessment tool to save grading time, convey feedback, and promote student learning*. Stylus Publishing, LLC, Sterling.
- Thaler, N., Kazemi, E. & Huscher, C. (2009). Developing a Rubric to Assess Student Learning Outcomes Using a Class Assignment. *Teaching of Psychology*, 36, 113-116.
- The National Center for Education Statistics Focus on NAEP (1998). NAEP and Dance: Framework and Field Tests, 3, (1).
- The Standards Site (2005). *The power of arts assessment in teaching and learning handbook*. Ανακτήθηκε στις 20 Αυγούστου του 2012 από την ηλεκτρονική διεύθυνση <http://www.standards.dfes.gov.uk>.
- Tinsley, H. E. A., & Weiss, D. J. (2000). Interrater reliability and agreement. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 95-124). San Diego, CA: Academic.
- Τυροβολά, Β. (1998). *Ελληνικοί Παραδοσιακοί Χορευτικοί Ρυθμοί*. Αθήνα: Gutenberg.
- Vars, F., G. (2002). Educational Connoisseurship, Criticism, and the Assessment of Integrative Studies. *Issues in Integrative Studies*, 20, 65-76.
- Warburton, C., E. (2002). From Talent Identification to Multidimensional Assessment: Toward new models of evaluation in dance education, *Research in Dance Education*, 3(2), 103-121.
- Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* (expanded 2nd edition). Baltimore, Association for Supervision & Curriculum Development.
- Zimmaro M. (2007). Using Rubrics to Grade Student Performance. Ανακτήθηκε στις 11 Ιουλίου του 2012 από την ηλεκτρονική διεύθυνση http://scholar.google.gr/scholar?q=Zimmaro+M.%2C+%282007%29.+Using+Rubrics+to+Grade+Student+Performance&btnG=&hl=el&as_sdt=0&as_vis=1.
- Ζωγράφου, Μ. (2003). *Ο χορός στην Ελληνική παράδοση*. Αθήνα: 'ART WORK'.