ScholarOne - Investigating Expert-Rater Agreement and Inter/Intra-Rater Reliability of Two Fundamental Movement Skills for the Locomotor Subscale of the FG-COMPASS

Lino Perez¹ and Ovande Furtado Jr²

¹Affiliation not available
²California State University Northridge College of Health and Human Development

August 08, 2023

Abstract

The Furtado-Gallagher Child Observational Movement Pattern Assessment System (FG-COMPASS) is an observational tool using sequential decisions to assess fundamental movement skill proficiency. The current version of the test has three locomotor and five manipulative skills. This study assessed expert-rater agreement and inter- and intra-rater reliability of two new scales to be added to the locomotor subscale. This study was divided into two phases. In Phase I, 60 children between the ages of 5 and 10 were filmed performing the gallop and vertical jump skills. An expert in motor behavior classified the videotapes using the newly created rating scales. Next, eight videos were selected for training purposes and 24 videos for testing purposes. In Phase II, 30 undergraduate students served as raters and underwent training before the testing session. The data were analyzed using weighted kappa (Kw) and the intra-class correlation coefficient (ICC). The results suggested a ‘very good’ agreement between the expert and raters for vertical jump (Kw = .96) and gallop (Kw = .89). The ICC values for vertical jump and gallop were .98 and .94, respectively. The mean kappa values for inter-rater reliability were considered ‘very good’ for vertical jump (MKw=.92) and ‘good’ for gallop (MKw=.78). The ICC values were .98 and .95 for vertical jump and gallop, respectively - which is considered ‘excellent’. The kappa values for intra-rater reliability were .96 and .85 for vertical jump and gallop, respectively. The ICC values were .98 and .92 for vertical jump and gallop, respectively. These results show that the proposed rating scales are reliable in assessing vertical jump and gallop. Future studies should focus on criterion-related validity and reliability evidence from live performances.

Hosted file


Hosted file
