

تعیین پایایی و روایی آزمون رشد حرکتی درشت در کودکان ۳ تا ۱۱ ساله شهر تهران

۸۵

تاریخ تصویب: ۸۸/۱۰/۱۲
تاریخ دریافت: ۸۸/۴/۲۱

❖ دکتر مهشید زارع زاده؛ عضو هیأت علمی دانشگاه شهید باهنر کرمان*
❖ دکتر احمد فرخی؛ استادیار دانشکده تربیت بدنی دانشگاه تهران
❖❖ دکتر انوشیروان کاظم نژاد؛ دانشیار دانشکده علوم پزشکی دانشگاه تربیت مدرس

چکیده:

هدف این مطالعه عبارت است از تعیین روایی سازه و سه جنبه از پایایی آزمون رشد حرکتی درشت (۲۸) در کودکان ۳ تا ۱۱ ساله تهرانی. این آزمون شامل دو خرده آزمون (جابه جایی و کنترل شیء) است. در سنجش الگوی حرکتی، ۱۲ مهارت بنیادی طراحی شدند. کل نمونه، شامل ۱۴۳۸ کودک بود که به طور مساوی بین هشت گروه سنی و دو جنسیت تقسیم شدند. ضریب پایایی همسانی درونی برای نمره جابه جایی و کنترل شیء، همچنین برای نمره مرکب کل، به طور میانگین، به ترتیب ۰/۷۸، ۰/۷۴، و ۰/۸۰ به دست آمد. پایایی همسانی درونی دختران و پسران نیز بالا بود. دامنه ضریب پایایی آزمون - آزمون مجدد از ۰/۶۵ تا ۰/۸۱ و پایایی درون نمره گذار بالای ۰/۹۵ حاصل شد. در ارزیابی روایی سازه، همان طرح دو عاملی اولریخ (۲۸) فرض شد و این فرضیه با تحلیل عاملی تأییدی آزموده شد. طبق این طرح، شش مهارت یا متغیر آزمون که توانایی حرکت کودک در محیط را می سنجد، روی یک عامل (جابه جایی) و شش متغیر دیگر که توانایی کنترل و دستکاری اشیاء را اندازه می گیرند روی عامل دیگر جای گرفته اند (کنترل شیء). با توجه به نتایج حاصل، در جامعه ما نیز دو عاملی بودن ساختار آزمون، همچنین درستی تخصیص مهارت ها به عامل های جابه جایی و کنترل شیء تأیید شد. به علاوه، ضرایب همبستگی بین نمرات آزمون با سن، از دیگر جنبه های روایی سازه یعنی ماهیت رشدی آزمون حمایت کرد. ضرایب حاصل گویای قابلیت آزمون در ایجاد تمایز بین گروه های سنی بود. در نتیجه، این مطالعه نشان داد می توان آزمون را در سنجش رشد حرکتی درشت جامعه تحقیق با اطمینان به کار گرفت.

واژگان کلیدی: آزمون - آزمون مجدد، آزمون رشد حرکتی درشت، پایایی درون نمره گذار، روایی سازه، همسانی درونی.

* E.mail : zarezade.mahshid@gmail.com

مقدمه

عضلات بزرگ بدن اند (۲۸). این مهارت ها که به الگوهای حرکتی هدف مند اشاره می کنند، شامل حرکات جابه جایی و کنترل شیء می شوند. الگوهای حرکتی جابه جایی، مهارت هایی را دربر می گیرند

عصر اصلی رشد حرکتی، مهارت های حرکتی بنیادی است (۲۴) که در زمره حرکات درشت محسوب می شوند؛ یعنی حرکاتی که وابسته به عمل

www.SID.ir

از ویژگی‌های آزمون استاندارد روایی است. آزمون روا آزمونی است که شواهد تجربی و نظری، سودمندی و مناسب بودن کاربردها و تفاسیر مبتنی بر نمرات آزمون را تأیید کند (۲۹). یکی از جنبه‌های روایی، روایی سازه است. سازه خصوصیتی است که در عملکرد فرد در آزمون منعکس می‌شود (۱۰)، مثل هوش. رشد حرکتی و روایی سازه به این موضوع اشاره دارند که به چه میزان سازه‌های اصلی (عامل‌ها) یا متغیرهایی که ادعا شده با آزمون اندازه‌گیری می‌شوند قابل شناسایی اند (۲۷) و تا چه حد نمرات آزمون، مفاهیم نظری مربوط به آن سازه را منعکس می‌کنند (۵). برای بررسی روایی سازه عموماً در مورد تعداد سازه‌ها، ساختار درونی، روابط اجزای آزمون، و ارتباط سازه‌ها با دیگر متغیرها فرضیه‌سازی و فرضیه‌ها آزمایش می‌شوند (۲۸).

برای مشخص شدن تعداد سازه‌ها و روابط درونی آزمون، از روش تحلیل عاملی اکتشافی و تأییدی استفاده می‌شود. در تحلیل عاملی اکتشافی از طریق تحلیل واریانس مشترک در ماتریس همبستگی بین مقادیر متغیرهای آزمون، آن دسته از متغیرهایی که تغییراتشان همبسته است با عنوان عامل دسته‌بندی می‌شوند. بدین ترتیب، سازه‌های نهفته در آزمون استخراج و ساختار عاملی آزمون مشخص می‌شوند (۴). اما ساختار عاملی به‌دست آمده منحصر به جامعه‌ای است که آزمون در آن طراحی شده است. بنابراین، مناسب بودن طرح عاملی آزمون در جامعه جدید از طریق تحلیل عاملی تأییدی و با محاسبه شاخص‌های نیکویی برازش (GFI)^۲ ارزیابی

که جهت انتقال و حرکت بدن در محیط استفاده می‌شوند، مانند دویدن و پرش. الگوهای حرکتی کنترل‌شده نیز مستلزم اعمال نیرو به شیء و یا دریافت نیرو از شیء‌اند، مانند پرتاب و دریافت توپ. این مهارت‌ها در اوایل کودکی بر اثر بالیدگی جسمانی و آموزش و تمرین رشد می‌کنند و پیش‌نیاز بازی‌ها و حرکات تخصصی‌ترند (۱۵).

تحقیقات نشان داده‌اند تبحر ناکافی در این مهارت‌ها، باعث عدم موفقیت و احساس بی‌کفایتی در بازی‌ها می‌شود (۳۰) که نتیجه آن حذف تدریجی فعالیت جسمانی در طول عمر (۱۳) و سطح نامطلوب آمادگی جسمانی است (۲۴). به علاوه، کودکانی که در انجام این مهارت‌ها با ضعف روبه‌رویند، ارتباط اجتماعی آن‌ها به راحتی کودکان ماهرتر ممکن نیست و سطح اضطراب آن‌ها بالاتر و عزت‌نفس آن‌ها پایین‌تر است (۲۳). همچنین، تأخیر در رشد حرکتی درشت اختلال در ادراک بینایی (۲۶) و ناتوانی در خواندن و مهارت‌های مربوط به زبان را به همراه خواهد داشت که اگر این تأخیر به موقع شناسایی نشود، باعث ناتوانی‌های یادگیری و شناختی خواهد شد (۲۰).

از این رو، یک جنبه از برنامه‌های غربالگری در ابتدای کودکی باید اندازه‌گیری رشد حرکات بنیادی باشد که ضرورت وجود آزمون‌های استاندارد در این مورد را نشان می‌دهد. یکی از مشهورترین آزمون‌ها در حوزه سنجش حرکتی «آزمون رشد حرکتی درشت-۲ (TGMD-۲)»^۱ است (۲۸) که به منظور اندازه‌گیری رشد حرکات بنیادی کودکان ۳ تا ۱۱ ساله طراحی شد و شش مهارت جابه‌جایی و شش مهارت کنترل‌شده را دربر می‌گیرد.

1. Test of Gross Motor Development

2. Goodness of Fit Indexes

می‌شود (۲۷).

آن تأخیر رشد حرکتی درشت این دسته از کودکان ذکر شد (۲۵).

پایایی، دیگر ویژگی آزمون‌های استاندارد و شرط لازم برای روایی است (۵). پایایی یعنی نمره‌های آزمون تا چه اندازه دقیق و عاری از خطایند. طبق نظریه کلاسیک آزمون، نمره مشاهده شده (X) مجموع نمره واقعی (T) و نمره خطاست (E) و مفهوم پایایی بدین معناست که چند درصد از تغییرات نمره مشاهده شده دربر دارنده تغییرات نمره واقعی است که این مفهوم به صورت نسبت واریانس این دو نمره بیان می‌شود $(\frac{\sigma_x^2}{\sigma_t^2})$ (۱۰). طبق این مفهوم آزمونی پایاست که کمترین واریانس خطای اندازه‌گیری را داشته باشد (۲۷).

خطای اندازه‌گیری از سه منبع نشأت می‌گیرد: محتوا و سؤالات آزمون، شرایط اجرای آزمون، و نمره‌گذاری (۶). خطای هر منبع به طرق مختلف برآورد می‌شود. برای تعیین پایایی محتوا، از روش همسانی درونی استفاده می‌شود. در این روش، پایایی با تأکید بر میزان همسانی و یکنواختی مواد آزمون ارزیابی می‌شود. خطای وابسته در شرایط اجرای آزمون از طریق آزمون-آزمون مجدد برآورد می‌شود. خصوصاً، در آزمون‌هایی که انتظار می‌رود نمرات سازه مورد اندازه‌گیری در آن‌ها در مدت زمان کوتاهی تغییر نکنند، با این روش میزان ثبات عملکرد در طول زمانی معین بررسی می‌شود (۵). در این شیوه، آزمون در دو زمان متفاوت و در شرایطی یکسان، اجرا و همبستگی بین نمرات دوبار اجرا پایایی آزمون محسوب می‌شوند که به پایایی ثبات نیز معروف است (۶). پایایی بین نمره‌گذاران و

در ابتدا، ساختار عاملی ۲-TGMD را اولریخ (۲۸) در کودکان آمریکایی بررسی کرد. تحلیل عاملی اکتشافی، وجود دو عامل (جابه‌جایی و کنترل شیء) یا سازه اصلی را نشان داد. مقدار شاخص‌های نیکویی برازش طرح دو عاملی آزمون نیز قابل قبول بود که گویای درستی این ساختار عاملی و همچنین درستی تخصیص هر یک از مهارت‌ها به خرده‌آزمون مربوط بود (۲۸).

در خارج از جامعه آمریکا مطالعات انگشت‌شماری در مورد ساختار عاملی آزمون انجام شده که عبارت‌اند از مطالعه اواجلیسو و همکاران (۱۴) در مورد کودکان یونانی، مطالعه ونگک و چونگک (۳۱) در مورد کودکان هنگ‌کنگ، و مطالعه سیمونز و همکاران (۲۵) در مورد کودکان عقب‌مانده ذهنی بلژیکی. در این مطالعات نیز، برازش یا تناسب معقول دو عاملی بودن ساختار آزمون در جوامع مورد نظر تأیید شد.

علاوه بر بررسی ساختار عاملی، یکی دیگر از جنبه‌های روایی سازه آزمون‌هایی که سازه مورد اندازه‌گیری آن‌ها ماهیت رشدی دارند بررسی قابلیت تمایز سنی یا پیشرفت نمرات با افزایش سن است (۵). بنابراین، یکی از فرضیه‌هایی که در ارزیابی روایی سازه آزمون‌هایی مانند ۲-TGMD تدوین شد این است که بین عملکرد کودکان در آزمون و متغیر سن همبستگی بالایی وجود دارد (۲۸). همبستگی بین نمره خرده‌آزمون‌های ۲-TGMD و سن در کودکان آمریکایی (۲۸) و هلندی (۲۲) قابل قبول و معنادار مشاهده شد. اما در کودکان عقب‌مانده ذهنی بلژیکی، قابلیت تمایز سنی آزمون تأیید نشد که علت

1. stability

و تنها محدود به چند جامعه است. از طرفی، با توجه به نظر متخصصان سنجش و اندازه گیری، ارزش روایی و پایایی آزمون‌ها مختص همان جامعه بررسی شده است و این اطلاعات قابل تعمیم به همه موقعیت‌های فرهنگی و جغرافیایی نیست (۳۲). همین امر کاربری TGMD-۲ را در کودکان ایرانی محدود می‌کند، زیرا هیچ مطالعه مشابهی که پایایی و روایی آزمون را در این جامعه بررسی کرده باشد یافت نشد. در نتیجه، به کارگیری آزمون در جامعه‌ای با ویژگی‌های متفاوت، ابتدا باید با ارزیابی روایی و پایایی همراه باشد تا درستی کاربردهای مبتنی بر نتایج آن تضمین شود (۱۲). بر همین اساس و با توجه به پیامدهای همه جانبه رشد تأخیر در پیشرفت مهارت‌های حرکتی بنیادی و ضرورت سنجش آن‌ها، هدف این مطالعه عبارت است از تعیین روایی سازه و سه جنبه از پایایی TGMD-۲ - یعنی پایایی همسانی درونی، ثبات، و عینیت - در کودکان ۳ تا ۱۱ ساله شهر تهران.

روش‌شناسی

الف) روش تحقیق و آزمودنی‌ها. روش تحقیق مطالعات پایایی و روایی از نوع توصیفی - همبستگی است (۷). جامعه آماری این تحقیق را کودکان هشت گروه سنی ۳ تا ۱۱ سال دبستان‌ها و مهد کودک‌های آموزش و پرورش شهر تهران تشکیل دادند. بخشی از مطالعه حاضر، هنجارسازی TGMD-۲ بود. حجم نمونه به منظور ارزیابی پایایی همسانی درونی و روایی سازه، همان حجم نمونه هنجاری در نظر گرفته شد. از آنجا که مطالعه مشابهی در ایران انجام نشده تا با برآورد میانگین یا

درون نمره گذار، دو شیوه بررسی پایایی نمره گذاری یا عینیت^۱ هستند. در این روش که مستلزم ثبت اجرا از طریق فیلم برداری است، اجرا با دو نمره گذار یا با یک نمره گذار در دو زمان متفاوت نمره گذاری می‌شود و سپس همبستگی بین نمره گذاری‌ها محاسبه می‌گردد (۱۰).

اندازه پایایی با ضریب همبستگی بیان می‌شود و معیار کلی برای مقبولیت آن، مقادیر بالای ۰/۷۰ ذکر شده است. در محاسبه ضریب پایایی، روش آماری همبستگی درون طبقه‌ای (ICC)^۲ پیشنهاد می‌شود. ICC که در اصل بر مبنای تحلیل واریانس است، با بررسی تغییرات در میانگین و انحراف استاندارد از کوششی به کوشش دیگر یا در بین نمرات، برآورد دقیق‌تری از مقدار واریانس خطا در مقایسه با تکنیک همبستگی گشتاوری پیرسون یا همان همبستگی بین طبقه‌ای فراهم می‌کند. مشهورترین شکل پایایی بر مبنای ICC، ضریب آلفای کرونباخ است (۲۱).

در راهنمای TGMD-۲، ضرایب پایایی همسانی درونی، ثبات، و عینیت آزمون در کودکان آمریکایی بالا و نزدیک به ۱ گزارش شده است (۲۸). اوآجلینو و همکاران (۲۰۰۲) آلفای همسانی درونی خرده آزمون جابه‌جایی و کنترل شیء را در کودکان یونانی به ترتیب ۰/۷۷ و ۰/۷۵ به دست آوردند (۱۴). اما در مطالعه‌ای دیگر، آلفای همسانی درونی خرده آزمون جابه‌جایی در کودکان یونانی ۰/۶۷ به دست آمد. محققان آن را کمترین مقدار قابل قبول تفسیر کردند (۲۶). در مورد کودکان هلندی نیز آلفای پایایی ثبات و عینیت بالای ۰/۸۰ گزارش شد (۱۶).

هر چند که شواهد قابل قبولی از پایایی و روایی TGMD-۲ در دست است، اما این اطلاعات اندک

1. Objectivity

2. Intraclass Correlation Coefficient

یورتمه رفتن، لی لی کردن، گام کشیده جهیدن، پرش طول، و سُرخوردن. خرده آزمون کنترل شی عبارت است از ضربه به توپ ثابت با دست، دریل درجا، دریافت توپ، ضربه به توپ ثابت با پا، پرتاب توپ از بالای شانه، غلتاندن توپ از زیر شانه.

هر مهارت از طریق چند ملاک عملکردی سنجیده می شود. محتوای هر ملاک، بخشی از الگوی پیشرفته اجرای مهارت در مورد عمل اندامها و تنه است. هر خرده آزمون مجموعاً ۲۴ ملاک عملکردی دارد. جهت نمره گذاری، کودک هر مهارت را طی دو کوشش اجرا می کند. پس از هر کوشش به هر یک از ملاک های عملکردی مهارت، اگر همان طور که آزمون مشخص کرده انجام شده باشد، نمره ۱ و در غیر این صورت نمره صفر داده می شود. مدت زمان اجرای آزمون ۱۵ تا ۲۰ دقیقه تخمین زده شد. هر خرده آزمون نمره خام کلی دارد که از جمع نمره مهارت های مربوط به دست می آید و حداکثر ۴۸ است. نمرات هنجاری آزمون شامل نمره استاندارد خرده آزمونها به ترتیب با میانگین و انحراف استاندارد ۱۰ و ۳ و همچنین بهره حرکتی درشت (GMQ) است که نمره ای است مرکب بر پایه نمره استاندارد هر دو خرده آزمون. GMQ که میانگین و انحراف استاندارد آن به ترتیب ۱۰۰ و ۱۵ است بهترین اندازه در مشخص کردن توانایی کلی کودک در مهارت های حرکتی درشت تعریف شده است (۱۱).

ج) روش های آماری. پایایی خرده آزمونها

و GMQ به صورت پایایی همسانی درونی در هر یک از گروه های سنی، آزمون- آزمون مجدد، و پایایی درون نمره گذار ارزیابی شد. به منظور از بین بردن اثر سسن در محاسبه ضریب پایایی آزمون-

واریانس مطالعه و از طریق فرمول های نمونه گیری، حجم نمونه به منظور هنجارسازی محاسبه شود، با استناد به منابع روش تحقیق با این محتوا اندازه نمونه ۱۴۳۸ نفر در نظر گرفته شد زیرا هر چه در مطالعات توصیفی نمونه با حجم بزرگ تری انتخاب شود، پارامترهای جامعه دقیق تر برآورد می شوند (۳). این حجم، به طور مساوی در بین دو جنسیت و گروه های سنی تقسیم شد. در نمونه گیری، ابتدا پنج منطقه آموزش و پرورش از مناطق جغرافیایی شهر تهران به طور تصادفی انتخاب، سپس با روش نمونه گیری تصادفی خوشه ای طبقه ای چند مرحله ای، آزمودنی ها از این پنج منطقه انتخاب شدند.

از این حجم نمونه هنجاری، ۶۳ نفر و ۳۲ نفر به ترتیب برای ارزیابی پایایی آزمون- آزمون مجدد (ثبات) و پایایی درون نمره گذار (عینیت) انتخاب شدند. آزمون مجدد با فاصله ۱۲ روز و در شرایط یکسان انجام شد و محقق برای پایایی نمره گذاری، فیلم عملکرد کودکان در آزمون را طی دو بار با فاصله ۱۲ روز نمره گذاری کرد.

ب) ابزار جمع آوری اطلاعات. آزمون

رشد حرکتی درشت- ۲ (۲۸) ابزار جمع آوری داده های مطالعه بود که نسخه تغییر یافته TGMD اولریخ در سال ۱۹۸۵ است. این آزمون را اولریخ در سال ۲۰۰۰ (۲۸) با نمونه ای ۱۲۰۸ نفری از کودکان آمریکایی اعتباریابی و هنجارسازی کرد. ۲-TGMD که رشد الگو یا شکل اجرای ۱۲ مهارت حرکتی بنیادی کودکان هشت گروه سنی ۳ تا ۱۱ سال را می سنجد، شامل دو خرده آزمون است: خرده آزمون جابه جایی، و خرده آزمون کنترل شی. خرده آزمون جابه جایی عبارت است از دویدن،

1. Gross Motor Quotient

طریق نرم افزار آماری LISREL ۸/۸ آزموده شد. همچنین، با محاسبه ضریب همبستگی گشتاوری پیرسون بین متغیر سن و نمره کل خرده آزمون‌ها، قابلیت تمایز سنی آزمون ارزیابی گردید.

یافته‌ها

پایایی ۲-TGMD

پایایی همسانی درونی: جدول ۱ ضرایب آلفای پایایی همسانی درونی دو خرده آزمون و GMQ و همچنین میانگین آلفا را نشان می‌دهد. در انتهای جدول نیز آلفا برای دو زیرگروه دختر و پسر آورده شده است.

آزمون مجدد و درون نمره گذار خرده آزمون‌ها، نمره کل هر خرده آزمون به نمره استاندارد تبدیل شد. میانگین ضرایب پایایی همسانی درونی گروه‌های سنی، به روش تبدیل Z محاسبه شد که روش آماری معدل گیری چند ضریب همبستگی از طریق تبدیل مقادیر همبستگی به مقادیر Z است (۲۷). ضرایب پایایی خرده آزمون‌ها بر اساس ضریب همبستگی درون طبقه‌ای و به شیوه آماری آلفای کرونباخ، از طریق نرم افزار آماری SPSS ۱۴ محاسبه شدند. اما آلفا برای پایایی همسانی درونی GMQ، از طریق فرمول زیر محاسبه شد که آلفای طراحی شده برای نمرات مرکب است (۱۹).

$$\alpha = 1 - \frac{\sum \delta_{ii}^2 (1 - \rho_{iii})}{\delta_{ii}^2}$$

δ_{ii}^2 = واریانس نمره استاندارد خرده آزمون، ρ_{iii} = ضریب پایایی خرده آزمون، δ_{ii}^2 = واریانس GMQ

همان‌طور که جدول ۱ نشان می‌دهد، دامنه ضریب پایایی همسانی درونی ۲-TGMD در هشت گروه سنی از ۰/۶۵ تا ۰/۸۸ است. میانگین آلفا نیز که ضریب پایایی کل را نشان می‌دهد بالای ۰/۷۴ است. به علاوه، آلفای خرده آزمون‌ها و GMQ در هر یک از دو جنسیت بالای ۰/۸۵ است. هر چند ضریب پایایی بالای ۰/۷۰ عموماً معیار تفسیر پایایی است، اما حداقل مقدار قابل قبول برای آلفا ۰/۶۵ نیز ذکر شده است (۹). همچنین، هیر و همکاران (۱۹۹۸) در تحقیقات توصیفی نشان دادند این مقدار تا ۰/۶۰ هم کاهش می‌یابد (۲۶). بر این اساس، پایایی همسانی درونی ۲-TGMD در کودکان ۳ تا ۱۱ ساله شهر تهران تأیید می‌شود.

پایایی ثبات و عینیت. جدول ۲ آلفای کرونباخ

روایی سازه ۲-TGMD به صورت تحلیل عاملی تأییدی و قابلیت تمایز سنی آزمون ارزیابی شد. به منظور بررسی درستی ساختار و روابط درونی آزمون، همان طرح دو عاملی اولریخ که در کودکان آمریکایی تأیید شده بود در کودکان ۳ تا ۱۱ ساله شهر تهران نیز فرض شد. چنین فرض شد که عامل جابه‌جایی است و شش متغیر دویدن، یورتمه رفتن، گام کشیده جهیدن، لی لی کردن، پرش طول، و سر خوردن را دربر می‌گیرد و شش متغیر باقیمانده (ضربه به توپ با دست، دریبل درجا، دریافت توپ، ضربه به توپ با پا، پرتاب توپ، و غلتاندن توپ) به عامل کنترل شیء مربوط می‌شوند. این فرضیه با تحلیل عاملی تأییدی و به شیوه تخمین بیشترین درست‌نمایی، از

است که ثبات کافی نمرات ۲-TGMD را نشان می‌دهد. همچنین، ضریب پایایی درون نمره‌گذار یا عینیت آزمون بالا و نزدیک به ۱ است. در نتیجه پایایی نمره‌گذاری نیز در جامعه تحقیق تأیید می‌شود.

روایی ۲-TGMD

تحلیل عاملی تأییدی: شکل ۱ طرح شماتیک ساختار دو عاملی آزمون و روابط درونی اجزای آن را در کودکان این تحقیق نشان می‌دهد.

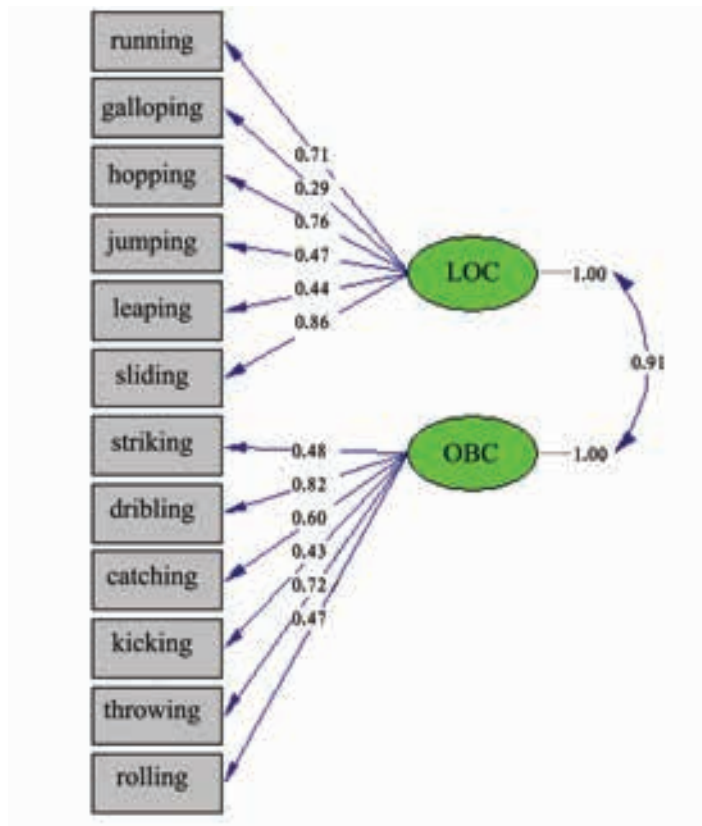
را برای پایایی آزمون-آزمون مجدد (ثبات) و پایایی درون نمره‌گذار (عینیت) نشان می‌دهد. برای کنترل اثر سن در این محاسبه از نمره استاندارد خرده‌آزمون‌ها استفاده شد. در این جدول میانگین (و انحراف استاندارد) نمره کل خرده‌آزمون‌ها طی دو بار آزمون و دو بار نمره‌گذاری نیز آورده شده است. محتوای جدول ۲ نشان می‌دهد، اختلاف میانگین نمره خام هر خرده‌آزمون در طی دو بار اجرای آزمون، به اندازه یک نمره و مقادیر α ۰/۶۵ تا ۰/۸۱

جدول ۱. ضرایب پایایی همسانی درونی خرده‌آزمون‌ها و بهره حرکتی درشت به تفکیک سن و جنسیت

GMQ	خرده‌آزمون		سن (تعداد)
	کنترل شی	جابه‌جایی	
۰/۸۶	۰/۷۲	۰/۸۸	۳ (۱۵۹)
۰/۸۳	۰/۶۹	۰/۸۷	۴ (۱۷۸)
۰/۸۰	۰/۶۹	۰/۸۳	۵ (۱۹۰)
۰/۸۱	۰/۷۵	۰/۷۷	۶ (۱۹۰)
۰/۷۷	۰/۷۶	۰/۷۵	۷ (۱۸۴)
۰/۷۹	۰/۷۷	۰/۶۶	۸ (۱۷۷)
۰/۸۴	۰/۷۸	۰/۷۴	۹ (۱۸۷)
۰/۷۲	۰/۷۱	۰/۶۵	۱۰ (۱۷۳)
۰/۸۰	۰/۷۴	۰/۷۸	میانگین α
۰/۹۱	۰/۸۵	۰/۹۲	دختر (۷۱۹)
۰/۹۲	۰/۸۹	۰/۹۱	پسر (۷۱۹)

جدول ۲. ضریب پایایی آزمون-آزمون مجدد و درون نمره‌گذار ۲-TGMD

پایایی درون نمره‌گذار			پایایی آزمون-آزمون مجدد			نمرات ۲-TGMD
ضریب عینیت (α)	نمره‌گذاری دوم	نمره‌گذاری اول	ضریب ثبات (α)	آزمون دوم	آزمون اول	
۰/۹۵	۳۱(۳)	۳۲(۳)	۰/۶۵	۳۳(۳)	۳۲(۳)	خرده‌آزمون جابه‌جایی
۰/۹۹	۲۴(۴)	۲۴(۵)	۰/۸۵	۲۱(۵)	۲۰(۶)	خرده‌آزمون کنترل شی
۰/۹۷	۱۰۰(۱۵)	۱۰۰(۱۵)	۰/۸۱	۱۰۰(۱۵)	۱۰۰(۱۵)	GMQ



شکل ۱. طرح شماتیک ساختار عاملی ۲- TGMD

همان‌طور که شکل ۱ نشان می‌دهد، همبستگی بین دو عامل کاملاً بالاست. این همبستگی بالا ممکن است گویای این باشد که به جای طرح دو عاملی، طرح یک عاملی با داده‌های به دست آمده تناسب دارد. بنابراین، در بررسی بیشتر صحت طرح عاملی مفروض، شاخص‌های نیکویی برآزش ملاحظه می‌شوند که جدول ۳ مقدار شاخص‌های نیکویی برآزش طرح دو عاملی آزمون را در کودکان این تحقیق نشان می‌دهد.

طبق شکل ۱، دامنه بار عاملی یا در واقع همبستگی مهارت‌ها بر پایه دو عامل جابه‌جایی (LOC) و کنترل شیء (OBC) از ۰/۲۹ تا ۰/۸۶ است. کمترین مقدار متعلق به یورتمه رفتن و بیشترین مقدار متعلق به سُرخوردن است. با توجه به اینکه منابع آمار بار عاملی بزرگ‌تر از ۰/۵۰ را بار عاملی بالا (۴) و بار عاملی کمتر از ۰/۳۰ را بار عاملی ضعیف (۸) بیان کرده‌اند، بارگیری مهارت‌ها روی عامل یا سازه مربوط مطلوب است.

جدول ۳. مقادیر شاخص‌های برازش در طرح دو عاملی ۲-TGMD

							شاخص برازش
NNFI	RMSEA	AGFI	GFI	χ^2 / df	df	χ^2	طرح عاملی
۰/۹۷	۰/۰۵۷	۰/۹۵	۰/۹۷	۵/۷۰	۵۳	۳۰۳/۹*	طرح دو عاملی

* $p \leq 0/05$

قابلیت تمایز سنی. جدول ۴، ضریب همبستگی بین متغیر سن و نمره دختران و پسران در خرده‌آزمون‌های ۲-TGMD را نشان می‌دهد. در این جدول میانگین (و انحراف استاندارد) نمره خرده‌آزمون‌ها نیز آورده شده است. با توجه به مبانی نظری رشد حرکتی، عملکرد حرکتی درشت با افزایش سن کودک بهبود می‌یابد. بنابراین فرض شد که همبستگی بین سن و نمره کل دو خرده‌آزمون بالا و معنادار است. جدول ۴ نشان می‌دهد میانگین عملکرد نمونه در خرده‌آزمون‌ها با بالا رفتن سن پیشرفت می‌کند که گویای ماهیت رشدی آزمون است. ضرایب به دست آمده نیز وجود رابطه‌ای معنادار و مثبت را نشان می‌دهند و مقدار آن‌ها به اندازه کافی بالا هست که از قابلیت تمایز سنی ۲-TGMD در جامعه کودک‌کان تحقیق حمایت می‌کند. اما مقادیر ضرایب حاصل نشان می‌دهند قابلیت تمایز سنی خرده‌آزمون کنترل شیء از خرده‌آزمون جابه‌جایی بیشتر است.

بحث

پایایی ۲-TGMD. در بررسی پایایی همسانی درونی ۲-TGMD، میانگین مقادیر ضریب آلفا در

تناسب کلی داده‌ها در ارزیابی طرح عاملی فرض شده، نخست بر مبنای معنادار نبودن χ^2 بررسی می‌شود. همان‌طور که در جدول ۳ می‌بینید، این شاخص در داده‌های این تحقیق معنادار است. اما χ^2 به شدت تحت تأثیر حجم نمونه قرار می‌گیرد و در نمونه‌های بزرگ احتمال معنادار بودن آن زیاد است. بنابراین، از شاخص‌های دیگر که معتبرتر از χ^2 اند استفاده می‌شود (۱۴). مقادیر دو تا پنج نسبت χ^2 به درجه آزادی (df)، نشان‌دهنده برازش معقول طرح عاملی مفروض است (۲۸). در شاخص نیکویی برازش (GFI)، شاخص نیکویی برازش اصلاح شده (AGFI) و شاخص تناسب هنجار نشده (NNFI) مقادیر بزرگ‌تر از ۰/۹۰ گویای برازش خوب طرح عاملی است و مقادیر کمتر از ۰/۰۸ ریشه میانگین مجذور خطای تقریبی (RMSEA) نیز برازش معقول طرح عاملی ذکر شده است (۱۴). با توجه به شاخص‌های برازش در جدول ۳ و معیار مقبولیت آن‌ها که در بالا ذکر کردیم، می‌توان گفت هر چند شاخص χ^2 معنادار است که آن هم به دلیل حجم بالای نمونه تحقیق است، اما مقدار شاخص‌های دیگر کاملاً معقول است. بر این اساس، درستی یا برازش دو عاملی بودن طرح ساختاری ۲-TGMD در جامعه این تحقیق تأیید می‌شود.

1. Adjusted Goodness of Fit Index
2. Non-Normed Fit Index
3. Root Mean Square Error of Approximation

جدول ۴. میانگین نمره کل (و انحراف استاندارد) خرده‌آزمون‌های ۲-TGMD به تفکیک سن و جنسیت

کنترل شیء		جابه‌جایی		سن
پسر	دختر	پسر	دختر	
۱۳/۵(۴)	۱۲(۳)	۱۷(۶/۵)	۱۶(۶/۵)	۳
۱۷(۴)	۱۵(۳/۵)	۲۵/۵(۶)	۲۴/۵(۶)	۴
۱۹(۴)	۱۷(۴)	۲۹(۵)	۲۸/۵(۵)	۵
۲۲(۵)	۱۸/۵(۴)	۳۱(۳/۵)	۳۱/۵(۴)	۶
۲۶/۵(۵)	۲۱(۴/۵)	۳۳(۳/۵)	۳۲(۳/۵)	۷
۲۸(۵)	۲۴(۵)	۳۳(۳)	۳۳/۵(۳)	۸
۳۰(۵/۵)	۲۵(۵)	۳۴(۴)	۳۴(۳)	۹
۳۲/۵(۴)	۲۷(۴)	۳۴/۵(۲/۵)	۳۴(۳/۵)	۱۰
۰/۷۹*	۰/۷۶*	۰/۶۶*	۰/۶۸*	همبستگی با سن (r)

* ۰/۰۵ p

ضریب پایایی اثرگذارند تغییرپذیری گروه است. با توجه به مفهوم پایایی در نظریه کلاسیک آزمون که پایایی را نسبت واریانس بین نمره واقعی و نمره مشاهده شده تعریف می‌کند، هرچه گروه از لحاظ توانایی مورد اندازه‌گیری ناهمگون‌تر باشد، یعنی هرچه دامنه تغییر نمرات یا پراکندگی در گروه بیشتر باشد، ضریب پایایی بزرگ‌تر است (۱۷). با مقایسه مشاهده‌ای انحراف استاندارد نمرات کودکان ایرانی با انحراف استاندارد همسالان آمریکایی خود که در راهنمای آزمون آورده شده، دیده شد در بیشتر گروه‌های سنی، انحراف استاندارد نمرات کودکان آمریکایی در دو خرده‌آزمون بیشتر بوده است (۲۸). بررسی پایایی آزمون- آزمون مجدد ۲-TGMD نشان داد ضریب ثبات دو خرده‌آزمون جابه‌جایی و کنترل شیء، همچنین نمره مرکب

دو خرده‌آزمون جابه‌جایی و کنترل شیء و نمره مرکب GMQ به ترتیب ۰/۷۸ و ۰/۷۴ و ۰/۸۰ به دست آمد. بر اساس اینکه مقادیر آلفای بالای ۰/۷۰، به صورت خوب و مقادیر بالای ۰/۸۰ به صورت خیلی خوب تفسیر می‌شوند (۹)، یافته‌های به دست آمده گویای این مطلب‌اند که پایایی همسانی درونی ۲-TGMD در کودکان ۳ تا ۱۱ ساله شهر تهران قابل قبول است. این بدین معناست که پاسخ آزمون‌شوندگان در طول محتوای دو خرده‌آزمون تجانس و همگنی کافی دارد.

در مقایسه ضرایب حاصل با آنچه اولریخ (۲۸) در مورد کودکان آمریکایی گزارش کرده، مشاهده شد ضرایب پایایی همسانی درونی دو خرده‌آزمون و GMQ در کودکان آمریکایی بالاتر است (به ترتیب ۰/۸۵، ۰/۸۸ و ۰/۹۱). از عواملی که بر اندازه

نمره گذاری، یا عینیت آزمون بالای ۰/۸۰ گزارش شده است. این امر منعکس کننده شیوه اجرا و نمره گذاری روشن و بدون ابهام آزمون رشد حرکتی درشت است.

روایی ۲-TGMD. بررسی درستی طرح دو عاملی ساختار آزمون اولریخ (۲۸) در کودکان ۳ تا ۱۱ ساله تهرانی، از طریق روش آماری تحلیل عاملی تأییدی مشخص کرد به جز مهارت یورتمه رفتن با بار عاملی ۰/۲۹، بار عاملی دیگر مهارت‌ها بر عامل یا خرده آزمون مربوط بالای ۰/۳۰ و قابل قبول است. در بحث چالشی اندازه مقادیر همبستگی در بررسی روایی آزمون‌ها، آناستازی و اورینا (۲۸) استدلال کردند در شرایطی معین، همبستگی‌هایی به کوچکی ۰/۲۰ نیز تعلق یک ماده یا یک خرده آزمون به یک آزمون را تأیید می‌کنند. بنابراین، بار عاملی مهارت یورتمه رفتن بر عامل جابه‌جایی نیز ممکن است قابل قبول در نظر گرفته شود.

در این قسمت، همبستگی بین دو عامل جابه‌جایی و کنترل شیء، ۰/۹۱ مشاهده شد. این همبستگی بالا نشان می‌دهد این دو عامل هر دو مهارت‌های حرکتی درشت را می‌سنجند و در واقع دو زیرسازه از سازه اصلی رشد حرکتی درشت‌اند. اما در مواردی که ضریب همبستگی، رابطه بین خرده آزمون‌های یک سازه را نشان می‌دهد، پیشنهاد شده که مقدار آن نباید خیلی بالا باشد (۵)، چون در این صورت همه خرده آزمون‌ها یک چیز را اندازه می‌گیرند، نه جنبه‌های متفاوت از یک سازه را. بنابراین، همبستگی مشاهده شده ممکن است منعکس کننده این باشد که دو خرده آزمون متمایز از یکدیگرند و ساختار دو عاملی ۲-TGMD با داده‌های نمونه این تحقیق

GMQ، به ترتیب ۰/۶۵، ۰/۸۵ و ۰/۸۱ است. ضریب به دست آمده برای خرده آزمون جابه‌جایی به صورت «قابل قبول» و برای خرده آزمون کنترل شیء و GMQ به صورت «خیلی خوب» تفسیر می‌شود. نتایج حاصل گویای این است که نمرات خرده آزمون کنترل شیء در مقایسه با نمرات خرده آزمون جابه‌جایی از ثبات بیشتری برخوردار است. اولریخ (۲۸) نیز چنین تفاوتی را مشاهده کرد (جابه‌جایی= ۰/۸۸، کنترل شیء= ۰/۹۳). به منظور توجیه این تفاوت می‌توان به مطالعه لی و همکاران (۱۸) استناد کرد. این محققان بر مبنای اصول نظری سؤال- پاسخ، ویژگی‌های روان‌سنجی ۲-TGMD را بررسی کردند و به این نتیجه دست یافتند که مستقل از ویژگی‌های گروه مورد اندازه‌گیری، مواد خرده آزمون کنترل شیء نسبت به خرده آزمون جابه‌جایی دقیق‌تر و دشوارترند. هر چه سؤالات آزمون دقیق‌تر باشد، دامنه وسیعی از توانایی مورد اندازه‌گیری را می‌سنجد، در نتیجه پایایی بالاتری نیز دارد (۲). در نتیجه می‌توان گفت پایایی بیشتر خرده آزمون کنترل شیء در آزمون- آزمون مجدد به دلیل دقیق‌تر بودن مواد آن در مقایسه با خرده آزمون جابه‌جایی است. از طرفی، به نظر می‌رسد به واسطه دشوارتر بودن مهارت‌های کنترل شیء، آزمون‌های حافظه یا یادداری کمتری از این مهارت‌ها در آزمون مجدد دارند. همین امر باعث تغییرپذیری کمتر عملکرد آن‌ها در آزمون مجدد و ثبات بیشتر نمره کنترل شیء آن‌ها شده است.

این تحقیق نشان داد نمره گذاری ۲-TGMD پایایی کاملاً بالایی دارد. در تمام مقالات مروری پایایی آزمون رشد حرکتی درشت، ضریب پایایی

(۲۲) نیز این مورد را مشاهده کردند. طبق جدول ۴، به نظر می‌رسد دلیل این تفاوت پیشرفت کندتر نمره جابه‌جایی در گروه‌های سنی بالاتر از هفت سال باشد. همان‌طور که جدول نشان می‌دهد، در همه گروه‌های سنی نمره جابه‌جایی از نمره کنترل شی بالاتر است که این امر، طبق مطالعه لی و همکاران (۱۸)، به دلیل دشوارتر بودن مواد خرده‌آزمون کنترل شی است. بهتر بودن عملکرد کودکان در خرده‌آزمون جابه‌جایی گویای این است که نمره آن‌ها به بالاترین حد نمره ممکن (۴۸) نزدیک‌تر است. بنابراین، به دلیل اثر سقف^۱ (۱)، پیشرفت نمره جابه‌جایی در گروه‌های سنی بالاتر، در مقایسه با نمره کنترل شی با کندی صورت می‌گیرد.

نتیجه‌گیری

نتایج این تحقیق از ساختار عاملی ۲-TGMD و مفاهیم نظری که آزمون بر مبنای آن‌ها طراحی شده حمایت کرد. مطالعه ابعاد مختلف پایایی نیز نشان داد نمرات آزمون از دقت کافی برخوردار است. بنابراین، می‌توان ۲-TGMD را آزمونی پایا و روا در سنجش مهارت‌های حرکتی بنیادی کودکان ۳ تا ۱۱ ساله شهر تهران به کار گرفت.

تناسب ندارد. اما مقدار شاخص‌های نیکویی برازش طرح دو عاملی مفروض از جمله GFI، NGFI، و NNFI نزدیک به یک و کاملاً مطلوب مشاهده شد که بیانگر تناسب معقول دو عاملی بودن ساختار ۲-TGMD و درستی روابط درونی آزمون در جامعه این مطالعه است.

بررسی روایی رشدی یا تمایز سنی ۲-TGMD، با محاسبه ضریب همبستگی بین سن و نمره دختران و پسران در خرده‌آزمون‌ها نشان داد این همبستگی معنادار و دامنه آن بین ۰/۶۶ و ۰/۷۹ است. در مطالعه اولریخ (۲۸)، توانایی تمایز سنی آزمون تأیید و دامنه همبستگی نمرات با سن از ۰/۶۹ تا ۰/۷۵ گزارش شد. نیمایجر (۲۲) نیز با مشاهده این همبستگی بین ۰/۶۶ تا ۰/۸۱، روایی رشدی آزمون را در کودکان هلندی تأیید کرد. هر چند مک‌ایچرون (۲۸) معیار تأیید این نوع همبستگی را ضرایب بالای ۰/۷۰ ذکر کرده، اما طبق تحقیقات ذکر شده در بالا، از مقادیر ضرایب همبستگی حاصل از قابلیت تمایز سنی ۲-TGMD در کودکان جامعه تحقیق حمایت می‌کند.

اما، در تحقیق حاضر، همبستگی بین سن و نمره جابه‌جایی کمتر از این همبستگی در خرده‌آزمون کنترل شی به دست آمد. اولریخ (۲۸) و نیمایجر

۱. اثر سقف بیان می‌کند همان‌طور که فرد به حد بالای امتیاز نزدیک می‌شود، ارتقا و بهبود عملکرد مشکل‌تر خواهد شد.

منابع

۱. اشمیت، ریچارد، ای.، ۱۳۷۵، مطالعه‌ی اساسی و شیوه‌های مطالعه‌ی رفتار و یادگیری حرکتی، ترجمه‌ی سید محمد کاظم واعظ موسوی. انتشارات دانشگاه امام حسین.
۲. بوام گارتنر، تد ای.؛ و جکسون، آندرو اس.، ۱۳۸۰، سنجش و اندازه‌گیری در تربیت بدنی (جلد اول)، ترجمه‌ی حسین سپاسی و پروین نوربخش، چاپ اول، تهران: انتشارات سمت.
۳. دلاور، علی، ۱۳۸۳، مبانی نظری و عملی پژوهش در علوم انسانی و اجتماعی، چاپ سوم، تهران: انتشارات رشد.
۴. سرمد، زهره؛ بازرگان، عباس؛ و حجازی، الهه، ۱۳۸۳، روش‌های تحقیق در علوم رفتاری، چاپ دهم، تهران: آگه.
۵. سیف، علی اکبر، ۱۳۸۳، اندازه‌گیری، سنجش و ارزشیابی آموزشی، چاپ شانزدهم (ویرایش سوم)، تهران: نشر دوران.
۶. شریفی، حسن پاشا، ۱۳۸۳، اصول روان‌سنجی و روان‌آزمایی، چاپ نهم، تهران: انتشارات رشد.
۷. نادری، عزت‌الله؛ و سیف‌نراقی، مریم، ۱۳۸۱، روش‌های تحقیق و چگونگی ارزشیابی آن در علوم انسانی (با تأکید بر علوم رفتاری)، چاپ بیست و یکم، تهران: بدر.
۸. نیکوس، تومانیس، ۱۳۸۵، راهنمای گام به گام SPSS برای تربیت بدنی و علوم ورزشی، ترجمه‌ی اسفندیار خسروی زاده، چاپ اول، تهران: بامداد کتاب.
9. Armitage, P.; Theodore, C. (1998). *Encyclopedia of Biostatistics*, John Wiley & Sons Ltd.
10. Burton, A.W. & Miller, D.E. (1998). *Movement skill assessment*, Human Kinetics.
11. Burton, A.W. & Rodgerson, R.W. (2001). "New perspectives on the assessment of movement and motor abilities", *Adapted Physical Activity Quarterly*, 18, 374-365.
12. Chow, S.M.K.; Hsu, Y.; Henderson, S.E.; Barnett, A.L. & Lo, S.K. (2006). "The movement ABC: A cross-cultural comparison of preschool children from Hong Kong, Taiwan, and the USA", *Adapted Physical Activity Quarterly*, 23, 31-48.
13. Deflandre, A.; Lorant, J.; Gavarry, O. & Falgairrette, G. (2001). "Determinants of physical activity and physical and sports activities in French school children", *Perceptual and Motor Skills*, 92, 399-411.
14. Evaggelinou, C.; Tsigilis, N. & Papa, A. (2002). "Construct validity of the test of gross motor development; A cross-validation approach". *Adapted Physical Activity Quarterly*, 19, 483-495.
15. Gallahue, D.L. & Ozmun, J.C. (2006). *Understanding motor development: Infant, children, adolescent, adult*, 6th ed. McGraw-Hill International Edition.
16. Houwen, S.; Visscher, C.; Hartman, E. & Lemink, K.A.P.M. (2007). "Gross motor skills and sport participation of children with visual impairments", *Research Quarterly for Exercise & Sport*, 78(1), 16-23.
17. Kubiszyn, T. & Borich, G. (1990). *Educational testing and measurement*, 3th ed. Harper Collins Publisher.
18. Lee, M.; Zhu, W. & Ulrich, D.A. (2005). "Many-Faceted Rasch calibration of TGMD-2", *Internet Article*.
19. Linn, L.L. (editor) (1988). *Educational measurement*. 3th ed. American Council on Education & National Council on Measurement in Education. Part I: Reliability.
20. Merriman, W.; Barnet, B.E. & Isenberg, D. (1995). "A preliminary investigation of the relationship between language and gross motor skills in preschool children", *Perceptual and Motor Skills*, 81, 1211-1216.
21. Morrow, Jr. J.R.; Jackson, A.W.; Disch, J.G.; Mood, D.P. (2005). *Measurement and evaluation in human performance*, 3th ed. Human Kinetics.
22. Niemeijer, A.S. (2007). *Neuromotor task training for children with developmental coordination disorder*,

Doctoral Thesis, University of Groningen, the Netherlands, p.16.

23. Piek, J.P.; Baynam, G.B. & Barrett, N.C. (2006). "The relationship between fine and gross motor ability, self-perceptions and self-worth in children and adolescents", *Human Movement Science*, 25, 65-75.
24. Reeves, L.; Broeder, C.E.; Kennedy-Honeycutt, L. & East, C. (1999). "Relationship of fitness and gross motor skills for five to six yr. old children", *Perceptual and Motor Skills*, 89, 739-477.
25. Simons, J.; Daly, D.; Theodorou, F.; Caron, C.; Simons, J. & Andoniadou, E. (2008). "Validity and reliability of TGMD-2 in 7-10 yr. Flemish children with intellectual disability", *Adapted Physical Activity Quarterly*, 25(1).
26. Skordilis, E.K.; Douka, A.; Spartali, I. & Koutsouki, D. (2004). "Depth perception of elementary school students with qualitatively evidenced locomotor impairments", *Perceptual and Motor Skills*, 99, 501-518.
27. Thomas, J.R. & Nelson, J.K. (1996). *Research methods in physical activity*. 3rd ed. Human Kinetics.
28. Ulrich, D.A. (2000). *Test of Gross Motor Development*. 2nd ed. Austin, TX: Pro-Ed.
29. Van Waelvelde, H.; De Weerd, W.; De Cock, P. & Smits-Engelsman, B.C.M. (2004). "Aspects of the validity of movement assessment battery for children", *Human Movement Science*, 23, 49-60.
30. Vira, R. & Ruadsepp, L. (2003). "Psychological correlates of physical activity among seven through eight grades", *Human Movement Studies*, 44, 501-517.
31. Wong, K.Y.A. & Cheung, S.Y. (2007). "Structural Validity of the Test of Gross Motor Development-2", AAHPERD National Convention and Exposition (March13-17) Baltimore, MD. (Poster presentation).
32. Yun, J. & Ulrich, D.A. (2002). "Estimating measurement validity: A tutorial", *Adapted Physical Activity Quarterly*, 19, 32-47.