

عینیت، اعتبار، و روایی آزمون عملی سال ۱۳۸۰، برای گزینش دانشجوی تربیت بدنی

❖ دکتر وحید ذوالاکتاف، دانشگاه اصفهان
 ❖ مرتضی ربیعی، کارشناس ارشد تربیت بدنی

فهرست :

۱۵	چکیده
۱۶	مقدمه
۱۷	روش شناسی تحقیق
۱۸	یافته های تحقیق
۲۱	بحث و نتیجه گیری
۲۳	منابع و مأخذ

چکیده: تحقیق حاضر در پی مشخص کردن میزان عینیت، اعتبار، و روایی آزمون عملی ورودی، برای گزینش دانشجویان رشته تربیت بدنی است که در سال ۱۳۸۰، برای اولین بار ابداع و به کار گرفته شد. از بین ۳۰ کلاس واحد (۲) تربیت بدنی پسران دانشگاه اصفهان، ۲ کلاس برای شرکت در تحقیق به صورت «نمونه گیری خوشه ای هدفدار» انتخاب شدند. ۳۰ داور از دانشجویان رشته تربیت بدنی انتخاب شدند که در آزمون عملی ۸۰ تجربه دآوری داشتند آنها پس از توجیه، بازآموزی و ایجاد هماهنگی، به طور مستقل در سه گروه ۱۰ نفره به آزمودنی ها امتیاز دادند. هرآزمودنی در فاصله یکساعته در روزآزمون، دوبار آزمون شد. هر آزمون به طور متوسط 94 ± 9 نایه به طول انجامید.

عینیت برای کیفیت اجرا، زمان آزمون و نمره کلی، به ترتیب برابر $0/84$ ، $0/99$ و $0/95$ به دست آمد. در محاسبه عینیت، از نمره های ۴۳ آزمودنی در ۳ گروه دآوری استفاده شد. اعتبار برای کیفیت اجرا، زمان آزمون و نمره کلی، به ترتیب برابر $0/67$ ، $0/71$ و $0/63$ به دست آمد. برای محاسبه اعتبار، از نمره های ۶۰ آزمودنی در آزمون ابتدایی و آزمون مجدد استفاده شد و برای روایی همزمانی آزمون، از معیار نمره تربیت بدنی استفاده شد. به طور کلی، همبستگی با آزمون ملاک به طور معناداری در آزمون مجدد، بالاتر از آزمون ابتدایی بود و حداکثر به $0/57$ ($P < 0/05$) رسید. بالا بودن سطح ضریبهای عینیت، نشان می دهند که آموزش و

هماهنگی ایجاد شده در سه گروه داوری، به خوبی صورت پذیرفته است و از این بابت، اشکالی متوجه آزمونها یا اجرای آن‌ها نیست. ضریب اعتبار نسبتاً پایین به دست آمده، سؤالات جدی در برابر به کارگیری این آزمون به عنوان ابزار گزینش دانشجو ایجاد می‌نماید. ارتقای سطح روایی آزمونها در آزمون مجدد نسبت به آزمون ابتدایی، اثر مثبت تمرین و آشنایی با آزمون را در منطقی‌تر نمودن تفاوت نمره‌ها نشان می‌دهند. به طور کلی، با توجه به بداعت تجربه‌ای از این نوع، نتایج به دست آمده دلگرم‌کننده است. تحقیق حاضر، ضمن آشکارسازی نقاط ضعف موجود، دلالت بر آن دارد که زمینه‌های مناسبی برای ارتقا و توسعه آزمون در آینده فراهم خواهد آمد.

واژه‌های کلیدی: عینیت، اعتبار، روایی، آزمون عملی و ورودی.

مقدمه

استانی برای رده‌های سنی ۱۰ تا ۱۷، به طور تفکیک شده صادر شد و در بسیاری از استان‌ها به مرحله اجرا درآمد. مطالعه نرم‌های تهیه شده برای اصفهان (۱۲) نشان می‌دهد که علی‌رغم تلاش صورت گرفته، نرم‌ها دارای اشکالات اساسی بودند و قابل اجرا نبودند. مهمترین انتقادات وارد بر این نرم‌ها عبارت بودند از: اجرای آزمون‌ها در شرایط غیراستاندارد، حجم نمونه کم (بین ۵۶ الی ۵۴ نفر برای هر گروه)، حذف جزء «ترکیب بدنی» از آزمونها، تغییر شرایط اجرایی آزمون دراز و نشست، کنترل نداشتن بر درستی توزیع داده‌ها، تبدیل دو ۱/۵ مایل به دو ۵۴۰ متر، نبودن برآورد روایی، اعتبار، عینیت آزمون‌ها و غیره.

در «آموزش عالی ایران»، پذیرش دانشجوی تربیت بدنی از سال ۱۳۳۵ شروع شد (۱۱). با توجه به تعداد زیاد داوطلبان و ظرفیت پایین پذیرش، متصدیان از ابتدا به دنبال بهره‌گیری از آزمونهایی بوده‌اند که آماده‌ترین افراد را از نظر جسمانی گزینش کنند. با این وجود در طول این دوران آزمون‌های آمادگی جسمانی، یا وجود نداشت یا با نظر متخصصان و مسئولان محلی ساخته

در آمریکا از دهه ۱۹۵۰، روشهای مختلف آزمون‌گیری برای اندازه‌گیری آمادگی جسمانی به وجود آمده‌اند. این روش‌های آزمون‌گیری استاندارد در طول ۶ دهه سیر تکاملی داشته‌اند. در آمریکا روشهای ملی آزمون‌گیری متنوع بوده‌اند و ده‌ها نوع از آن را می‌توان برشمرد. روش آزمون‌گیری مؤسسه «ایفرد»^۱ تا اواخر دهه ۱۹۸۰، بیشترین مورد استفاده در آمریکا و بسیاری از کشورهای دیگر، از جمله ایران را داشت. به هر حال، در سال ۱۹۹۲ برتری روش «فیتنس گرام»^۲ نشان داده شد و در سال ۱۹۹۴، کلاً روش ایفرد در فیتنس گرام همچنان ادامه دارد. در ایران، تلاشهای متفرقه‌ای برای توسعه روش‌های استاندارد اندازه‌گیری آمادگی جسمانی صورت گرفته است. در «آموزش و پرورش»، روش آزمون‌گیری مؤسسه مذکور الگو بود و با نام «ایفر ۳» یا «ایفرد» معرفی شد (وجه تسمیه نامگذاری مشخص نیست!). گامهای اول از سال ۱۳۶۳ برداشته شد، ولی ثمر چندانی در بر نداشت. در سال ۱۳۷۰ براساس اجرای آزمون‌ها در تهران، نرم‌هایی تهیه و برای اجرا به سراسر کشور ابلاغ شد. این موضوع مورد اعتراض معلمان شهرستانی قرار گرفت. در سال ۱۳۷۶، دستورالعمل تهیه نرم‌های

1. Aahperd

2. Fitnessgram

3. Aahper

دانشجوی گزینش شده تربیت بدنی را رقم می‌زند، ضرورت تحقیق حاضر کاملاً روشن می‌شود.

مطالعه درباره ضریبهای اعتبار و روایی در ایران و خارج از ایران، با دو برداشت نسبتاً متفاوت در حال اجراست. نمونه تحقیقات داخلی را می‌توان در کارهای سیاه کوهیان و گائینی (۱۰)، سازوار و دیگران (۹)؛ میناسیان و علیزاده (۱۵) ملاحظه کرد. در این تحقیقات به جای بررسی دقت وسایل ساخته شده، مطالعات از نوع بررسی پایایی، اعتبار و روایی است. مطالعه لوبین و همکاران می‌تواند الگوی خوبی برای این گونه مطالعات باشد. مطالعات اصیل اعتبار و روایی سنج، بیشتر برای آزمون‌های تازه ابداع یا اصلاح شده به کار می‌روند. چند مورد از آخرین نمونه‌های این نوع مطالعات معرفی می‌شوند: برادلی (۱)، چرچ و دیگران (۲)، هویی و یوین (۳) و استاگ بروگر و هائل (۷). مطالعه حاضر، با توجه به تکنیک‌های به کار رفته در این مطالعات، طراحی و اجرا گردیده است. آزمودنی‌ها:

روش شناسی تحقیق

آزمودنیها: از ۳۰ کلاس که در حال گذراندن واحد (۲) تربیت بدنی پسران دانشگاه اصفهان بودند، ۲ کلاس برای شرکت در تحقیق به صورت نمونه‌گیری خوشه‌ای هدفدار انتخاب شدند. این ۲ کلاس مجموعاً ۳۰ دانشجو داشتند که ۲۳ نفر آنها در آزمون شرکت کردند. از این ۲۳ نفر، ۲۰ نفر هم در آزمون ابتدایی و هم در آزمون مجدد شرکت کردند، اما ۳ نفر به دلیل اظهار خستگی، فقط در آزمون ابتدایی شرکت کردند. زمان آزمون به طور متوسط برابر $8/65 \pm 60/6$ / ۹۴ ثانیه و فاصله آزمون تا آزمون مجدد حدوداً یک ساعت بود. آزمودنیها همگی پسر، و دارای دامنه سنی ۱۸ تا ۲۴ سال بودند.

شده و به کار می‌رفتند. از سال ۱۳۶۹ به بعد، آزمون عملی به طور یکنواخت در سراسر کشور به اجرا درآمد. این مجموعه آزمایش شامل: دراز و نشست، پرش طول جفت، کشش بارفیکس، دو ۴×۹ و دو ۵۴۰ متر بود. در طول سالهای ۶۹ الی ۸۰، تغییرات و اصلاحات آزمون بسیار کم بوده است. بررسی مواد آزمونها نشان می‌دهد که آن‌ها نیز کم و بیش برگرفته از آزمونهای معرفی شده دهه ۱۹۶۰ مؤسسه ایفرد هستند. با نگاهی گذرا به متون‌های معاصر مترجم می‌شویم که تمایل بهره‌گیری از این آزمونها هنوز به طور گسترده‌ای هم در آموزش و پرورش و هم در آموزش عالی وجود دارد (۱۳، ۱۴ و ۱۷).

سال ۱۳۸۰، در آزمون عملی ورودی دانشجویان تربیت بدنی، چه از نظر هدفها و چه از نظر مواد امتحانی، دگرگونی اساسی ایجاد شد. مؤلفان آزمون جدید معتقد بودند، آزمونهای آمادگی سلامت سنج یا حرکت سنج، قابلیت پیش بینی استعداد و موفقیت در مهارت‌های ورزشی را ندارند. بنابراین، آزمون مرکبی را ساختند که شامل ۹ ایستگاه به هم پیوسته بود. در هر ایستگاه، یک یا چند مورد از اجزای آمادگی جسمانی سنجیده می‌شد. داوران به زمان کلی آزمون و کیفیت اجرا در هر ایستگاه امتیاز می‌داده‌اند و نمره نهایی از ترکیب آن‌ها به دست می‌آمد. سال ۱۳۸۱، اصلاحات اندکی در این آزمون ایجاد شد.

گزارش معرفی آزمون ۸۰ (۱۶)، فاقد مطالعات مربوط به عینیت، اعتبار، و روایی آزمون بود. ذوالاکتاف (۸) با توجه به پیچیدگی‌های اجرایی، به نا‌آشنایی و تمرین آزمودنی‌ها، همچنین پیچیدگیهای امتیاز دهی آزمون معتقد بود که ضریبهای عینیت، اعتبار و روایی آزمون نمی‌تواند در حد مطلوبی باشد. با توجه به شک منطقی پیش آمده و با توجه به آنکه آزمون فوق در هر سال، سرنوشت هزاران داوطلب و صدها

روش آماری: تمام داده‌ها به وسیله نرم افزار اس. پی. اس. اس (۹/۱) مورد تجزیه و تحلیل آماری قرار گرفتند. برای بررسی روایی، از ضریب همبستگی بین گروهی «پرسون» و برای بررسی عینیت و اعتبار داده‌ها از ضریب همبستگی «درون گروهی»^۱ استفاده شده است (۵).

یافته‌های تحقیق

اطلاعات توصیفی: جدول ۱، در بردارنده اطلاعات لازم آماری در مورد نتایج آزمون است. اطلاعات فوق، نشان از نزدیکی میانگین و میانه نمره‌ها، همگنی واریانسها و قابل قبول بودن چولگی‌ها و کورتزها، با توجه به خطای استاندارد آنهاست. این موارد، به معنای قبول توزیع طبیعی نمره‌هاست. بنابراین، در تحلیل‌های استنباطی، از آزمونهای پارامتریک استفاده خواهد شد. جدول ۱، نقاط درصدی را برای هر مورد گزارش کرده است.

عینیت آزمون: جدول ۲، ضریبهای عینیت آزمون را به تفکیک برای زمان اجرا، کیفیت اجرا و نمره کلی نشان می‌دهد. همانطور که ملاحظه می‌شود، عینیت کلی آزمون به اندازه کافی بالاست و این بیشتر به دلیل نمره زمان اجرای آزمون است تا نمره کیفیت اجرا. برای محاسبه ضریبهای عینیت، از نمره‌هایی استفاده شده است که هر گروه داوری در آزمون ابتدایی و آزمون مجدد به آزمودنیها داده‌اند.

اعتبار آزمون: جدول ۳ ضریبهای اعتبار آزمون را

*. برای دریافت راهنمای کامل آزمون ۸۰، می‌توان با آدرس پژوهشکده تربیت بدنی و علوم ورزشی تماس گرفت: تهران، میدان ولی عصر، خیابان شقایق، پلاک ۴۱، تلفن‌های ۶۴۰۹۰۳۷ و ۶۴۰۸۹۱۳-۰۲۱ و نمابر ۶۴۰۹۰۳۷-۰۲۱.

1. Intraclass

داوری: ۳ گروه داوری، به طور مستقل به کیفیت و زمان اجرای آزمودنیها به طور همزمان نمره دادند. با توجه به پیچیدگیهای آزمون در هر گروه مستقل، داوران عبارت بودند از: ۹ داور ایستگاه و یک سرداور. این سرداور هم وقت رانگه می‌داشت و هم نمره‌های اعلام شده داوران ایستگاه را ثبت می‌کرد. داوران همگی از دانشجویان رشته تربیت بدنی بودند که در آزمون عملی ۸۰ تجربه داوری داشتند. علاوه بر این، آنها دوباره توجیه و بازآموزی شدند و هماهنگی‌های لازم در هر گروه حدود یک ساعت قبل از آزمون انجام شد.

آزمون عملی* ۸۰: آزمون شامل ۹ ایستگاه به این قرار بود: بارفیکس: دریل پا، دریل دست، تعادل (چوب، موازنه، غلت و پرش ماهی) مانع (پرش و گذر از زیر)، دراز و نشست، ۴×۹ متر رفت و برگشت، پرتاب توپ پزشکی و سرعت (۲۰ متر). نمره فرد دارای دو بخش بود: یکی زمان کلی اجرای آزمون که ۵۰٪ از نمره نهایی بود. دیگری، نمره کیفیت اجرا که ۵۰٪ باقیمانده نمره نهایی فرد بود. در هر ایستگاه، آزمودنی می‌توانست یکی از علائم پنج گانه کارت‌های سفید، زرد، سبز، قرمز و سیاه را با امتیازی خاص بگیرد. جمع امتیاز کارت‌ها، نمره پنجگانه (کارت‌های سفید، زرد، سبز، قرمز و سیاه) را با امتیازی خاص بگیرد. جمع امتیاز کارت‌ها نمره کیفیت فرد را مشخص می‌کرد. شیوه محاسبه نمره نهایی فرد در گزارش تحقیق هادوی (۱۳۸۰) مشخص نشده است. در تحقیق حاضر، برای محاسبه نمره کلی، نمره‌های زمان و کیفیت اجرا به نمره $Z = (x-m)/sd$ و سپس به نمره $T = 10Z + 50$ تبدیل و از آنها معدل گیری می‌شود. در معادلات فوق، Z به معنای نمره استاندارد (با میانگین صفر و انحراف معیار ۱)، X نمره فرد، m معدل نمره‌ها sd انحراف استاندارد کل نمره‌ها و T نمره استاندارد (با میانگین ۵۰ و انحراف معیار ۱۰) است.

جدول ۱. اطلاعات توصیفی مربوط به آزمون

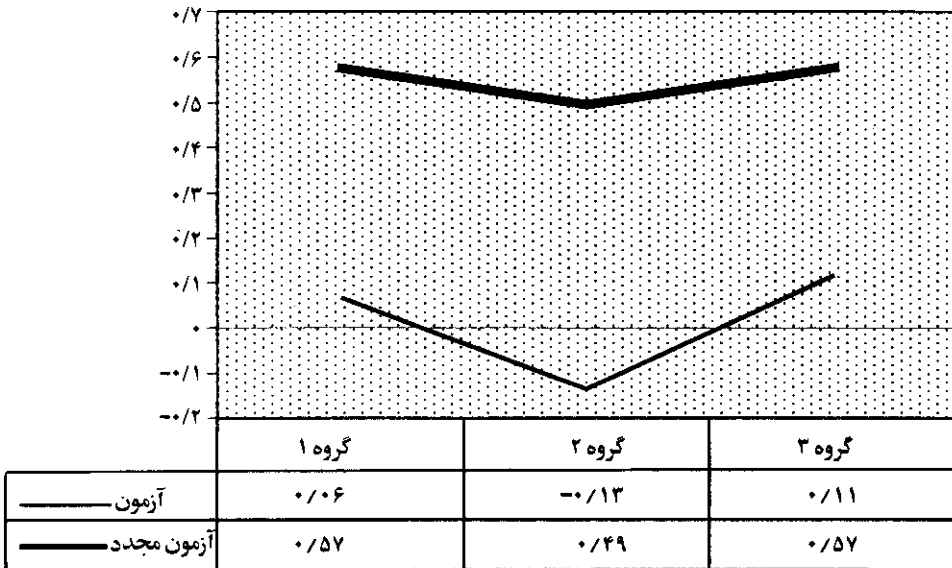
نمره کلی (آزمون مجدد)	نمره کلی	زمان (آزمون مجدد)	زمان	نمره کیفیت (آزمون مجدد)	نمره کیفیت	
۶۰	۶۹	۶۰	۶۹	۶۰	۶۹	تعداد
۵۳,۵۷۴۹	۴۶,۸۹۱۴	۸۴,۹۷۵۲	۹۴,۰۵۸۰	۴۲,۵۴۱۷	۴۱,۵۶۸۸	میانگین
۵۳,۷۱۴۳	۴۷,۸۳۵۸	۸۴,۲۳۰۰	۹۴,۶۳۰۰	۴۲,۱۲۵۰	۴۱,۷۵۰۰	میانه
۶,۷۶۶۳	۷,۲۶۴۸	۵,۸۵۵۴	۸,۶۴۶۶	۳,۲۲۶۳	۳,۲۸۹۹	انحراف معیار
۰,۰۸۳	-۰,۴۵۰	۰,۸۶۵	-۰,۰۶۱	-۰,۰۷۳	-۰,۰۴۳	چولگی
۰,۳۰۹	۰,۲۸۹	۰,۳۰۹	۰,۲۸۹	۰,۳۰۹	۰,۲۸۹	م.ا چولگی
۰,۵۲۰	-۰,۲۷۳	۰,۸۹۴	-۰,۵۸۱	-۰,۸۴۷	-۰,۷۶۴	کورتوز
۰,۶۰۸	۰,۵۷۰	۰,۶۰۸	۰,۵۷۰	۰,۶۰۸	۰,۵۷۰	م.ا کورتوز
۳۸,۶۷	۲۶,۸۲	۷۴,۵۱	۷۷,۶۵	۳۵,۵۰	۳۴,۵۰	حداقل
۶۶,۳۶	۵۹,۲۲	۱۰۰,۲۶	۱۱۱,۷۴	۴۸,۲۵	۴۷,۰۰	حداکثر
۳۸,۶۶۶۸	۲۶,۸۲۰۰	۱۰۰,۲۶۰۰	۱۱۱,۷۴۰۰	۳۵,۵۰۰۰	۳۴,۵۰۰۰	رتبه درصدی
۴۳,۲۵۲۸	۳۳,۴۹۷۶	۹۹,۹۷۳۵	۱۰۹,۱۷۵۰	۳۷,۲۸۷۵	۳۵,۸۷۵۰	۱
۴۵,۰۳۵۴	۳۶,۷۹۹۶	۹۵,۰۵۶۰	۱۰۴,۲۱۰۰	۳۸,۲۷۵۰	۳۷,۲۵۰۰	۵
۴۵,۹۵۶۶	۳۸,۹۵۳۲	۹۰,۳۶۱۰	۱۰۳,۴۷۵۰	۳۸,۷۸۷۵	۳۸,۳۷۵۰	۱۰
۴۷,۳۲۷۶	۳۹,۵۴۴۲	۸۹,۸۸۴۰	۱۰۰,۶۳۰۰	۳۹,۸۰۰۰	۳۸,۵۰۰۰	۱۵
۴۸,۶۶۲۷	۴۳,۲۰۴۴	۸۸,۰۹۰۰	۱۰۰,۱۹۰۰	۴۰,۰۶۲۵	۳۸,۷۵۰۰	۲۰
۴۹,۸۶۱۹	۴۳,۴۷۷۵	۸۶,۰۴۹۰	۹۹,۹۱۰۰	۴۰,۵۰۰۰	۳۹,۰۰۰۰	۲۵
۵۰,۵۸۹۳	۴۴,۶۰۲۷	۸۵,۹۴۹۰	۹۸,۳۶۰۰	۴۱,۰۸۷۵	۳۹,۸۷۵۰	۳۰
۵۱,۷۷۹۳	۴۵,۸۴۸۲	۸۵,۵۵۲۰	۹۶,۵۵۰۰	۴۱,۵۰۰۰	۴۰,۵۰۰۰	۳۵
۵۲,۴۸۷۶	۴۶,۵۱۷۳	۸۴,۸۲۶۵	۹۵,۸۱۵۰	۴۱,۷۵۰۰	۴۱,۳۷۵۰	۴۰
۵۳,۷۱۴۳	۴۷,۸۳۵۸	۸۴,۲۳۰۰	۹۴,۶۳۰۰	۴۲,۱۲۵۰	۴۱,۷۵۰۰	۴۵
۵۴,۵۸۸۶	۴۸,۸۷۷۴	۸۳,۷۳۹۰	۹۴,۴۸۰۰	۴۳,۳۸۷۵	۴۱,۸۷۵۰	۵۰
۵۴,۹۷۶۶	۴۹,۳۴۶۲	۸۲,۳۷۴۰	۹۱,۰۰۰۰	۴۳,۵۰۰۰	۵۲,۲۵۰۰	۵۵
۵۵,۵۲۷۱	۵۰,۱۸۶۹	۸۱,۷۳۰۰	۸۹,۴۶۵۰	۴۳,۷۵۰۰	۴۳,۲۵۰۰	۶۰
۵۷,۰۹۱۶	۵۱,۴۴۰۰	۸۱,۳۶۵۰	۸۹,۳۲۰۰	۴۵,۱۷۵۰	۴۳,۵۰۰۰	۶۵
۵۸,۰۰۴۶	۵۲,۷۳۰۲	۸۰,۹۹۲۵	۸۹,۲۲۰۰	۴۵,۲۵۰۰	۴۳,۷۵۰۰	۷۰
۵۹,۸۳۸۵	۵۳,۷۰۱۸	۸۰,۷۴۰۰	۸۷,۰۸۰۰	۴۵,۴۵۰۰	۴۵,۰۰۰۰	۷۵
۶۱,۶۷۱۰	۵۴,۴۲۱۷	۸۰,۳۹۵۵	۸۳,۵۴۵۰	۴۶,۷۵۰۰	۴۵,۲۵۰۰	۸۰
۶۳,۷۸۷۳	۵۶,۰۱۴۷	۷۸,۳۹۸۰	۸۱,۸۵۰۰	۴۷,۰۰۰۰	۴۶,۷۵۰۰	۸۵
۶۶,۲۷۲۱	۵۷,۶۹۲۴	۷۴,۸۰۷۰	۷۸,۶۳۵۰	۴۷,۰۰۰۰	۴۶,۸۷۵۰	۹۰
۶۶,۳۵۶۷	۵۹,۲۱۸۱	۷۴,۵۱۰۰	۷۷,۶۵۰۰	۴۸,۲۵۰۰	۴۷,۰۰۰۰	۹۵
						۹۹

جدول ۲. ضریبهای عینیت آزمون

ضریبهای آلفا			تعداد نمره‌های آزمودنیها در هر گروه	تعداد گروه‌های داوری
نمره کل	زمان اجرا	کیفیت اجرا		
۰/۹۴	۰/۹۹	۰/۸۴	۴۳	۳

جدول ۳. ضریبهای اعتبار آزمون

ضریب‌های آلفا			تعداد نمره‌های آزمودنیها در هر گروه	تعداد آزمون
نمره کل	زمان اجرا	کیفیت اجرا		
۰/۶۳	۰/۷۱	۰/۶۷	۶۰	۲



نمودار ۱. نمودار خطی ضریبهای همبستگی پیرسون، بین نمره‌های تربیت بدنی با آزمون و آزمون مجدد در سه گروه داوری.

و ضریب به دست آمده حداکثر ضریب ممکن بود. اگر برای مثال، نحوه آرایش اجزای آزمون در ۳ شهر مختلف را ۳ گروه متفاوت برگزار می کردند. قاعدتاً انتظار آن بود که ضریبهای عینیت پایین بیاید. این ضریبها، با متفاوت شدن کادر آموزش دهنده به ۳ گروه داوری و متفاوت شدن زمان اجرای آزمون ها، باز هم می توانست پایین تر بیاید. دلیل این موضوع آن است که آرایش آزمون نسبتاً پیچیده، فهم از ملاک های داوری نیز قابل تفسیر و تغییر پذیر است.

به هر حال، بالا بودن ضریبهای عینیت نشان می دهد که آموزش ها مؤثر بود کمترین خدشه ای متوجه کار داوران این مطالعه است.

اعتبار با وجود بالا بودن ضریبهای عینیت، ضریبهای اعتبار به طور محسوسی پایین است. اعتبار کمتر از ۰/۸۵، معمولاً برای آزمونهای جسمانی، نامناسب ارزشیابی می شود. قبلاً حدس زده شد که پیچیدگی آزمون از یک طرف و نبودن فرصت تمرین از طرف دیگر، اعتبار آزمون را مخدوش می کند. این هر دو عامل بر اجرای آزمون اثر گذارند. برای خنثی سازی این عامل، معمولاً با آزمون، فرصت تمرینی بیشتری به آزمودنی داده می شود. یکی از هدفهای طراحان این آزمون، گزینش افرادی است که برای پیشرفت، احتیاج به تدریسات کمتری دارند. بنابراین، فرصتهای تمرینی محدودیتهایی دارند. به نظر محقق، برای آزمونی به این پیچیدگی، دست کم ۱ یا ۲ فرصت تمرینی، برای بالا بردن ضریب اعتبار در حد مطلوب نیاز است. البته این در صورتی مؤثر خواهد بود که ملاک های داوری از پیچیدگی خارج شود تا تفسیرهای متفاوتی وجود نداشته باشد.

روایی: بنابر اظهار مؤلف، این آزمون دارای روایی «منطقی» و روایی «سازه» خوبی است. این

به تفکیک برای زمان اجرا، کیفیت اجرا، و نمره کلی نشان می دهد. همانطور که ملاحظه می شود، این ضریب برای نمره کلی آزمون ۰/۶۳ و غیرقابل قبول است. برای محاسبه ضریبهای اعتبار، از نمره های هر یک از ۲۰ آزمودنی در دو نوبت آزمون گیری استفاده شد که ۳ گروه داوری امتیاز داده بودند. برای هر نوبت از آزمون ۶۰ نمره وجود داشت که اعتبار بین آنها محاسبه شد.

روایی آزمون: در این تحقیق روایی همزمان مدنظر محقق بود و برای محاسبه آن، از ضریب همبستگی بین گروهی پرسون استفاده شده است. ضریبهای روایی به تفکیک برای هر یک از گروه های داوری محاسبه شد. نمره ملاک در این تحقیق، نمره درس تربیت بدنی (۲) است که معلم در آخر ترم به دانشجو داده است. نتایج به دست آمده از محاسبه ضریبهای روایی، در نمودار ۱ نشان داده شده است. همانطور که ملاحظه می شود، در هر ۳ گروه داوری ضریبهای روایی در آزمون مجدد به طور معناداری افزایش یافته است.

بحث و نتیجه گیری

عینیت: ضریب عینیت آزمون برای نمره کل، ۰/۹۴ و خوب است. این ضریب برای زمان اجرا ۰/۹۹ و برای کیفیت اجرا ۰/۸۴ است. همانطور که ملاحظه می شود، نمره عینیت برای کیفیت اجرا، ۰/۱۵ کمتر از زمان اجراست. با توجه به پیچیدگی نمره دهی به کیفیت اجرا، این موضوع قابل پیش بینی بود. به هر حال، توجه داشته باشید که آرایش آزمون برای همه آزمودنیها یکسان بوده و به طور همزمان ۳ گروه داوری به آنها امتیاز دادند. این موضوع به طور قابل ملاحظه ای در ارتقای ضریب عینیت مؤثر بوده

بشود. همان‌طور که قبلاً بیان شد در این آزمون، عاملهایی نظیر دقت، هماهنگی عصبی عضلانی و چابکی را در ۳ یا ۴ ایستگاه می‌سنجد، در صورتی که استقامت عضلانی، تعادل و اجرای غیرایستاده در ۲ ایستگاه و قدرت و سرعت در ۱ ایستگاه سنجیده می‌شوند. عاملهایی مانند: ترکیب بدنی، انعطاف‌پذیری، توان هوازی و زمان عکس‌العمل، در هیچ ایستگاهی سنجیده نمی‌شوند. همچنین، قابلیت‌های حرکتی اندازه‌گیری شده، فقط به رشته‌های ژیمناستیک، دو و میدانی، فوتبال و بسکتبال مربوط است. برای افزایش جامعیت آزمون دو راه وجود دارد: اول، باید همه عاملهای آمادگی جسمانی، به‌طور یکسان در یکی دو ایستگاه اندازه‌گیری شود و از همه رشته‌های متداول هم، حرکات اصلی در آزمون گنجانده شود. دوم، مواد آزمون به گونه‌ای طراحی شوند که نه به رشته ورزشی خاصی مربوط شود و نه به هیچ یک از عاملهای آمادگی جسمانی تأکید خاصی کند.

نتیجه‌گیری

آزمونهای موجود آمادگی سنج، چه از نوع سلامت محور و چه از نوع حرکت محور، هیچ یک به تنهایی قادر به استعدادیابی ورزشی نبوده و هنوز راه مؤثری برای نحوه ترکیب مفید این آزمونها برای استعدادیابی بیان نکرده است. آزمون طراحی شده برای گزینش دانشجوی تربیت بدنی در سال ۸۰، گامی ضروری و مؤثر در راه پر کردن این خلأ است. تحقیق حاضر ضمن بررسی عینیت، اعتبار و روایی آزمون ۸۰ راه‌های مؤثری را برای ارتقای آزمون و ضریبهای مورد بحث پیشنهاد می‌کند.

بدان معنی است که اجزای آزمون منطقیاً به خوبی انتخاب شده‌اند و آزمون برای انتخاب نهایی وسیله خوبی برای رتبه‌بندی داوطلبان از نظر استعداد و آمادگی برای یادگیری ورزشی است. کم و کیف این ادعا، به وضوح در مقاله‌ای دیگر نقد شده است. به هر حال، چنانچه آزمونی دارای روایی منطقی و روایی سازه خوبی باشد، انتظار است که روایی همزمان آن هم از نظر آماری در حد قابل قبولی باشد. در این مقاله از نمره واحد تربیت بدنی دانشجویان، به عنوان نمره ملاک استفاده شده است. نکته جالب آن است که آزمون ابتدایی هیچ همبستگی معناداری با نمره تربیت بدنی ندارد. ولی آزمون مجدد، حدود ۵۰ الی ۶۰ صدم در هر ۳ گروه داوری همبستگی نشان می‌دهد. به عبارتی، یک بار فرصت تمرین، باعث شده است تا رتبه‌های افراد با رتبه‌های نمره تربیت بدنی آنها هماهنگی بیشتری پیدا کند، اما بدون فرصت تمرین، هیچ نظم منطقی با نمره ترتیب بدنی ایجاد نشد. این احتمال وجود دارد که ۳ فرصت تمرینی ضریب همبستگی با نمره تربیت بدنی بالاتر برود که این، به دلیل پیچیدگی آزمون است. احتمال دارد بیش از ۲ یا ۳ فرصت تمرینی در یک جلسه، باعث افزایش خستگی شود و اثر معکوس بر ترتیب منطقی نمره‌ها بگذارد. یافتن دفعات تمرین مناسب برای این آزمون، نیاز به تحقیقی مستقل دارد.

در حال حاضر، نمره تربیت بدنی (۲)، صرفاً براساس توانایی‌های ورزشی داده نمی‌شود. مسایل انضباطی، حضور بر سر کلاس، درصد پیشرفت، اخلاق ورزشی، میزان تلاش، پوشاک ورزشی و غیره، ممکن است بر نمره دانشجویان تأثیر گذاشته باشد. چنانچه نمره ورزش، صرفاً از آزمون‌های ورزشی به دست آید، ممکن است ضریب روایی آن با آزمون‌هایی از نوع آزمون مورد مطالعه خیلی بیشتر

منابع و مأخذ

1. Bradley J. 2000 Reliability, repeatability, and sensitivity of the modified shuttle test in adult. Cystic Fibrosis CHEST (117). 1666-71.
2. Church TS, Kampert JB, Wilkinson WJ, Dunn AL, and Blair SN. (2001). Evaluating the reproducibility and validity of the Aerobic Adaptation Test, Medicine & Science in Sports and Exercise - (33) 10: 1770-13.
3. HUISS, and Yuen PY. (2000). Validity of the modified back-saver sit-and- reach test, Medicine & science in sports and Exercise. (32) 9: 1655-9.
4. Levine JA, Baukol, PA, and Westertrep KR. (2001). Validation of the tracmor triaxial accelerometer system for walking, Medicine & Science in Sports and Exercise. (33)9: 1593-4.
5. Morrow JR, Jackson AW, Disch JG, Mood DP. (1995). Measurement and evaluation in health performance, Human Kinetics. PP. 85-8.
6. Safrit MJ. (1995). Complete guide to youth fitness testing, Human Kinetics, chapter 2.
7. Stockbrugger BA, and Haennel RG. (2001). Validity and reliability of a medicine ball explosive power test. Journal of Strength and Conditioning Research. (15)4:431-8.
۸. ذوالاکناف و، ۱۳۸۰، نقدی بر آزمون ورودی تربیت بدنی ۸۰، مجموعه مقالات هفته پژوهش ۱۳۸۰، دانشکده تربیت بدنی، دانشگاه اصفهان
۹. سازوار، ناظم ف، و فرهمپور ن، ۱۳۸۰، سنجش اعتبار آزمون میدانی هوازی، بیشینه شاتل ران برای برآورد حداکثر اکسیژن مصرفی، سومین همایش بین المللی تربیت بدنی و علوم ورزشی، خلاصه مقالات، ۴-۷۳.
۱۰. سیاه کوهیان م، گاتینی ۱۳۸۰، ارزیابی روایی و پایایی چربی سنج های ساخت داخل در اندازه گیری چربی زیر پوستی، سومین همایش بین المللی تربیت بدنی و علوم ورزشی خلاصه مقالات، ۷۷.
۱۱. کریمی ح، ۱۳۶۴، نگاهی به تاریخ تربیت بدنی جهان، مرکز نشر دانشگاهی، ۱۹۵.
۱۲. مجتهدی ح، ۱۳۷۷، تهیه و تدوین نرم استاندارد آزمون آمادگی جسمانی مرتبط با سلامتی ایفرد در دانش آموزان دختر و پسر ۱۰ تا ۱۷ سال (استان اصفهان)، جزوه پژوهشی - شورای تحقیقات اداره کل آموزش و پرورش استان اصفهان.
۱۳. محرم زاده م، محمدزاده ح، ۱۳۷۹، بررسی سطح آمادگی جسمانی دانشجویان دانشگاه ارومیه و تهیه نرم امتحانی، حرکت (۴) ۱۷-۱۰۳.
۱۴. میرکاظمی س ع، همتی نژاد م، رضایی نژاد ر، ۱۳۸۰، تهیه نرم آمادگی جسمانی دانش آموزان دختر سال اول نظام جدید شهرستان بیرجند و مقایسه آن با نرمهای موجود، حرکت (۸)، ۲۶-۱۱۹.
۱۵. میناسیان و، علیزاده ح، ۱۳۸۰، تعیین میزان روایی دستگاه اندازه گیری توان بی هوازی محقق ساخته، سومین همایش بین المللی تربیت بدنی و علوم ورزشی، خلاصه مقالات، ۹-۱۰۸.
۱۶. هادوی ف، ۱۳۸۰، طرح آزمایشی آزمون عملی داوطلبان ورود به دوره کارشناسی تربیت بدنی ۸۱-۸۰ گزارش تحقیقی. پژوهشکده ورزش.
۱۷. همتی نژاد م، رضایی نژاد ر، ۱۳۸۰، مقایسه آمادگی جسمانی دانشجویان پسر رشته تربیت بدنی ورودی سالهای ۷۵، ۷۶ دانشگاه گیلان در ابتدا و انتهای دوره تحصیل، دومین کنگره علمی ورزشی دانشگاه های آسیا، خلاصه مقالات، ۹-۸۸.