پژوهش در حیطه کودکان استثنایی 71/ سال پنجم، شماره 3، پاییز 1384
Research on Exceptional Children, Fall 2005, vol.17, No 3, 295-322

انطباق و استانداردسازی آزمون غیرکلامی هوش استایدرز- اومان برای کودکان 7/5- سال: یک مطالعه مقدماتی

اصغر مینایی

چکیده
پژوهش حاضر به هدف بررسی و بهبود آزمون غیرکلامی هوش استایدرز اومان برای کودکان 7-5 سال (SON-R 2) انجام گرفت. آزمون ساده- پایه 2 (IRT) با جمعیت 12 نفر (4 دختر، 8 پسر) که با روش نمونه برداری تصادفی انتخاب شدند اجرا گردید. برای تحلیل سوالات از مدل کلاسیک آزمون در چهار فرآیند آزمون و مرحله 3 برآمد. نتایج نشان دهنده علائم خاص از آزمون می‌باشد که به این شکل هستند: تفاوت در میانگین مقدار پرس، تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگین مقدار پرس و تفاوت در میانگ

واژه‌های کلیدی: آزمون غیرکلامی هوش، اعتبار، روابط

* Email: as_minaei@yahoo.com

در این جراحی مطالب در اینجا ذکر نشده چنین شکلی است که نمایش آزمون غیرکلامی هوش و مفاهیم تازه و ویژه مقدماتی نشان می‌دهد.
مقدمه

از آزمونهای سنی، هش و کسر و استفاده‌ای (GI)، مانند آزمون‌هایی که در حیطه کودکان استثنایی (LP) مورد استفاده قرار می‌گیرند (GI)، و توصیف طرفداران آزمون‌های طرف‌بینی یافته‌ای (LP) مورد استفاده قرار می‌گیرد. این آزمون‌ها به دلیل میزان بازه‌ای یافته‌ای پیشین را اندام‌های دارای و این‌طوری در طول زمان و از این راه، با پایداری افراد مرتبط با تعیین اجتماعی اقتصادی پایین و افراد مبتلا به مشکلات یافته‌ای و اختلالات رفتاری و همچنین افرادی که فرصت‌های اهدافی جهت کسب دانش و مهارت‌های مورد نیاز برای موقعیت آزمون‌هاشان نقش داشته‌اند، کم برآورد می‌کنند. نکته‌ای که تلویح‌ها یا این استفاده‌ای املایی است که آزمون‌هایی هیچ‌گاه امکان‌پذیری نمی‌باشد، این است که اسلام‌نژاد دو نماینده در شرایط بهره‌گیری که چه میزان می‌توان انتظار رشد در عملکرد را داشته، فراهم نمی‌گردد. در نتیجه، این آزمون‌ها نخواهند توانست به اندازه کافی بین کودکان عصب‌بندی دهنه و کودکان مبتلا به ناتوانی‌های یادگیری تمایز قابل شوند.

ویژگی مشترک آزمون‌های LP، آن‌ها مورد در زمان‌های ایجاد آزمون است. هدف از آزمون‌ها، حذف تفاوت‌های ناشی از فرصت‌های فرهنگی و آموزشی پیشین و بهینه کردن شرایط یادگیری است. اگرچه تحقیق درخصوص طرف‌بینی یافته‌ای یادگیری برای چندین دهه است که ادامه دارد، ولی ابزارهای عملی برای سنجش، فقط در جنگ سال‌ها دیده است که اثرات آن شدیدانده‌اند.

از آزمون‌های ای امکان‌پذیری از طرفداران آزمون‌های هش نابینایه به فرهنگ هنر خاطر محتوایی مورد استفاده قرار و فقدان این آزمون‌ها از لحاظ محتوای دستورالعمل‌ها اغلب به مشارکت‌های زیانی خاصی وابسته است. لذا، باعث می‌گردد تا افراد اقلیت‌های فرهنگی و قومی و همچنین افراد مبتلا به مشکلات شناوی، گفتاری و زبانی در وضعیت نامطلوب قرار گیرند. عملکرد پایین این گروه‌ها در یک آزمون ممکن است اساساً انعکاس داشته کلمات به‌عنوان یادگیری (نخستین لاروس، 1993). می‌توان گفت که این آزمون‌ها بر هوش مبتلا یک طبقه دادن با هوش سیال (کتاب، 1971). این انتقادها نسبت به آزمون‌های باعث گردیده تا آزمون‌های غیرکلامی هوش مانند مانی‌سیسای پیشروندل ریون (ریون، 1986).
انطباق و استاندارسی آزمون گرایش‌العملی

۹۳۸؛ و آزمون هوش نابسته به فرهنگ کلی ۹۵۰) با هدف به حداکثر رساندن اتکا
به دانش اکتسابی و توانایی گرامی برورش پایند.

در اوایل دهه چهل میلادی، خانم نان استادیروی - اولمک (۱۹۴۲) که در یک موسسه
مربوط به کودکان ناشنا در هلند کار می‌کرد اولین آزمون گرایش‌العملی، وسیع‌النمایه ی
را پیروی داد. او هم‌اکنون توانایی یادگیری، یعنی میزان توانایی کودکان در یادگیری از
اموزش‌ها در مدرسه تعریف کرد. این آزمون، اولین آزمون بود که حوزه وسیعی از هوش
را پیدا کرد و استفاده به این زبان می‌پوشد. آزمون‌های گرایش‌العملی
موجود در آن زمان، برای بررسی طبقه وسیعی از توانایی‌های یادگیری مناسب‌نیبودند.

زیرا آنها اساساً از آزمون‌های عملی که به توانایی‌های فضایی مربوط می‌شدند، تشکیل
می‌بودند. اولین آزمون از خرده آزمون‌های گرایش‌العملی تشکیل می‌پذیرد که به
استدلال انتزاعی و غیرمرتب می‌شدند و دارایچراغ‌های برای کودکان ناشناوار
سال بود. در حال حاضر، نسل جدید آزمون‌های SON-R در دو نسخه وجود دارد. یک
نسخه برای کودکان سنین پایین با عنوان ۴/۱-۷ و نسخه دیگر
برای کودکان سنین بالا با عنوان ۵/۱-۱۷. مقاله حاضر مطالبی را توصیف
می‌کند که با آزمون ۵/۷-۷ صورت گرفته است.

آزمون مستلزم توانایی‌های زبانی عصبی نسبتی، تاثیر، اجرای آن به شیوه انتقالی، صورت
می‌گیرد. هدف شیوه‌های انتقالی محکوم تعیین سوابق ارائه شده همراه با
کاهش نسبتاً کم اعتبار است (ویس، ۱۹۸۲) در ۴/۱-۷. نقطه شروع با سطح
ورودی هر خرده آزمون براساس شرایط کودک و سالی حضور در مهد کودک، تعیین
می‌شود و اجرای هر خرده آزمون نیز پس از ۳ پاسخ نادرست متناسب و در یک دوم
برخی از خرده آزمون‌ها نیز پس از ۲ پاسخ نادرست متوالی متوقف می‌گردد. نتایج، در
مورد درست‌بود این است که به آزمون‌های فرضی داده می‌شود. مزیت اصلی
ار از اینکه برای آزمون ۵/۷-۷ صورت گرفت که آزمون ۵/۷-۷
برخی از آزمون‌های هوش نابسته به فرهنگ و آزمون‌های نظری-یادگیری شیب‌دار
تا آزمون‌های سنی هوش کلی.
در ایران نیاز زیادی به آزمون‌های روایی روز، مناسب و استاندارد، بوضیه آزمون‌های \textit{هوش برای کودکان سنین پایین و جنوب هارب} به منظور برآورد ساختن این نیاز و با توجه به مطالعاتی که در خصوص مزایای آزمون 2 R-SON، 7/9 و انتقادات وارده بر آزمون‌های سنین هون در کلی ذکر گردیده، آزمون 2 R-SON به‌عنوان استانداردی برای کودکان ایرانی انتخاب گردید. با این حال، پیش از استانداردسازی و هنگامی که لازم است بررسی شود چه آیا موارد استفاده در این آزمون برای کودکان ایرانی اشنا است؟ آیا آزمون از اعتبار و روایی ۱ مطلوب برخوردار است؟ و آیا کاربرد این آزمون در ایران، عملی است؟ به منظور بدست‌آوردن چنین شواهدی است که مطالعه حاضر صورت گرفت. بنابراین، هدف مطالعه حاضر بررسی این موضوع است که به منظور سنجش عدالتیه و منصفانه سازه هوش در کودکان ایرانی با این آزمون، آیا و تا چه اندازه لازم است که اطمینانی در موارد آزمون‌های صورت گیرد. این هدف با رهنمودهای کمیسیون بین‌المللی آزمون ۱۱ در خصوص استفاده از آزمون، ماهنگ و مطلق است (ون دی، ویگار و همکار ۹۶). این است که بروز ده‌هندگان و ناشن آزمون باید شواهدی در خصوص اینکه سوالات و موارد آزمون برای جامعه مورد نظر، مناسب و آشنا است فراموش نگذرد.

روشن
نمونه و روش نمونه‌گیری
نمونه‌های مناسب به‌دست‌آمده در این تحقیق (۱۲۵ کودک نر و ۵۹ دختر) با دامنه سینی ۱ تا ۸ سال و ۹ ماه تا ۷ سال و ۹ ماه با میانگین ۱۴ و احراز استاندارد ۰.۱ است که در تاپستان سال ۱۳۸۷ در مهدوسی کودک تحت نظارت به‌پاییزی مشغول به آموزش بودند. جهت انتخاب نمونه از روش نمونه‌گیری تصادفی چند مرحله ای استفاده گردید. برای این منظور ز میان مهدوسی کودک تحت نظارت به‌پاییزی مناطق شمال جنوب، شرق و غرب تهران، تعد ۱۲ مهد کودک از هر منطقه (مهد کودک) بصورت کامل تصادفی انتخاب گردید. در گام بعد با استفاده از دفاتر تهیه‌نامه، از میان کودکان هر مهد تعداد ۲ کودک دختر و پسر با استفاده از روش تصادفی منظم انتخاب شدند.
این مورد استفاده در این پژوهش، آزمون غیرکلامی هوش استاندارد - اومان برای کودکان ۲۱ سال است (تختخ و همکاران ۱۹۹۸). این آزمون دارای ۶۰ فرم کوتاه و بلند است. فرم بلند آزمون از ۶ خرده آزمون تشکیل یافته است که به ترتیب اجرا
عبارتند: - مورال‌های ۱۰ سوال - طبیعت بندی ۱۵ سوال - پژوهش ۱۶ سوال - قیاس ۱۷ سوال - موضعیت ۱۶ سوال - الگوه ۱۶ سوال. استانداردسازی این آزمون در هند بر اساس یک نمونه ملی از کودکان هند
این موجود در مجموع دارای ۱۱ سوال است و در مدت ۱۰ دقیقه در یک یا دو جلسه اجرا می‌گردد. خرده آزمون‌های ۲ سون ر-۲ (PS) و سون ای‌کو (RS) است که مقیاس عملي ۱۱/۷ به دو گروه کلی طبیعی بندی می‌گردد. گروه اول شامل خرده آزمون‌های الگوه، مورال‌های کو (PS) و پژوهش (RS) است. سون ای‌کو، الگوه و سون ر-۲ (PS) این مقیاس است. در این آزمون برای هر کودک، علاوه بر نمره خرده آزمون، ۳ نمره دیگر به نام‌های SON-IQ، (به‌عنوان نیم‌یافته)، SON-RS و SON-R (به‌عنوان یافته) و COTAN (به‌عنوان کل) به دست می‌آید.
اگر دو فرم این آزمون توسط کمیته آزمون موسسه روان شناسان هند (COTAN) به دو فرم استفاده شده اند، نتایج این ارزیابی در زیر آراشده است (اقتباس از تختخ و همکاران ۱۹۹۸).

<table>
<thead>
<tr>
<th>نمره</th>
<th>سایتهای مربوط</th>
<th>نمره</th>
<th>سایتهای مربوط</th>
</tr>
</thead>
<tbody>
<tr>
<td>خوب</td>
<td>چاپی</td>
<td>علی‌اکبر تفتی</td>
<td>روانشناسی</td>
</tr>
<tr>
<td>خوب</td>
<td>چاپی</td>
<td>علی‌اکبر تفتی</td>
<td>روانشناسی</td>
</tr>
<tr>
<td>خوب</td>
<td>چاپی</td>
<td>علی‌اکبر تفتی</td>
<td>روانشناسی</td>
</tr>
<tr>
<td>نمره</td>
<td>نمره</td>
<td>نمره</td>
<td>نمره</td>
</tr>
<tr>
<td>اعتبار</td>
<td>اعتبار</td>
<td>اعتبار</td>
<td>اعتبار</td>
</tr>
<tr>
<td>روایت سازه</td>
<td>روایت سازه</td>
<td>روایت سازه</td>
<td>روایت سازه</td>
</tr>
<tr>
<td>روان‌آموز</td>
<td>روان‌آموز</td>
<td>روان‌آموز</td>
<td>روان‌آموز</td>
</tr>
<tr>
<td>کافی</td>
<td>کافی</td>
<td>کافی</td>
<td>کافی</td>
</tr>
</tbody>
</table>
جدول 1- ضرایب اعتبار خرده آزمون‌ها، مقیاس‌ها و کل آزمون به تفکیک گروه‌های سنی

| سن | مردان | زنان | کل اخلاقی | کل آزمون | کل استدلال | متوسط مقایسه‌ها | سه‌گانه‌ای | گردهمایی‌ها | علوم مختلف | کل کودکان | کلشناسی | کل نظریه‌پردازی | کل استباقی | کل فرهنگی
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

ضرایب اعتبار نمرات مقياس عملی، مقياس استدلال و کل آزمون با استفاده از فرمول آلفرای طبقه‌ای ۰۸ برآورد شده است. این فرمول جهت برآورد ضرایب اعتبار ترکیبات خطي مورد استفاده قرار مي‌گيرد. (ناتالی، 1978؛ ناتالی و برنشاين، 1994؛ آسرون، 2000.)

تلمخن و همکارانش (1998) مطالعات متعددی در زمینه روایی آزمون ۲ ۱/۲ SON-R ۷-۲ انجام داده‌اند. نتایج این مطالعات حاکی از روایی خوب نمرات آزمون است. ازجمله این مطالعات می‌توان به همبستگی بین این آزمون بایستی آزمون‌های شناختی اشاره کرد.

جدول 2- خلاصه‌ای از ضرایب همبستگی بین ۲ ۱/۲ SON-R ۷-۲ با سایر آزمون‌های شناختی را نشان می‌دهد.

خرده آزمون‌های ۲ ۱/۲ SON-R ۷-۲ را می‌توان بر اساس نوع مواد که مورد استفاده قرار می‌گیرد به دو دسته تقسیم کرد: اول - خرده آزمون‌هایی که از مواد تصویری استفاده می‌کنند (طبیعی یا دستی موافقت‌شده و یا لازمه)؛ دوم - خرده آزمون‌هایی که از مواد تصویری بی‌مورد مانند اشکال هندسی استفاده می‌نمایند (موزاییک‌ها، الگوهای و قیاسی‌ها).
جدول ۲- ضرایب همبستگی آزمون ۲ SON-R ۷- با سایر آزمون‌های شناختی

<table>
<thead>
<tr>
<th>آزمون ملاک</th>
<th>N</th>
<th>۰.۸</th>
<th>۹</th>
<th>۱۵۲</th>
<th>۵۰</th>
<th>۷۲</th>
</tr>
</thead>
<tbody>
<tr>
<td>TONI-۲*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WPPSI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>مقیاس کلی</td>
<td></td>
<td></td>
<td></td>
<td>۵۳</td>
<td>۴۵</td>
<td>۱۱۵</td>
</tr>
<tr>
<td>مقیاس عمل</td>
<td></td>
<td></td>
<td></td>
<td>۳۹</td>
<td>۲۹</td>
<td>۱۱۵</td>
</tr>
<tr>
<td>BOS</td>
<td></td>
<td></td>
<td></td>
<td>۴۴</td>
<td>۴۹</td>
<td>۱۱۵</td>
</tr>
<tr>
<td>GOS*</td>
<td></td>
<td></td>
<td></td>
<td>۴۸</td>
<td>۳۹</td>
<td>۱۱۵</td>
</tr>
<tr>
<td>RAKIT</td>
<td></td>
<td></td>
<td></td>
<td>۴۳</td>
<td>۳۸</td>
<td>۱۱۵</td>
</tr>
</tbody>
</table>

مطلب‌باید در مورد افزایش برخی اینکه یا به نحوی که با هم مطابقت یا استانداردی باشد، ممکن است به این شکل نشان دهد که باید هرگونه مطالعه، تحقیق یا بررسی چنین جامعه‌ای به استانداردهای استاندارد و هنری‌بایندی‌گردد. از این رو در این مطالعات، فرم کوتاه آزمون مورد استفاده قرار گرفت و گزارش حاصل انستیتیت حاصل از فرم کوتاه آزمون مربوط می‌شود.

شیوه اجرای

گام نخست در این مطالعه، ترجمه دستورالعمل اجرای آزمون به زبان فارسی و دیگر متون و مطالب تسخیر بر نحوه اجرایی و دستورالعمل مورد نظر جهت برگزاری آزمون و مقایسه میزان مشارکت از چنین فرم کوتاه آزمون استفاده گردید. در این انتخاب مهدکار کودک، مجوز لازم برای حضور در مهارت بهبودیستی اخذ گردید. کلیه موسسه‌های توسط مؤسسه زیست‌شناسی مورد آزمون قرار گرفتند. جم‌آوری داده‌ها تا به طور تمام انجامید.

همانطور که گفته شد اجرای آزمون به شیوه انتظاری صورت می‌گیرد. در این مطالعه نیاز از شیوه‌های انتظاری استفاده کرده‌دی. برای اجرای آزمون بر روی هر کودک، بطور متوسط ۵۰ دقیقه وقت صرف گردید. بعد از هر سوال به کودک گفت‌گو می‌شد که آیا پاسخ یا درست است یا نادرست. ارائه پاسخ‌های صحیح به کودک یک بخش مهم از شیوه اجرای
استاندارد آزمون 2/7 SON-R تا 2/7 باید باعث می‌گردد تا دستورالعمل‌های روشن و واضح گردد و به کودک فرض می‌دهد تا خطاها و وپاره‌شدن‌ها خود درس بگیرد و استراتژی حل مسئله خود را تعیین نماید. بعد از هر پاسخ نادرست، بلافاصله از کودک پرسیده می‌شود که چگونه اجازه دهد که الکترونیک را به سوال را درست کند. نوازیت نام آنها را یادآوری می‌کند. اجرای براورد آزمون بعد از 3 پاسخ نادرست متناوب و یا 2 پاسخ نادرست متوالی متوقف می‌گردد. از توابع گردد که خوده‌ایت باشند. از مردانه گردد که خوده‌ایت در طول ساعت کلاس در یک میانگین خوب و متغیر متوسط ضعیف از شیبی نامیدند.

تحليل دادها

رای تحلیل و درآوردن و نزدیک‌سازی روا سنگی سوالات (درد. ۹۰ درصدی و درد ۵ درصدی) از مدل چند پارامترهای IRT، نمودار نیست. این گردد که جهت پاسخ و پاسخ‌های داده‌هایه و پاسخ‌های پاسخ‌های داده‌هایه در محل IRT از این پاسخ‌هایه ۹۸۸ (استاندارد گردد) و ۹۸۸ (انتشار گردد) و ۹۸۸ (انتشار گردد) به دلیل این داده‌هایه، ۹۸۸ (انتشار گردد) و ۹۸۸ (انتشار گردد) از داده‌هایه. در بخش چندین نمودار نیست. در آزمون‌های کوتاه به انداده فرمول‌های کروکورس، کم‌برادر نمک کنی (تریبرگ و زگرس)، جهت پاسخ‌های آزمون‌های ماشینه و کل آزمون از فرمول‌های طبقه ای نیست. در بررسی این اثبات، از نظر لازم است استاندارد آزمون ۳۰۰۰ (انتشار گردد) در بررسی نیست.

برای بررسی تفاوت بین میانگین نمرات دختران و پسران در خود آزمون‌ها از تحلیل واریانس چند متغیری (MANOVA) به کار برده و از آزمون T، همیشه ۰۰ (استیونس ۲/۰۰۰۰) استفاده شد و برای بررسی اختلاف بین میانگین نمرات این دو گرده در مقایسه (عملی و استاندارد) و گرده در نتیجه آزمون ۴ بار دو گرده مستقل (هالو ۲۰۰) استفاده گردد. در انجام این تحلیل‌ها از نمرات خام استفاده گردد. زیرا در نمرات خام تمام واریانس یا پراکندگی موجود در نمرات به حساب می‌آید (اختلاف و
نتایج

میانگین و انحراف استاندارد نمرات خام و نمرات استاندارد و همچنین جدولی و کشیدگی خرده آزمون‌ها، مقیاس‌ها و کل آزمون به تفکیک جنس (جوان و ) و کل گروه (جدول 1) ارائه شده است.

جدول 3: مشخصه‌های آماری خرده آزمون‌ها، مقیاس‌ها و کل آزمون در گروه پسران (N = 66)

<table>
<thead>
<tr>
<th>نمرات استاندارد</th>
<th>نمرات خام</th>
<th>خرده آزمون‌ها و مقیاس‌ها</th>
</tr>
</thead>
<tbody>
<tr>
<td>میانگین</td>
<td>انحراف استاندارد</td>
<td>میانگین</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>9</td>
</tr>
</tbody>
</table>

جدول 4: مشخصه‌های آماری خرده آزمون‌ها، مقیاس‌ها و کل آزمون در گروه دختران (N = 59)

<table>
<thead>
<tr>
<th>نمرات استاندارد</th>
<th>نمرات خام</th>
<th>خرده آزمون‌ها و مقیاس‌ها</th>
</tr>
</thead>
<tbody>
<tr>
<td>میانگین</td>
<td>انحراف استاندارد</td>
<td>میانگین</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>
جدول ۵ منشوه‌های آماری خرده آزمون‌ها، مقیاس‌ها و کل آزمون در کل گروه (N=125)

<table>
<thead>
<tr>
<th>خرده آزمون‌ها و مقیاس‌ها</th>
<th>نمرات خام</th>
<th>نمرات استاندارد</th>
<th>کشیدگی</th>
<th>جولگی</th>
</tr>
</thead>
<tbody>
<tr>
<td>موزاییکها</td>
<td>7 9</td>
<td>2 4</td>
<td>7 9</td>
<td>2 4</td>
</tr>
<tr>
<td>طلا ایران</td>
<td>7 3</td>
<td>2 3</td>
<td>7 3</td>
<td>2 3</td>
</tr>
<tr>
<td>انسان‌شناسی</td>
<td>7 3</td>
<td>2 3</td>
<td>7 3</td>
<td>2 3</td>
</tr>
<tr>
<td>انگلها</td>
<td>7 9</td>
<td>2 4</td>
<td>7 9</td>
<td>2 4</td>
</tr>
</tbody>
</table>

(نتیجه آزمون: نشان داده که از میانگین نمرات دختران و پسران در مقیاس عملی، مقياس استدلال و کل آزمون تفاوت ام‌ام‌اوا درامای معنادار وجود ندارد. مقادیر مشاهده شده برای مقیاس عملی، مقياس استدلال و کل آزمون به ترتیب [0.11, 0.06] است. 

شاخص‌های آماری سوالات بر اساس مدل کلاسیک اعتبار و مدل ۳ پرامتری می‌شود. TR به ترتیب در جدول ۶ و ۷ اندازه است. درجه نشانه‌ی سوال در مدل کلاسیک، که به نماد P نشان داده می‌شود، بیانگر نسبت افرادی است که به سوال پاسخ درست دادند. قدرت تشخیص سوال در این مدل، که گاهی اوقات روابط سوال نیز نامیده می‌شود، همواره به صورت همبستگی دو رشته‌ای یا دو رشته‌ای از سوال به نمره کل آزمون یا خرده آزمون بین می‌شود (ترنداکی، ترجمه هر، 2019). این شاخص معلوم می‌نماید که سوال می‌تواند گروه‌های مختلف افراد را از هم جدا کند. در واقع این شاخص نشان دهنده همسویی و هماهنگی سوال با آزمون یا خرده آزمون است. از آنگا که در برنامه PSS، همبستگی دو رشته‌ای از نوع ای وجود می‌کند. )
ندارد، جهت برآورده کردن میزان آزمون از همبستگی گشتاوری پرسشنامه استفاده گردید. همبستگی دو روشی ای نماید که برای محاسبه معادل با همبستگی گشتاوری پرسشنامه است (کلایر، 2000). از انجا که نمره سوال در نمره کل خرده آزمون لحاظ می‌شود، این امر باعث افزایش گرفتن همبستگی سوال با نمره کل می‌گردد. از این رو، به تاثیر نمره سوالان بعنوان پیش‌بینی نمره کل خرده آزمون، یک اصلاح صورت گرفت (کلایر، 2000؛ ناتالی و برنشتاین، 1994).

جدول ۶- درجه دشواری و قدرت تشخیص سوالات بر اساس مدل کلاسیک اعتبار

<table>
<thead>
<tr>
<th>سوال</th>
<th>درجه دشواری</th>
<th>قدرت تشخیص</th>
<th>موارسیکا</th>
<th>طبقه‌بندی‌ها</th>
<th>درجه انطباق</th>
<th>قدرت تشخیص</th>
</tr>
</thead>
<tbody>
<tr>
<td>۱</td>
<td>۳</td>
<td>۷</td>
<td>۳</td>
<td>۷</td>
<td>۵</td>
<td>۰</td>
</tr>
<tr>
<td>۲</td>
<td>۵</td>
<td>۹</td>
<td>۵</td>
<td>۹</td>
<td>۷</td>
<td>۰</td>
</tr>
<tr>
<td>۳</td>
<td>۷</td>
<td>۱۱</td>
<td>۷</td>
<td>۱۱</td>
<td>۹</td>
<td>۰</td>
</tr>
<tr>
<td>۴</td>
<td>۹</td>
<td>۱۳</td>
<td>۹</td>
<td>۱۳</td>
<td>۱۱</td>
<td>۰</td>
</tr>
<tr>
<td>۵</td>
<td>۱۱</td>
<td>۱۵</td>
<td>۱۱</td>
<td>۱۵</td>
<td>۱۳</td>
<td>۰</td>
</tr>
<tr>
<td>۶</td>
<td>۱۳</td>
<td>۱۷</td>
<td>۱۳</td>
<td>۱۷</td>
<td>۱۵</td>
<td>۰</td>
</tr>
<tr>
<td>۷</td>
<td>۱۵</td>
<td>۱۹</td>
<td>۱۵</td>
<td>۱۹</td>
<td>۱۷</td>
<td>۰</td>
</tr>
<tr>
<td>۸</td>
<td>۱۷</td>
<td>۲۱</td>
<td>۱۷</td>
<td>۲۱</td>
<td>۱۹</td>
<td>۰</td>
</tr>
<tr>
<td>۹</td>
<td>۱۹</td>
<td>۲۳</td>
<td>۱۹</td>
<td>۲۳</td>
<td>۲۱</td>
<td>۰</td>
</tr>
<tr>
<td>۱۰</td>
<td>۲۱</td>
<td>۲۷</td>
<td>۲۱</td>
<td>۲۷</td>
<td>۲۳</td>
<td>۰</td>
</tr>
<tr>
<td>۱۱</td>
<td>۲۳</td>
<td>۲۹</td>
<td>۲۳</td>
<td>۲۹</td>
<td>۲۷</td>
<td>۰</td>
</tr>
<tr>
<td>۱۲</td>
<td>۲۵</td>
<td>۳۱</td>
<td>۲۵</td>
<td>۳۱</td>
<td>۲۹</td>
<td>۰</td>
</tr>
<tr>
<td>۱۳</td>
<td>۲۷</td>
<td>۳۳</td>
<td>۲۷</td>
<td>۳۳</td>
<td>۳۱</td>
<td>۰</td>
</tr>
<tr>
<td>۱۴</td>
<td>۲۹</td>
<td>۳۵</td>
<td>۲۹</td>
<td>۳۵</td>
<td>۳۳</td>
<td>۰</td>
</tr>
<tr>
<td>۱۵</td>
<td>۳۱</td>
<td>۳۷</td>
<td>۳۱</td>
<td>۳۷</td>
<td>۳۵</td>
<td>۰</td>
</tr>
</tbody>
</table>

نتایج جدول فوق نشان می‌دهدکه درجه دندانی و قدت تشخیص سوالات در سطح مطلوب قرار دارد. اظهار می‌دارند که متوسط درجه دندانی سوالات باید تقریباً ۳۰ باشد و همچنین بايد پراکندگی نسبتی ژیادی از دندانی سوالات وجود داشته باشد.

ارقام ارائه شده در جدول ۶ حاکی از این است که اکثر سوالات از قدرت تشخیص مطلوب و بالایی بر دندان و صرفه ۸ سوال یا گروه بسته و برنشتاب (۹۴۶) اظهار می‌دارند که اکثر ضریب همبستگی سوال - نمره کل در دامنه ۰ تا ۱۰۰ قرار دارد. گزاره ۰ ۹۶۵ به نقل از نیوگام و هامیل (۹۷۸) اظهار کرده است که در ژیستیر ضریب همبستگی ۰ ۰۳ با الگو در صورتی که آزمون نسبتی طولانی باشد می‌تواند درای روابط تلقی شود. آنالیز (۹۹۰) و آنالیز و بررسی (۹۹۷) توصیه می‌کند که ضریب همبستگی بیش از ۰ ۰۳ می‌تواند قابل قبول تلقی گردد.

در مدل ۳ پارامتری، هر سوال براساس ۴ پارامتر ۰۰۱، و توصیف گردید. در مدل ۲ پارامتری و فرض بر این است که امکان باش درست از طریق حدس وجود ندارد و مقدار این برای با سفر در نظر گرفته می‌شود. به پایان تقریباً در این مدل پارامتر ۶ برای سوال بیان شد و سوال صرفه بر اساس ۲ پارامتر و توصیف می‌شود. نماد ۵ در پارامتر بانگ پارامتر دندانی سوال است و با توجه به مقدار توانایی مسکن (۱۰۱). به راه فنی و رضایت می‌توان گفت که این پارامتر نقطه عطف تابع ویژگی سوال را توصیف می‌کند و معمولاً در زمان زیر (۵۰ تا ۵۰۰۰ می‌گذرد.

پارامتر قدرت تشخیص در مدل ری IRT با نماد ۳ نمایش داده می‌شود و بیانگر این است که سوال تا چه انداد، می‌تواند دیگر افرادی که تنها آنها پایین‌تر از جایگاه سوال است با افرادی که تنها آنها بالاتر از این جایگاه قرار دارد، تماشای ایجاد شده. به بیان دیگر پارامتر ۴ نشان می‌دهدکه وقتی سطح تنایی مکن ۷۷ بیا می‌رود احتمال موفقیت در سوال با چه سرعتی افزایش می‌یابد (ترندیده، ترجیح هرمز ۱۳۷۵).

به هنگام نظیر این پارامتر در مقیاس (۵ تا ۵۰۰۰) تعیین می‌شود. وی سلاتی که دارای قدرت تشخیص منفی استند از آزمونها حذف می‌گردد. همچنین بهبود آورده
مقدار بزرگتر از 1 برای این پارامتر غیر معنی‌دار است. با این حال، دامنه این پارامتر معمولاً بین 0.5 تغییر می‌کند.

پارامتر ت پیانگر احتمال بدست آوردن پاسخ درست سوال از طریق حدس محض است. به بیان دیگر، این پارامتر نشان می‌دهد که در سطح خیلی پایین از توانایی قرار دارد. هر چند دامنه تغییر پارامتر حدس پارامتر ت تا 0.5 است اما در عمل مقدار بالاتر از 0.5 برای این کاندیدات نیست (بیک، 135). جدول 7 پارامترهای سوالات را بر اساس مدل‌های IRT، که به استفاده از برنامه CL (هاسور، 2000) بر اساس شد اند تماش می‌دهد.

جدول 7: پارامترهای a، b و c سوالات بر اساس مدل‌های IRT (۱۲۵-۷)
جدول 8- همبستگی پیشنهادی سوالات در مدل کلاسیک و مدل‌های IRT

<table>
<thead>
<tr>
<th>همبستگی پیشنهادی در مدل کلاسیک و IRT</th>
<th>خریده آزمونها</th>
</tr>
</thead>
<tbody>
<tr>
<td>مدل کلاسیک و IRT 0.7</td>
<td>موزاییکها</td>
</tr>
<tr>
<td>0.5</td>
<td>طبقه بندی</td>
</tr>
<tr>
<td>0.3</td>
<td>موفقیت‌ها</td>
</tr>
<tr>
<td>0.1</td>
<td>الگوها</td>
</tr>
</tbody>
</table>

پژوهش در حیطه کودکان استثنایی، 71/ سال پنجم، شماره3/ پاییز 1387

جدول 9- همبستگی بین نمره خام و برآورد‌های ثبت در خریده آزمونها

<table>
<thead>
<tr>
<th>نمره خام و برآورد‌های ثبت</th>
<th>خریده آزمونها</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAP و MLE</td>
<td>EAP و MLE</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>موزاییکها</td>
<td>طبقه بندی</td>
</tr>
<tr>
<td>موفقیت‌ها</td>
<td>الگوها</td>
</tr>
</tbody>
</table>

اعتبار نمرات خریده آزمونها که بر اساس فرمول PS (فلدتر و برنز: 988؛ اسپانژ، 2000)، برآورد شده‌اند به تفکیک جنس (جدول 1)، و گروه سنی (جدول 2) ارائه شدند. در این جداول همچنین ضرایب اعتبار نمای ت مقیاس عمیل، MLE، MLE (جدول 9) و كل آزمون (IQ) که با استفاده از فرمول آللفای طبیعی (نانالی و برنشتایزر، 1994؛ اسپانژ، 2000) برآورد شده‌اند نشان داده شده است. از آنجا که ضرایب اعتبار PS براساس نمونه‌هایی که از حاضری سن ناهمگن هستند برآورد شده،
بنابراین جهت حذف تأثیر سرژ، یک اصلاح صورت گرفت. ضرایب اربه شده در جدول 10 ضرایب اصلاح شده هستند.

جدول 10- ضرایب اعتبار نمرات خرده آزمون‌ها، مقیاس‌ها و کل آزمون به تفکیک جنس و کل گروه

<table>
<thead>
<tr>
<th>خرده آزمون‌ها</th>
<th>دختران</th>
<th>پسران</th>
</tr>
</thead>
<tbody>
<tr>
<td>کل گروه</td>
<td></td>
<td></td>
</tr>
<tr>
<td>کل آزمون</td>
<td>موزاییک‌ها</td>
<td>طبیه‌بندی‌ها</td>
</tr>
<tr>
<td>مقياس عملي</td>
<td>0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>مقياس استدلال</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>کل آزمون</td>
<td>0.11</td>
<td>0.04</td>
</tr>
</tbody>
</table>

همانطور که در جدول 10 و 11 ملاحظه می‌کنید، ضرایب اعتبار نمرات خرده آزمون‌ها، مقیاس‌ها و کل آزمون برای هر جنس، کل گروه و گروه‌های سنی مختلف تمام‌اً در سطح مطلوب قرار دارند. البته بودن مقدار این ضرایب حاکی از این است که آزمون مورد بحث آزمون معتمد است و می‌توان نتایج آن را با اطمینان مورد استدلال قرار داد.

جدول 11- ضرایب اعتبار نمرات خرده آزمون‌ها، مقیاس‌ها و کل آزمون به تفکیک گروه سنی

<table>
<thead>
<tr>
<th>خرده آزمون‌ها و مقیاس‌ها</th>
<th>0-10</th>
<th>11-20</th>
<th>21-30</th>
</tr>
</thead>
<tbody>
<tr>
<td>کل گروه</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>کل آزمون</td>
<td>موزاییک‌ها</td>
<td>طبیه‌بندی‌ها</td>
<td></td>
</tr>
<tr>
<td>مقياس عملي</td>
<td>0.1</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>مقياس استدلال</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>کل آزمون</td>
<td>0.11</td>
<td>0.04</td>
<td>0.06</td>
</tr>
</tbody>
</table>
ضرائب تعمیم پذیری

ضرایب تعمیم پذیری نماینده کل آزمون و مقیاس‌های عملی و استدلال نیز مورد بررسی قرار گرفت. این ضرایب نشان می‌دهد که تنها نتایج خرید آزمون‌ها را به کل حیطه مورد سنجش خرده آزمون‌ها تعمیم داده‌اند (تلهخ و مکارز ۹۸). تعمیم پذیری با استفاده از فرمول ضریب آلفا محاسبه گردید. که در آن به جای نمرات سوال از نمرات خرده آزمون یا به‌عنوان واحد تحلیل استفاده شده است. از آنها که شاخصی از هم‌سایی درونی برای آزمون‌ها یا خرده آزمون‌های همگن است می‌توان برای بروز اعتبار استفاده مود. با این حال ضریب آلفا برای نمره کل خرده آزمون‌ها که هر کدام از آنها واریانس مختصر خود را دارند در ایالت معیار متقابل است. در این مورد ضریب آلفا ر می‌توان به‌عنوان داشت شاخص تعمیم پذیری تفسیر نمود. قدر ضریب آلفا همبستگی مورد انتظار بین نمره کل آزمون و نمره کل در یک ترکیب متقابل اما با حجم یکسان (مساوی) از خرده آزمون‌های دو جهت خرده آزمون‌های غیرکلاسی مشابه. تصویر کرد. ضریب آلفا همبستگی مورد انتظار بین نمره کل آزمون و نمره کل در یک ترکیب متقابل اما با حجم یکسان (مساوی) از خرده آزمون‌های دو جهت خرده آزمون‌های غیرکلاسی مشابه. تصویر کرد.

جدول ۱۲: ضرایب تعمیم پذیری نماینده کل آزمون، مقیاس عملی و استدلال به تفکیک جنس، گروه سنی و کل گروه

<table>
<thead>
<tr>
<th>جنس</th>
<th>کلم آزمون</th>
<th>مقیاس عملی</th>
<th>مقیاس استدلال</th>
<th>گروه سنی</th>
<th>پسران</th>
<th>دختران</th>
<th>کم آزمون</th>
<th>اجمالی</th>
</tr>
</thead>
<tbody>
<tr>
<td>کل گروه</td>
<td>۰/۹</td>
<td>۰/۶</td>
<td>۰/۵</td>
<td>۰/۷</td>
<td>۰/۵</td>
<td>۰/۳</td>
<td>۰/۹</td>
<td>۰/۹</td>
</tr>
<tr>
<td>۲۰-۳۰</td>
<td>۰/۹</td>
<td>۰/۶</td>
<td>۰/۵</td>
<td>۰/۷</td>
<td>۰/۵</td>
<td>۰/۳</td>
<td>۰/۹</td>
<td>۰/۹</td>
</tr>
<tr>
<td>۳۱-۴۰</td>
<td>۰/۹</td>
<td>۰/۶</td>
<td>۰/۵</td>
<td>۰/۷</td>
<td>۰/۵</td>
<td>۰/۳</td>
<td>۰/۹</td>
<td>۰/۹</td>
</tr>
<tr>
<td>۴۱-۵۰</td>
<td>۰/۹</td>
<td>۰/۶</td>
<td>۰/۵</td>
<td>۰/۷</td>
<td>۰/۵</td>
<td>۰/۳</td>
<td>۰/۹</td>
<td>۰/۹</td>
</tr>
</tbody>
</table>

رابطه بین نمرات خرده آزمون‌ها با یکدیگر و نمرات هر خرده آزمون با مقیاس عملی، استدلال و کل آزمون و همچنین رابطه بین نمرات هر خرده آزمون با مجموع ۳ خرده آزمون دیگر مورد بررسی قرار گرفت. همبستگی درونی نمرات خرده آزمون‌ها با یکدیگر...
برای کل گروه و برای دختران و پسران در جدول 3، ارائه شده است. به منظور کنترل تأثیر سن از روی همبستگی تفکیکی استفاده گردید (گیلفورد و فرومبت. 1578؛ جن و پایوی. 2002).

جدول 13- ضرایب همبستگی درونی نمرات خرده آزمون‌ها در کل گروه و به تفکیک پسران و دختران

<table>
<thead>
<tr>
<th>خرده آزمون‌ها</th>
<th>پسران</th>
<th>دختران</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mos</td>
<td>Cat</td>
<td>Sit</td>
</tr>
<tr>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
</tr>
</tbody>
</table>

میانگین ضرایب همبستگی بین خرده آزمون‌ها که با استفاده از روش تبدیل ۲ فیشر بسته آماده برای کل گروه و گروه پسر و دختران به ترتیب ۵۸.۸، ۳.۳ و ۳.۰ است. کلیه ضرایب همبستگی در سطح ۵٪ هندسی آماری معناداری هستند. تفاوت بین ضرایب همبستگی درونی خرده آزمون‌ها در دو گروه دختران و پسران با استفاده از آزمون Z (جن و پایوی. 2002) مورد بررسی معناداری آماری قرار گرفت. نتایج نشان داد که تفاوت بین هیچ یک از ضرایب همبستگی در سطح ۵٪ به لحاظ آماری معنادار نیست.

جدول 14- همبستگی بین نمرات خام خرده آزمون‌ها با نمرات خام مقیاس عملی (PS)

مقیاس استدلال (RS) و نمره کل آزمون را برای کل گروه و به تفکیک بیرای پسران و دختران نشان می‌دهد. همبستگی نمرات خرده آزمون‌ها با نمره کل آزمون از طریق محاسبه همبستگی بین نمرات هر خرده آزمون با مجموع ۳ خرده آزمون دیگر صورت می‌گیرد. به منظور کنترل تأثیر سن از روی همبستگی تفکیکی استفاده گردید (گیلفورد و فرومبت. 1578؛ جن و پایوی. 2002). کلیه ضرایب در سطح ۱٪ به لحاظ آماری معنادار محسون. تفاوت بین ضرایب همبستگی در دو گروه دختران و پسران با استفاده از آزمون Z (جن و پایوی. 2002) مورد بررسی معناداری آماری قرار گرفت. نتایج نشان داد که در سطح ۵٪ بین هیچ یک از ضرایب همبستگی تفاوت آماری معنادار وجود ندارد.
جدول 14 - ضرایب همبستگی بین خرده آزمون‌ها و مقیاس‌ها و نمره کل آزمون برای کل گروه و به تفکیک جنس

<table>
<thead>
<tr>
<th>دختران</th>
<th>پسران</th>
</tr>
</thead>
<tbody>
<tr>
<td>نمره کل RS</td>
<td>نمره کل PS</td>
</tr>
<tr>
<td>64/05</td>
<td>54/80</td>
</tr>
<tr>
<td>78/00</td>
<td>64/00</td>
</tr>
<tr>
<td>89/00</td>
<td>75/00</td>
</tr>
<tr>
<td>59/00</td>
<td>45/00</td>
</tr>
</tbody>
</table>

جدول 15 - میانگین نمرات استاندارد خرده آزمون‌ها، مقیاس‌ها و نمره کل آزمون به تفکیک گروه‌های سنی

<table>
<thead>
<tr>
<th>گروه سنی</th>
<th>خرده آزمون‌ها و مقیاس‌ها</th>
</tr>
</thead>
<tbody>
<tr>
<td>فردی</td>
<td>کل آزمون</td>
</tr>
<tr>
<td>6-7</td>
<td>8-9</td>
</tr>
<tr>
<td>همبستگی با سن</td>
<td></td>
</tr>
<tr>
<td>موزاییکا</td>
<td>طبیعی بندی</td>
</tr>
<tr>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

از آنجا که هوش ماهیت ویژه ویسته به سن است لذا با استناد به یک آزمون با سن تقویمی قوی همبستگی داشته باشند. میانگین نمرات استاندارد گروه‌های سنی مختلف در خرده آزمون‌ها، مقیاس‌ها و کل آزمون در صورتی، ارائه شده است. نمونه آخر این جدول که عملکرد در خرده آزمون‌ها، مقیاس‌ها و کل آزمون با سن رابطه دارد، زیرا که با فاز ویژه، میانگین نمرات خرده آزمون‌ها، مقیاس‌ها و کل آزمون آغاز به یادبود. این نتیجه‌گیری با این واقعیت که تمامی ضرایب همبستگی ارائه شده در سن‌گروه‌های چپ جدول به لحاظ آماری در سطح 0.01 معنادار هستند مورد تایید قرار می‌گیرد. این یافته‌ها بیانگر این است که آزمون مورد بحث از قدرت تماشایگذاری سنی بسیار خوبی بیانگر است.
همانطور که پیش از این نیز گفته شد به هنگام اجرای آزمون از مربیان مهندسی کودک در خواست کردیم تا کودک را از لحاظ میزان انگیزه، توجه، همگرایی و درک دستورات و اموزش‌های در یک مقياس ۴ امتیازی از تهیه و ضعیف، ارزیابی نمایند. علاوه بر این، پژوهشگر نیز بر اساس مشاهده خود در طول اجرا از آزمون هر کودک ای به همان شیوه مورد ارزیابی قرار می‌داد. بین عملکرد کودکان در خرده آزمون و مقیاس‌ها و کل آزمون با ارزیابی مربیان و پژوهشگر ضریب همبستگی محاسبه گردید. از آنجا که خرده آزمون در مقیاس‌ها و کل آزمون از اعتبار کامل (یعنی برخورد نیستند، از این رو جهت کنترل تأثیر بر اعتبار از قسمت اصلاح برای کاهش (گیلفورد و فروچی، ۹۷۸ استفاده گردید. ضرایب ارائه شده در جدول ۱۶ ضرایب اصلاح شده هستند.

میانگین ضرایب همبستگی مربوط به ارزیابی‌های پژوهشگر، که با استفاده از روش تبدل Z تی‌پ‌د، به دست آمد، ۰.۱۰ است که یک ضریب همبستگی متوسط (نقطه ۷۷۱) محسوب می‌گردد. علاوه بر این، بیش از ضریب همبستگی خرده آزمون موضوعی‌با میزان همکاری که ۵.۰ است، کلیه ضرایب همبستگی در سطح ۱.۰ به لحاظ آماری معنادار هستند.

متوسط ضرایب همبستگی مربوط به ارزیابی‌های مریب ۴:۵ است که یک ضریب همبستگی بسیار ضعیف (نقطه ۳۷) در حساب می‌آید. این باعث شد اکثر ما و با یافته‌های حاصل از تحقیقات تلنخ و همکاران (۱۹۹۸) و تلنخ و لاروس (۲۰۰۴) همکاران ندارند. اینها در پژوهش‌های خود دریافتند که بین عملکرد کودکان در خرده آزمون با ارزیابی‌های مختلف، همبستگی و سطح وجود دارد. یکی از دلایل احتمالی عدم همکاران دریافت‌ها ممکن است این باشد که در ایران تعداد کودکانی که تحت اموزش و نظارت یک مربی قرار دارند نسبتاً زیاد است و این مسئله باعث می‌گردد تا مربی فرضی یا انرژی‌گران که شناخت کرده بهتر کودکان و ارزیابی دقیق آن‌ها نداشته باشد.
جدول 16 - ضرایب همبستگی بین ممکن در خرده آزمونها، مقیاسها و کل آزمون با ارزیابی‌های مربی و پژوهشگر

<table>
<thead>
<tr>
<th>خرده آزمونها/مقايسه‌ها/کل آزمون</th>
<th>ارزیابی مربی</th>
<th>ارزیابی پژوهشگر</th>
<th>انگیزش نهایی همکاری</th>
<th>نهایی همکاری دستورالعمل‌ها</th>
</tr>
</thead>
<tbody>
<tr>
<td>مراقبت‌های طبی بدنیها</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>مراقبت‌های پیش‌بینی</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>آزمون‌های اخلاقی</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>کل آزمون</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
</tr>
</tbody>
</table>

آخرین شاهدی که در مورد روایی مورد بررسی قرار گرفت، قدرت تمایزگذاری است. این شاخص که به واریانس یا پراکندگی نمرات مربوط است، مانی به حداکثر می‌رسد که توزیع نمرات به شکل مستطیلی باشد. حداقل تمایزگذاری (سفر) زمانی به دست خواهد آمد که تمام افراد نمره یکسانی دریافت کنند. شاخص قدرت تمایزگذاری آزمون، دلیل فرگسون است که مقدار آن بیز 10 تغییر نمی‌کند (فرگسون 949؛ به نقل از کلازی، 1980). مقدار این شاخص در توزیع‌های نرمال 0.3 است. بطور کلی، آزمون‌های اخلاقی همکاری به حداقل افتاده است. ضرایب قدرت تمایزگذاری خرده آزمون‌ها می‌باشد که به کل آزمون به تغییر جنس و کل گروه ارائه شده است. کلیه ضرایب در سطح بالای 0.1 قرار دارد و با ملاک ارائه شده توسط کل آزمون همخوانی دارد.

جدول 17 - ضرایب قدرت تمایزگذاری خرده آزمون‌های مقیاس‌ها و کل آزمون

<table>
<thead>
<tr>
<th>خرده آزمون‌های مقیاس‌ها/کل آزمون</th>
<th>پسران</th>
<th>دختران</th>
<th>کل گروه</th>
</tr>
</thead>
<tbody>
<tr>
<td>مراقبت‌های طبی بدنیها</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>مراقبت‌های پیش‌بینی</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>آزمون‌های اخلاقی</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>کل آزمون</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Downloaded from joec.ir at 19:11 +0430 on Wednesday August 21st 2019
شناخت تصاویر

همانطور که قبلاً گفته شد، خرده آزمونهای طبقه بندی پد و موقتیتی از تصاویر معنادار تشکیل یافته اند. از این رو احتمال تورش فرتهانگی ۷ در این نوع خرده آزمون‌ها بیشتر از خرده آزمون‌های است که از تصاویر پد معد، سایر اشکال هندسی، تشکیل یافته‌اند (جنسی‌تبار ۹۸۰). بنابراین شیوه دیگری که جهت بررسی مناسب و روا بودن آزمون ۲-۷ برای کودکان ایرانی مورد استفاده قرار گرفت بر شناخت تصاویر خرده آزمون‌های طبقه بندی پد و موقتیتی توسط کودکانی که به سوال پاسخ نادرست داده‌اند، مفیدی بود. اصلی که در پاره‌ای از این شیوه قرار دارد، این است که کودک نباید به خاطر اینکه دو یا چند تا از تصاویر یک سوال را تمی شناسد در دادن پاسخ درست به سوال دچار شکست گردد. جدول ۱۸ تا از سوالات خرده آزمون طبقه بندی پد را نشان می‌دهد که شامل تصاویری هستند که برای حداقل ۵۰ درصد آزمون مورد مطالعه، که به سوال پاسخ نادرست داده‌اند، ناشانه‌هستند. بستون اول، شماره سوال و شماره تصویر را نشان می‌دهد. در بستون دوم فراوانی پاسخ‌های نادرست به سوال این شده است و در بستون سوم نریصدی از این کودکان که تصویر را نمی‌شناختند نشان داده شده است.

بر طبق نتایج جدول ۱۸ تصاویر مورد استفاده در سوالات ۸ ۱ ۸، ۱۳، ۱۰، ۱۹۸۱ خرده آزمون طبقه بندی پد بطور متوسط برای ۱۰۰ درصد کودکان مطالعه ناشانه‌هستند. در خرده آزمون موقتیتی، کودکان مورد مطالعه هیچ مشکلی در تشکیل تصاویر مورد استفاده در این خرده آزمون نداشتد. بطور کلی، نتایج این روش نشان داد که از بین ۶۰ سوال این آزمون، بنظر می‌آید تعداد ۸ تا سوالات خرده آزمون طبقه بندی پد رای ورش هستند و لازم است برای آشنایی جایگزین آنها گردد.
جدول 18 سوالات خرده آزمون طبقه‌بندی‌شده شامل تصاویری است که برای حداقل 500 از کودکان ناشتا هستند.

<table>
<thead>
<tr>
<th>سوال</th>
<th>درصد پاسخ‌های نمی شناسیم</th>
<th>درصد پاسخ‌های نمی شناسیم</th>
<th>درصد پاسخ‌های نمی شناسیم</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 (تصویر شماره 1: اصلی)</td>
<td>27%</td>
<td>8%</td>
<td>27%</td>
</tr>
<tr>
<td>8 (تصویر شماره 5: چاپ‌گری)</td>
<td>12%</td>
<td>12%</td>
<td>12%</td>
</tr>
<tr>
<td>1 (تصویر شماره 1: چاپ‌گری)</td>
<td>34%</td>
<td>34%</td>
<td>34%</td>
</tr>
<tr>
<td>1 (تصویر شماره 5: چاپ‌گری)</td>
<td>22%</td>
<td>22%</td>
<td>22%</td>
</tr>
<tr>
<td>0 (تصویر شماره 1: چاپ‌گری)</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>0 (تصویر شماره 3: چاپ‌گری)</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>0 (تصویر شماره 4: چاپ‌گری)</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>0 (تصویر شماره 6: چاپ‌گری)</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>0 (تصویر شماره 1: چاپ‌گری)</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>0 (تصویر شماره 3: چاپ‌گری)</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>0 (تصویر شماره 4: چاپ‌گری)</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>0 (تصویر شماره 6: چاپ‌گری)</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>0 (تصویر شماره 1: چاپ‌گری)</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>0 (تصویر شماره 3: چاپ‌گری)</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>0 (تصویر شماره 4: چاپ‌گری)</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>0 (تصویر شماره 6: چاپ‌گری)</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

بحث و نتایج گیری

در این پژوهش، ویژگی‌های روان سنی و عملی بودن آزمون 2-R-SON-R-7 برای کودکان ایرانی مورد بررسی قرار گرفت. نتایج این مطالعه (جدول 2) نشان می‌دهد که نمرات خرده آزمون‌ها، مقیاس‌ها و کل آزمون ز هم‌سانی درونی بالایی برخوردارند. این یافته که با یافته‌های حاصل از مطالعات تلخ و همکاران ۹۹۸ : و تلخ و لاروس ۹۹۸ هم‌خوانی دارد بنابراین این است که اعتبار نمرات آزمون به اندازه‌ای است که بتوان با اطمینان از آن استفاده کرد.

بر اساس یافته‌های مربوط به تماشای گذاری سنی (جدول 13) همبستگی سوال-نمره کل (جدول 14) همبستگی درونی خرده آزمون‌ها (جدول 13) همبستگی خرده آزمون‌ها با
مقیاس‌ها و نمره کل (جدول ۴) و همچنین همبستگی بین نمرات چندکان در خرده آزمون، مقیاس‌ها و کل آزمون با ارزش‌یابی‌های پژوهشگر و مربی از جای انگیزش به دست وظیفه و موانع کنن نتیجه گرفت که این آزمون یک ابزار روا جهت سنجش هوش کلی است. می‌توان با اطلاعات خاطر از ان استفاده کرد. شاهد دیگری که در خصوص روایت نمرات آزمون مورد بررسی قرار گرفت، قدرت تایگن‌گرایی (جدول ۴) است. مقدار این شاخص برای کلیه خرده آزمون، مقیاس‌ها و کل آزمون در سطح بالایی قرار دارد و حاکی از روابط نمرات آزمون است. علاوه برای، با فاقد‌های حاصل از مطالعات بین‌فرهنگی که درخصوص روابط این آزمون در کشورهای آمریکا، استرالیا، و انگلیس، صورت گرفته است (الاروس و تلخر ۹۹۹) حاکی از این است که همبستگی این آزمون با مقیاس‌های عملی تعدادی از BAS70 و WPPSI-R، قوی‌تر از همبستگی آن با مقیاس‌های کلاسیک این آزمون است. این همبستگی‌ها از روابط واگر و همگرای آزمون مورد مطالعه حمایت می‌کند.

تحلیل‌هایی که در خصوص شاخص‌های آماری سوالات (درجه دشواری و قدرت تشخیص) با استفاده از مدل کلاسیک اعتبار (جدول ۲) و مدل ۲ ۳ پارامتری مبتنی بر RT صورت گرفت (جدول ۲) نشان داد که کلیه سوالات از شاخص‌های آماری مطلوبی برخوردارند. بر ورودی مربوط به بارامتر ۶ در مدل‌های RT نشان می‌دهد که سوالات از لحاظ دشواری در دامنه ۳-۷ + پترن تقریباً یکنواخت توزیع شده. علاوه برای براو این آزمون در اثر می‌توان یافته‌ها از قدرت تشخیص مطلوبی برخوردارند.

پژوهش‌های می‌توان گفت که آره ن ۲ SON-R ۷-۲ با به نظر الگو یافته‌های مربوط به اعتبار، روایت، مدت اجرای سهولت نمر اغذی و تفسیر که از مهم‌ترین جهت‌های عملی بودن آن به حساب می‌آید (هومون ۳۸۱) یک ابزار کاملاً مناسب برای سنجش هوش کلی کودکان ۷-۲ سال ایران است و از آن می‌توان بعنوان وسیله‌ای معیار و روا یا برای (الف) شناسایی کودکانی که بهتر قابل ملاحظه ای از لحاظ شناختی، ضعیف تر از همسالان خود هستند، ب) مستند کردن پیشرفته کارکرد
شناختی کودکان در نتیجه برنامه‌های مداخله‌ای ویژه، ج. مشخص کردن نقاط قوت و ضعف کودکان در مهارت‌های شناختی و انداز گیری هر گونه کلی در مطالعات پژوهشی استفاده کرد.

با این حال، با ایده‌ها خاطر نشان کرد که قبل از استانداردیسازی و هنجاربندی آزمون در مقياس وسیع باید برخی از تصاویر سوالات 8، 14 و 15 خرده آزمون طبق به یا با تصاویری که مناسب با فرهنگ ایران و آشنا برای کودکان ایرانی هستند، جایگزین گشتند. علاوه برای، ترتیب سوالات نیز با باید بر اساس دشواری آنها اصلاح گردد. به نظر می‌رسد که با انجام این اصلاحات، می‌توان مطمئن بود که استانداردیسازی و هنجاربندی آزمون مورد بحث در مقياس وسیع می‌تواند سودمند خواهد بود.

یادداشت‌ها:

1) General Intelligence 2) Learning Potential
3) Tellegen & Laros 4) Cattel, R.B.
5) Raven 6) Nan Snijders - Oomen
7) Snijders-Oomen Nonverbal Intelligence Test 8) Reliability
9) Weiss 10) Validity
11) International Test Commission 12) Van de Vijver & Hambleton
13) Multi – Stage Random Sampling 14) Mosaics
15) Categories 16) Puzzles
17) Analogies 18) Situations
19) Patterns 20) Performance Scale
21) Reasoning scale
22) the test commission of the NetherLands Institute for Psychologists
23) Guttman 24) Stratified Alpha
25) Nunnally 26) Nunnally & Bernstein
27) Osmburn 28) Meaningful
29) Non-Meaningful 30) Bayley Developmental Scales
31) Groningen Developmental Scales
32) Revision of Amsterdam Intelligence Test for Children
33) Test of Nonverbal Intelligence
34) Item–Response Theory
35) IRT Command Language
36) Hanson
37) Feldt & Brennan
38) Ten Berge & Zegres
39) Two-Group Multivariate Analysis of Variance
40) Hotelling’s T²
41) Stevense
42) Howell
43) Achenbach & Rescorla
44) Item Validity
45) Kline, P.
46) Garrett
47) Newcomer & Hammill
48) Baker
49) Latent Ability
50) Maximum Likelihood (MLE)
51) Bayes
52) Expected a posterior (EAP)
53) Generalizability Coefficient
54) Partial Correlation
55) Guilford & Fruchter
56) Chen & Popovich
57) Fisher’s Z Transformation
58) Correction for Attenuation
59) Discrimination Power
60) Cultural Bias
61) Jensen
62) British Ability Scales
63) McCarthy Scales of Children’s Abilities


