انطباق و هنجاریابی آزمون رشد زبان برای کودکان فارسی زبان تهرانی (بخش اول)

TOLD-P:3

سید حسن زاده

اصغری‌پور

چکیده:
پژوهش حاضر با هدف انطباق، اعتباری‌پایی، ورای بررسی و هنجاریابی آزمون رشد زبان، روش نمونه‌ای با حجم ۴۲۲ نفر (۲۴ نفر دختر و ۱۸۵ نفر گویندگان آزمون ت지도ئین) از کودکان ۱۲ به ۶ سال (متوسط سن) از جمعیت ۷–۱۱ سال و ۸–۱۱ سال) شهر تهران که به روش نمونه برداری چند مرحله‌ای انتخاب شدند، اجرا گردید. این آزمون مشتمل بر ۶ خرده آزمون (۳ خرده آزمون اصلی و ۳ خرده آزمون نکتمپلیک) است. عملکرد این ترکیب نمرات خرده آزمون به‌وسط نمره ۶ کنونی می‌باشد. برای انطباق و هنجاریابی آزمون این طرح دستورالعمل اجرا به ویژه هنگامی که به زبان فارسی ترجمه شد. اگرچه به توجه به ویژگی‌های نویز، معنی شناختی و واحدهای زبان فارسی و با مد نظر قرار دادن مبانی نظری آزمون، نسبتی از کوچک‌ها مجدداً حاصلی گردید با اجرای مراحل بیشتر تجربی و تجربی و تجزیه و تحلیل داده‌ها بر پایه مدل کلاسیک اعتبار قرنیه به آزمون راه‌یافته از آزمون با استفاده از فرمول النگار (نمرات مکمل) با استفاده از فرمول گینسون که برای برآورد نمرات مکمل طراحی شده است در آزمون و ترکیب‌ها با استفاده از روش‌های مناسب بوده و برای شناسایی البنا به دست آمده در زمینه اعتبار و روابط آزمون، اطمنان کافی برای استفاده از این ابزار را برای ارزیابی مهارت‌های زبانی کودکان فارسی زبان تهرانی به‌دست می‌دهد.

کلید واژه‌ها: هنجاریابی، زبان دریافتی، زبان بینایی، اعتباری‌پایی، ورای بررسی

این تحقیق در پژوهشکده کودکان استثنایی انجام شده است.
عضو هیأت علمی پژوهشکده کودکان استثنایی
مقدمه

نسخه اولیه آزمون رشد زبانould در سال 1977 میلادی توسط نیوگامبر و همکار برای ارزیابی رشد زبان کودکان انگلیسی زبان - 40 سال تمام (8 سال و 11 ماه) در آمریکا تنظیم شد. این نسخه شامل 5 خرده آزمون اصلی: آزمون تصوری، آزمون شفاهی، درک دستوری، تقدم جمله و تکمیل دستوری و 6 خرده آزمون تکمیلی: تماشای گذاری کلمه و تولید کلمه بود. با جمع اندازه خرده آزمون های اصلی بهره بزایی محسوب می شد. در سال 1982 میلادی آموزش به TOLD-P، نرم‌افزار مهارت‌های زبانی کودکان 6-12 تا 8-11 همراهی دیگر خود یعنی TOLD-P، آموزش به ارزیابی مهارت‌های زبانی کودکان 6-12 تا 8-11 ساله تهیه شده بود، مجزا گردید. علاوه بر تغییر نام، چند تغییر قابل ملاحظه دیگر نیز در این نسخه از آزمون صورت گرفت. پیوست آنکه از آزمون‌نگری می توانست از طریق ترکیب تفاسیر نتایج حاصله از خرده آزمون‌ها، نمرات مرکب یک شاخص نماینده گویش کرد. صحبته کردن، نحو و معنی شناسی می شدند. با به دست آوردن دوم آنکه در دستورالعمل اجرایی آزمون توجه ویژه‌ای به مبانی نظری آزمون و همچنین نحوه تفسیر نتایج شده بود و بالاخره آنکه جمعیت نمونه هنجری به 1014 نفر افزایش پیدا کرده بود. همچنین شواهدی دال بر روایت محتملاً ارائه شده بود ولی با این حال محتوای اصلی و گویه‌های این نسخه با نسخه قبلی یکسان بود. در سال 1988 میلادی ویرایش دوم این آزمون با عنوان 2 منتشر شد. در این نسخه خرده آزمون‌های تماشای گذاری کلمه و تولید کلمه مجدداً طراحی شدند. از ترکیب ویژه‌ای این در خرده آزمون، نمره ترکیبی آوازشناسی به دست می آمد. نمونه هنجری این نسخه از آزمون به ۲۴۵۳ نفر رسید. برای افزایش اعتبار آزمون تعداد گویه‌های برخی از خرده آزمون‌ها نیز افزایش ییدا کرد. آخرین نسخه آزمون مورد نظر در سال 1994 هنجری در ویرایش دو در سال 1997 انتشار یافت. در مطالعه حاضر، آخرین نسخه آزمون برای انتباه و هنجری به زبان فارسی مورد استفاده قرار گرفته است. در این نسخه نسبت به نسخه‌های قبلی علاوه بر آنکه مجددًا هنجری بر روی گروه‌های 1000 نفری از 28 ایالت آمریکا انتخاب شده است، شواهدی بیشتر عالم تورش ناشی از جنسیت و تعداد ارائه شده است و ضریب اعتبار برای زیرگروه‌های نمونه
هنجرای نیز محاسبه شده است. گذشته از این مطالعات بسیاری در زمینه روانی آزمون انجام شده است تا نشان دهد آزمون علاوه بر آنکه برای جمعیت کل روانی دارد در ارتباط با زیر گروه‌ها نیز از روایی کانی برخوردار است. در خرده آزمون تحلیل واحی و واگذاری به مجموعه آزمون اضافه شده و به خرده آزمون تماشاخ‌گذاری کلیه، تحلیل واحی و تولید کلمه نیز به عنوان خرده آزمون‌های تکمیلی معرفی شده اند. بنیان نظری علمی آزمون با توجه به نظریات جاری زبانشناسی روز آمد شده و کلیه تصورات خرده آزمون‌های مصور رنگی شده اند.

چارچوب نظری آزمون رشد زبان


1977 بهره بردن (نیکارد و هامیل، 1997).
است. کاری که توسط ویلک (1989) و نیبولد (1988) در مورد مهارت‌های فرا زبانشناسی نوجوانان انجام شده، نشان دهنده این تأکید و توجه است. اگر چه ارزش این نوع مطالعات و نظریات زبانشناسی مشخص است، ولی این اطلاعات برای تهیه آزمون‌های هنجر مرجع استاندارد شده مناسب نیستند. بهتر است از این نوع تحقیقات زبانی، همراه با دیگر رویکردهای غیررسمی ارزیابی-آموزشی استفاده شود (نیکوکار و هامیل، 1997).

در اینجا چهار یک از مختصات زبانشناسی و نیز هر یک از نظام‌های زبانشناسی مختصراً توضیح داده می‌شود تا خواننده بتواند منطق حاکم بر خود را آزمون‌ها را بهتر درک کند. البته خواننده با پایتخت بی‌خاطر داشته باشد که اگر چه این مختصات به صورت جداگانه و مستقل توضیح داده می‌شوند، اما در واقع یکپارچه و مرتبط با هم هستند. در ادامه به بررسی نظام‌ها و مختصات زبانی و رابطه آنها با الگویی که برای مفهوم پیکایشیدن به این آزمون به کارگرفته شده است، می‌پردازیم.

اشخاصی که از قواعد نحوی آگاهی ضمنی ندارند، نمی‌توانند جملات دستوری قابل قبولی را بیان کنند. با این حال به دلایل مختلف، برخی از افرادی که آگاهی مناسب از نحو دارند هنوز هم در دستور زبان دچار مشکلاتی هستند. این کاربرد دستوری نامناسب یا استاندارد نشده کمتر مورد توجه زبانشناسان می‌باشد، زیرا آنها عمداً به درک این موضوع علاقه‌مند هستند که چرا متفکران یک زبان دارای توانایی و یک یک سری جملات نامحدود هستند (نیکوکار و هامیل، 1997).

معناشناسی: اصلاح معناشناسی به مطالعه معنا در زبان اطلاق می‌شود (پنی رابطه میان زبان و تفکر). دانش معناها مستلزم درک نحوه ارتباط معنا با اشكال تکواز زبانی یا اجرای زبان است و همچنین قوانینی که چگونگی انتقال معنا را از طریق جملات نشان می‌دهد. به عبارت دیگر، لغات به خصوصی از معناشناسی محصول می‌شوند، زیرا به توانایی درک و کاربرد معنای کلمات مستقل مانند "توب" مرتبط هستند. به علاوه، کاربرد کلمات مستقل مانند "توب" در یک جمله معنی دار به آگاهی معناشناسی ویژه ی نیاز دارد. برای نمونه، جمله "توب، پسر را زد" از نظر نحوی، تکواز زبانی و اجرای زبانی صحیح است اما از نظر معناشناسی غیر ممکن است. همچنین معناشناسی فراتر از محدوده جمله بوده و به بانی وسیعتر مرتبط می‌شود.
نظام‌های زبان‌نواختی

علاوه بر تقسیم زبان به مختصات آن، بر اساس نظریه پردازان زبان‌نواختی مانند بلوم (1973)، چی‌پیه (1974) و ایستقرام (1976) زبان را بر اساس نظام‌هایی که گوشه کردن (دیوانی)، سازمان‌دهی (ترکیبی-ویژه‌ای) و صحبت کردن (بینایی) نامیده می‌شوند، طبقه‌بندی می‌کنند. به طور کلی، با یادگیری این نظام‌های نشان‌دهنده چگونگی کاربرد و درک زبان هستند. به طور دقیق‌تر، گوشه کردن شامل عملیات رمزگذاری است که توسط آن معنی نمادین و افتراقی به گفتار شنیده شده، اختراعات می‌باشد. صخبت کردن، توصیف کننده عملیات رمزگذاری است که در طی آن گفتار معنی در فرمول‌بندی و تولید می‌شود. ظاهراً بسیاری از فعالیت‌های زبانی، این دو نظام را در بر می‌دارند. با این وجود، از نظر کمیتی تنها تکلیف زبانی را به آسانی به دو شکل صخبت کردن و یا گوشه کردن تقسیم کرد، محققان علاقومند به الگوهای پردازش اطلاعات در حوزه رشد زبان همانند سارکمی (1971) و رابینویتز و کاسر (1983) نظام‌های زبانی را گسترش‌دهنده کردند. تا یک نظام سازمان‌دهی را پیش آن اضافه کنند. این نظام به ظرفیت منجر می‌شود تا تفکیک یا سازمان‌دهی اطلاعات زبان‌نواختی وارد حساب گوشنده بکند تا در اشاره دارد. کاس (1983) ثابت کرد که طور نظری ظرفیت حافظه (یعنی توانایی فراخوانی اطلاعاتی که قبل از گرفته شده‌اند) به کارآیی پردازش سازمانی بستگی دارد. دو راهکار اساسی سازمان‌دهی عبارتند از سازگاری و تکه‌گیری گروه‌های که در آن یک نماد به نماد دیگر مثلا می‌شود، همانگونه که "گرم و سرد" یا "پدر و مادر" به هم مرتب نهادند.

الگوی دوبردی

چارچوب نظری زبان‌نواختی که در قسمت قبل به آن پرداخته شد را می‌توان به سورت یک الگوی دوبردی به دو کل مختصات زبان‌نواختی و نظام‌های زبان‌نواختی است. نشان داد (به جدول 1 نگاه کنید). این جدول، ارتباط میان مفاهیم موجود در الگو و خرده آزمون‌های آزمون رشد زبان اولیه را نیز نشان می‌دهد.
جدول 1-الگوی دو بعدی ساختار زبانی که در ساختن خرده‌آزمون‌های آزمون‌ریزی شما از آن استفاده شده است

<table>
<thead>
<tr>
<th>مهارت‌های زبان‌شناختی</th>
<th>مهارت‌های پیکره‌چسبی (ویژه ویزیگزایی)</th>
<th>مهارت‌های دریافت حرف</th>
<th>نظام‌های زبان‌شناختی</th>
</tr>
</thead>
<tbody>
<tr>
<td>ویژگان شفاهی</td>
<td>ویژگان تصویری</td>
<td>تحلیل واژه</td>
<td>معناشناختی</td>
</tr>
<tr>
<td>ویژگان ریتمی</td>
<td>تکمیل دستوری</td>
<td>تحلیل واژه</td>
<td>نحو</td>
</tr>
<tr>
<td>تکمیل دستوری</td>
<td>تحلیل واژه</td>
<td>تمايزگذاری کلمه</td>
<td>واج شناسی</td>
</tr>
<tr>
<td>تحلیل واژه</td>
<td>تحلیل واژه</td>
<td>تمايزگذاری کلمه</td>
<td>واج شناسی</td>
</tr>
</tbody>
</table>

مختصات زبان‌شناختی

مخصص مبانی وسیله‌های خرده‌آزمون‌های ارائه‌شده است. خرده‌آزمون ویژگان تصویری که نظام گوش کردن را دربر می‌گیرد؛ خرده‌آزمون ویژگان ریتمی که شامل نظام سازمانی است و خرده‌آزمون ویژگان شفاهی که نظام صحت کردن را در بر دارد. همانگونه که از اساسی این سه خرده‌آزمون برداشت می‌شود آنها خزانه لغات را منسجم می‌کند. اگر چه خزانه لغات تناها جنبه محکم‌تر از مبانی وسیله‌های ارائه‌شده است و برای آموزه‌گر و والدین نیز آشنا و مفهوم‌دار می‌باشد.

مختصات نحوی نیز به وسیله به خرده‌آزمون‌های ارائه‌شده است. خرده‌آزمون درک دستوری، عمداً نظام گوش کردن، خرده‌آزمون تکمیل جمله، نظام سازماندهی و خرده‌آزمون تکمیل دستوری نیز نظام صحت کردن را دربر می‌گیرد. خرده‌آزمون درک دستوری به اندازه‌گیری توانایی فهم نظم و ترتیب کلمات از دست داده. خرده‌آزمون تکمیل دستوری، آگاهی و اطلاعات کودکان را از تکاورانشناختی (بعنی توانایی آنها را در به کارگیری تکاورانشناختی مناسب در بانک جمله) می‌سنجد.

مختصات واج شناسی نیز به وسیله به خرده‌آزمون اندازه‌گیری می‌شود که هر یک از آنها نشان دهنده یک نظام خاص زبان‌شناختی است. خرده‌آزمون تمايزگذاری کلمه، نظام گوش کردن، خرده‌آزمون تحلیل واژه، نظام سازماندهی و خرده‌آزمون تولید کلمه نیز نظام صحت کردن را در بر می‌گیرد. محتوای واقعی خرده‌آزمون‌های تمايزگذاری کلمه و تولید کلمه بیشتر به اندازه‌گیری گفتار می‌پردازد تا زبان (نیوکامو و همایل، ۱۹۹۷).  محتوای خرده‌آزمون تحلیل واژه به مهارت منحصر به فردی احتیاج دارد که پیش‌بینی
کننده بسیار خوبی از توانایی خواندن است (البته درک و ناشنوایی از این امر به نظر می‌رسد). در کل، تماسی این سه خرده آزمون یادهای مهمی از رشد زبانی کودکان و یادگیری خرده سال را می‌سنجند.

عموماً کودکان یک ژنرالی از ۶ تا ۷ ساله، به‌طور موقت آمیزی اکثر توانایی‌های واژگانی را در نظام‌های زبانی خود گنجانده‌اند.

نظام‌های زبانشناسی

خرده آزمون‌های که نظام گوش کردن را می‌سنجند عبارتند از: تمازگذاری کلمه (واژه‌شناسی)، و ارگان تصویری (معناشناسی) و درک دستوری (نحو) که یادهایی از عمليات رمزگشایی موجود در درک گفتار را می‌سنجند.

نظام سازمانده توسط خرده آزمون‌های تحمل واجی (واژه‌شناسی) و ارگان ربطی (معناشناسی) و تکلیف جمله (نحو) سنجیده می‌شود. این خرده آزمون‌ها یادهایی از راهکارهای تداعی گروه و واسطه‌گری را که برای سازمانده یازده گفتار استفاده می‌شوند، اندازه‌گیری کرده و آنها را به نمادهای دیگر متصل می‌کنند.

نظام صحت کردن توسط سه خرده آزمون تولید کلمه (واژه‌شناسی) و ارگان شفاهی (معناشناسی) و تکمیل دستوری (نحو) سنجیده می‌شود. تمامی این سه خرده آزمون، یادهایی از رمزگذاری را که برای تولید گفتار معنی‌دار استفاده می‌شوند، ارژیابی می‌کنند.

توصیف خرده آزمون‌های ترکیب‌ها

الگوی دو بعدی که مورد بررسی قرار گرفت اساس نظریه ۹ خرده آزمون است. خرده آزمون‌های مربوط به معناشناسی و نحو، خرده آزمون‌های اصلی و سه خرده آزمون مربوط به واژه‌شناسی جزئی خرده آزمون‌های تکمیلی هستند.

خرده آزمون‌های اصلی

خرده آزمون‌های اصلی شامل شش خرده آزمون هستند که نتایج آنها ترکیب شده و نمرات ترکیبی (یعنی برای نمونه، گوش کردن، سازمانده، صحت کردن، معناشناسی و نحو) را به شکل می‌دهند. این خرده آزمون‌ها یک سوست معنا شناسی و نحو به کار می‌روند. خرده‌
آزمون‌هایی که با شناسایی را انداده‌گیری می‌کنند، اجزای تکمیلی آزمون رشته زبان اولیه هستند. جداسازی نمرات واحش نشان‌رسان از نمرات تکیه‌گاه گوش کردن، سازمان‌دهی، صحبت کردن و زبان گفتاری باعث تمایزی واضح و آشکار می‌شود و فناوری نتایج (یعنی مهرات های ارتباطی) اصلی که توسط خردهآزمونهای واحش نشان‌رسان انداده‌گیری می‌شوند و توانش زبانی شده و موجب گردیده است تا اختلالات ویژه در هر یک از این حیطه‌ها، به خوبی و با کارآیی بالا مشخص شوند.

خردهآزمون 1-واژگان تصوری: این خردهآزمون مکانی‌شناختی دارای 30 گویه است و میزان درک کودک را از معانی مرتبط با کلمات فارسی انداده‌گیری می‌کند.

خردهآزمون 2-واژگان ربطی: این خردهآزمون مکانی‌شناختی که 30 گویه دارد به سنجش توانایی کودک در درک و بیان شفاهی ارتباط می‌باشد.

خردهآزمون 3-واژگان شفاهی: این خردهآزمون مکانی‌شناختی 28 گویه دارد و به انداده‌گیری توانایی کودک در ارائه تئاتری شفاهی از کلمات رابطه‌ای فارسی می‌پردازد.

خردهآزمون 4-درک دستوری: این خردهآزمون نحوی، 25 گویه دارد و توانایی کودک را در درک و فهم معنای جمله‌ها می‌سنجد.

خردهآزمون 5-تملیق جمله: این خردهآزمون نحوی که به صورت 30 گویه‌ای طراحی شده است، به انداده‌گیری توانایی کودکان در تولید صحیح جمله‌های فارسی می‌پردازد.

خردهآزمون 6-درک دستوری: این خردهآزمون نحوی، 28 گویه دارد و به سنجش توانایی کودکان در شناختن، فهمیدن و به کارگیری اشکال تکواز شناختی رابطه زبان فارسی می‌پردازد.

خردهآزمون‌های تکمیلی

این گروه از خردهآزمون‌ها شامل 3 خردهآزمون زیر است:

خردهآزمون 7-تماپی‌گذاری کلمه: این خردهآزمون واحش نشان‌رسان با داشتن 20 گویه به سنجش توانایی کودکان در تشخیص تفاوت‌های موجود در صدادهای گفتاری مشخص می‌پردازد.

خردهآزمون 8-تحلیل واحش: این خردهآزمون 14 گویه دارد و یک تکنیک سازمان‌دهی است که به بررسی توانایی کودک در مورد نظام واحش نشان‌رسان زبان فارسی می‌پردازد. این خردهآزمون
یک جنبه از مهارت پردازش شنیداری، به ویژه توانایی تجزیه کلمات به واحدهای واجی کوچک‌تر را می‌سنجد.

خوده آزمون 9- تولید کلمه: این خردآزمون واج شناختی، 20‌گویه دارد و توانایی کودک را در تولید صداها و گفتاری مهم زبان فارسی اندازه‌گیری می‌کند.

ترکیب‌ها

همان‌گونه که قبلا گفتیم، اساس تهیه خرده‌آزمون‌ها یک الگوی دو بعدی بوده است. هر خرده‌آزمون، یک نظام زبانی (گویش کردن، سازماندهی، صحت کردن) و یک مختصه زبانی (معناشناختی، نحو، واج شناختی) را شامل می‌شود. بنابراین، خرده‌آزمون‌ها با توجه به نظام‌ها و یا مختصه‌های مشترک می‌توانند گروه‌بندی شوند و ترکیب‌های زیر را به وجود آورند:

- گویش کردن (واژگان تصویری + درک دستوری)
- سازماندهی (واژگان ربطی + تکمیل جمله)
- صحت کردن (واژگان شفاهی + تکمیل دستوری)
- معنی‌شناختی (واژگان تصویری + واژگان ربطی + واژگان شفاهی)
- نحو (درک دستوری + تکمیل جمله + تکمیل دستوری)
- زبان گفتاری (واژگان تصویری + واژگان ربطی + واژگان شفاهی + درک دستوری + تکمیل جمله + تکمیل دستوری)

مراحل هنگامی همه‌پرسی

۱- مرحله بیش تجربی: برای بررسی کیفی گویه‌ها و آگاهی نسبت به مسائل و مشکلات احتمالی اجرایی، آزمون ابتدا بر روی ۹۶ نفر (۴۹ پسر و ۴۷ دختر) از کودکان ۴-۹ ساله مراکز پیش دستاپ و پایه‌های اول تا سوم مدارس ابتدا اجرای گردید و براساس نتایج حاصله تغییراتی در برخی از گویه‌ها داده شد و همچنین تعدادی از گویه‌های مهم حذف و گویه‌های مناسب تری جایگزین آنها شد.
۲- مراحل تجریبی: در این مرحله، آزمون بر روی ۲۸۰ چودک (۱۹۰ دختر و ۱۹۰ پسر) در گروه سنی ۱۰ - ۰۴ تا ۱۱ - ۰۷ ساله اجرا گردید. نمونه فون به صورت تصادفی از بین کودکان مراکز پیش دستی و دسته‌های مناطق گانه شهرتهران انتخاب گردید. برای انتخاب نمونه از بین کودکانی که به مرکز پیش دستی و یا مدرسه نمی‌رفتند از نمونه‌های در دسترس استفاده شد.

اجرای این مرحله، پس از آموزش آزمونگران طرح که حداقل دارای مدرک لیسانس روانشناسی بودند انجام گرفت. پس از استخراج نتایج، تجزیه و تحلیل داده‌ها بر اساس مدل کلاسیک اعتبار صورت گرفت. با استفاده از این مدل برای هر گروه، ۲ یا کمتر درجه دستوری و قدرت تشخیص برآورد گردید. براساس نتایج حاصل از این مرحله تعدادی از گروه‌ها حذف گردید و فرم نهایی آزمون براساس گروه‌های باقی مانده شکل گرفت.

۳- اجرای نهایی: در این مرحله برای انتخاب نمونه برای تشکیل جداول هنری از روی نمونه‌گیری چند مرحله‌ای استفاده گردید. به این صورت که ابتدا شهر تهران به مس منطقه جغرافیایی شمال، مرکز و جنوب تقسیم گردید. آنگاه یا به‌هورگیری از اطلاعات منتشر شده توسط مرکز آمار ایران، نمونه لازم برای هر منطقه جغرافیایی براساس متغیرهای جنسیت و سن کودکان و همچنین وضعیت اقتصادی - اجتماعی و دلیل تبعیض گردید. در گام بعدی با توجه به حجم جمعیت در هر منطقه جغرافیائی، تعدادی مدرسه‌های کودکان و پسران و نیز مرکز پیش دستی و بوادان داشته باشند. سپس از بین دانش‌آموزان مدارس و مراکز پیش دستی و نمونه لازم استخراج و آزمون بر روی آنها اجرا شد. برای انتخاب نمونه از بین کودکانی که به مرکز پیش دستی و یا مدرسه نمی‌رفتند از نمونه‌های در دسترس استفاده شد. در نمونه انتخاب شده عدادی کودک که توان ذهنی، دارای اختلالات گفتاری، ناتوانایی یادگیری و سایر معلولیت‌ها نیز نگاهندگانش داد. در نهایت آزمون بر روی ۱۲۵ کودک اجرا و جداول هنری براساس نتایج حاصل از این تعداد به دست آمد. آزمون‌گران این مرحله، هم‌مان آزمون‌گران مرحله تجربی بوادان که در اجرای آزمون تجلیل کافی داشتند. در نهایت نمونه نهایی برای تشکیل جداول هنری به تفکیک متغیرهای جمعیت شناختی در جدول ۲ آمده است.
جدول ۲ - ویژگی‌های جمعیت نشات‌یابی گروه‌های هنری (N = ۱۳۳۵)

<table>
<thead>
<tr>
<th>فراوانی</th>
<th>درصد</th>
</tr>
</thead>
<tbody>
<tr>
<td>چندهتی</td>
<td>۵۶</td>
</tr>
<tr>
<td>هفتتی</td>
<td>۶۹</td>
</tr>
<tr>
<td>ناحیه جغرافیایی تهران</td>
<td></td>
</tr>
<tr>
<td>شمال</td>
<td>۳۱۷</td>
</tr>
<tr>
<td>مرکز</td>
<td>۴۸۶</td>
</tr>
<tr>
<td>جنوب</td>
<td>۴۲۲</td>
</tr>
<tr>
<td>وضعیت روانی</td>
<td></td>
</tr>
<tr>
<td>عادی</td>
<td>۱۱۳۷</td>
</tr>
<tr>
<td>دارای اختلال واداری</td>
<td>۲۷</td>
</tr>
<tr>
<td>دارای اختلال انتقالی</td>
<td>۲۶</td>
</tr>
<tr>
<td>کمتر از دیالگی</td>
<td>۱۷</td>
</tr>
<tr>
<td>سایر معلولت‌های</td>
<td>۸</td>
</tr>
<tr>
<td>سطح تحصیلات مادر</td>
<td></td>
</tr>
<tr>
<td>فوق لیسانس</td>
<td>۲۵</td>
</tr>
<tr>
<td>دیپلم تا لیسانس</td>
<td>۶۹۲</td>
</tr>
<tr>
<td>کمتر از دیپلم</td>
<td>۵۱۹</td>
</tr>
<tr>
<td>سطح تحصیلات پدر</td>
<td></td>
</tr>
<tr>
<td>فوق لیسانس</td>
<td>۷۷</td>
</tr>
<tr>
<td>دیپلم تا لیسانس</td>
<td>۵۹۳</td>
</tr>
<tr>
<td>کمتر از دیپلم</td>
<td>۵۶۹</td>
</tr>
<tr>
<td>سن</td>
<td></td>
</tr>
<tr>
<td>۱ ساله</td>
<td>۱۹۷</td>
</tr>
<tr>
<td>۲ ساله</td>
<td>۳۲۷</td>
</tr>
<tr>
<td>۳ ساله</td>
<td>۲۷۳</td>
</tr>
<tr>
<td>۴ ساله</td>
<td>۲۳۰</td>
</tr>
<tr>
<td>۵ ساله</td>
<td>۲۲۰</td>
</tr>
<tr>
<td>۶ ساله</td>
<td>۲۸۲</td>
</tr>
<tr>
<td>۷ ساله</td>
<td>۲۷۳</td>
</tr>
<tr>
<td>۸ ساله</td>
<td>۲۸۲</td>
</tr>
</tbody>
</table>

نمرات هنری

نمرات استاندارد برای خرده‌آزمون‌ها و ترکیب‌ها جداول هنری برای نمرات خاص هر یک از خرده‌آزمون‌ها در ۱۰ گروه‌سنی به فاصله ۶ ماه از یکدیگر (از ۰ تا ۱۱-۸) نهایی گردیده. با توجه به اینکه در آزمون‌هایی که به صورت انفرادی اجرا می‌شوند، عوامل زمان و هزینه، تعداد افراد، مورد نیاز برای هنری‌هایی‌های را محدود
می‌کنند، لذا نوعی از نمونه‌هایی با حجم ۱۰۰ تا ۲۰۰ نفر برای هر گروه سنی یا کلاسی برای هدف‌های هنجاری استفاده می‌گردد. نوسانات نمونه‌گیری ناشی از حجم نسبتاً اندازه‌گیری می‌شوند. بنابراین با توجه به کمی حجم نمونه در هر گروه سنی (تقییاً ۱۰۰ نفر) و به منظور کاهش مشکلات ناشی از این موضوع، از روشن‌کارها و راپرتون (۱۹۸۷) ارائه‌دهنده‌ای استفاده می‌شود. در این روش، پس از تعیین تفاوت‌های موجود در میانگین‌ها و انحراف استانداردهای گروه‌های سنی متوالی، داده‌های گروه‌های سنی مختلف به‌کار می‌رود. کاربرد این روش منجر به هنگام‌های می‌گردد که در آن مقایسه‌گروه‌های سنی مختلف به‌طور می‌گردد. پس از تشکیل توزیع کلی، نمرات خام به نمرات استاندارد هنگام تبدیل شدهند. نتایج به‌دست آمده برای هر خریده آزمون با استفاده از یک تبدیل خطي به نمرات استاندارد با میانگین ۱۰ و انحراف استاندارد ۳، تبدیل گردید. نمره استاندارد متناهی با آزمون‌های ترکیبی که بهره نامیده می‌شود با استفاده از روش گیلفرد و فروچر (۱۹۷۸) برای ادغام واریانس‌ها محاسبه شد. این فرمول برای ادغام نمرات استاندارد خریده آزمون‌هایی که آزمون‌های ترکیبی را تشکیل می‌دهند، بر روی تبدیل شده نمرات استاندارد آزمون‌های ترکیبی دارای میانگین ۱۰۰ و انحراف استاندارد ۱۵ است.

معادله‌های سنی

معادله سنی خریده آزمون‌ها نیز تهیه گردید. این ارزش‌ها، بیانگر سطح با سن توانایی زبانی متناهی با نمره خام آزمون‌های است. برای تعیین معادله‌های سنی، میانگین نمرات خام هر خریده آزمون برای هر گروه سنی محاسبه و با استفاده از پهنه نرم‌افزار استاندارد گراف، نمودار پراکندگی که در محور y های آن سن بر حسب می‌آورد و در محور x های آن میانگین نمرات خام قرار داشتند. نمودار بسته آمده همواره گردید و در مرحله بعد برای تعیین نمرات خام متناهی با گروه‌های سنی خارج از محدوده‌های هنجاری‌ای و سابقه‌های سنی موجود در محدوده‌های هنجاری‌ای که به‌طور مستقیم مورد آزمون قرار گرفته بودند از عمل برونزیاری و درون‌یابی استفاده گردید.
نتایج

1- اعتبار آزمون

مفهوم اعتبار آزمون به این معناست که چگونه افرادی که در نمرات آزمون به تفاوت‌های واقعی در خصوصیات مورد مطالعه و چگونه ادنزه به خاطرای تصادفی قابل استند است (آناستزاسی، 1988، ص 91). به همین دلیل سازندگان یک آزمون به ویژه آزمون‌های تخصصی مانند آزمون مورد بحث بايد شاخص قابل توجهی مبنی بر این که آزمون‌شان از اعتبار کافی برخوردار است و می‌توانند با اطمینان برای هدف‌های عنوان شده مورد استفاده قرار گیرند. ارائه بهره‌مندی مطالعه اعتبار یک آزمون بر اساس مقدار واریانس خطای مربوط به نمرات آن تأکید دارد. این خطای می‌تواند باعث افزایش نمره‌های محتوایی و نمونه‌برداری زمانی و تفاوت‌های نمره‌گذاران باشد. صرف نظر از این که گذاشتن واریانس خطای مورد بررسی قرار گرفته باشد، نتایج معمولاً برحسب ضریب اعتبار که کاربرد ویژه‌ای از ضریب همبستگی متناول است گزارش می‌شوند. برای این که آزمونی به عنوان یک آزمون معتبر مورد توجه قرار گیرد ضریب اعتبار نمرات آن باید نزدیک به 0/8 و یا بیشتر از آن باشد. ضریب اعتبار 0/8 توسط بیشتر صاحب‌نظران برای تعیین بین آزمون‌های معترف و غيرمعترف پذیرفته شده است (آناستزاسی، 1987). در اینجا وضعیت خرد آزمون‌های اصلی و تکمیلی و ترکیب‌های در ارتباط با دو منبع خاطای محتوایی و زمانی مورد بحث قرار می‌گیرد.

1-1- نمونه‌برداری محتوایی

خطای ناشی از نمونه‌برداری محتوایی، به میزان قابل توجهی به درجه همگنی بین مواد یک آزمون یا خرد آزمون ارتباط دارد. از آن جهت که هدف آزمون، ادنزه‌گیری ویژگی، توانایی و یا محتوای خاصی است، لذا هر چقدر مواد آزمون ارتباط بیشتری با یکدیگر داشته باشند، خطا آزمون کمتر خواهد بود. هرگاه مواد آزمون، با یکدیگر ارتباط نداشته باشند، ویژگی‌های متفاوتی را ادنزه می‌گیرند و در نتیجه مقدار خطای ناشی از نمونه‌برداری محتوایی نزولی خواهد گرفت (نیوکام و هامیل، 1997، ص 55)
جدول 3- ضرایب آلфа در پنج گروه سنی

<table>
<thead>
<tr>
<th>سن به سال</th>
<th>میانگین ضرایب</th>
<th>خردآموزی‌ها</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>60</td>
<td>76</td>
<td>74</td>
</tr>
<tr>
<td>61</td>
<td>79</td>
<td>78</td>
</tr>
<tr>
<td>62</td>
<td>89</td>
<td>88</td>
</tr>
<tr>
<td>63</td>
<td>90</td>
<td>88</td>
</tr>
<tr>
<td>64</td>
<td>92</td>
<td>87</td>
</tr>
<tr>
<td>65</td>
<td>94</td>
<td>84</td>
</tr>
<tr>
<td>66</td>
<td>97</td>
<td>82</td>
</tr>
<tr>
<td>67</td>
<td>99</td>
<td>77</td>
</tr>
<tr>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(اعشار حذف شده است)

خطای نیم‌نبرداری محتمالی (اعتبار همسایی درونی) با استفاده از روش ضریب آلфа کروناخ مورد بررسی قرار گرفت. ضرایب آلфа برای خرده‌آموزی‌ها و ترکیب‌ها در پنج محدوده سنی با استفاده از داده‌های حاصل از کل نمونه هنجاریایی محاسبه شد. ضرایب آلфа برای ترکیب‌ها با استفاده از فرمول گیلگورد (1954، ص 293) که برای برآورد آلфа نمرات مرکب طراحی شده است به دست آمد. این نتایج در جدول 3 گزارش شده است. ضرایب آلфа برای کل نمونه از طریق میانگین گیری ضرایب آلфа سطوح سنی مختلف با استفاده از روش تبدیل Z به دست آمد.

همیشه نمی‌توان گفت که چون یک آزمون برای کل یک جامعه معنی‌دار است، پس برای هر یک از زیرگروه‌های درون آن جامعه نیز از اعتبار یکسانی برخوردار است. بنابراین سازندگان آزمون با یک ثابت کننده که آزمونشان در واقع برای گروه‌های فرعی به‌صورت زیرگروه‌هایی که
احتمالاً مورد ارژیابی قرار می‌گیرند یا کسی که به‌دلیل اختلاف نژادی، معلولیتی و زبانی، ممکن است مورد مناسبی برای سنگش نباشد نیز معیار است. در جدول ۴، ضرایب آلفا برای زیرگروه‌گزارش شده است.

زیرگروه‌هایی مورد مطالعه را پسران، دختران، کودکانی که مبتلا به نقص توجه/بیش فعالی، اختلال پاپاژیر، تأخیر در گفتار و زبان و کم‌توانی ذهنی تشخیص داده شده بودند. تشکیل می‌دهند.

جدول ۴- ضرایب اعتبار خردآموزان‌ها و ترکیب‌های دوگرهای اصلی و چهارگره فرعی

| ترکیب‌ها |
|---|---|
| خردآموزان‌ها | یرگه‌ی بصورتی یرگه‌ی بیش‌طوری یرگه‌ی تکمیل‌داری کلمه تحلیل واژه تولدکنیه | یرگه‌ی گوینده یرگه‌ی سایه‌مانده یرگه‌ی صحبتگردان معمول‌شناسی نحو زبان کتائار | |
|\(\text{پسران}\) | \(\text{دختران}\) | \(\text{کلاس‌های دفتری}\) | \(\text{کلاس‌های دفتری}\) | \(\text{کلاس‌های دفتری}\) | \(\text{کلاس‌های دفتری}\) | \(\text{کلاس‌های دفتری}\) | \(\text{کلاس‌های دفتری}\) |
| ۳۲ | ۶۵ | ۳۵ | ۶۵ | ۳۵ | ۶۴ | ۳۵ | ۶۴ |
| ۳۳ | ۶۸ | ۳۵ | ۶۸ | ۳۵ | ۶۴ | ۳۵ | ۶۴ |
| ۳۶ | ۶۸ | ۳۵ | ۶۸ | ۳۵ | ۶۴ | ۳۵ | ۶۴ |
| ۳۷ | ۶۸ | ۳۵ | ۶۸ | ۳۵ | ۶۴ | ۳۵ | ۶۴ |
| ۳۸ | ۶۸ | ۳۵ | ۶۸ | ۳۵ | ۶۴ | ۳۵ | ۶۴ |
| ۳۹ | ۶۸ | ۳۵ | ۶۸ | ۳۵ | ۶۴ | ۳۵ | ۶۴ |
| ۴۰ | ۶۴ | ۳۵ | ۶۴ | ۳۵ | ۶۴ | ۳۵ | ۶۴ |
| ۴۱ | ۶۴ | ۳۵ | ۶۴ | ۳۵ | ۶۴ | ۳۵ | ۶۴ |

۱- نمونه‌گیری زمانی

خطای ناشی از نمونه‌گیری زمانی به میزان ثبات عملکرد فرد در آزمون، در طول زمان
امکان‌دار. این خط معمولاً با روش آزمون - بازار آزمون بی‌اورد می‌گردد. نتایج زمانی آزمون حاضر نیز با استفاده از روش آزمون - بازار آزمون مورد مطالعه فارغ‌گرفت. تعادل افراد مورد مطالعه در این مرحله ۹۹ نفر بودند که به طور تصادفی از بین کل افراد مورد استفاده در هنجارهای آزمون به انتخاب گردیدند. فاصله بین دو اجرای آزمون ۳ ماه بود. نتایج حاصل در جدول ۵ ارائه گردیده است. در این جدول میانگین و انحراف استاندارد خردهآزمون‌ها و ترکیب‌ها در هر اجرای آزمون و همچنین ضرایب همبستگی (r) که نشان دهنده میزان رابطه بین دو اجرای آزمون است، ارائه گردیده است.

جدول ۵- نتایج اعتبار بازارآزمون

<table>
<thead>
<tr>
<th>جریان اول</th>
<th>جریان دوم</th>
<th>خردهآزمون‌ها</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r )</td>
<td>( SD )</td>
<td>( M )</td>
</tr>
<tr>
<td>0/86</td>
<td>7/1</td>
<td>8/15</td>
</tr>
<tr>
<td>0/88</td>
<td>6</td>
<td>15/6</td>
</tr>
<tr>
<td>0/85</td>
<td>9/1</td>
<td>11/3</td>
</tr>
<tr>
<td>0/86</td>
<td>3/1</td>
<td>18/3</td>
</tr>
<tr>
<td>0/86</td>
<td>7/6</td>
<td>17/5</td>
</tr>
<tr>
<td>0/85</td>
<td>5/8</td>
<td>15/7</td>
</tr>
<tr>
<td>0/86</td>
<td>3</td>
<td>24/1</td>
</tr>
<tr>
<td>0/84</td>
<td>1/9</td>
<td>10/9</td>
</tr>
<tr>
<td>0/84</td>
<td>1/6</td>
<td>18/8</td>
</tr>
</tbody>
</table>

درک و ترکیب‌ها

- خردهآزمون‌ها
- جریان اول
- جریان دوم
- خردهآزمون‌ها

خلاصه‌ای از ضرایب اعتبار در جدول ۶، ارائه شده است. این جدول وضعیت آزمون را در ارتباط با دو منبع خطای المحتویاتی و زمانی نشان می‌دهد. میانگین ضرایب با استفاده از روش تبدیل Z فیشر (گیلمور و فروچر، ۱۹۷۸) و درصد آماری است.
جدول ۶- خلاصه‌ای از اعتبار آزمون در ارتباط با دو منبع خطای آزمون

<table>
<thead>
<tr>
<th>منابع خطای آزمون</th>
<th>نمونه‌برداری زمانی</th>
<th>متوسط‌ضایبی</th>
<th>نمونه‌برداری جملاتی</th>
<th>خرده‌آزمون‌ها</th>
</tr>
</thead>
<tbody>
<tr>
<td>۷۷</td>
<td>۷۸</td>
<td>۷۶</td>
<td>۸۱</td>
<td>واژگان تصویری</td>
</tr>
<tr>
<td>۸۴</td>
<td>۸۸</td>
<td>۸۹</td>
<td>۸۲</td>
<td>واژگان ربطی</td>
</tr>
<tr>
<td>۸۴</td>
<td>۸۸</td>
<td>۸۹</td>
<td>۹۰</td>
<td>واژگان شفاهی</td>
</tr>
<tr>
<td>۷۸</td>
<td>۸۲</td>
<td>۷۴</td>
<td>۸۸</td>
<td>درک دستوری</td>
</tr>
<tr>
<td>۸۸</td>
<td>۹۰</td>
<td>۹۰</td>
<td>۸۸</td>
<td>تقلید جمله</td>
</tr>
<tr>
<td>۸۳</td>
<td>۸۲</td>
<td>۸۲</td>
<td>۸۳</td>
<td>تکمیل دستوری</td>
</tr>
<tr>
<td>۸۷</td>
<td>۸۵</td>
<td>۸۷</td>
<td>۸۷</td>
<td>نمای‌گذاری کلمه</td>
</tr>
<tr>
<td>۸۹</td>
<td>۸۹</td>
<td>۹۴</td>
<td>۸۹</td>
<td>تحلیل راجح</td>
</tr>
<tr>
<td>۸۳</td>
<td>۸۲</td>
<td>۸۲</td>
<td>۸۲</td>
<td>نتلایه‌کلمه</td>
</tr>
<tr>
<td>۸۵</td>
<td>۸۸</td>
<td>۸۲</td>
<td>۸۵</td>
<td>کلمه‌کنترل‌گر</td>
</tr>
<tr>
<td>۸۸</td>
<td>۸۸</td>
<td>۹۲</td>
<td>۹۰</td>
<td>گوش‌کردن</td>
</tr>
<tr>
<td>۸۸</td>
<td>۸۸</td>
<td>۹۰</td>
<td>۸۵</td>
<td>سن‌بندی</td>
</tr>
<tr>
<td>۸۸</td>
<td>۸۲</td>
<td>۹۲</td>
<td>۸۸</td>
<td>صحبت کردن</td>
</tr>
<tr>
<td>۸۹</td>
<td>۷۷</td>
<td>۹۱</td>
<td>۸۹</td>
<td>معنی‌سازی</td>
</tr>
<tr>
<td>۹۲</td>
<td>۸۸</td>
<td>۹۶</td>
<td>۹۲</td>
<td>نحو</td>
</tr>
</tbody>
</table>

۴) اعتبار حذف شده است.

یادداشت‌ها

1) Test of Language Development
2) Test of Language Development- Primary
3) Test of Language Development - Intermediate
4) Composite score
5) Integrating- Mediating
6) Mediation
7) Associative
8) Pretryout
9) Tryout
10) Age equivalents
11) Extrapolation
12) Interpolation
13) Reliability