

## انطباق و هنجاریابی آزمون رشد زبان TOLD-P:3 برای کودکان فارسی زبان تهرانی\* (بخش دوم)

سعید حسن زاده\*\* اصغر مینایی\*\*

### چکیده

پژوهش حاضر با هدف انطباق، اعتباریابی، رواسازی و هنجاریابی آزمون رشد زبان، روی نمونه‌ای با حجم ۱۲۳۵ نفر (۶۲۶ پسر و ۶۰۹ دختر) از کودکان ۰-۴ (۴ سال تمام) تا ۸-۱۱ (۸ سال و ۱۱ ماه) شهر تهران که با روش نمونه‌برداری چند مرحله‌ای انتخاب شدند، اجرا گردید. آزمون TOLD-P:3 مشتمل بر ۹ خرده آزمون (۶ خرده آزمون اصلی و ۳ خرده آزمون تکمیلی) است. علاوه بر این، از ترکیب نمرات خرده آزمون‌های اصلی، ۶ نمره مرکب نیز به دست می‌آید. برای انطباق و هنجاریابی این آزمون ابتدا دستورالعمل اجرایی و گویه‌های آزمون به زبان فارسی ترجمه شد. آنگاه با توجه به ویژگی‌های نحوی، معنی شناختی و واج شناختی زبان فارسی و با مد نظر قرار دادن مبانی نظری آزمون، بسیاری از گویه‌ها مجدداً طراحی گردید. با اجرای مراحل پیش تجربی و تجزیه و تحلیل داده‌ها بر پایه مدل کلاسیک اعتبار، فرم نهایی آزمون شکل گرفت. ضریب همسانی درونی خرده آزمون‌ها با استفاده از فرمول آلفای کرونباخ و ثبات زمانی خرده آزمون‌ها و ترکیب‌ها با استفاده از روش آزمون - بازآزمون به دست آمد. ضرایب آلفا برای ترکیب‌ها (نمرات مرکب) با استفاده از فرمول گیلفورد که برای برآورد آلفای نمرات مرکب طراحی شده است، به دست آمد. روایی محتوایی، ملاکی و سازه خرده آزمون‌ها و ترکیب‌ها با استفاده از روش‌های مناسب مورد بررسی قرار گرفت. نتایج به دست آمده در زمینه اعتبار و روایی آزمون، اطمینان کافی برای استفاده از این ابزار را برای ارزیابی مهارت‌های زبانی کودکان فارسی زبان تهران به دست می‌دهد. ادامه این پژوهش در این شماره ارائه می‌گردد.

واژه‌های کلیدی: هنجاریابی، زبان دریافتی، زبان بیانی، اعتباریابی، رواسازی

\* این تحقیق در پژوهشکده کودکان استثنایی انجام شده است  
\*\* عضو هیات علمی پژوهشکده کودکان استثنایی

## ۲- روایی<sup>۱</sup> آزمون

مقصود از روایی آزمون پاسخ دادن به این پرسش است که آزمون چه چیزی را اندازه‌گیری می‌کند و تا چه اندازه از این لحاظ کارآیی دارد. در این مورد ذکر این نکته ضروری است که نام آزمون نمی‌تواند نشان دهد که آزمون چه چیزی را اندازه می‌گیرد. اسامی آزمون‌ها در واقع برچسب‌های کوتاه و مناسبی برای شناسایی آزمون‌ها هستند (آناستازی، ترجمه براهنی، ۱۳۶۴).

اغلب کارشناسانی که با اندازه‌گیری آموزشی و روان‌شناختی سر و کار دارند توصیه می‌کنند افرادی که دست‌اندرکار ساخت آزمون‌ها هستند بایستی شواهدی در زمینه سه نوع روایی ارائه دهند: روایی محتوا، روایی ملاک مرجع و روایی سازه (آیکن، ۱۹۹۴؛ آناستازی، ۱۹۸۸؛ گرونلند و لین، ۱۹۹۰؛ سلویا و یسلدیک، ۱۹۹۱؛ والاس، لارسن و الکسنین، ۱۹۹۲).

### ۲-۱ روایی محتوا

روایی محتوا به طور کلی متضمن بررسی منظم محتوای آزمون به منظور پاسخ دادن به این پرسش است که آیا آزمون نمونه‌های معرفی از حوزه رفتاری مورد اندازه‌گیری را در بر می‌گیرد یا نه (آناستازی، ۱۹۸۸، ص ۱۴۰).

در این جا دو دسته شاهد در زمینه روایی محتوایی این آزمون مورد بحث قرار می‌گیرد. ابتدا دلایل منطقی برای گویه‌های آزمون ارائه می‌دهیم. در گام بعدی، روایی گویه‌ها را با استفاده از نتایج شیوه‌های کلاسیک تحلیل گویه مورد بررسی قرار می‌دهیم.

#### ۲-۱-۱ مبنای منطقی گویه‌ها

در زمینه مبنای منطقی گویه‌ها سازندگان نسخه اصلی آزمون بر این نکته تأکید دارند که قالب مورد استفاده در خرده آزمون‌ها از سایر آزمون‌های رایج که چندین دهه با موفقیت مورد استفاده قرار گرفته‌اند، انتخاب شده‌اند و همین امر شاهی بر وجود روایی در آن‌هاست (نیوکامر و هامیل، ۱۹۹۷، ص ۶۴). در انطباق گویه‌های آزمون با توجه به نبود آزمون‌های زبانی استاندارد به زبان فارسی، ابتدا با در نظر گرفتن ویژگی‌های نحوی، معنی شناختی و واج شناختی زبان فارسی و همچنین رعایت کامل مبانی نظری و دلایل منطقی که سازندگان آزمون در تدوین خرده آزمون‌ها و انتخاب گویه‌ها در نظر داشته‌اند،

مخزنی از گویه‌ها برای هر خرده آزمون تهیه گردید و سپس با استفاده از روش‌های کلاسیک تحلیل گویه، گویه‌هایی که از ملاک‌های روان‌سنجی مطلوبی برخوردار بودند انتخاب شدند.

## ۲-۱-۲ تحلیل کلاسیک گویه‌ها

شیوه‌های کلاسیک انتخاب گویه بر قدرت تشخیص و درجه دشواری گویه تأکید دارند. برای تعیین قدرت تشخیص گویه (که گاهی اوقات روایی گویه نیز نامیده می‌شود) روش همبستگی دو رشته‌ای نقطه‌ای<sup>۲</sup> که در آن همبستگی هر گویه با نمره کل محاسبه می‌شود، مورد استفاده قرار گرفت. گارت (۱۹۶۵)؛ به نقل از نیوکامر و هامیل، ۱۹۹۷، ص ۷۲، عنوان کرده است که گویه‌های دارای ضریب همبستگی  $0/۲۰$  یا بالاتر در صورتی که آزمون نسبتاً طولانی باشد، می‌توانند دارای روایی تلقی شوند. در یک آزمون کوتاه لازم است که گویه‌ها از روایی بالاتری برخوردار باشند. به منظور تفسیر قدرت تشخیص گویه‌ها، اصول کلی تفسیر ضرایب روایی به کار گرفته شد. آناستازی (۱۹۸۸، ص ۲۰۳) توصیه می‌کند که ضرایب همبستگی  $0/۲۰$  یا  $0/۳۰$  می‌توانند قابل قبول تلقی گردند. به همین منظور مقدار  $0/۲۰$  به عنوان حداقل سطح قابل قبول برای گویه‌هایی که باید در فرم نهایی گنجانیده شوند انتخاب گردید.

درجه دشواری گویه<sup>۳</sup> (یعنی درصد آزمودنی‌هایی که به گویه پاسخ صحیح داده‌اند)، جهت تشخیص گویه‌های بسیار آسان یا بسیار دشوار و همچنین به منظور تعیین ترتیب گویه‌ها در فرم نهایی، محاسبه گردید. آناستازی (۱۹۸۸) می‌نویسد که دشواری متوسط گویه‌ها باید تقریباً  $0/۵۰$  باشد و همچنین باید پراکندگی نسبتاً زیادی در دشواری گویه‌ها وجود داشته باشد. گویه‌هایی با درجه دشواری بین  $0/۱۵$  تا  $0/۸۵$  عموماً قابل قبول در نظر گرفته می‌شوند.

براساس آماره‌های قدرت تشخیص و درجه دشواری، گویه‌های نامناسب از نسخه آزمایشی آزمون حذف شدند. گویه‌های انتخابی برای فرم نهایی از آسان به دشوار مرتب گردید. جدول ۷، خرده آزمون، تعداد گویه‌های هر خرده آزمون، تعداد گویه‌های حذف شده از فرم آزمایشی و همچنین دلیل حذف آن‌ها را نشان می‌دهد.

در نهایت برای نشان دادن خصوصیات گویه‌های نسخه نهایی، با استفاده از کل نمونه هنجاریابی یک تحلیل گویه صورت گرفت. ضرایب قدرت تشخیص و درجه دشواری به دست آمده برای گویه‌ها در جداول ۸ و ۹ ارائه شده است.

جدول ۷- تعداد گویه‌های فرم تجربی و نهایی

تعداد گویه‌های حذف شده (دشواری)	تعداد گویه‌های حذف شده (آسان)	تعداد گویه‌ها (فرم نهایی)	تعداد گویه‌ها (فرم تجربی)	خرده آزمون‌ها
۱۰	۲۰	۳۰	۶۰	واژگان تصویری
۵	۱۰	۳۰	۴۵	واژگان ربطی
۱۵	۱۷	۲۸	۶۰	واژگان شفاهی
۱۷	۱۸	۲۵	۶۰	درک دستوری
۳۰	۳۵	۳۰	۹۵	تقلید جمله
۲۱	۱۱	۲۸	۶۰	تکمیل دستوری
۲۰	۳۴	۲۶	۸۰	تمایزگذاری کلمه
۲۹	۱۷	۱۴	۶۰	تحلیل واجی
۶	۱۴	۲۰	۴۰	تولید کلمه

جدول ۸- میانه ضرایب تشخیص خرده آزمون‌ها در ۵ گروه سنی\*

سن					خرده آزمون‌ها
۸	۷	۶	۵	۴	
۳۳	۳۳	۲۵	۲۳	۳۲	واژگان تصویری
۴۷	۵۲	۴۶	۴۶	۴۲	واژگان ربطی
۴۷	۵۶	۴۶	۴۶	۴۵	واژگان شفاهی
۲۹	۲۸	۲۷	۲۷	۲۴	درک دستوری
۵۴	۵۴	۴۸	۴۸	۴۴	تقلید جمله
۳۷	۳۷	۲۸	۲۸	۳۴	تکمیل دستوری
۴۱	۴۵	۴۹	۴۹	۵۵	تمایزگذاری کلمه
۴۵	۳۹	۷۹	۷۸	۷۰	تحلیل واجی
۴۳	۴۲	۵۴	۵۴	۴۷	تولید کلمه

\*اعشار حذف شده است

جدول ۹- میانه درجه دشواری خرده‌آزمون‌ها در ۵ گروه سنی\*

سن						
۸	۷	۶	۵	۴	خرده‌آزمون‌ها	
۷۴	۷۱	۶۷	۶۸	۶۶	واژگان تصویری	
۶۰	۵۲	۳۹	۳۹	۳۲	واژگان ربطی	
۵۱	۴۵	۳۷	۳۷	۲۲	واژگان شفاهی	
۷۵	۷۱	۶۴	۶۴	۵۳	درک دستوری	
۶۵	۶۱	۳۹	۳۹	۲۵	تقلید جمله	
۶۵	۶۰	۴۲	۴۲	۳۲	تکمیل دستوری	
۹۵	۹۳	۸۴	۸۴	۷۷	تمایزگذاری کلمه	
۹۱	۸۴	۴۴	۴۴	۲۶	تحلیل واجی	
۹۴	۹۲	۹۰	۹۰	۸۸	تولید کلمه	

\*اعشار حذف شده است

## ۲-۲ روایی وابسته به ملاک<sup>۴</sup>

آناستازی (۱۹۸۸، ص ۱۴۵) روایی وابسته به ملاک را به عنوان «کارآمدی آزمون در پیش‌بینی رفتار یک فرد در فعالیت‌های خاص» توصیف می‌کند. وی اظهار می‌دارد که عملکرد فرد در آزمون با یک ملاک مقایسه می‌شود، ملاکی که خود نوعی شاخص مستقیم یا غیر مستقیم از آن چیزی است که آزمون برای پیش‌بینی آن طراحی شده است. بنابراین، اگر آزمونی مانند آزمون رشد زبان که فرض می‌شود توانایی زبانی را اندازه‌گیری می‌کند، دارای روایی باشد، بایستی با سایر آزمون‌های شناخته شده‌ای که تصور می‌رود همان توانایی را اندازه می‌گیرند همبستگی خوبی داشته باشد. از آن جا که در زبان فارسی آزمونی که بتواند همانند این آزمون، رشد زبان را در ابعاد مختلف اندازه بگیرد وجود نداشت، لذا توانستیم صرفاً روایی ملاکی خرده‌آزمون‌های تولید کلمه، تمایزگذاری کلمه، واژگان ربطی و واژگان شفاهی را به ترتیب با استفاده از آزمون اطلاعات آوایی، آزمون تشخیص شنوایی و پیمن، خرده‌آزمون شباهت‌ها و واژگان و کسلر (شهیم، ۱۳۷۳) بررسی نماییم.

در جدول ۱۰، ضرایب همبستگی بین خرده‌آزمون‌های انتخابی و آزمون‌های ملاک ارائه شده است.

جدول ۱۰- ضرایب همبستگی بین خرده آزمون‌های انتخابی و آزمون‌های ملاک

آزمون‌های ملاک				
خرده آزمون‌ها	خرده آزمون واژگان	خرده آزمون شایه‌ها	آزمون اطلاعات آوایی	آزمون ویمن
واژگان شفاهی	۰/۵۷			
واژگان ربطی		۰/۷۱		
تولید کلمه			۰/۴۲	
تمایزگذاری کلمه				۰/۷۰

ضرایب همبستگی به دست آمده، تماماً در سطح  $P < ۰/۰۱$  معنی‌دار هستند که این خود شاهدهی بر روایی وابسته به ملاک خرده‌آزمون‌های انتخابی است.

### ۳-۲- روایی سازه<sup>۵</sup>

روایی سازه با این مسأله سر و کار دارد که یک آزمون خاص تا چه اندازه سازه یا صفت بخصوصی را اندازه می‌گیرد. گرونلند و لین (۱۹۹۰) برای نشان دادن این نوع روایی یک روش سه مرحله‌ای را پیشنهاد می‌کنند. ابتدا چند سازه که فرض می‌شود عملکرد در آزمون را تبیین می‌کنند مشخص می‌شود. در مرحله دوم، براساس این سازه‌ها، فرضیه‌هایی شکل می‌گیرد و در مرحله سوم صحت و سقم این فرضیه‌ها با استفاده از شیوه‌های منطقی یا تجربی مورد بررسی قرار می‌گیرد. سازه‌های اساسی که فرض می‌شود زیر بنای عملکرد در این آزمون را تشکیل می‌دهند به صورت فرضیه‌هایی به صورت زیر مورد بررسی قرار گرفت:

- ۱- از آن جا که زبان محاوره‌ای، ماهیتاً وابسته به رشد است، لذا عملکرد در این آزمون باید با سن تقویمی همبستگی بالایی داشته باشد.
- ۲- از آن جا که این آزمون، زبان را اندازه می‌گیرد، بنا براین، نتایج آن باید بین گروه‌هایی از افراد که در توانایی‌های زبانی، متوسط به پایین شناخته شده‌اند، تمایز قائل شود.
- ۳- از آن جا که خرده آزمون‌ها، زبان را (به طرق مختلف) اندازه می‌گیرند، لذا بایستی همبستگی معنی‌داری با یکدیگر داشته باشند.

۴- با توجه به اینکه گویه‌های یک خرده آزمون خاص، ویژگی یکسانی را اندازه می‌گیرند، لذا گویه‌های هر خرده آزمون بایستی همبستگی بالایی با نمره کل خرده آزمون داشته باشند.

۵- از آن جا که هر یک از خرده آزمون‌ها بعد خاصی از زبان (اما نه متفاوت) را اندازه می‌گیرند، لذا عملکرد در آزمون، به مقدار قابل ملاحظه‌ای بایستی توسط یک عامل قابل تبیین باشد.

### ۱-۳-۲ تمایزگذاری سنی<sup>۶</sup>

در جدول ۱۱، میانگین‌ها و انحراف استانداردهای خرده آزمون‌ها در ۵ گروه سنی و همچنین ضرایب همبستگی عملکرد در خرده آزمون‌ها با سن ارائه گردیده است. محتوای جدول نشان می‌دهد که خرده آزمون‌ها با سن رابطه دارند، زیرا که با افزایش سن، میانگین خرده آزمون‌ها نیز افزایش می‌یابد. این نتیجه‌گیری با این واقعیت که تمامی ضرایب همبستگی ارائه شده در ستون سمت چپ جدول به لحاظ آماری معنی‌دار هستند، مورد تأیید قرار می‌گیرد. رابطه خرده آزمون تمایزگذاری کلمه و تولید کلمه با سن کمتر است، چرا که اغلب کودکان در ۷ سالگی به این قابلیت‌ها دست می‌یابند. این یافته‌ها فرضیه اول ما را درباره رابطه نمرات با سن تأیید می‌کند.

### ۲-۳-۲ تمایزگذاری گروهی<sup>۷</sup>

یکی از راه‌های ارزیابی روایی یک آزمون، مطالعه عملکرد گروه‌های مختلف در آزمون است. بایستی بین نتایج عملکرد گروه‌هایی که فرض می‌شود در خصیصه مورد اندازه‌گیری آزمون متفاوت هستند، اختلاف وجود داشته باشد. بنابراین، در مورد این آزمون انتظار می‌رود که عملکرد کودکان دارای تأخیر در گفتار و زبان، کم‌تر از کودکان مبتلا به اختلالات یادگیری و همچنین کودکان مبتلا به نقص توجه/ بیش‌فعالی<sup>۸</sup> در مقایسه با کودکان عادی ضعیف‌تر باشد. در واقع، آزمونی که نتواند بین این گروه‌ها و گروه‌های عادی تمایز قائل شود، هیچ‌گونه ارزش بالینی یا تشخیصی نخواهد داشت و در واقع فاقد روایی سازه است. میانگین‌های ارائه شده در جدول ۱۲، قویاً از روایی سازه آزمون مورد بحث حمایت می‌کنند.

جدول ۱۱ - میانگین‌ها و انحراف‌استانداردهای خرده‌آزمون‌ها در پنج گروه سنی و همبستگی آن‌ها با سن

همبستگی با سن	سن					خرده‌آزمون‌ها
	۸	۷	۶	۵	۴	
سن	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	
۰/۳۸	۲۳/۹۶(۳/۶۷)	۲۲/۲۸(۴/۳۵)	۲۱/۹۵(۳/۸۲)	۲۰/۳۸(۴/۰۲)	۱۸/۷۲(۵/۲۱)	واژگان تصویری
۰/۳۸	۱۷/۳۷(۶/۵۵)	۱۵/۴۰(۷/۲۴)	۱۲/۷۲(۶/۷۱)	۱۱/۷۲(۶/۴۲)	۹/۶۳(۵/۵)	واژگان ربطی
۰/۵۲	۱۷/۲۰(۶/۵۰)	۱۳/۷۱(۷/۳۵)	۱۲/۲۴(۶/۴۱)	۱۱/۷۲(۶/۴۲)	۸/۱۱(۵/۱۴)	واژگان شفاهی
۰/۵۰	۱۹/۳۵(۳/۴۸)	۱۷/۶۵(۳/۹۳)	۱۷/۱۳(۳/۳۹)	۱۵/۱۱(۴/۱۶)	۱۲/۶۱(۴/۲۷)	درک دستوری
۰/۴۷	۲۰/۱۹(۷/۲۶)	۱۷/۰۵(۷/۴۶)	۱۴/۳۶(۶/۷۶)	۱۳/۰۶(۶/۳۳)	۹/۰۵(۵/۵۳)	تقلید جمله
۰/۶۱	۲۰/۱۶(۴/۸۱)	۱۷/۲۳(۵/۴۱)	۱۴/۸۶(۴/۷۸)	۱۳/۰۶(۴/۶۶)	۹/۳۲(۴/۶۸)	تخمیل دستوری
۰/۳۵	۱۹/۶۰(۲/۷۱)	۱۸/۵۴(۳/۱۳)	۱۸/۲(۲/۸۲)	۱۶/۵۷(۴/۴۴)	۱۵/۴۱(۵/۴۱)	تمایزگذاری کلمه
۰/۶۰	۱۲/۶۲(۳/۰۶)	۱۱/۸۱(۳/۵۰)	۹/۶۵(۴/۷۷)	۶/۰۴(۵/۶۵)	۳/۵۶(۴/۴۶)	تحلیل واجی
۰/۲۷	۱۹/۱۱(۱/۸۴)	۱۸/۸۹(۱/۸۷)	۱۸/۴۷(۳/۴۸)	۱۷/۱۰(۳/۰۱)	۱۷/۱۰(۳/۲۶)	تولید کلمه

M= میانگین

SD= انحراف استاندارد

بررسی نتایج خرده‌آزمون‌ها و ترکیب‌های زیر گروه‌های فرعی نیز جالب است. فرض بر این است که زیر گروهی که بیشترین اختلال زبانی را دارد، کمترین بهره‌زبان گفتاری<sup>۹</sup> و زیر گروهی که اعضای آن کمترین مقدار اختلال زبانی را دارند، بالاترین بهره‌زبان گفتاری را در میان زیر گروه‌های فرعی به دست آورند. انتظار می‌رود که سایر زیر گروه‌ها در بین این دو حد قرار گیرند. براساس نتایج حاصل از تحقیقات صورت گرفته و نظریه‌های موجود در این زمینه، پیش‌بینی ما این است که زیر گروه کم‌توانان ذهنی، کمترین عملکرد را در خرده‌آزمون‌ها و ترکیب‌ها داشته باشند و سایر زیر گروه‌ها نیز به ترتیب شامل اختلالات زبانی، ناتوانی یادگیری و نقص توجه/بیش‌فعالی باشند. با نگاه به جدول ۱۲ می‌توان دریافت که به طور کلی نتایج به دست آمده، با نتایج پیش‌بینی شده هماهنگی و همخوانی دارند.

جدول ۱۲- میانگین نمرات استاندارد زیر گروه‌ها در خرده‌آزمون‌ها و ترکیب‌ها

خرده‌آزمون‌ها و ترکیب‌ها	کل نمونه N=۱۲۳۵	بصران N=۶۲۶	دختران N=۶۰۹	زیر گروه‌ها		
				اختلال در تذکیری N=۴۷	تاخیر در گفتار و زبان N=۳۶	کم‌توان دهنی N=۱۷
خرده‌آزمون‌ها						
واژگان تصویری	۱۰	۱۰	۱۰	۹	۸	۶
واژگان ربطی	۱۰	۱۰	۱۰	۸	۷	۸
واژگان شفاهی	۱۰	۹/۷	۱۰	۷	۸	۷
درک دستوری	۱۰	۹/۷	۱۰	۹	۸	۷
تقلید جمله	۱۰	۹/۷	۱۰	۸	۶	۵
تکمیل دستوری	۱۰	۹/۸	۱۰	۹	۸	۴
تمایزگذاری کلمه	۱۰	۹/۷	۹/۹	۹	۸	۷
تحلیل واجی	۱۰	۹/۷	۱۰	۹	۸	۷
تولید کلمه	۱۰	۹/۷	۱۰	۹	۶	۸
ترکیب‌ها						
گوش‌دادن	۱۰۰	۹۹	۱۰۰	۹۳	۷۹	۷۹
سازماندهی	۱۰۰	۹۸	۱۰۲	۸۶	۷۹	۷۹
صحبت کردن	۱۰۰	۹۹	۱۰۱	۹۱	۸۱	۷۹
معنی‌شناسی	۱۰۰	۹۹	۱۰۱	۸۷	۸۸	۸۰
نحو	۱۰۰	۹۸	۱۰۲	۹۳	۸۵	۸۳
زبان‌گفتاری	۱۰۰	۹۹	۱۰۱	۸۹	۸۶	۷۴

### ۳-۳-۲ روابط درونی خرده‌آزمون‌ها

همبستگی درونی نمرات خام خرده‌آزمون‌ها، با استفاده از کل نمونه هنجاریابی محاسبه گردید. به منظور کنترل تأثیر سن از روش همبستگی تفکیکی<sup>۱۰</sup> استفاده شد. ضرایب همبستگی تفکیکی به دست آمده در جدول ۱۳ ارائه شده است. ضرایب همبستگی ارائه شده در این جدول به سه بخش قابل تقسیم است: ضرایبی که روابط درونی خرده‌آزمون‌های اصلی را با دامنه‌ای از ۰/۴۴ تا ۰/۷۹ با میانه ۰/۵۵ نشان می‌دهند که این مقدار به لحاظ آماری معنی‌دار است. از آن جا که این شش خرده‌آزمون همبستگی درونی مطلوبی با یکدیگر دارند، تصور می‌رود که تمامی آن‌ها جنبه‌های مرتبطی از یک سازه (زبان‌گفتاری) را اندازه می‌گیرند؛ ضرایبی که روابط میان سه خرده‌آزمون تکمیلی را با شش خرده‌آزمون اصلی نشان می‌دهند. این ضرایب دارای دامنه‌ای از ۰/۱۸ تا ۰/۴۶ و میانه ۰/۳۱ می‌باشند که این مقدار، در مقایسه با روابط درونی خرده

آزمون‌های اصلی اندک است. از این رو این سه خرده آزمون در طبقه خرده آزمون‌های تکمیلی جایگزین می‌شوند.

مقدار ضرابی که روابط درونی سه خرده آزمون تکمیلی را نشان می‌دهند در سطح پایینی قرار دارد و این امر نشانگر این است که نبایستی این خرده‌آزمون‌ها را برای تشکیل یک نمره مرکب با یکدیگر ترکیب نمود.

جدول ۱۳- همبستگی درونی خرده‌آزمون‌ها برای کل نمونه هنجاریابی\*

خرده آزمون‌های تکمیلی			خرده آزمون‌های اصلی					
تولید کلمه	تحلیل واجی	نمایر گذاری کلمه	تکمیل دستوری	تقلید جمله	درک دستوری	واژگان شفاهی	واژگان ربطی	واژگان تصویری
<b>خرده‌آزمون‌ها اصلی</b>								
۲۱	۳۲	۳۰	۵۵	۴۶	۵۵	۵۵	۵۲	—
۲۶	۳۵	۱۸	۵۹	۴۵	۴۴	۷۹	—	واژگان تصویری
۲۶	۳۹	۲۶	۶۳	۵۲	۵۱	—	—	واژگان ربطی
۲۱	۴۶	۳۸	۵۶	۴۶	—	—	—	واژگان شفاهی
۳۱	۴۲	۳۱	۶۴	—	—	—	—	درک دستوری
۲۵	۴۶	۳۶	—	—	—	—	—	تقلید جمله
—	—	—	—	—	—	—	—	تکمیل دستوری
<b>تکمیلی</b>								
۲۰	۳۳	—	—	—	—	—	—	نمایر گذاری کلمه
۲۴	—	—	—	—	—	—	—	تحلیل واجی
—	—	—	—	—	—	—	—	تولید کلمه

\*اعشار حذف شده است

#### ۴-۳-۲ تحلیل عاملی<sup>۱۱</sup>

روایی سازه، همچنین به میزان قابلیت تشخیص صفات زیر بنایی آزمون و این که این صفات تا چه اندازه، مدل نظری را که آزمون بر آن مبتنی است، منعکس می‌کنند، مرتبط است. برای بررسی این نوع روایی، عملکرد افراد گروه هنجار در خرده آزمون‌های اصلی با استفاده از روش مؤلفه‌های اصلی<sup>۱۲</sup> مورد تحلیل قرار گرفت (کلاین، ترجمه صدر السادات و مینایی، ۱۳۸۰). از آن جا که هر یک از این شش خرده‌آزمون، جنبه

خاصی (اما نه متفاوتی) از زبان را اندازه می‌گیرند، انتظار داشتیم تمام این خرده آزمون‌ها تنها در یک عامل که توانایی کلی زبان گفتاری را اندازه می‌گیرد، بار شوند. نتیجه تحلیل عاملی در جدول ۱۴ ارائه شده است. نتیجه به دست آمده نشان می‌دهد که تمام این آزمون‌ها در یک عامل بار می‌شوند. این عامل کلی را می‌توان بهره‌زبان گفتاری نامید.

جدول ۱۴- تحلیل عاملی خرده آزمون‌های اصلی

خرده آزمون‌ها	عامل
واژگان تصویری	۰/۶۴
واژگان ربطی	۰/۶۹
واژگان شفاهی	۰/۷۷
درک دستوری	۰/۶۵
تقلید جمله	۰/۶۴
تکمیل دستوری	۰/۷۹
ارزش ویژه <sup>۱۳</sup>	۴/۱۹
واریانس کل تبیین شده	۰/۶۸

### ۵-۳-۲ روایی سؤال

گیلفورد و فروچر (۱۹۷۸) اظهار می‌دارند که از طریق همبسته نمودن عملکرد افراد در گویه‌ها با نمره کل مقیاس می‌توان اطلاعاتی در زمینه روایی سازه مقیاس به دست آورد. ضرایب همبستگی که در جدول ۸ ارائه شده است قویاً دلالت بر روایی سازه دارند.

### ۳- قدرت تمایزگذاری<sup>۱۴</sup>

آخرین شاهد بر روایی که مورد بررسی قرار دادیم، قدرت تمایزگذاری است. این شاخص که به واریانس یا پراکندگی نمرات مربوط است، زمانی به حداکثر می‌رسد که توزیع نمرات به شکل مستطیلی باشد. حداقل تمایزگذاری (صفر) زمانی به دست خواهد آمد که تمام افراد نمره یکسانی دریافت کنند. شاخص قدرت تمایزگذاری آزمون، دلتای فرگسون است که مقدار آن بین ۱ تا ۰ تغییر می‌کند (فرگسون، ۱۹۴۹؛ به نقل از کلاین، ۲۰۰۰). مقدار این شاخص در توزیع‌های نرمال ۰/۹۳ است. به طور کلی آزمون روا، آزمونی

است که قدرت تمایزگذاری آن بالاتر از ۰/۹۰ باشد. در جدول ۱۵، ضرایب قدرت تمایزگذاری خرده آزمون‌های اصلی و تکمیلی ارائه شده است.

جدول ۱۵- ضرایب قدرت تمایزگذاری خرده آزمون‌های اصلی و تکمیلی

خرده آزمون‌ها	δ (دلتا)
اصلی	
واژگان تصویری	۰/۹۵
واژگان ربطی	۰/۹۷
واژگان شفاهی	۰/۹۲
درک دستوری	۰/۹۷
تقلید جمله	۰/۹۰
تکمیل دستوری	۰/۹۲
تکمیلی	
تمایزگذاری کلمه	۰/۸۰
تحلیل واجی	۰/۸۷
تولید کلمه	۰/۷۹

ضرایب به دست آمده برای خرده آزمون‌های اصلی همگی در سطح بالایی قرار دارند. لیکن ضرایب خرده آزمون‌های تکمیلی در مقایسه با ملاک ارائه شده توسط کلاین، در سطح نسبتاً پایینی قرار دارند چرا که اغلب کودکان در ۷ سالگی به این قابلیت‌ها دست می‌یابند و در واقع واریانس نمرات کاهش می‌یابد.

### بحث و نتیجه‌گیری

بر پایه یافته‌های این پژوهش موارد زیر را می‌توان استنباط کرد:

- ۱- نتایج حاصل از بررسی اعتبار آزمون نشان می‌دهد که آزمون از اعتبار مطلوبی برخوردار است. به بیان دیگر ضرایب اعتبار آزمون به اندازه‌ای است که کاربران بتوانند به نتایج آن در اندازه‌گیری‌ها اطمینان داشته باشند.
- ۲- براساس نتایج حاصل از بررسی روایی آزمون، می‌توان نتیجه گرفت که این آزمون یک ابزار روا، در سنجش زبان است و می‌توان با اطمینان خاطر از آن استفاده کرد.

۳- نتیجه حاصل از اجرای تحلیل مؤلفه‌های اصلی نشان می‌دهد که کلیه خرده آزمون‌های اصلی در یک عامل باردارند. به عبارت دیگر می‌توان گفت که این آزمون از یک عامل کلی اشباع شده است. این عامل کلی را می‌توان بهره‌زبان گفتاری نامید. این نتیجه با یافته نیوکامر و هامیل (۱۹۹۷) همخوانی دارد.

به‌طور خلاصه، نتایج حاصل از بررسی ویژگی‌های روان‌سنجی آزمون حاکی از معتبر بودن و روا بودن آزمون است. لذا از این آزمون می‌توان به‌عنوان ابزاری برای الف) شناسایی کودکانی که به‌طور قابل ملاحظه‌ای در مهارت‌های زبانی، ضعیف‌تر از همسالان خود می‌باشند، ب) مشخص کردن نقاط قوت و ضعف کودکان در مهارت‌های زبانی، ج) مستند کردن پیشرفت زبان کودکان در نتیجه برنامه‌های مداخله‌ای ویژه و د) اندازه‌گیری زبان در مطالعات پژوهشی، استفاده کرد.



#### یادداشت‌ها

- |                             |  |
|-----------------------------|--|
| 1) Validity                 | 2) Point biserial correlation                |
| 3) Item difficulty          | 4) Criterion- related validity               |
| 5) Construct validity       | 6) Age discrimination                        |
| 7) Group discrimination     | 8) Attention- deficit hyperactivity disorder |
| 9) Spoken language quotient | 10) Partial correlation                      |
| 11) Factor analysis         | 12) Principal component                      |
| 13) Eigen value             | 14) Discrimination power                     |

#### منابع

- آناستازی، آنا (۱۳۶۴). *روان‌آزمایی*. مترجم: محمد نقی براهنی. تهران: انتشارات دانشگاه تهران.
- شهیم، سیما (۱۳۷۳). *انطباق و هنجاریابی مقیاس تجدید نظر شده هوشی وکسلر برای کودکان*. انتشارات دانشگاه شیراز.
- کلاین، پل (۱۳۸۰). *راهنمای آسان تحلیل عاملی*. مترجمان: سید جلال صدرالسادات، اصغر مینایی. تهران: سمت.

- Aiken, L. R. (1994). *Psychological testing and assessment*. Needham Heights, MA: Allyn & Bacon.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- Angoff, W. H., & Robertson, G. J. (1987). A Procedure for standardizing individually administered tests, normed by age or grade level. *Applied psychological measurement*, 11, 33- 46.
- Bates, E. (1976). *Language and context: The acquisition of pragmatics*. New York: Academic Press.
- Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York: Wiley.
- Bloom L., (1983). Of continuity, nature, and magic. In R. Golinkoff (Ed.), *The transition from preverbal to verbal communication*. Hillsdale, NJ: Erlbaum.
- Bloom L., Tackeff, J., & Lahey, M. (1984). Learning to speak in complement constructions. *Journal of Child Language*, 11, 391- 406.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Case, R. (1978). Intellectual development from birth to adulthood: A neo- Piagetian interpretation. In R. Siegler (Ed.), *Children's thinking: What develops?* (PP. 37-72). Hillsdale, NJ: Erlbaum.
- Chapman, R. (1974). Discussion summary- Developmental relationships between receptive and expressive language. In R. L. Schiefelbusch & L. L. Lloyd (Eds.), *Language perspectives- Acquisition, retardation, and intervention* (PP. 335- 334). Austin, TX: PRO-ED.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1968). *Language and mind*. New York: Harcourt, Brace & World.
- Chomsky, N. (1969). *Deep structure, surface structure, and semantic interpretation*. Bloomington: Indiana University Linguistics Club.

- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Holland: Foris.
- DeVito, J. A. (1970). *The psychology of speech and language*. New York: Random House.
- Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York: Macmillan.
- Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw- Hill.
- Guilford, J. P. & Fruchter, B. (1978). *Fundamental statistics in psychology and education*. New York: McGraw- Hill.
- Ingram, D. (1974). The relationship between comprehension and production. In R. L. Schiefelbusch & L. L. Lloyd (Eds.), *Language perspectives- Acquisition, retardation, and intervention*. Austin, TX: PRO-ED.
- Jakobson, R., Fant, C., & Halle, M. (1963). *Preliminaries to speech analysis*. Cambridge, MA: MIT Press.
- Kline, P. (2000). *The handbook of psychological testing* (2nd ed.) Routledge.
- Leonard, L. & Loeb, D. (1988). Government binding theory and some of its applications: A tutorial. *Journal of Speech and Hearing Research*, 31, 515- 524.
- Lieberman, I., Shankweiler, D., & Liberman, A. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (PP. 1-33). Ann Arbor: University of Michigan Press.
- Markman, E. M. (1981). Comprehension monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills* (PP. 61- 84). New York: Academic Press.
- Newcomer, P. L., & Hammill, D. D. (1997). *Test of language development- primary*. (3rd ed.) Austin, TX: PROED.
- Ninio, A., & Snow, C. (1988). Language acquisition through language use: The functional sources of children's early utterances. In Y. Levy, I. Schlesinger, & M. Braine (Eds.), *Categories and processes in language acquisition* (PP. 196- 219). Hillsdale, NJ: Erlbaum.
- Nippold, M. (1988). *Later language development: Ages nine through nineteen*. Austin, TX: PRO- ED.
- Owens, R. (1995). *Language disorders: A functional approach to assessment and intervention*.

- Needham Heights, MA: Allyn & Bacon.
- Rabinowitz, M., & Glaser, R. (1985). Cognitive structure and process in highly competent performance. In F. D. Horowitz, & M. O'Brien (Eds.), *The gifted and talented: Developmental perspectives* (PP.rj 75-98). Washington, DC: American Psychological Association.
- Rees, N. (1978). Pragmatics of language. In R. Schiefelbusch (Ed.), *Bases of language intervention* (PP. 3-22). Baltimore: University Park Press.
- Salvia, J., & Ysseldyke, J. E. (1995). *Assessment* (6th ed.). Boston: Houghton Mifflin.
- Schlesinger, I. (1977). The role of cognitive development and linguistic input in language acquisition. *Journal of Child Language*, 4, 153- 169.
- Vygotsky, L. S. (1977). *Thought and language*. Cambridge, MA: MIT Press.
- Wallace, G., Larsen, S. C. & Elksnin, L. (1992). *Educational assessment of learning problems*(2nd ed.). Needham Heights, MA: Allyn & Bacon.
- Wiig, E. (1989). *Steps to language competence: Developmental metalinguistic strategies*. San Antonio: Psychological Corp.