پژوهش در حیطه کودکان استثنایی۲/ سال اول، شماره۲/ زمستان ۱۳۸۰/۱۳٤-۱۱۹
Research on Exceptional Children, Winter 2002, Vol.1, No.2, 119-134

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 ۱۱-۸ (۸ سال و ۱۱ماه) 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 نمرهٔ 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 پایهٔ 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 سازهٔ خرده آزمون‌ها و ترکیب‌ها با استفاده از روش‌های مناسب مورد بررسی قرار گرفت. نتایج به دست آمده در زمینهٔ اعتبار و روایی آزمون، اطمینان کافی برای استفاده از این ابزار را برای ارزیابی مهارت‌های زبانی کودکان فارسی زبان تهران به دست می‌دهد. ادامهٔ 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 هیأت علمی پژوهشکده کودکان استثنایی

[Downloaded from joec.ir on 2026-10-30]
[DOR: 20.1001.1.16826612.1380.1.2.1.1]

۲- روایی[1] 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، ترجمهٔ 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 زمینهٔ 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 زمینهٔ 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 زمینهٔ مبنای منطقی گویه‌ها سازندگان نسخهٔ 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 شاهدی 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،

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

مخـزنی از گویـه‌هـا بـرای هـر خـرده آزمـون تهیه گردید و سپس با استفاده از روش‌های کلاسـیک تحـلیل گویـه، گویـه‌هایی که از ملاک‌های روان سنجی مطلوبی برخوردار بودند انتخاب شدند.

۲-۱-۲ تحلیل کلاسیک گویه‌ها

شیوه‌های کلاسیک انتخاب گویه بر قدرت تشخیص و درجهٔ دشواری گویه تأکید دارند. بـرای تعییـن قدرت تشخیص گویه (که گاهی اوقات روایی گویه نیز نامیده می‌شود) روش همبسـتگی دو رشـته‌ای نقطـه‌ای[۲] کـه در آن همبسـتگی هـر گویـه بـا نمرهٔ کل محاسبه مـی‌شـود، مـورد استفاده قرار گرفت. گارت(۱۹۶۵؛ به نقل از نیوکامر و هامیل، ۱۹۹۷، ص ۷۲)، عنوان کرده است که گویه‌های دارای ضریب همبستگی ۰/۲۰ یا بالاتر در صورتی که آزمـون نسـبتاً طولانـی باشـد، می‌توانند دارای روایی تلقی شوند. در یک آزمون کوتاه لازم اسـت کـه گویـه‌هـا از روایـی بالاتـری بـرخوردار باشـند. به منظور تفسیر قدرت تشخیص گویـه‌هـا، اصـول کـلی تفسیر ضرایب روایی به‌کار گرفته شد. آناستازی (۱۹۸۸، ص ۲۰۳) توصـیه مـی‌کند که ضرایب همبستگی ۰/۲۰ یا ۰/۳۰ می‌توانند قابل قبول تلقی گردند. به همیـن مـنظور مقـدار ۰/۲۰ به عنوان حداقل سطح قابل قبول برای گویه‌هایی که باید در فرم نهایی گنجانیده شوند انتخاب گردید.

درجـهٔ دشـواری گویـه [۳] (یعنی درصد آزمودنی‌هایی که به گویه پاسخ صحیح داده‌اند)، جهـت تشـخیص گویه‌های بسیار آسان یا بسیار دشوار و همچنین به منظور تعیین ترتیب گویـه‌هـا در فـرم نهایی، محاسبه گردید. آناستازی (۱۹۸۸) می‌نویسد که دشواری متوسط گویـه‌ها باید تقریباً ۰/۵۰ باشد و همچنین باید پراکندگی نسبتاً زیادی در دشواری گویه‌ها وجـود داشـته باشـد. گویـه‌هایی با درجهٔ دشواری بین ۰/۱۵ تا ۰/۸۵ عموماً قابل قبول در نظر گرفته می‌شوند.

براسـاس آمـاره‌هـای قـدرت تشـخیص و درجـهٔ دشـواری، گویه‌های نامناسب از نسخهٔ آزمایشـی آزمـون حـذف شدند. گویه‌های انتخابی برای فرم نهایی از آسان به دشوار مرتب گـردید. جـدول ۷، خـرده آزمـون، تعداد گویه‌های هر خرده آزمون، تعداد گویه‌های حذف شده از فرم آزمایشی و همچنین دلیل حذف آن‌ها را نشان می‌دهد.

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

در نهـایت بـرای نشان دادن خصوصیات گویه‌های نسخهٔ نهایی، با استفاده از کل نمونهٔ هـنجاریابی یـک تحـلیل گویـه صـورت گـرفت. ضـرایب قدرت تشخیص و درجهٔ دشواری به‌دست آمده برای گویه‌ها در جداول ۸ و ۹ ارائه شده است.

جدول ۷- تعداد گویه‌های فرم تجربی و نهایی

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

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

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

*اعشار حذف شده است

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

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

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

*اعشار حذف شده است

۲-۲ روایی وابسته به ملاک[۴]

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 خرده آزمون‌های تولید کلمه، تمایزگذاری کلمه، واژگان ربطی و واژگان شفاهی را به ترتیب با استفاده از آزمون اطلاعات آوایی، آزمون تشخیص شنوایی ویپمن، خرده آزمون شباهت‌ها و واژگان وکسلر (شهیم، ۱۳۷۳) بررسی نماییم.

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

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

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

خرده آزمون‌ها	آزمون‌های ملاک: خرده‌آزمون واژگان	خرده‌آزمون شباهتها	آزمون اطلاعات آوایی	آزمون ویپمن
واژگان شفاهی	۰/۵۷			
واژگان ربطی		۰/۷۱		
تولید کلمه			۰/۴۲	
تمایزگذاری کلمه				۰/۷۰

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

۳-۲- روایی سازه[۵]

روایی سازه با این مسأله سر و کار دارد که یک آزمون خاص تا چه اندازه سازه یا صفت بخصوصـی را انـدازه می‌گیرد. گرونلند و لین (۱۹۹۰) برای نشان دادن این نوع روایی یک روش سـه مـرحله‌ای را پیشـنهاد مـی‌کنـند. ابتدا چند سازه که فرض می‌شود عملکرد در آزمـون را تـبیین مـی‌کنـند مشـخص مـی‌شـود . در مرحـلهٔ دوم، براسـاس ایـن سازه‌ها، فرضـیه‌هـایی شـکل مـی‌گیـرد و در مرحلهٔ سوم صحت و سقم این فرضیه‌ها با استفاده از شـیوه‌های منطقی یا تجربی مورد بررسی قرار می‌گیرد. سازه‌های اساسی که فرض می‌شود زیـر بـنای عمـلکرد در این آزمون را تشکیل می‌دهند به صورت فرضیه‌هایی به صورت زیر مورد بررسی قرار گرفت:

۱ - از آن جـا کـه زبـان محاوره‌ای، ماهیتاً وابسته به رشد است، لذا عملکرد در این آزمون باید با سن تقویمی همبستگی بالایی داشته باشد.

۲- از آن جـا 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 جـا کـه خـرده آزمـون‌هـا، زبان را (به طرق مختلف) اندازه می‌گیرند، لذا بایستی همبستگی معنی‌داری با یکدیگر داشته باشند.

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

۴- بـا توجه به اینکه گویه‌های یک خرده آزمون خاص، ویژگی یکسانی را اندازه می‌گیرند، لـذا گویـه‌هـای هـر خرده آزمون بایستی همبستگی بالایی با نمرهٔ کل خرده آزمون داشته باشند.

۵- از آن جـا کـه هـر یـک از خـرده آزمون‌ها بعد خاصی از زبان (اما نه متفاوت) را اندازه مـی‌گیـرند، لذا عملکرد در آزمون، به مقدار قابل ملاحظه‌ای بایستی توسط یک عامل قابل تبیین باشد.

۱-۳-۲ تمایزگذاری سنی[۶]

در جـدول ۱۱، میـانگین‌هـا و انحـراف اسـتانداردهای خرده آزمون‌هادر ۵ گروه سنی و همچـنین ضرایب همبستگی عملکرد در خرده آزمون‌ها با سن ارائه گردیده است. محتوای جـدول نشـان مـی‌دهـد کـه خـرده آزمـون‌هـا با سن رابطه دارند، زیرا که با افزایش سن، میـانگین خـرده آزمـون‌هـا نیـز افزایش می‌یابد. این نتیجه‌گیری با این واقعیت که تمامی ضرایب همبستگی ارائه شده در ستون سمت چپ جدول به لحاظ آماری معنی‌دار هستند، مـورد تـأیید قرار می‌گیرد. رابطه خرده‌آزمون تمایزگذاری کلمه و تولید کلمه با سن کمتر اسـت، چـرا کـه اغـلب کودکـان در ۷ سالگی به این قابلیت‌ها دست می‌یابند. این یافته‌ها فرضیهٔ اول ما را دربارهٔ رابطهٔ نمرات با سن تأیید می‌کند.

۲-۳-۲ تمایزگذاری گروهی[۷]

یکـی از راه‌هـای ارزیابی روایی یک آزمون، مطالعهٔ عملکرد گروه‌های مختلف در آزمون است. بایستی بین نتایج عملکرد گروه‌هایی که فرض می‌شود در خصیصهٔ مورد اندازه‌گیری آزمـون مـتفاوت هسـتند، اخـتلاف وجـود داشته‌باشد. بنابراین، در مورد این آزمون انتظار مـی‌رود که عملکرد کودکان دارای تأخیر در گفتار و زبان، کم توانان ذهنی، کودکان مبتلا بـه اختلالات یادگیری و همچنین کودکان مبتلا به نقص توجه/ بیش فعالی[۸] در مقایسه با کودکان عادی ضعیف‌تر باشد. در واقع، آزمونی که نتواند بین این گروه‌ها و گروه‌های عادی تمایـز قـائل شود، هیچ‌گونه ارزش بالینی یا تشخیصی نخواهد داشت و در واقع فاقد روایی سـازه اسـت. میـانگین‌هـای ارائه شده در جدول ۱۲، قویاً از روایی سازه آزمون مورد بحث حمایت می‌کنند.

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

جدول ۱۱- میانگین‌هاوانحراف‌استانداردهای‌خرده‌آزمون‌هادرپنج‌گروه سنی و همبستگی آن‌ها با سن

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

M= میانگین

SD= انحراف استاندارد

بررسی نتایج خرده آزمون‌ها و ترکیب‌های زیر گروه‌های فرعی نیز جالب است. فرض بر این است که زیر گروهی که بیشترین اختلال زبانی را دارد، کمترین بهرهٔ زبان گفتاری[۹] و زیر گروهی که اعضای آن کمترین مقدار اختلال زبان را دارند، بالاترین بهرهٔ زبان 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.

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

جدول۱۲- میانگین نمرات استاندارد زیر گروه‌ها در خرده‌آزمون‌ها و ترکیب‌ها

خرده‌آزمون‌ها و ترکیب‌ها	زیر گروه‌ها						
	کل نمونه N=۱۲۳۵	پسران N=۶۲۶	دختران N=۶۰۹	اختلال در یادگیری N=۴۷	تأخیر در گفتاروزبان N=۲۶	کم‌توان ذهنی N=۱۷	نقص توجه/ بیش فعالی N=۸
خرده آزمون‌ها							
واژگان تصویری	۱۰	۱۰	۱۰	۹	۸	۶	۸
واژگان ربطی	۱۰	۱۰	۱۰	۸	۷	۸	۹
واژگان شفاهی	۱۰	۹/۷	۱۰	۷	۸	۷	۹
درک دستوری	۱۰	۹/۷	۱۰	۹	۸	۷	۶
تقلید جمله	۱۰	۹/۷	۱۰	۸	۶	۵	۶
تکمیل دستوری	۱۰	۹/۸	۱۰	۹	۸	۴	۵
تمایزگذاری کلمه	۱۰	۹/۷	۹/۹	۹	۸	۷	۸
تحلیل واجی	۱۰	۹/۷	۱۰	۹	۸	۷	۷
تولید کلمه	۱۰	۹/۷	۱۰	۹	۶	۸	۹
ترکیب‌ها							
گوش‌دادن	۱۰۰	۹۹	۱۰۰	۹۳	۷۹	۷۹	۸۴
سازماندهی	۱۰۰	۹۸	۱۰۲	۸۶	۷۹	۷۹	۸۵
صحبت کردن	۱۰۰	۹۹	۱۰۱	۹۱	۸۱	۷۹	۹۰
معنی شناسی	۱۰۰	۹۹	۱۰۱	۸۷	۸۸	۸۰	۹۱
نحو	۱۰۰	۹۸	۱۰۲	۹۳	۸۵	۸۳	۸۳
زبان گفتاری	۱۰۰	۹۹	۱۰۱	۸۹	۸۶	۷۴	۹۰

۳-۳-۲ روابط درونی خرده آزمون‌ها

همبستگی درونی نمرات خام خرده‌آزمون‌ها، با استفاده از کل نمونهٔ هنجاریابی محاسبه گردید. به منظور کنترل تأثیر سن از روش همبستگی‌تفکیکی[۱۰] 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 خرده آزمون همبستگی درونی مطلوبی با یکدیگر دارند، تصور می‌رود که تمامی آن‌ها جنبه‌های مرتبطی از یک سازه (زبان گفتاری) را اندازه می‌گیرند؛ ضرایبی که روابط میان سه خرده آزمون تکمیلی را با شش خرده آزمون 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

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

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

مقدار ضرایبی 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 نمرهٔ مرکب با یکدیگر ترکیب نمود.

جدول۱۳- همبستگی درونی خرده‌آزمون‌ها برای کل نمونهٔ هنجاریابی*

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

***اعشار حذف شده است**

٤-۳-۲ تحلیل عاملی[11]

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[12] مورد تحلیل قرار گرفت (کلاین، ترجمه صدر السادات و مینایی، ۱۳۸۰). از آن‌جا که هر یک از این شش خرده‌آزمون، جنبهٔ

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

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. نتیجهٔ تحلیل عاملی در جدول ۱۴ ارائه شده است. نتیجهٔ به دست آمده نشان می‌دهد که تمام این آزمون‌ها در یک عامل بار می‌شوند. این عامل کلی را می‌توان بهرهٔ زبان گفتاری نامید.

جدول۱۴- 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[۱۳]	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 نمرهٔ کل مقیاس می‌توان اطلاعاتی در زمینهٔ روایی سازه مقیاس به دست آورد. ضرایب همبستگی که در جدول ۸ ارائه شده است قویاً دلالت بر روایی سازه دارند.

۳- قدرت تمایزگذاری[۱۴]

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 نمرهٔ 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 طورکلی آزمون روا، آزمونی

[Downloaded from joec.ir on 2026-10-30]
[DOR: 20.1001.1.16826612.1380.1.2.1.1]

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

جدول۱۵- ضرایب‌قدرت‌تمایزگذاری‌خرده 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 پایهٔ 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 آزمون‌یک‌ابزار روا، در سنجش زبان است و می‌توان با اطمینان خاطر از آن استفاده کرد.

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

۳- نتیجهٔ حاصل از اجرای تحلیل مؤلفه‌های اصلی نشان می‌دهد که کلیهٔ 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 بهرهٔ زبان گفتاری نامید. این نتیجه با یافتهٔ نیوکامر و هامیل(۱۹۹۷) همخوانی دارد.

به‌طور خلاصه، نتایج حاصل از بررسی ویژگی‌های روان سنجی 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 نتیجهٔ 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 (۱۳٦٤). *روان آزمایی*. مترجم: محمد نقی براهنی . تهران: انتشارات دانشگاه تهران.

شهیم، سیما (۱۳۷۳). *انطباق و هنجاریابی مقیاس تجدید نظر شدهٔ هوشی وکسلر برای کودکان*. انتشارات دانشگاه شیراز.

کلاین ، پل (۱۳۸۰) . *راهنمای آسان تحلیل عاملی*. مترجمان: سید جلال صدرالسادات، اصغر مینایی. تهران: سمت.

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

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.

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

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.

Liberman, 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 commu nication 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.*

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]

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 perspectivs* (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). *Tbought 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 Antionio: Psychological Corp.

[Downloaded from joec.ir on 2026-10-30]

[DOR: 20.1001.1.16826612.1380.1.2.1.1]