

# تدوین نسخه فارسی آزمون غربالگر سنجش تکامل کودکان «بیلی» و تعیین اعتبار و روایی آن

فرین سلیمانی<sup>۱</sup>، نادیا آذری<sup>۲</sup>، آدیس کراسکیان موجمباری<sup>۳</sup>، روشنگ وامقی<sup>۴</sup>، سهیلا شهشهانی پور<sup>۵</sup>، فیروزه ساجدی<sup>۱</sup>

- ۱- متخصص کودکان، دانشیار مرکز تحقیقات توانبخشی اعصاب اطفال، دانشگاه علوم بهزیستی و توانبخشی، تهران، ایران
- ۲- متخصص کودکان، مرکز تحقیقات توانبخشی اعصاب اطفال، دانشگاه علوم بهزیستی و توانبخشی، تهران، ایران
- ۳- دکترای آمار زیستی، استادیار گروه مشاوره و راهنمایی، دانشگاه آزاد اسلامی، واحد کرج، البرز، ایران
- ۴- متخصص کودکان، استادیار مرکز تحقیقات توانبخشی اعصاب اطفال، دانشگاه علوم بهزیستی و توانبخشی، تهران، ایران

دریافت مقاله: ۹۲/۰۹/۲۰

پذیرش مقاله: ۹۲/۱۲/۰۱

\* آدرس نویسنده مسئول:

تهران، اوین، بلوار دانشجو، خیابان کودکان، دانشگاه علوم بهزیستی و توانبخشی، مرکز تحقیقات توانبخشی اعصاب اطفال.

\* تلفن: ۰۰۹۹ ۲۲۱۸۰۰ (۲۱) ۹۸+

\* رایانامه:

azarinadia@yahoo.com

## چکیده

**هدف:** این مطالعه به منظور تدوین نسخه فارسی روا و دارای اعتبار آزمون غربالگر سنجش تکامل کودکان بیلی انجام گرفت.

**روش بررسی:** مراحل ترجمه، باز ترجمه، تعیین روایی صوری و محتوایی و تطبیق فرهنگی و زبانی برای پنج حیطه تکاملی آزمون به ترتیب انجام شد. سپس یک مطالعه پایلوت بر روی ۴۵ کودک فارسی زبان ۱ تا ۴۲ ماهه در تهران در سال ۱۳۹۱ اجرا گردید تا میزان وضوح عناوین، تطابق فرهنگی و تشخیص سوالات مبهم مشخص گردد. در ارزیابی اعتبار آزمون از سه روش برآورد همسانی درونی، بازآزمایی و اعتبار ارزیاب‌ها در ۳۴ کودک استفاده شد. سپس برای تعیین خصوصیات روان‌سنجی آزمون، تعداد ۲۶۰ کودک ۱ تا ۴۲ ماهه از مراکز سلامت و مراقبتی واقع در پنج منطقه شهر تهران، به صورت نمونه‌گیری مستمر انتخاب شدند و آزمون برای ایشان اجرا گردید.

**یافته‌ها:** در مناسب‌سازی فرهنگی و زبانی که توسط تیم تخصصی انجام شد، برخی تغییرات غیرقابل اجتناب بود. در تمامی گروه‌های سنی در حیطه‌های ارتباطی تغییراتی داده شد که به طور عمده منشا زبانی یا فرهنگی داشتند. ضریب همسانی درونی در خرده‌آزمون‌های شناختی، ارتباط درکی و بیانی، حرکات ظریف و درشت به ترتیب ۰/۹۶، ۰/۹۵، ۰/۹۵، ۰/۹۵ و ۰/۹۴ به دست آمد و روایی سازه آزمون با روش تحلیل عاملی و مقایسه نمره‌ها در گروه‌های سنی تأیید شد.

**نتیجه‌گیری:** نسخه فارسی آزمون غربالگر سنجش تکامل کودکان «بیلی» دارای اعتبار و روایی است.

**کلیدواژه‌ها:** اعتباریابی، تکامل، غربالگری، کودک



## مقدمه

در حال حاضر اهمیت تشخیص زودرس اختلالات تکاملی و ارائه خدمات مداخله‌ای در شیرخواران و نوپایان ارزش روزافزونی پیدا کرده است. ارزیابی به موقع و دوره‌ای تکامل امکان تشخیص و درمان زودرس را فراهم کرده و مانع از دست رفتن پتانسیل‌های تکاملی کودک می‌شود (۱، ۲). دلیل تلاش افزایش درجهت تشخیص در سنین پایین‌تر آنست که برنامه‌های مداخله‌گرانه مقرون به صرفه هستند و چنانچه در سنین پایین‌تر شروع شوند تأثیر بیشتری دارند (۳). پیش آگهی تکاملی کودکان را بهبود می‌بخشند (۴) و دارای اثرات مثبت کوتاه مدت و بلند مدت هستند (۵-۷). در حال حاضر اغلب متخصصین بر اساس قضاوت بالینی خود این پایش را انجام می‌دهند ولی مطالعات متعددی نشان داده است که قضاوت بالینی به تنهایی در تشخیص تأخیر تکامل کارآیی لازم را ندارد (۸، ۹، ۳). به طور متوسط تنها ۳۰٪ از کودکان دارای مشکلات تکاملی یا رفتاری توسط پزشک بدون استفاده از ابزار غربالگر تشخیص داده می‌شوند (۱۱، ۱۰، ۸، ۵).

بهینه‌سازی دقت در ارزیابی‌های اولیه تکاملی اهمیت زیادی در تشخیص تأخیر تکاملی و اختلالات خفیف دارد. ابزارهای غربالگری تکاملی ذهنی<sup>۱</sup> خصوصاً برای پیدا کردن اختلالات تکاملی خفیف ارزش تشخیصی کمی دارند. جمعیت‌هایی که در خطر ایجاد مشکلات تکاملی هستند (کودکان درخطر)، به طور اولیه نیازمند ارزیابی بسیار دقیق‌تری هستند. به عبارت دیگر، ابزارهای دقیق‌تر عینی<sup>۲</sup> و مشتق از آزمون‌های تشخیصی برای ارزیابی کودکان درخطر مورد نیاز است (۱۴-۱۲).

انجمن متخصصین کودکان آمریکا که یک مرجع جهانی قابل قبول است توصیه می‌کند غربالگری تکاملی در هر ویزیت پایش رشد کودکان انجام شود (۸، ۵). از طرفی می‌دانیم تکامل کودک یک جریان پویا است. کودکان به طور متفاوتی مهارت‌ها را کسب می‌کنند و پس از اینکه برای اولین بار موفق به انجام مهارت جدیدی شدند، ممکن است به طور مداوم آن را بروز ندهند. بنابراین انجام تنها یک آزمون در یک مقطع زمانی خاص تنها یک صحنه از یک جریان پویا را نشان می‌دهد. از طرفی کودکانی که در ابتدا تکامل طبیعی دارند ممکن است در سنین بالاتر دچار مشکل تکاملی یا رفتاری شوند. شرایط محیطی نیز ممکن است در جهت تأثیر مثبت یا منفی بر تکامل عمل نماید. بنابراین سیر تکاملی کودک نیز ممکن است پیشرفت یا پسرفت داشته باشد (۱۵)، لذا برای تشخیص اختلالاتی که ممکن است

با رشد کودک ظاهر شوند، انجام غربالگری به صورت دوره‌ای ضرورت دارد (۱۶، ۱۵). از طرفی اغلب ابزارهای غربالگر تکاملی از حساسیت و ویژگی حدود ۷۰ تا ۸۰٪ برخوردارند که در مقایسه با سایر تست‌های غربالگر پزشکی کمتر است، ولی این مشکل با بررسی‌های دوره‌ای و مکرر قابل رفع است (۱۸-۱۶، ۱۴، ۶).

آزمون غربالگر تکاملی بیلی فرم خلاصه‌شده نسخه سوم آزمون تشخیصی تکامل بیلی است که تنها آزمون تشخیصی تکاملی بوده و همبستگی بالایی با عنوان‌های این آزمون تشخیصی و همچنین نسخه دوم بیلی دارد (۱۹). این آزمون در کشور آمریکا در سال ۲۰۰۶ با بررسی روی ۱۷۰۰ نمونه پس از نسخه اول (در سال ۱۹۶۹) و نسخه دوم (در سال ۱۹۹۳) تدوین گردید. این آزمون وضعیت تکاملی را از دوره نوزادی تا ۴۲ ماهگی، در ۴ گروه سنی A (۱ تا ۶ ماهه)، B (۷ تا ۱۲ ماهه)، C (۱۳ تا ۲۴ ماهه)، و D (۲۵ تا ۴۲ ماهه) و در پنج حیطه تکاملی (شناختی، ارتباط درکی، ارتباط بیانی، حرکات ظریف و حرکات درشت) ارزیابی و در مقایسه با نقاط برش تعیین شده اعلام می‌نماید. در حیطه شناختی دارای ۳۳ عنوان (از ۹۱ عنوان آزمون تشخیصی)، در حیطه زبانی دارای ۲۴ عنوان از ۴۹ عنوان ارتباط درکی و ۲۴ عنوان از ۴۸ عنوان ارتباط بیانی، و در حیطه حرکتی در دو زیر گروه دارای ۵۵ عنوان (۲۷ عنوان از ۶۶ عنوان حرکات ظریف و ۲۸ عنوان از ۷۲ عنوان حرکات درشت) آزمون تشخیصی می‌باشد.

این آزمون توسط کارشناس دارای مدرک روان‌سنجی یا دوره‌دیده انجام شده و اجرای آن توسط مشاهده رفتار کودک در حین بازی و تعامل با وی ارزیابی می‌گردد. دستورالعمل آن برای اجرا دارای زبانی صریح و ساده بوده و اجرای آن بسته به گروه سنی بین ۱۵ تا ۲۵ دقیقه طول می‌کشد (۱۹). برای اجرای آزمون پس از ثبت اطلاعات دموگرافیک، ابتدا باید سن کودک را محاسبه و در صورت نارس بودن (تولد کمتر از ۳۶ هفته جنینی) تا دو سالگی، سن وی را اصلاح نمود؛ سپس گروه سنی کودک را مشخص نمود. در انتهای هر خرده‌آزمون با جمع امتیازات به دست آمده، نمره خام آن خرده‌آزمون محاسبه شده و سپس با استفاده از جداول موجود، کودک از نظر ریسک اختلال تکامل در سه دسته درخطر<sup>۳</sup> (کمتر از ۲ انحراف معیار)، در حال بروز<sup>۴</sup> (بین منهای ۱ تا منهای ۲ انحراف معیار)، و توانمند<sup>۵</sup> (نرمال) قرار می‌گیرد. چنانچه پس از انجام آزمون غربالگر مشخص شود کودک در یکی از حیطه‌های تکاملی دچار تأخیر است (یعنی در دسته درخطر باشد)، برای ارزیابی بیشتر جهت آزمون تشخیصی بیلی ارجاع خواهد شد. اگر نمره



## روش بررسی

مطالعه به روش مقطعی در شهر تهران اجرا گردید. مراحل ترجمه، باز ترجمه، تعیین روایی صوری و محتوایی و تطبیق فرهنگی و زبانی برای پنج حیطه تکاملی آزمون به ترتیب انجام شد.

سپس یک مطالعه مقدماتی با استفاده از نسخه فارسی اصلاح شده بر روی ۴۵ کودک فارسی زبان بدون تأخیر تکاملی شناخته شده انجام شد. این مرحله با هدف تعیین میزان وضوح سؤالات و دستورالعمل برای آزمونگر، تطابق فرهنگی، تشخیص سؤالات مبهم و تشخیص مشکلات در روند کاربرد آزمون بود.

سپس از طریق فراخوان از کارشناسان رشته‌های کاردرمانی، روان‌شناسی و روان‌سنجی و آسیب‌شناس گفتار و زبان جهت همکاری دعوت به عمل آمد و پس از برگزاری آموزش تئوری (شامل توضیح محتوای آزمون و دستورالعمل، گروه‌های سنی، چگونگی تعیین گروه سنی و نقاط شروع و خاتمه و تصحیح سن جنینی در نوزادان نارس و پر کردن پرسشنامه دموگرافیک) آموزش عملی (چگونگی بررسی رفتار کودک و تعامل با وی، مشاهده ده آزمون توسط فرد مورد نظر و سپس انجام ده آزمون توسط وی و در صورت انجام صحیح آزمون، تأیید وی توسط مربی) سه کارشناس در رشته کاردرمانی با گرایش جسمی و ذهنی کودکان و یک کارشناس روان‌سنجی مورد تأیید قرار گرفتند. فاز میدانی در ۴ مرکز سلامت شهرداری و یک مرکز بهداشتی-درمانی در ۵ نقطه شهر تهران (شمال، جنوب، غرب، شرق و مرکز) پس از کسب رضایت آگاهانه از والدین کودکان آغاز گردید.

جهت ارزیابی خصوصیات روان‌سنجی آزمون اعتبار<sup>۲</sup> و روایی<sup>۳</sup> اندازه‌گیری شد. در ارزیابی اعتبار آزمون از سه روش برآورد همسانی درونی<sup>۴</sup>، بازآزمایی<sup>۵</sup> و اعتبار ارزیاب‌ها<sup>۶</sup> استفاده شد. همسانی درونی آزمون با استفاده از ضریب آلفای کرونباخ برای پنج حیطه اندازه‌گیری شد. پایایی آزمون با استفاده از بازآزمایی بر ۱۳ کودک ۴ تا ۷ روز پس از آزمون اول توسط همان آزمونگر و اعتبار ارزیاب‌ها با استفاده از اجرای آزمون توسط دو ارزیاب همزمان بر ۳۴ کودک بررسی شد.

برای بررسی روایی سازه از دو روش تحلیل عاملی<sup>۷</sup> و مقایسه میانگین نمره‌ها در گروه‌های سنی استفاده شد. برای تعیین حجم نمونه پژوهش به تبعیت از منابع موجود در زمینه مدل آماری تحلیل عاملی، از جمله Comrey (۲۸) ملاک تعیین حجم نمونه، یک حجم نمونه ۲۰۰ تا ۳۰۰ نفری در ۴ گروه سنی

کودک در قسمت در حال بروز قرار گیرد، می‌توان کودک را تحت نظر گرفت و بعد از دو تا چهار هفته (بسته به گروه سنی) دوباره آزمون غربالگر را انجام داد و در صورتی که نمره کودک در گروه توانمند واقع شود، نیاز به اقدام خاصی نیست و طبق معمول پیگیری می‌شود.

تحقیقات گسترده همبستگی بالای خرده‌آزمون‌های غربالگر بیلی را با خرده‌آزمون‌های تشخیصی و نسخه دوم بیلی نشان داده است (۱۹-۲۲). اعتبار و یا قابلیت اجرای این آزمون حتی در جمعیت‌های مختلف در معرض خطر مانند کودکان دارای اختلالات تکاملی نظیر سندرم داون، کودکان نارس و خیلی نارس، کودکان کوچک جثه نسبت به سن حاملگی، کودکان دارای اختلالات زبانی خاص، اختلالات نافذ رشد، و کودکان با فلج مغزی نیز تأیید شده است (۱۹). بررسی روایی و پایایی نسخه سوم بیلی در تایوان با نسخه اصلی صورت گرفته و در این جمعیت مورد تأیید قرار گرفت (۲۲). در گزارش moore نیز در کودکان انگلیسی زبان اعتبار بالای آزمون گزارش گردیده است (۲۱).

در گزارش نسخه با زبان مالایی در شیرخواران ۱۳ تا ۹ ماهه نیز مورد تأیید بود (۲۳). بررسی توالی عناوین آزمون نسخه با زبان هلندی نیز مورد تأیید قرار گرفته است (۲۴).

تست عینی دیگر در غربالگری تکامل کودکان از تولد تا ۸ سالگی تست دنور<sup>۲</sup> است. در منابع مختلف حساسیت آن از ۴۰ تا ۸۳ درصد و ویژگی آن از ۴۰ تا ۸۰ درصد گزارش شده است (۲۵، ۱۷).

غربالگر ذهنی دیگر، پرسشنامه سنین و مراحل است و بر اساس پرسش از والدین بوده و در ایران استانداردسازی شده است و از طریق تحلیل عاملی روایی آن تأیید شده است، ولی به دلیل نداشتن آزمون استاندارد طلایی حساسیت و اختصاصی بودن آن به دست نیامده است (۲۶). میزان ویژگی آن در مطالعه نیدلمن ۰/۷۶ تا ۰/۹۶ و حساسیت آن ۰/۷۰ تا ۰/۹۰ غیر سن ۴ ماهگی گزارش شده است (۲۷).

با توجه به این حقیقت که اکثر کودکان در خطر در کشور ما به علل متعددی از جمله عدم وجود ابزارهای غربالگری عینی معتبر و استاندارد، مورد ارزیابی قرار نگرفته‌اند، این مطالعه با هدف تدوین نسخه فارسی آزمون غربالگر سنجش تکامل کودکان «بیلی» و تعیین اعتبار و روایی آن برای کودکان فارسی زبان انجام شد.



مرحله اصلاحات و تطبیق‌های فرهنگی و زبانی نیز توسط تیم تخصصی انجام شد.

با وجود سعی در حفظ محتوای عناوین نسخه اصلی، برخی تغییرات به منظور افزایش وضوح مفهوم نسخه فارسی و تطبیق فرهنگی غیرقابل اجتناب بود. تقریباً تمامی تغییرات در حیطه "ارتباط درکی و بیانی"، منشاء زبانی داشت. برخی از این اصلاحات شامل موارد زیر است:

۱- تغییراتی که در بخش دستورالعمل خرده‌آزمون ارتباطی درکی داده شد: جایگزین نمودن بازی‌های غیرمرسوم با بازی‌های آشنا تر (مثل تاپ تاپ خمیر، لی‌لی-لی‌لی حوضک، دالی موشک، اتل متل توتوله، دس‌دسی-سرسری و کلاغ پر)، استفاده از کلمه لیوان به جای فنجان، کلمات توب به جای مکعب، شیرینی به کیک، پرنده به ماهی.

این تغییرات با استفاده از مطالعاتی که در زمینه رشد واژگان کودکان فارسی زبان ایرانی انجام شده بود (۳۰، ۲۹) تطبیق داده شد.

تغییر دیگر نوع بیان مالکیت بود. با توجه به وجود شکل کسره نشان دهنده مالکیت در زبان فارسی، این نوع از ضمائر نیز به دستورالعمل افزوده شد و با توجه به اینکه در زبان فارسی برای هر دو جنس دختر و پسر از یک ضمیر واحد "او" استفاده می‌شود، علاوه بر ضمیر، جنسیت نیز ذکر گردید.

۲- تغییراتی که در بخش دستورالعمل خرده‌آزمون ارتباطی بیانی داده شد: تغییرات در خصوص افعال کمکی (که عموماً در زبان فارسی به کار نمی‌روند)، استخراج بیان افعال استمراری در دستورالعمل، متفاوت بودن علائم افعال مستقبل و مضارع استمراری در دستور زبان فارسی و مشکل بودن استخراج بیان استفاده از جمع از کودک فارسی زبان بود، که در نحوه پرسش این سؤالات تغییراتی داده شد.

پس از تدوین نسخه فارسی تعداد ۲۶۰ کودک مورد بررسی قرار گرفتند که شامل ۱۲۶ نفر (۴۸/۵٪) دختر و ۱۳۴ نفر (۵۱/۵٪) پسر بودند.

فراوانی و درصد هر یک از گروه‌های سنی و میزان تحصیلات پدر و مادر در جدول ۱ نمایش داده شده است. ضریب همبستگی محاسبه شده در هر گروه سنی و به تفکیک حیطه‌ها در جدول ۲ نشان داده شده است.

و به صورت نمونه‌گیری مستمر از مراکز مذکور انتخاب شدند. در پژوهش حاضر به منظور پاسخ به این پرسش که آزمون دارای چند عامل معنی‌دار است، از روش تحلیل مؤلفه‌های اصلی<sup>۱</sup> استفاده شد.

قبل از اجرای تحلیل عاملی دو موضوع مورد مطالعه قرار گرفت: ۱- کفایت نمونه‌برداری ۲- اطمینان نسبت به این مطلب که ماتریس همبستگی زیربنای تحلیل عاملی در جامعه برابر صفر نیست. اندازه (KMO)<sup>۲</sup> ۰/۸۹۸ (تا ۰/۹۴۴) همراه با نتیجه اجرای آزمون کرویوت بارتلت<sup>۳</sup> برای پنج حیطه آزمون محاسبه گردید و سطح معنادار بودن مقدار مشخصه آماری آزمون بارتلت (برای هر پنج حیطه) کوچک‌تر از ۰/۰۰۰۱ بود. به این ترتیب مشخص گردید که علاوه بر کفایت نمونه‌برداری، اجرای تحلیل بر پایه ماتریس‌های همبستگی بین مواد پرسشنامه نیز قابل توجیه است.

برای تعیین این مطلب که مجموعه مواد تشکیل‌دهنده آزمون مورد مطالعه از چند عامل معنی‌دار اشباع شده، دو شاخص ارزش ویژه<sup>۴</sup> و نسبت واریانس تبیین شده توسط هر عامل اندازه‌گیری شد.

### یافته‌ها

در تدوین پرسشنامه از نظرات متخصصین و مرور متون داخلی و خارجی در زمینه رشد و تکامل کودک استفاده گردید. در مرحله اول عناوین آزمون در هر یک از حیطه‌های پنج‌گانه توسط پنج متخصص فارسی‌زبان آشنا به زبان انگلیسی و تکامل کودک، به فارسی ترجمه گردید و سپس توسط پنج متخصص دیگر آشنا به هر دو زبان و دارای تجربه در زمینه تکامل کودک بازترجمه شد. با مقایسه نسخه بازترجمه با نسخه اصلی آزمون، قسمت‌های ناهمخوان مشخص و توسط همان متخصصین در نسخه فارسی اصلاح گردید.

در مرحله بعد این نسخه از نظر اعتبار محتوایی در هر بخش توسط دو صاحب‌نظر دست‌اندرکار امور تکامل کودکان بررسی شد. این صاحب‌نظران متشکل بودند از دو متخصص کودکان، دو آسیب‌شناس گفتار و زبان، و سه نفر متخصص کاردرمانی کودکان با گرایش جسمی و ذهنی که همگی دارای دانش و تجربه در زمینه تکامل کودک بودند. در این



جدول ۱. توزیع فراوانی و درصد گروه نمونه بر اساس اطلاعات جمعیت شناختی

درصد	فراوانی	سن	
۱۳/۸	۳۶	۱ ماه - ۶ ماه	A
۲۰/۰	۵۲	۷ ماه - ۱۲ ماه	B
۳۵/۴	۹۲	۱۳ ماه - ۲۴ ماه	C
۳۰/۸	۸۰	۲۵ ماه - ۴۲ ماه	D
جنسیت			
۵۱/۵	۱۳۴	پسر	
۴۷/۵	۱۲۶	دختر	
میزان تحصیلات پدر			
۸/۱	۲۱	دیپلم و زیر دیپلم	
۷۷/۵	۱۴۱	فوق دیپلم و لیسانس و فوق لیسانس	
۱۲/۳	۳۲	دکتری تخصصی	
۲/۳	۶	بدون پاسخ / نامشخص	
میزان تحصیلات مادر			
۵/۱	۱۳	دیپلم و زیر دیپلم	
۸۷/۴	۲۲۷	فوق دیپلم و لیسانس و فوق لیسانس	
۵/۴	۱۴	دکتری تخصصی	
۲/۳	۶	بدون پاسخ / نامشخص	

جدول ۲. همسانی درونی آزمون به تفکیک حیطه‌های پنجگانه در گروه‌های سنی و بر اساس جنسیت

گروه	شناختی	ارتباط درکی	ارتباط بیانی	حرکات ظریف	حرکات درشت
سن (غربالگری)	A	۰/۸۴	۰/۷۸	۰/۸۰	۰/۸۱
	B	۰/۷۰	۰/۷۲	۰/۸۱	۰/۸۸
	C	۰/۸۱	۰/۸۴	۰/۸۵	۰/۷۶
	D	۰/۷۹	۰/۶۶	۰/۸۴	۰/۷۷
جنسیت	میانگین ضرایب	۰/۷۹	۰/۷۶	۰/۸۳	۰/۸۱
	پسر	۰/۹۵	۰/۹۴	۰/۹۴	۰/۹۴
	دختر	۰/۹۶	۰/۹۶	۰/۹۶	۰/۹۵
	تمام سنین	۰/۹۶	۰/۹۵	۰/۹۵	۰/۹۴

پایایی آزمون غربالگری با استفاده از بازآزمایی در جدول ۳ و جدول ۴ گزارش شده است. اعتبار ارزیابی‌ها با استفاده از اجرای آزمون توسط دو ارزیاب در

جدول ۳. ضریب بازآزمایی در اجرای نخست و دوم آزمون

خرده‌آزمون	شاخص d کوهن	ضریب همبستگی گشتاوری پیرسون	ضریب همبستگی رتبه‌ای اسپیرمن
شناختی	۰/۲۷۰	**۰/۹۹۱	**۰/۹۴۷
ارتباط درکی	۰/۳۵۶	**۰/۹۹۲	**۰/۹۶۸
ارتباط بیانی	۰/۰۲۰	**۰/۹۸۵	**۰/۹۷۷
حرکات ظریف	۰/۱۴۹	**۰/۹۹۸	**۰/۹۷۵
حرکات درشت	۰/۲۶۵	**۰/۹۹۸	**۰/۹۹۳

\*\*P<۰,۰۱



جدول ۴. اعتبار ارزیاب‌ها برای آزمون در حیطه‌های پنجگانه

پایایی ارزیاب‌ها		شاخص d کوهن	خرده‌آزمون
ضریب همبستگی بین نمره‌های دو ارزیاب (ضریب همبستگی گشتاوری)	ضریب همبستگی رتبه‌ای		
پیرسون	اسپیرمن		
**۰/۹۹۱	**۰/۹۶۷	۰/۰۰۲	شناختی
**۰/۹۹۹	**۰/۹۹۱	۰/۰۶۱	ارتباط درکی
**۰/۹۹۹	**۰/۹۹۴	۰/۰۳۵	ارتباط بیانی
**۰/۹۹۸	**۰/۹۹۳	۰/۰۴۰	حرکات ظریف
**۰/۹۹۰	**۰/۹۷۷	۰/۰۷۱	حرکات درشت

P&lt;۰/۰۱\*\*

در جدول ۵ میزان ارزش ویژه و درصد واریانس تبیین شده توسط عامل نخست با اجرای تحلیل مؤلفه‌های اصلی نشان داده شده است. با توجه به میزان ارزش ویژه و درصد واریانس تبیین شده توسط عامل نخست مشخص می‌گردد که در هر حیطه مطلوب‌ترین شرایط برای اجرای تحلیل عاملی در مجموعه عناوین هر حیطه یک مدل تک‌عاملی می‌باشد.

در جدول ۵ میزان ارزش ویژه و درصد واریانس تبیین شده توسط عامل نخست با اجرای تحلیل مؤلفه‌های اصلی نشان داده شده است. با توجه به میزان ارزش ویژه و درصد واریانس تبیین شده توسط عامل نخست مشخص می‌گردد که در هر حیطه مطلوب‌ترین شرایط برای اجرای تحلیل عاملی در مجموعه عناوین هر حیطه یک مدل تک‌عاملی می‌باشد.

جدول ۵. ارزش ویژه و درصد واریانس تبیین شده توسط عامل نخست با اجرای تحلیل مؤلفه‌های اصلی

حیطه	ارزش ویژه عامل نخست	درصد واریانس تبیین شده توسط عامل نخست
شناختی	۱۴/۰۹۱	۴۲/۷۰۰
ارتباط درکی	۱۱/۰۷۴	۴۶/۱۴۲
ارتباط بیانی	۱۱/۳۹۸	۴۷/۴۹۰
حرکات ظریف	۱۱/۳۹۱	۴۲/۱۹۰
حرکات درشت	۱۱/۶۷۰	۴۱/۶۸۰

در جدول ۶ بارهای عاملی پرسش‌های آزمون به تفکیک حیطه‌های پنج‌گانه نشان داده شده است. به منظور تحقیق درباره ماهیت روابط بین عناوین آزمون و دستیابی به تعاریف عامل‌ها فرض بر این قرار گرفت که ضرایب بیشتر از ۰/۳ در تعریف عامل‌ها سهم با معنا دارند و بنابراین ضرایب کمتر از این مقدار به عنوان عامل‌های تصادفی در نظر گرفته شد. با توجه به دامنه بین ۰/۳ تا ۰/۸ در عناوین مختلف در حیطه‌های پنج‌گانه، بنابراین می‌توان چنین استنتاج کرد که بهترین ساختار برای آزمون غربالگری بیلی یک مدل تک‌عاملی می‌باشد و به عبارت دیگر این آزمون برای سنجش سازه رشد تکاملی شیرخواران ۱ تا ۴۲ ماهه شهر تهران دارای روایی است.

در جدول ۶ بارهای عاملی پرسش‌های آزمون به تفکیک حیطه‌های پنج‌گانه نشان داده شده است. به منظور تحقیق درباره ماهیت روابط بین عناوین آزمون و دستیابی به تعاریف عامل‌ها فرض بر این قرار گرفت که ضرایب بیشتر از ۰/۳ در تعریف عامل‌ها سهم با معنا دارند و بنابراین ضرایب کمتر از این مقدار به عنوان عامل‌های تصادفی در نظر گرفته شد. با توجه به دامنه بین ۰/۳ تا ۰/۸ در عناوین مختلف در حیطه‌های پنج‌گانه، بنابراین می‌توان چنین استنتاج کرد که بهترین ساختار برای آزمون غربالگری بیلی یک مدل تک‌عاملی می‌باشد و به عبارت دیگر این آزمون برای سنجش سازه رشد تکاملی شیرخواران ۱ تا ۴۲ ماهه شهر تهران دارای روایی است.



جدول ۶. بارهای عاملی پرسش‌های آزمون به تفکیک حیطه‌های پنج‌گانه

حیطه شناختی		حیطه ارتباط درکی		حیطه ارتباط بیانی		حیطه حرکات ظریف		حیطه حرکات درشت	
شماره پرسش	بار عاملی	شماره پرسش	بار عاملی	شماره پرسش	بار عاملی	شماره پرسش	بار عاملی	شماره پرسش	بار عاملی
۱*	۰/۱۷۹	۱*	۰/۱۸۱	۱*	۰/۲۴۸	۱*	۰/۱۱۲	۱*	۰/۲۲۲
۲	۰/۳۲۳	۲*	۰/۲۷۹	۲*	۰/۱۵۶	۲*	۰/۲۷۲	۲*	۰/۲۵۲
۳	۰/۳۲۳	۳*	۰/۲۴۵	۳	۰/۳۳۲	۳*	۰/۲۳۵	۳	۰/۳۴۲
۴*	۰/۲۸۳	۴	۰/۳۱۷	۴	۰/۳۶۵	۴	۰/۳۳۶	۴	۰/۳۰۵
۵	۰/۵۱۷	۵	۰/۴۸۰	۵	۰/۵۰۵	۵	۰/۴۰۵	۵	۰/۴۵۵
۶	۰/۵۴۵	۶	۰/۴۹۴	۶	۰/۶۲۰	۶	۰/۴۵۱	۶	۰/۵۵۹
۷	۰/۵۳۶	۷	۰/۵۹۹	۷	۰/۶۶۷	۷	۰/۴۸۳	۷	۰/۵۷۱
۸	۰/۵۸۶	۸	۰/۶۳۰	۸	۰/۶۷۷	۸	۰/۵۳۰	۸	۰/۶۴۱
۹	۰/۶۷۱	۹	۰/۷۵۵	۹	۰/۶۸۶	۹	۰/۶۷۶	۹	۰/۶۰۵
۱۰	۰/۶۹۵	۱۰	۰/۶۶۵	۱۰	۰/۷۷۴	۱۰	۰/۷۱۸	۱۰	۰/۶۲۷
۱۱	۰/۷۳۶	۱۱	۰/۷۷۸	۱۱	۰/۸۲۰	۱۱	۰/۷۲۱	۱۱	۰/۷۱۸
۱۲	۰/۶۹۶	۱۲	۰/۸۲۷	۱۲	۰/۸۲۷	۱۲	۰/۷۶۱	۱۲	۰/۷۸۵
۱۳	۰/۷۱۰	۱۳	۰/۸۳۳	۱۳	۰/۸۴۲	۱۳	۰/۷۷۹	۱۳	۰/۸۰۸
۱۴	۰/۷۴۵	۱۴	۰/۸۰۸	۱۴	۰/۸۱۸	۱۴	۰/۸۱۳	۱۴	۰/۸۱۷
۱۵	۰/۷۹۸	۱۵	۰/۸۵۷	۱۵	۰/۸۵۳	۱۵	۰/۸۲۱	۱۵	۰/۸۳۲
۱۶	۰/۸۰۰	۱۶	۰/۸۵۰	۱۶	۰/۸۴۴	۱۶	۰/۸۱۲	۱۶	۰/۸۳۹
۱۷	۰/۸۰۸	۱۷	۰/۸۴۰	۱۷	۰/۸۲۴	۱۷	۰/۷۷۱	۱۷	۰/۸۱۶
۱۸	۰/۷۶۵	۱۸	۰/۸۴۶	۱۸	۰/۸۳۸	۱۸	۰/۸۱۹	۱۸	۰/۷۶۸
۱۹	۰/۸۱۷	۱۹	۰/۸۰۴	۱۹	۰/۸۰۱	۱۹	۰/۷۸۹	۱۹	۰/۷۸۵
۲۰	۰/۷۹۶	۲۰	۰/۷۹۳	۲۰	۰/۷۵۸	۲۰	۰/۷۷۰	۲۰	۰/۶۵۹
۲۱	۰/۸۱۲	۲۱	۰/۷۲۳	۲۱	۰/۷۵۳	۲۱	۰/۷۸۰	۲۱	۰/۷۷۷
۲۲	۰/۷۸۲	۲۲	۰/۷۲۴	۲۲	۰/۷۲۳	۲۲	۰/۷۳۶	۲۲	۰/۷۴۹
۲۳	۰/۷۷۳	۲۳	۰/۶۰۸	۲۳	۰/۶۴۶	۲۳	۰/۷۲۹	۲۳	۰/۶۱۵
۲۴	۰/۷۴۴	۲۴	۰/۵۹۶	۲۴	۰/۴۲۱	۲۴	۰/۶۵۹	۲۴	۰/۶۳۹
۲۵	۰/۷۳۹			۲۵		۲۵	۰/۶۰۸	۲۵	۰/۶۰۱
۲۶	۰/۷۱۹			۲۶		۲۶	۰/۵۶۰	۲۶	۰/۵۶۹
۲۷	۰/۶۴۲			۲۷		۲۷	۰/۵۵۰	۲۷	۰/۶۱۶
۲۸	۰/۶۳۴			۲۸				۲۸	۰/۴۰۱
۲۹	۰/۶۳۴								
۳۰	۰/۸۷۲								
۳۱	۰/۵۸۱								
۳۲	۰/۳۷۷								
۳۳	۰/۴۹۹								

\* ضرایب بیشتر از ۰/۳ در تعریف عامل‌ها که دارای سهم با معنا هستند

صفر مبنی بر برابری میانگین نمره آزمودنی‌ها در چهار گروه سنی با اطمینان ۹۹ درصد رد می‌شود. با توجه به رد فرض صفر و برای مقایسه دو به دو میانگین‌ها از آزمون تعقیبی شفه<sup>۱</sup> استفاده شد و سپس تفاوت میانگین‌ها محاسبه گردید که نتایج آن در جدول ۷ آمده است. معنی‌داری تفاوت بین میانگین‌های محاسبه‌شده نشان‌دهنده وجود رابطه بین سن و نمره‌های آزمون در حیطه‌های پنج‌گانه است و با افزایش سن، نمره آزمودنی‌ها در هر پنج حیطه افزایش می‌یابد. این یافته‌ها فرضیه ارائه شده در فوق و در نتیجه روایی سازه آزمون را مجدداً مورد تأیید قرار داد.

از آنجا که ماهیت و محتوای آزمون، رشد تکاملی است، لذا عملکرد در این آزمون باید با سن تقویمی رابطه داشته باشد. برای آزمون این فرضیه نمره حاصل از پنج حیطه آزمون در گروه‌های سنی با استفاده از تحلیل واریانس یک‌راهه با یکدیگر مقایسه شد. شاخص F محاسبه‌شده به ترتیب در خرده‌آزمون‌های شناختی، ارتباط درکی و بیانی، حرکات درشت و ظریف ۵۳۹/۷۷، ۴۴۶/۶۱، ۳۴۱/۶۱، ۳۶۲/۳۸ و ۳۹۵/۷۰ با سطح معنی‌داری  $P < 0/01$  بود. با توجه به اینکه شاخص F محاسبه شده (در هر پنج حیطه) بزرگ‌تر از ۰/۰۱ با درجات آزادی ۳ و ۲۵۶ (۳/۳۸) است، بنابراین فرض

1- Scheffe post hoc test



جدول ۷. مقایسه دو به دو میانگین‌ها در چهار گروه سنی آزمون غربالگری

تفاوت میانگین‌ها و سطح معناداری						
حرکات درشت	حرکات ظریف	ارتباط بیانی	ارتباط درکی	شناختی	گروه سنی ۲	گروه سنی ۱
**۶/۰۳	**۴/۳۸	**۳/۸۲	**۴/۱۹	**۶/۵۸	B	A
**۱۲/۱۶	**۱۰/۶۷	**۱۰/۵۶	**۱۰/۸۷	**۱۴/۲۷	C	
**۱۷/۹۸	**۱۷/۴۱	**۱۶/۸۷	**۱۶/۹۰	**۲۲/۵۹	D	
**۶/۱۳	**۶/۳۰	**۶/۷۴	**۶/۶۸	**۷/۶۹	C	B
**۱۱/۹۶	**۱۳/۰۳	**۱۳/۰۵	**۱۲/۷۱	**۱۶/۰۱	D	
**۵/۸۲	**۶/۸۳	**۶/۳۰	**۶/۰۳	**۸/۳۲	D	C

P &lt; ۰/۰۱ \*\*

نسخه دوم بیللی نشان داده است (۲۲، ۲۱، ۱۹).

بررسی روایی و پایایی نسخه سوم بیللی در تایوان با نسخه اصلی صورت گرفته و در این جمعیت مورد تأیید قرار گرفت (۲۲). در تایوان ضریب همبستگی در دفعات آزمون در خرده‌آزمون شناختی ۰/۹۲ تا ۰/۹۹، در خرده‌آزمون زبانی ۰/۸۸ تا ۰/۹۸ و در حرکتی ۰/۸۵ تا ۰/۹۸ گزارش شده است و میزان ضریب همبستگی در بین آزمونگران در خرده‌آزمون‌های شناختی ۰/۸۷ تا ۰/۹۷، زبانی ۰/۷۶ تا ۰/۹۵ و حرکتی ۰/۷۷ تا ۰/۸۹ گزارش شده است. در گزارش moore نیز در کودکان انگلیسی زبان اعتبار بالا آزمون گزارش گردیده است (۲۱).

میزان اعتبار این تست در آمریکا در بخش همسانی درونی در خرده‌آزمون شناختی به‌طور متوسط ۰/۸۵، درکی ۰/۸۸، ارتباط بیانی ۰/۸۸، حرکات ظریف ۰/۸۲ و حرکات درشت ۰/۸۶ بود. در ثبات زمانی در فاصله زمانی متوسط ۷ روز بر ۲۰۳ نمونه در گروه‌های سنی متفاوت بین ۰/۸۰ تا ۰/۸۳ به‌دست آمد (۱۹).

در نسخه مالایی میزان آلفای کرونباخ در ارزیابی همبستگی درونی از ۰/۹۲۶ در خرده‌آزمون شناختی تا ۰/۷۵۱ در خرده‌آزمون حرکات ظریف بوده است (۲۳) و در زبان هلندی نیز توالی عناوین آزمون مورد تأیید قرار گرفت (۲۴).

این آزمون، کودکان را از نظر خطر برای اختلالات تکاملی به سه گروه پر خطر یا در خطر، در حال بروز، و توانمند تقسیم می‌نماید.

بررسی آمریکا دقت طبقه‌بندی<sup>۴</sup> گروه کودکان در خطر را متوسط ذکر کرده است؛ یعنی کودکانی که در آزمون غربالگر بیللی در خرده‌آزمون شناختی به‌عنوان در خطر مشخص شده‌اند، ۵۴/۳۵ درصد در آزمون تشخیصی در کمتر از منهای دو انحراف معیار بوده و در ۲/۹۲ درصد موارد بین منهای ۱ و منهای ۲ انحراف معیار بوده‌اند و هیچکدام در آزمون تشخیصی طبیعی نبوده‌اند.

## بحث

بررسی ارتباط بین محتوی آزمون و ساختار مورد نظر برای اندازه‌گیری، یک منبع مهم برای شواهد روایی آزمون می‌باشد. شواهد روایی محتوایی بر پایه آزمون‌های آماری یا تجربی نمی‌باشد؛ بلکه بر پایه این است که موارد آزمون به‌اندازه کافی گویا و مرتبط با صفات و عملکردهایی که در حال اندازه‌گیری است، باشد. محتوای آزمون شامل کلمات، شکل و قالب موارد آزمون و همچنین روند توزیع و نمره‌دهی آزمون است. در این مطالعه روایی محتوای آزمون توسط صاحب‌نظران مورد تأیید قرار گرفت و ۷ نفر از متخصصین آشنا با رشد و تکامل کودک شامل متخصص کودکان، کاردرمان، و آسیب‌شناس گفتار و زبان، محتوای نسخه فارسی آزمون را مورد تأیید قرار دادند.

در ارزیابی اعتبار آزمون از سه روش برآورد همسانی درونی، بازآزمایی و اعتبار ارزیابی‌ها استفاده شد. در این مطالعه همسانی درونی آزمون در خرده‌آزمون شناختی ۰/۹۶؛ در خرده‌آزمون ارتباط درکی و بیانی ۰/۹۵ و در حرکات ظریف و درشت ۰/۹۵ و ۰/۹۴ به‌دست آمد که بسیار خوب است. پایایی آزمون-بازآزمون (ضریب بازآزمایی) در پنج حیطه در حدود ۰/۹۹ و پایایی بین آزمونگران ۰/۹۹ بود. بنابراین با توجه به اعتبار بالای این آزمون می‌توان به این نتیجه دست یافت که پس از آموزش مناسب می‌توان به ارزیابی پایا در نسخه فارسی دست یافت.

آزمون غربالگر تکاملی بیللی فرم خلاصه‌شده آزمون تشخیصی است؛ به این ترتیب که ۳۳ مورد از ۹۱ عنوان مقیاس شناختی، ۲۴ مورد از ۴۹ عنوان ارتباط درکی، ۲۴ مورد از ۴۸ عنوان ارتباط بیانی، ۲۷ مورد از ۶۶ عنوان حرکات ظریف و ۲۸ مورد از ۷۲ عنوان حرکات درشت آزمون تشخیصی به‌صورت عناوینی در پنج حیطه در آزمون غربالگر خلاصه شدند. تحقیقات گسترده همبستگی بالای این عناوین را با عنوان‌های آزمون تشخیصی و





داون به ترتیب ۰/۹۵ و ۰/۸۷ گزارش شده است (۱۹). این میزان حساسیت و اختصاصی بودن در اختلالات نافذ رشد در آزمون شناختی ۰/۸۲ و ۰/۸۶ در ارتباط درکی ۰/۹۲ و ۰/۸۳ و در ارتباطی بیانی ۰/۹۲ و ۰/۸۳ ذکر می‌شود. در اختلالات زبان خاص میزان حساسیت و اختصاصی بودن در ارتباط درکی ۰/۴۱ و ۰/۷۷ و در ارتباط بیانی ۰/۵۴ و ۰/۷۸ گزارش شده است (۱۹). در فلج مغزی میزان دقت (حساسیت و اختصاصی بودن) در خرده‌آزمون حرکات ظریف ۰/۷۵ و ۰/۷۸ و در خرده‌آزمون حرکات درشت ۰/۹۶ و ۰/۹۲ گزارش شده است (۱۹).

در مطالعه حاضر نیز روایی سازه با دو روش تحلیل عاملی و مقایسه نمره‌ها در گروه‌های سنی مورد تأیید قرار گرفت. این آزمون در مطالعات تایوان، کانادا و انگلیس نیز روا شناخته شده است (۲۰-۲۲).

در مطالعه جاکسون و همکاران در سال ۲۰۱۲ در نبراسکای آمریکا مشخص گردید که خرده‌آزمون حرکتی بیلی در ارجاع به موقع شیرخواران در خطر به خدمات مداخله به‌هنگام در مقایسه با مقیاس حرکتی شیرخواران آلبرتا همبستگی بالایی داشته است (۳۱)، بنابراین به‌صورت مناسبی می‌تواند در این خصوص مورد استفاده قرار گیرد.

ابزارهای غربالگری غالباً به دو صورت کلی طراحی شده‌اند:

۱- آزمون‌های تکاملی که در آن آزمونگران به‌طور مستقیم رفتارهای کودک را مورد مشاهده و ارزیابی قرار می‌دهند (آزمون‌های عینی) مانند تست‌های دنور II، مقیاس شاخص‌های زبانی (ELMS)، Brigance، Battelle و Bayley (۳۲).

۲- پرسشنامه‌های تکاملی که والدین آنها را پر می‌کنند (آزمون‌های ذهنی) مانند پرسشنامه ارزیابی والدین از وضعیت تکاملی<sup>۱</sup> و پرسشنامه سنین و مراحل<sup>۲</sup>. در این پرسشنامه‌ها از اطلاعات والدین استفاده می‌شود. توجه به نظر والدین در مورد وضعیت تکاملی کودکانشان از سال‌ها قبل مناسب و قابل اعتماد شناخته شده است (۳۳-۳۵). والدین کم‌سواد ممکن است مشکلاتی در خواندن پرسشنامه‌ها داشته باشند. سؤال از والدین به‌روشی مناسب می‌تواند بر این مشکلات فائق آید. از طرفی برخی پزشکان معتقدند والدینی با سطح بالای آموزشی، ممکن است بیش از حد نسبت به تکامل کودکان خود حساس شوند و استفاده از پرسشنامه‌های والد محور می‌تواند سبب افزایش میزان ارجاع شود (۳۳) و ممکن است در مورد صحت اطلاعات گرفته شده از والدین شبهه وجود داشته باشد. باید به این نکته توجه داشت که پرسشنامه‌ها بیشتر برای غربالگری‌های دو مرحله‌ای

در خرده‌آزمون حرکات ظریف ۴۱/۸۲٪ گروه در خطر در تست تشخیصی کمتر از منهای ۲ انحراف معیار، ۲/۰۷ در بین منهای ۱ تا منهای ۲ انحراف معیار بوده و هیچکدام در گروه طبیعی قرار نداشتند. در خرده‌آزمون ارتباط درکی ۶۵/۹۱٪ گروه در خطر کمتر از منهای ۲ انحراف معیار بوده و هیچکدام از این کودکان به‌عنوان توانمند شناسایی نشده‌اند (۱۹). دقت طبقه‌بندی برای کودکانی که در آزمون تشخیصی مابین منهای ۱ و منهای ۲ انحراف معیار قرار داشته و در تست غربالگر به‌عنوان در حال بروز دسته‌بندی شدند بیشتر بود. درصد کودکان این گروه که به‌درستی به‌عنوان در حال بروز تشخیص داده شدند از ۶۳/۸۷٪ در خرده‌آزمون شناختی تا ۷۷/۷۸٪ در خرده‌آزمون ارتباط درکی بود. درصد کودکانی که در این گروه به اشتباه به‌عنوان گروه در خطر بودند، بسیار کم است و حدوداً بین ۰/۸۲٪ تا ۵/۲۱٪ قرار دارد (۱۹). برای گروه توانمند دقت طبقه‌بندی خیلی دقیق بوده است. درصد کودکانی که در این گروه به‌عنوان توانا مشخص شده‌اند، بین ۸۳/۸۴٪ در خرده‌آزمون شناختی تا ۹۲/۱۱٪ در خرده‌آزمون‌های ارتباط درکی می‌باشد و هیچکدام از این کودکان به اشتباه در گروه در معرض خطر تشخیص داده نشدند (۱۹). نکته قابل ذکر اینکه هیچ کودکی در آزمون تشخیصی نیست که نمره پایین‌تر از دو انحراف استاندارد را کسب کرده باشد، ولی در آزمون غربالگری در گروه توانمند باشد و هیچ کودکی در تست غربالگری در گروه پرخطر نبوده که همزمان در آزمون تشخیصی نمره بالاتر از منهای ۱ انحراف معیار (نرمال) کسب کرده باشد (۱۹).

چنانچه پس از انجام آزمون غربالگر مشخص شود کودک در یکی از حیطه‌های تکاملی (شامل شناختی، ارتباط درکی و بیانی، حرکات درشت و ظریف) دچار تأخیر (یعنی در گروه در خطر باشد) برای ارزیابی بیشتر جهت آزمون تشخیصی بیلی ارجاع خواهد شد. اگر نمره کودک در گروه در حال بروز قرار گیرد، می‌توان کودک را تحت نظر گرفت و بعد از دو تا چهار هفته (بسته به گروه سنی) دوباره آزمون غربالگر بیلی را انجام داد و در صورتی که نمره کودک در گروه توانمند واقع شود، نیاز به اقدام خاصی نیست و طبق معمول پیگیری می‌شود. بدین ترتیب نیاز به بررسی کودک با تست تشخیصی بیلی که نیاز به صرف وقت بیشتری داشته و هزینه بیشتری در بر دارد کاسته خواهد شد (۱۳-۱۱).

در آمریکا روایی آزمون با مطالعه بر روی کودکان با نیازهای خاص مثل سندرم داون، فلج مغزی، اختلالات نافذ رشد و اختلالات خاص زبان نیز به‌دست آمده است. میزان حساسیت و اختصاصی بودن در خرده‌آزمون شناختی در تشخیص درست کودکان سندرم



که بر خلاف آزمون‌های عینی نیاز به همکاری کودک نداشته؛ بر اساس شاخص‌های تکاملی تهیه شده و آنها را به والدین آموزش می‌دهد، مقرون به صرفه بوده و زمان انجام آن کوتاه است. از نقاط ضعف آن به این مسائل می‌توان اشاره کرد که هر پرسشنامه شامل ۵-۴ صفحه است که نگهداری آن به فضای زیادی نیاز دارد و همچنین ممکن است والدین کم سواد در تکمیل آن دچار مشکل شوند و در ۱۳٪ موارد قادر به تشخیص تأخیر تکاملی نیست (۲۷)؛ بنابراین در گروه پرخطر استفاده از آن مطمئن نمی‌باشد. قابل توجه اینکه نسخه فارسی این آزمون نیز که روایی آن توسط تحلیل عاملی به دست آمد با مدل تک عاملی تبیین گردید و قابل ذکر است که در حال حاضر استفاده از دنور ۲ به دلیل طیف گسترده میزان حساسیت و اختصاصی بودن، پیشنهاد نمی‌گردد (۱۶).

یکی از محدودیت‌های مطالعه حاضر بررسی روایی آزمون با استفاده از مدل تحلیل عاملی است که به دلیل نداشتن نسخه فارسی استاندارد آزمون تشخیصی در زمان انجام مطالعه بود و به همین دلیل حساسیت و اختصاصی بودن تست مانند نسخه اصلی قابل ارزیابی نبود.

### نتیجه گیری

این نتایج نشان می‌دهد که نسخه فارسی آزمون غربالگر بیلی دارای پایایی و روایی مناسب برای استفاده کودکان فارسی زبان بوده و با توجه به عینی بودن تست و محدوده سنی بسیار مناسب (دوره نوزادی تا ۴۲ ماهگی)، و پس از مقایسه نتایج آن با آزمون تشخیصی تکاملی بیلی و تعیین میزان حساسیت و اختصاصی بودن آن، به عنوان آزمون غربالگر مرجع در کودکان در خطر ایرانی تا ۴۲ ماهگی توصیه شود.

به کار می‌روند و موارد مشکوک و ناموفق آن‌ها باید با آزمون‌های تشخیصی و یا غربالگرهای عینی مانند Brigance و Baylay Infant Neurodevelopmental Screener (BINS) که نیاز به وقت و مهارت بیشتری دارند ارزیابی شوند (۳۳).

همان‌طور که قبلاً نیز گفته شد ابزاری که به صورت همه گیر برای تمام جوامع و همه سنین کاربرد داشته باشد موجود نیست و در بسیاری از کشورهای در حال رشد ابزارهای غربالگر متناسب با فرهنگ وجود ندارد و ابزارهای موجود خارجی جهت استفاده، باید ابتدا برای جمعیت هر کشور استانداردسازی شوند (۳۵، ۳۴). در میان این ابزارها آزمون دنور و پرسشنامه سنسن و مراحل دارای نسخه فارسی است ولی روایی ملاکی آنها ارزیابی نشده است. آزمون دنور ۲ یک تست عینی در غربالگری تکامل کودکان از تولد تا ۸ سالگی است در منابع مختلف حساسیت آن از ۴۰-۸۳ درصد و ویژگی آن از ۴۰-۸۰ درصد گزارش شده است (۲۵، ۱۶). در ایران خصوصیات روانسجی دنور در مقایسه با پرسشنامه سنین و مراحل در نمونه ۱۹۷ نفری بررسی شد که ضریب توافق کاپا بین دو آزمون ضعیف (۰/۲۱) و ضریب توافق با معاینه پزشک ۰/۱۷ گزارش گردید که نویسندگان به این نتیجه رسیدند که ضریب توافق کاپا در مورد آزمون دنور ضعیف بوده و در نتیجه پیشنهاد مقایسه نتایج آزمون دنور ایرانی را با یک استاندارد طلایی تشخیصی دادند (۳۶).

پرسشنامه سنین و مراحل، در ایران استاندارد شده است ولی به علت عدم دسترسی به آزمون تشخیصی تکاملی استاندارد ایرانی، امکان ارزیابی روایی ملاکی و تعیین حساسیت و ویژگی در نسخه فارسی میسر نبوده است (۲۶). میزان ویژگی آن در مطالعه نیدلمن ۰/۷۶-۰/۹۱ و حساسیت آن ۰/۷۰-۰/۹۰ غیر سن ۴ ماهگی گزارش شده است (۲۷). از نقاط قوت آن می‌توان گفت

### منابع

- 1-Briggs Gowan MJ, Carter AS, Irwin JR, Wachtel K, Cicchetti DV. The brief infant toddler social & emotional assessment: screening for social emotional problems & delays in competence. *J Pediatr Psychol*. 2004; 29(2): 143-155.
- 2-Halfon N, Regalado M, Sareen H, Inkelas M, Reuland CH, Glascoe FP, et al. Assessing development in the pediatric office. *Pediatrics*. 2004; 113(6 suppl): 1926-33.
- 3-Rydz D, Srour M, Oskoui M, Marget N, Shiller M, Birnbaum R, et al. Screening for developmental delay in setting of a community pediatric clinic: A prospective assessment of Parent -Report Questionnaires. *Pediatrics*. 2006; 118(4): e1178-86.
- 4-Mayson TA, Harris SR, Bachman CL. Gross Motor Development of Asian and European Children on Four Motor Assessments: A Literature Review. *Pediatr Phys Ther*. 2007; 19(2): 148-53.
- 5-Soleimani F, Dadkhah A. Validity and reliability of Infant Neurological International Battery for detection of gross motor developmental delay in Iran. *Child Care Health Dev*. 2007; 33(3): 262-65.
- 6-Levine DA. Guiding parents through behavioral issues affecting their child's health: the primary care provider's role. *Ethn Dis*. 2006; 16(2 Suppl 3): S3-21-8.
- 7-Wagner J, Jenkins B, Smith JC. Nurses' Utilization of Parent Questionnaires for Developmental Screening: Developmental Screening Tools. *Pediatr Nurs*. 2006; 32(5): 409-12.
- 8-Glascoe FP. Early detection of developmental and behavioral problems. *Pediatr Rev*. 2000; 21(8): 272-9; quiz 280.
- 9-Marks K, Hix-Small H, Clark K, Newman J. Lowering Developmental Screening Thresholds and Raising Quality Improvement for Preterm Children. *Pediatrics*. 2009; 123(6): 1516-23.
- 10-King T, Glasco F. Developmental surveillance of infants and young children in pediatric primary care. *Curr Opin Pediatr*. 2003; 15(6): 624-29.
- 11-Sand N, Silverstein M, Glascoe FP, Gupta VB, Tonniges TP, O'Connor KG. Pediatricians' reported practices regarding developmental screening: do guidelines work? *Pediatrics*. 2005; 116(1): 174-9.
- 12-Meade VA, Sweeney JK, Chandler LS, Woodward BJ. Identifying 4-month-old infants at risk in community screening. *Pediatr Phys Ther*. 2009; 21(2): 150-7.
- 13-Soleimani F, Vameghi R, Hemmati S, Salman-Roghani R. Perinatal



- and Neonatal Risk Factors for Neurodevelopmental Outcome in Infants in Karaj. Arch Iran Med. 2009; 12(2): 135-9.
- 14-Soleimani F, Vameghi R, Biglarian A. Antenatal and Intrapartum Risk Factors for Cerebral Palsy in Term and Near-term Newborns. Arch Iran Med. 2013; 16(4): 213-6.
- 15-Glascoe FP. Screening for developmental and behavioral problems. Ment Retard Dev Disabil Res Rev. 2005;11(3):173-9.
- 16-Kliegman R. Nelson textbook of pediatrics. 18th ed. Philadelphia (Pa.): Elsevier Saunders; 2007.
- 17-Turchi RM, Antonelli RC, Norwood KW, Adams RC, Brei TJ, Burke RT, et al. Patient- and Family-Centered Care Coordination: A Framework for Integrating Care for Children and Youth Across Multiple Systems. Pediatrics. 2014;133(5):e1451-e1460.
- 18-Committee on Children With Disabilities. Developmental Surveillance and Screening of Infants and Young Children. Pediatrics. 2001;108(1):192-5.
- 19-Bayley N. Bayley scales of infant and toddler development. San Antonio, Tex: Harcourt Assessment; 2006.
- 20-Acton BV, Biggs WS, Creighton DE, Penner KA, Switzer HN, Thomas JH, et al. Overestimating neurodevelopment using the Bayley-III after early complex cardiac surgery. Pediatrics. 2011; 128(4): e794-800.
- 21-Moore T, Johnson S, Haider S, Hennessy E, Marlow N. Relationship between test scores using the second and third editions of the Bayley scales in extremely preterm children. 2012. J Pediatr. 160(4): 553-8.
- 22-Yu Y-T, Hsieh W-S, Hsu C-H, Chen L-C, Lee W-T, Chiu N-C, et al. A psychometric study of the Bayley Scales of Infant and Toddler Development - 3rd Edition for term and preterm Taiwanese infants. Res Dev Disabil. 2013;34(11):3875-83.
- 23-Suwaibah Z, Chua Bee S, Agnis S, Mohd Sharani A, Shazia Iqbal H. Reliability and Validity for Malay Version of Bayley Scales of Infant and Toddler Development-Third Edition (Bayley-III): Preliminary Study. International Proceedings of Economics Development & Research; 2012, Vol. 40, p116.
- 24-Steenis LJP, Verhoeven M, Hessen DJ, van Baar AL. First steps in developing the Dutch version of the Bayley III: Is the original Bayley III and its item sequence also adequate for Dutch children? European Journal of Developmental Psychology. 2014;11(4):494-511.
- 25-Dworkin PH. 2003 C. Anderson Aldrich award lecture: enhancing developmental services in child health supervision--an idea whose time has truly arrived. Pediatrics. 2004;114(3):827-31.
- 26-Sajedi F, Vameghi R, Keraskian Mojembari A, Habibolahi A, Lornezhad H, Delavar B. [Standardization and validation of the ASQ developmental disorders screening tool in children of Tehran city (Persian)]. Tehran Univ Med J. 2012;70(7):436-46.
- 27-Needelman RD. Developmental assessment. In: Kliegman R. Nelson textbook of pediatrics. 18th ed. Philadelphia (Pa.): Elsevier Saunders; 2007. , pp: 62-6.
- 28-Comrey AL. A first course in factor analysis. New York: Academic Press; 1973.
- 29-Hayati L, Babazadeh M, Solaymanzadeh F, Farrokhi H. [Sound Growth of Persian Speaking Children Aged 6 to 24 Months in Four Sections (Persian)]. Ph.D. Dissertation. 1997. University of Social Welfare and rehabilitation Sciences.
- 30-Mehdipour N, Shirazi TS. [Expressive Vocabulary of Persian Speaking Children 8 to 24 Months. (Persian)]. Final Report of Research Project No. 230. University of Social Welfare and rehabilitation Sciences, Speech Therapy group. 2007.
- 31-Jackson BJ, Needelman H, Roberts H, Willet S, McMorris C. Bayley Scales of Infant Development Screening Test-Gross Motor Subtest: efficacy in determining need for services. Pediatr Phys Ther. 2012;24(1):58-62.
- 32-Katherine TeKolste, Center on Human Development and Disability, University of Washington, Developmental Surveillance and Screening, Monitoring to Promote Optimal Development. Utah Consortium, 5/2004.
- 33-Al-Ansari SS, Bella H. Translation and adaptation of the revised Denver pre-screening developmental questionnaire for Madinah children, Saudi Arabia. Ann Saudi Med. 1998;18(1):42-6.
- 34-Drotar D, Stancin T, Dworkin P. Pediatric developmental screening: understanding and selecting screening instruments. New York, NY: Commonwealth Fund [Internet]. 2008 [cited 2014 Jan 11]; Available from: <http://onceinaweil.net/sites/default/files/abcd/abcd.pediatric.developmental.screening.pdf>
- 35-Oberklaid F, Efron D. Developmental delay--identification and management. Aust Fam Physician. 2005;34(9):739-42.
- 36-Shahshahani S, Vameghi R, Azari A, Sajedi F, Kazemnejad A. [Developmental screening of 4-60 months children in Tehran city using Denver Developmental Screening Test II and Ages and Stages Questionnaires (Persian)]. Journal of Rehabilitation. 2011; 12(3): 65-71.

# Developing of the Persian Version of Bayley Scales of Infant and Toddlers Development Screening Test and Determine its Validity and Reliability

Soleimani F. (M.D.)<sup>1</sup>, \*Azari N. (M.D.)<sup>2</sup>, Kraskian Mojembari A. (Ph.D.)<sup>3</sup>, Vameghi R. (M.D.)<sup>4</sup>, Shahshahanipour S. (M.D.)<sup>2</sup>, Sajedi F. (M.D.)<sup>1</sup>

Receive date: 11/12/2013  
Accept date: 20/02/2014

1-Pediatrician, Associate Professor,  
Pediatric Neurorehabilitation  
Research Center, University of  
Social Welfare and Rehabilitation  
Sciences, Tehran, Iran

2-Pediatrician, Pediatric  
Neurorehabilitation Research Center,  
University of Social Welfare and  
Rehabilitation Sciences, Tehran, Iran

3-Ph.D. in Biostatistician, Assistant  
Professor, Department of Counseling  
and Guidance, Azad Islamic  
University, Karaj Unit, Alborz, Iran

4-Pediatrician, Assistant Professor,  
Pediatric Neurorehabilitation  
Research Center, University of  
Social Welfare and Rehabilitation  
Sciences, Tehran, Iran

\*Correspondent Author Address:  
Pediatric Neurorehabilitation  
Research Center, University of Social  
Welfare and Rehabilitation Sciences,  
Koodakyar St., Daneshjoo Blvd.,  
Evin, Tehran, Iran.

\*Tel: +98 (21) 22180099

\*E-mail: azarinadia@yahoo.com

## Abstract

**Objective:** The aim of the study was to develop a valid and reliable Persian version of the Bayley Scales of Infant and Toddlers Development Screening Test.

**Materials & Methods:** Forward and backward translation, face and content validity determination, and cultural and linguistic adaptations of the questionnaires were performed, consecutively. Clarity of items of the translated version of the scale were examined by a study that carried out on 45 Persian 1- to 42-month-old children, recruited by sampling in Tehran, during the year 2012. In order to determine the reliability; internal consistency, inter-rater and test-retest, 34 tests carried out. For determining the psychometric properties of the tests, later a representative sample of 260, 1- to 42-months-old children recruited through consecutive sampling from health -care centers in five main districts in Tehran.

**Results:** Performing cultural and lingual adaptations, our expert team made some inevitable changes to the test items. Almost in all age groups, cultural or linguistic changes were made in items in communication domains. The Cronbach alpha in cognitive, communication (receptive and expressive), fine and gross motor subtests were 0.96, 0.95, 0.95 and 0.94 respectively. The construct validity of the tests by factor analysis and comparison of mean scores in age groups was confirmed.

**Conclusion:** The Persian version of the Bayley Scales of Infant and Toddlers Development Screening test is a valid and reliable tool.

**Keywords:** Development, Screening, Validation, Child