

تعیین اعتبار و روایی آزمون مقیاس‌های تکاملی شیرخواران و نوبیان Bayley

نادیا آذری^۱ (M.D)، فرین سلیمانی^{۱*} (M.D)، روشنک وامقی^۱ (M.D)، فیروزه ساجدی^۱ (M.D)، سهیلا شهشهانی^۱ (M.D)، حسین کریمی^۱ (M.D)، آدیس کراسکیان^۱ (Ph.D)، امین شاهرخی^۱ (M.D)، رباب تیموری^۱ (Ph.D)، مسعود غریب^۳ (Ph.D Candidate)

۱- مرکز تحقیقات توانبخشی اعصاب اطفال، دانشگاه علوم بهزیستی و توانبخشی، تهران، ایران

۲- گروه مشاوره و راهنمایی، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران

۳- دانشکده پیراپزشکی، دانشگاه علوم پزشکی مازندران، ساری، ایران

چکیده

سابقه و هدف: به منظور تشخیص کودکانی که در آزمون‌های غربالگر، مشکوک به تاخیر تکاملی شناخته می‌شوند، نیازمند استفاده از آزمون‌های تشخیصی تکاملی روا و معتبر هستیم. آزمون Bayley یک آزمون استاندارد طلایی جهانی است که وضعیت تکامل کودکان را در حیطه‌های شناختی، ارتباطی، و حرکتی ارزیابی می‌نماید. هدف از این مطالعه تعیین روایی و اعتبار آزمون فوق بود.

مواد و روش‌ها: روش مطالعه توصیفی - تحلیلی بود. ابتدا آزمون با فرهنگ فارسی منطبق سازی گردید. سپس روایی محتوایی و صوری آن، با استفاده از نظرات خبرگان ارزیابی شد. همچنین تعداد ۲۶۰ کودک ۱ تا ۴۲ ماهه، توسط آزمون فوق مورد ارزیابی تکاملی قرار گرفتند. اعتبار آزمون از طریق سه روش ارزیابی همسانی درونی، اعتبار ارزیاب‌ها و اعتبار باز آزمون و روایی سازه با استفاده از روش‌های تحلیل عاملی و مقایسه میانگین نمرات بررسی گردید.

یافته‌ها: تعداد آزمودنی‌ها ۲۶۰ کودک ۱۵ روزه تا ۴۲ ماهه بودند که از این تعداد ۱۳۴ نفر پسر (۵۱/۵٪) بودند. میانگین ضرایب آلفای کرونباخ برای تمامی حیطه‌ها بالاتر از ۰/۷۶ بود. در روش بازآزمایی، ضریب همبستگی پیرسون در حیطه‌های مختلف، حداقل ۰/۹۸۷ ($P < ۰/۰۰۱$) و در روش اعتبار ارزیاب‌ها ضریب همبستگی پیرسون در حیطه‌های مختلف حداقل ۰/۹۹۱ ($P < ۰/۰۰۱$) گزارش گردید. روایی صوری و محتوایی آزمون با استفاده از نظرات خبرگان تایید شد. همچنین روایی سازه آزمون توسط روش‌های تحلیل عاملی و مقایسه میانگین نمرات تایید گردید.

نتیجه‌گیری: آزمون تشخیصی تکاملی Bayley از اعتبار و روایی مناسبی در کودکان فارسی زبان برخوردار است.

واژه‌های کلیدی: تکامل کودک، شیرخوار، تکرارپذیری نتایج پژوهش، آزمون‌های روانشناسی اعصاب

مقدمه

اختلالات تکاملی از شایع‌ترین مشکلات طب کودکان در سراسر دنیا بوده و روند آن رو به افزایش می‌باشد. یک مطالعه گسترده که در ایالات متحده آمریکا انجام شد نشان داد که در فاصله سال‌های ۲۰۰۶ تا ۲۰۰۸، از هر ۶ کودک یک نفر

دچار اختلالات تکاملی (شامل اختلالات خفیفی چون مشکلات گفتار و زبان یا مشکلات جدی‌تر مانند ناتوانی‌های هوشی، فلج مغزی، و اوتیسم) بوده است [۱]. شیوع کلی اختلالات تکاملی کودکان در ایران حدود ۰/۱۴۶ برآورد شده است [۲]. شیوع اختلالات گفتاری و آسیب ویژه زبانی

ارتباطی درکی، ارتباطی بیانی، حرکات ظریف و حرکات درشت مورد بررسی قرار می‌دهد.

با توجه به این‌که در کشور ما ابزار مناسبی برای تشخیص تاخیر تکاملی در شیرخواران و کودکان کم‌تر از ۴ سال در دسترس مراقبین بهداشتی کشور قرار نداشت، در این مطالعه بر آن شدیم تا ضمن ترجمه و انطباق فرهنگی آزمون معیارهای تکاملی شیرخواران و کودکان Bayley، اعتبار و روایی آن را جهت استفاده در کودکان ایرانی مورد بررسی قرار دهیم.

مواد و روش‌ها

این مطالعه به روش توصیفی - تحلیلی انجام پذیرفت. در این مطالعه از آزمون تشخیصی معیارهای تکاملی شیرخواران و نوپایان Bayley استفاده گردید.

آزمون مقیاس‌های تکاملی شیرخواران و نوپایان Bayley، یک ابزار ارزیابی فردی از عمل‌کرد تکاملی شیرخواران و نوپایان است که حیثه‌های شناختی، ارتباطی درکی، ارتباطی بیانی، حرکات ظریف و حرکات درشت را ارزیابی می‌نماید. این آزمون علاوه بر ترسیم پروفایل عمل‌کرد تکاملی کودک، امکان تشخیص تاخیر تکاملی، تعیین نقاط ضعف و قوت کودک برای استفاده در نقشه درمانی، و طراحی روشی برای پایش پیشرفت تکاملی کودک را فراهم می‌آورد. این آزمون برای کودکان سنین یک تا ۴۲ ماه طراحی شده است [۸].

به منظور تهیه نسخه فارسی از روش ترجمه - بازترجمه (Translation-backtranslation) استفاده گردید. این روش دارای ۶ مرحله شامل بر ترجمه؛ سنتز؛ بازترجمه؛ بازبینی توسط کمیته تحقیق؛ اجرای پایلوت و تهیه فرم نهایی می‌باشد [۹]. ابتدا، فرم سوالات و کتابچه دستورالعمل اجرای آزمون به منظور ترجمه در اختیار گروه تحقیق، که آشنایی کافی با زبان انگلیسی و تکامل کودکان داشتند، قرار داده شد. سپس اعضای دیگری از گروه تحقیق سوالات ترجمه شده را به زبان انگلیسی باز ترجمه نمودند و در نهایت سوالات باز ترجمه شده با متن اصلی انگلیسی سوالات در آزمون توسط کمیته تحقیق مقایسه شد و اشکالات ترجمه استخراج گردید. با

فراون‌تر بوده و تا حدود ۴۴/۳ گزارش شده است [۳].

سازمان بهداشت جهانی تشخیص اختلالات تکاملی کودکان زیر سه سال را در اولویت قرار داده است، زیرا اولاً پاسخ به مداخلات درمانی در این گروه سنی مطلوب‌تر است و ثانیاً به دلیل تواتر بالای مراجعات بهداشتی درمانی در این گروه سنی امکان تشخیص زودرس مشکلات تکاملی در این سنین محتمل‌تر است.

شناسایی عوامل خطر و اثرات متقابل آن‌ها به منظور تشخیص زودرس اختلالات تکاملی امری ضروری برای سلامت کودک و خانواده بوده و یک وظیفه اصلی برای مسئولین مراقبت‌های اولیه و تمامی پزشکان اطفال محسوب می‌شود [۴].

پژوهش‌ها نشان می‌دهند که با تکیه بر قضاوت بالینی و بدون استفاده از ابزار، تنها ۳۰ درصد کودکانی که دارای مشکلات ذهنی، اختلالات یادگیری، گفتاری، و سایر اختلالات تکاملی هستند تشخیص داده می‌شوند. هم‌چنین قضاوت بالینی کم‌تر از ۵۰ درصد مشکلات عاطفی و رفتاری جدی را تشخیص می‌دهد [۵]. بنابراین به منظور یافتن کودکان مبتلا به اختلالات تکاملی که باید مداخلات زودرس دریافت دارند می‌بایست از یک ابزار تشخیصی معتبر استفاده نمود [۶]. یک چالش مهم در تشخیص زودرس اختلالات تکاملی، در دسترس داشتن ابزارهایی است که متناسب با هر منطقه و منطبق بر فرهنگ آن جامعه باشد. ابزارهای موجود در کشورهای غربی برای استفاده در آن کشورها معتبر می‌باشند ولی وقتی این ابزارها در کشورهای دیگر مورد استفاده قرار می‌گیرند محدودیت‌هایی در تفسیر نمرات و سهولت استفاده از ابزار بروز می‌نماید [۷]. به همین دلیل لازم است که اعتبار و روایی این ابزارها در این کشورها مورد بررسی مجدد قرار گیرد و تطبیق فرهنگی متناسب اعمال گردد.

آزمون مقیاس‌های تکاملی شیرخواران و نوپایان Bayley، یک ابزار معتبر جهانی است که عمل‌کرد تکاملی شیرخواران و نوپایان یک تا چهل و دو ماه را در حیثه‌های شناختی،

توسط همان آزمونگر مورد آزمون مجدد قرار گرفتند. هم‌چنین به منظور تعیین اعتبار به روش اعتبار بین ارزیاب‌ها تعداد ۳۴ نمونه از نمونه اصلی هم‌زمان توسط آزمونگر دیگری نمره‌دهی شدند.

نمرات خام به دست آمده در اختیار متخصص آمار قرار داده شد. اعتبار (Reliability) آزمون از طریق سه روش ارزیابی همسانی درونی با استفاده از ضریب آلفای کرونباخ (Cronbach alpha coefficient)، اعتبار ارزیاب‌ها (Inter-rater reliability) و اعتبار باز آزمون (Test-retest reliability) محاسبه گردید. روایی (Validity) سازه آزمون با استفاده از روش‌های تحلیل عاملی (Factor analysis) و مقایسه میانگین نمرات، به شرح ذیل تحلیل گردید.

برای تعیین روایی سازه به روش تحلیل عاملی، در ابتدا مقدار KMO (Kaiser-Meyer-Olkin) در همه حیطه‌ها با هدف تعیین کفایت تعداد نمونه‌ها محاسبه شد. هم‌چنین، به منظور اثبات این که ماتریس همبستگی زیر بنای تحلیل عاملی، در جامعه برابر صفر نیست آزمون کرویت بارتلست (Bartlett's test of sphericity) انجام شد. به منظور پاسخ به این پرسش که آزمون از چه عواملی اشباع است، از روش تحلیل مولفه‌های اصلی استفاده شد. در ادامه، برای تعیین این مطلب که مجموعه مواد تشکیل‌دهنده آزمون مورد مطالعه از چند عامل معنادار اشباع شده، سه شاخص ارزش ویژه، نسبت واریانس تبیین شده توسط هر عامل و نمودار ارزش‌های ویژه (نمودار سنگ‌ریزه‌ای) به کار گرفته شدند.

برای بررسی ارتباط عمل‌کرد در آزمون با سن تقویمی، نمرات حاصل از حیطه‌های پنج‌گانه در گروه‌های سنی آزمون با استفاده از تحلیل واریانس یک طرفه با یک‌دیگر مقایسه شدند و در ادامه برای مقایسه دو به دو میانگین‌ها و مشخص نمودن زیرگروه‌های همگن از آزمون تعقیبی شفه (Scheffé's method) استفاده شد.

پژوهش حاضر با مجوز کمیته اخلاق دانشگاه علوم پزشکی و توان‌بخشی و کسب رضایت آگاهانه از والدین شرکت‌کننده انجام شد.

توجه به این‌که هنگام انجام ترجمه یک آزمون به یک زبان دیگر تفاوت‌های فرهنگی نیز باید لحاظ شوند [۱۰]، در مواردی که توسط کمیته تحقیق لازم دانسته شد، انطباق فرهنگی لازم اعمال گردید. جهت بررسی روایی صوری (Face validity)، محتوایی (Content validity) و روایی سازه (Construct validity) آزمون ترجمه شده [۱۰]، در اختیار ۵ نفر از متخصصین مربوط در هر یک از حیطه‌های تکامل شامل حیطه‌های شناختی، ارتباطی (گفتاری)، حرکتی (ظریف و درشت)، قرار داده شد تا روایی صوری و محتوایی سوالات ارزیابی شوند. اجرای طرح پایلوت بر روی ۳۴ کودک با طیف سنی ۱ تا ۴۲ ماه انجام پذیرفت و پس از استخراج مشکلات اجرایی و اصلاح آن‌ها، فرم نهایی آزمون تهیه گردید.

شایان ذکر است که با توجه به تفاوت‌های فرهنگی و زبانی موجود، عمده‌ترین بخشی که دست‌خوش تغییر گردید، حیطه ارتباطی بود؛ که طی برگزاری جلسات مکرر با متخصصین این حیطه، تغییرات و انطباقات فرهنگی-زبانی لازم اعمال گردید. تغییرات و انطباقات اعمال شده به طور خلاصه شامل جایگزین نمودن بازی‌های متداول ایرانی به جای بازی‌های نامبرده شده در آزمون و جایگزین نمودن بعضی از تصاویر آزمون با تصاویر آشنا برای کودکان ایرانی بود. تغییرات عمده در بخش گفتاری شامل تغییر در سوالات مربوط به ضمایر ملکی، درک حروف اضافه، و زمان‌های افعال با توجه به زبان فارسی بودند.

به منظور بررسی اعتبار و روایی سازه، پس از تهیه ابزار به تعداد لازم و آموزش آزمون‌گران به تعداد ۴ نفر (۳ نفر دارای مدرک کارشناسی ارشد کاردرمانی و یک نفر کارشناس روانشناسی)، و استقرار آزمون‌گران در چند خانه سلامت شهرداری و مرکز بهداشتی درمانی اجرای مرحله اصلی طرح بر روی ۲۶۰ کودک ۱ تا ۴۲ ماه در ۱۷ گروه سنی، آغاز گشت.

به منظور تعیین اعتبار به روش آزمون-باز آزمون تعداد ۴۵ نمونه از نمونه اصلی ظرف ۴ الی ۷ روز پس از آزمون اول

نتایج

تعداد آزمودنی‌ها ۲۶۰ کودک ۱۵ روزه تا ۴۲ ماهه بود که از این تعداد ۱۳۴ نفر پسر (۵۱/۵٪) بودند. روایی صوری و محتوایی آزمون با استفاده از نظرات خبرگان حیطه‌های مختلف تکامل کودکان مورد تایید قرار گرفت.

نتایج بررسی همسانی درونی خرده‌آزمون‌های شناختی، ارتباطی درکی و بیانی، حرکات ظریف و درشت از طریق محاسبه آلفای کرونباخ در جدول ۱ گزارش شده است.

نتایج بررسی پایایی آزمون به روش بازآزمایی در جدول ۲ و نتایج بررسی پایایی آزمون به روش اعتبار ارزیاب‌ها در جدول ۳ گزارش شده است.

در بررسی روایی سازه، مقدار KMO در حیطه‌های مختلف بین ۰/۹۴۷ الی ۰/۹۷۶ گزارش شد که نشانگر کفایت تعداد نمونه‌ها جهت انجام تحلیل عاملی بود. هم‌چنین، آزمون کرویت بارتلت (Bartlett's test of sphericity) به منظور اثبات این که ماتریس همبستگی زیر بنای تحلیل عاملی، در جامعه برابر صفر نیست انجام شد که $p < ۰/۰۰۰۱$ به دست آمد؛ و مجذور خی کفایت نمونه‌برداری و اجرای تحلیل بر پایه

ماتریس‌های همبستگی بین مواد آزمون را تایید نمود. در ادامه به منظور استخراج یک عامل از مجموعه پرسش‌های هر حیطه، ارزش ویژه و درصد واریانس تبیین شده توسط عامل نخست در هر حیطه محاسبه گردید (جدول ۴). با توجه به ارزش ویژه و درصد تبیین شده توسط عامل نخست در هر پنج حیطه چنین برداشت شد که مطلوب‌ترین شرایط برای اجرای تحلیل عاملی در مجموعه پرسش‌های هر حیطه یک مدل تک عاملی می‌باشد.

علاوه بر این در طرح سنگ‌ریزه‌ای آزمون تشخیصی بیلی، به تفکیک حیطه‌های پنج‌گانه، نیز چنین استنباط شد که سهم عامل نخست در واریانس کل پرسش‌های هر حیطه آزمون چشمگیر و از سهم بقیه عامل‌ها متمایز است.

به منظور مقایسه میانگین نمره‌ها در گروه‌های سنی، ابتدا نشان داده شد که فرض صفر مبنی بر برابری میانگین نمره آزمودنی‌ها در ۱۷ گروه سنی با اطمینان ۹۹ درصد رد می‌شود و سپس برای مقایسه دو به دو میانگین‌ها و تعیین گروه‌های همگن از آزمون تعقیبی شفه استفاده شد. آنگاه میانگین نمرات گروه‌های مختلف سنی با هم مقایسه شدند (جدول ۵).

جدول ۱. همسانی درونی (آلفای کرونباخ) تست تشخیصی Bayley به تفکیک خرده آزمون‌ها

گروه	شناختی	ارتباط درکی	ارتباط بیانی	حرکات ظریف	حرکات درشت
میانگین ضرایب	۰/۸۳	۰/۷۶	۰/۸۴	۰/۷۷	۰/۸۰
پسر	۰/۹۸	۰/۹۷	۰/۹۷	۰/۹۷	۰/۹۸
دختر	۰/۹۸	۰/۹۷	۰/۹۸	۰/۹۷	۰/۹۸
تمام سنین	۰/۹۸	۰/۹۷	۰/۹۸	۰/۹۷	۰/۹۸

جدول ۲. ضریب بازآزمایی تست تشخیصی Bayley به تفکیک خرده آزمون‌ها (n=۴۵)

شاخص d کوهن (ضریب همبستگی بین دو نوبت اجرا)	اجرای نخست		اجرای دوّم		پایایی
	SD	M	SD	M	
ضریب همبستگی گشتاوری پیرسون					
شناختی	۲۴/۵۵	۴۷/۹۲	۲۵/۹۵	۴۶/۶۹	**۰/۹۹۳
ارتباط درکی	۱۴/۱۰	۲۳/۳۱	۱۵/۲۴	۲۳/۰۰	**۰/۹۸۷
ارتباط بیانی	۱۷/۱۰	۲۳/۶۲	۱۸/۱۵	۲۳/۰۰	**۰/۹۹۰
حرکات ظریف	۱۶/۵۶	۳۱/۱۷	۱۷/۳۸	۳۱/۸۵	**۰/۹۹۷
حرکات درشت	۲۱/۶۰	۴۳/۰۸	۲۱/۱۴	۴۲/۳۱	**۰/۹۹۷

** P < ۰/۰۰۱

جدول ۳. ضریب اعتبار ارزیاب‌ها برای تست تشخیصی Bayley به تفکیک خرده آزمون‌ها (n=۳۴)

پایایی ارزیاب‌ها (ضریب همبستگی بین نمره‌های دو ارزیاب)	شاخص <i>d</i> کوهن ضریب همبستگی گشتاوری پیرسون	ارزیاب دوم		ارزیاب نخست		
		SD	M	SD	M	
		**۰/۹۹۶	۰/۰۷۵	۱۶/۲۵	۶۱/۹۱	
**۰/۹۹۴	۰/۱۳۳	۱۰/۷۴	۲۸/۷۴	۱۱/۱۳	۲۹/۰۹	ارتباط درکی
**۰/۹۹۸	۰/۰۱۰	۱۲/۶۳	۳۰/۲۶	۱۲/۸۱	۳۰/۲۹	ارتباط بیانی
**۰/۹۹۱	۰/۱۳۰	۱۰/۲۹	۴۰/۸۲	۱۰/۱۸	۴۰/۵۰	حرکات ظریف
**۰/۹۹۷	۰/۰۳۳	۱۱/۲۷	۵۴/۳۵	۱۱/۰۲	۵۴/۲۶	حرکات درشت

** $P < ۰/۰۰۱$

جدول ۴- ارزش ویژه و درصد واریانس تبیین شده توسط عامل نخست با اجرای تحلیل عاملی مولفه‌های اصلی

حیطه	ارزش ویژه عامل نخست	درصد واریانس تبیین شده توسط عامل نخست
شناختی	۳۳/۸۱۷	۳۷/۱۶۱
ارتباطی درکی	۱۹/۷۵۸	۴۰/۲۳۲
ارتباطی بیانی	۲۲/۹۸۷	۴۷/۸۸۹
حرکات ظریف	۲۲/۶۶۲	۳۴/۳۶
حرکات درشت	۳۰/۴۶۵	۴۲/۱۲

جدول ۵- نتایج تجزیه و تحلیل واریانس برای مقایسه میانگین نمرات تست تشخیصی Bayley به تفکیک خرده آزمون‌ها در هفده گروه سنی

گروه‌های سنی	شناختی		ارتباط درکی		ارتباط بیانی		حرکات ظریف		حرکات درشت	
	SD	M	SD	M	SD	M	SD	M	SD	M
A	۳/۱۷	۵/۵۶	۱/۶۹	۳/۶۱	۰/۷۰	۱/۶۱	۱/۸۵	۳/۶۱	۱/۴۳	۳/۹۴
B	۲/۹۹	۸/۷۰	۱/۱۴	۴/۴۰	۱/۲۵	۳/۷۵	۱/۶۸	۵/۵۳	۳/۲۰	۸/۴۵
C	۱۰/۷۶	۱۴/۳۸	۲/۸۲	۶/۶۲	۳/۰۷	۵/۵۴	۸/۱۹	۸/۸۵	۸/۵۱	۱۳/۴۶
D	۴/۰۳	۱۴/۷۴	۱/۱۸	۷/۰۵	۱/۱۷	۵/۳۴	۸/۲۶	۱۲/۷۴	۲/۷۵	۱۶/۴۵
E	۲/۹۵	۲۳/۴۰	۲/۴۴	۸/۸۰	۱/۷۰	۶/۷۰	۲/۵۷	۱۷/۸۰	۳/۰۱	۲۰/۲۰
F	۳/۳۴	۲۶/۱۵	۱/۹۲	۹/۶۵	۲/۲۳	۷/۱۲	۲/۶۱	۱۹/۶۲	۳/۵۰	۲۴/۶۲
G	۲/۰۱	۳۰/۱۰	۱/۸۰	۱۱/۰۳	۲/۶۲	۹/۲۵	۱/۸۹	۲۱/۶۹	۴/۴۱	۲۹/۶۹
H	۳/۱۰	۳۳/۹۷	۱/۸۱	۱۲/۰۶	۲/۵۷	۹/۸۴	۲/۲۳	۲۵/۸۷	۴/۳۴	۳۴/۶۸
I	۳/۷۹	۴۰/۷۱	۲/۵۶	۱۴/۹۷	۳/۱۳	۱۴/۱۶	۲/۵۲	۲۸/۰۶	۴/۳۷	۴۲/۶۳
J	۳/۹۱	۴۸/۹۱	۲/۹۰	۱۹/۴۹	۳/۳۸	۱۹/۵۴	۲/۴۵	۳۲/۳۴	۳/۶۴	۵۰/۷۱
K	۴/۲۰	۵۲/۴۷	۲/۹۷	۲۰/۷۸	۴/۲۷	۲۱/۲۸	۲/۲۴	۳۴/۵۳	۳/۴۶	۵۲/۱۹
L	۴/۰۶	۵۸/۵۰	۲/۹۳	۲۵/۲۰	۳/۲۷	۲۶/۸۵	۳/۳۵	۳۷/۸۵	۲/۱۶	۵۵/۴۰
M	۶/۰۷	۶۲/۵۰	۳/۸۶	۳۰/۲۲	۶/۵۰	۳۲/۷۸	۳/۰۱	۴۰/۱۱	۴/۰۶	۵۷/۲۸
N	۵/۵۷	۶۵/۸۲	۵/۱۹	۳۱/۷۱	۷/۴۰	۳۱/۷۶	۴/۴۱	۴۱/۶۷	۳/۶۳	۵۸/۸۱
O	۵/۵۴	۷۰/۳۹	۴/۲۱	۳۵/۲۶	۶/۱۹	۳۸/۵۲	۱۰/۱۹	۴۲/۸۷	۳/۳۷	۶۱/۲۷
P	۷/۸۰	۷۰/۷۹	۴/۴۲	۳۸/۲۴	۴/۷۶	۴۱/۴۱	۴/۰۳	۵۰/۳۳	۴/۴۵	۶۳/۰۰
Q	۴/۶۹	۷۶/۸۳	۴/۴۱	۴۲/۳۹	۷/۰۱	۴۳/۰۰	۱۲/۲۰	۵۱/۱۷	۲/۶۶	۶۶/۵۰
(۳۸۶,۱۶F)	**۳۴۸/۷۳	**۳۷۴/۳۵	**۲۵۹/۷۷	**۱۶۷/۹۳	**۶۲۴/۷۹					

** $P < ۰/۰۱$

بحث و نتیجه گیری

سازمان بهداشت جهانی تشخیص زودرس اختلالات تکاملی کودکان را در اولویت قرار داده است [۱۱]؛ زیرا مداخلات زودرس به عنوان کلید اساسی برای به حداقل رساندن اثرات درازمدت تاخیر تکاملی شناخته شده‌اند [۱۲].

استفاده از ابزارهای استاندارد ارزیابی تکاملی در تشخیص زودهنگام تاخیر تکاملی در کودکان اهمیت ویژه‌ای دارد. همچنین این ابزارها در ارزیابی اثربخشی مداخلات به کار برده می‌شوند [۱۳].

با توجه به نیاز سیستم بهداشتی کشور در جهت دسترسی به یک آزمون استاندارد ارزیابی تکاملی کودکان، آزمون Bayley جهت استانداردسازی برای کودکان ایرانی انتخاب شد. این آزمون در کشور آمریکا طراحی و هنجاریابی شده و به عنوان استاندارد طلایی ارزیابی تکاملی شیرخواران شناخته شده است [۱۴].

ابزارهای موجود در کشورهای غربی برای استفاده در آن کشورها معتبر می‌باشند ولی وقتی این ابزارها در کشورهای دیگر مورد استفاده قرار می‌گیرند محدودیت‌هایی در تفسیر نمرات و سهولت استفاده از ابزار بروز می‌نماید [۷]. به همین دلیل لازم است که اعتبار و روایی این ابزارها در این کشورها مورد بررسی مجدد قرار گیرد و تطبیق فرهنگی متناسب اعمال گردد.

مطالعه حاضر جهت تعیین روایی و اعتبار آزمون تشخیصی Bayley صورت پذیرفت و یافته‌های ما در این پژوهش، روایی و اعتبار نسخه فارسی آزمون Bayley را تایید نمود.

روایی محتوایی و صوری نسخه فارسی آزمون Bayley، توسط متخصصین حیطه‌های تکاملی مورد تایید واقع شد.

با توجه به نتایج ذکر شده در بخش یافته‌ها، که در آن ضریب آلفای کرونباخ در بررسی همسانی درونی، بالاتر از ۰/۷۶ و ضرایب همبستگی پیرسون (در روش‌های بازآزمایی و اعتبار بین ارزیاب‌ها) بالاتر از ۰/۹۸۷ ($P < ۰/۰۰۱$) گزارش شدند، اعتبار آزمون مورد تایید واقع می‌شود.

با توجه به نتایج به دست آمده در روش تحلیل عاملی، چنین استنتاج می‌شود که بهترین ساختار برای این آزمون، یک مدل تک عاملی می‌باشد. به عبارت دیگر روایی این آزمون جهت سنجش وضعیت تکاملی شیرخواران ۱ تا ۴۲ ماهه شهر تهران مورد تایید واقع می‌شود.

توجه به نتایج به دست آمده در روش مقایسه میانگین نمرات، نشان می‌دهد که میانگین نمرات محاسبه شده در گروه‌های سنی مختلف، تفاوت معنی‌داری داشته و نشان‌دهنده وجود رابطه بین سن و نمره‌های آزمون در حیطه‌های پنج‌گانه می‌باشد؛ به شکلی که با افزایش سن، نمره آزمودنی‌ها در هر پنج حیطه افزایش می‌یابد. این یافته‌ها فرضیه ارائه شده مبنی بر مرتبط بودن عمل‌کرد در این آزمون با سن و در نتیجه روایی سازه آزمون را مورد تایید قرار می‌دهد.

یافته‌های پژوهش حاضر نتایج به دست آمده توسط Godamunne P که اعتبار ابزار Bayley را به روش آزمون - بازآزمون، بررسی نمود و نشان داد که آزمون از اعتبار بسیار بالایی برخوردار می‌باشد [۱۵]، را مورد تایید قرار می‌دهد.

همچنین نتایج این مطالعه همسو با مطالعه Deroma L می‌باشد که اعتبار آزمون Bayley را به روش بررسی توافق بین دو آزمونگر (روانشناس) آموزش دیده مورد ارزیابی قرار داد و نتیجه‌گیری نمود که اعتبار آزمون Bayley به روش اعتبار بین ارزیاب‌ها، بسیار بالا بوده و نتایجی که توسط دو آزمونگر ارائه می‌گردند با یکدیگر قابل قیاس می‌باشند [۱۶].

بالا بودن اعتبار نسخه مالزیایی این آزمون هم‌چنین توسط Suwaibah با استفاده از بررسی ثبات داخلی و محاسبه آلفای کرونباخ نشان داده شد [۱۷]. او برای تهیه ترجمه مالزیایی آزمون از روش ترجمه - بازترجمه که توسط Brislin ارائه شده استفاده نمود که روش و نتایج این مطالعه نیز همسو با یافته‌های مطالعه موجود می‌باشند. وی هم‌چنین روایی آزمون را به روش همگرایی بین خرده‌مقیاس‌ها بررسی نمود و نشان داد که بین ۵ خرده‌مقیاس، همبستگی مثبت و بالایی موجود می‌باشد [۱۷].

استفاده نشده است؛ که می‌تواند به صورت یک پیشنهاد پژوهشی برای سایر محققین این حیطه مطرح شود.

یکی از چالش‌های موجود تعیین سطح تکامل کودکان با نیازهای ویژه است. استفاده از آزمون Bayley برای برآورد سطح تکاملی در کودکان با نیازهای ویژه با در نظر گرفتن ملاحظات خاصی امکان‌پذیر می‌باشد، Linda visser در سال ۲۰۱۴، به منظور تهیه یک ابزار مناسب جهت ارزیابی تکاملی کودکان با نیازهای ویژه، آزمون Bayley را به زبان هلندی ترجمه نمود و مناسب بودن و روایی فرم تطابق داده شده این آزمون برای کودکان با نیازهای ویژه را مورد بررسی قرار داد. نتیجه مطالعات وی نشان داد که انطباقات انجام شده باعث ارتقاء روایی آزمون در کودکانی شد که اختلال تکاملی خفیف تا متوسط حرکتی یا بینایی داشتند؛ و این افزایش روایی خصوصاً در حیطه شناختی بیش‌تر مشهود بود [۲۲].

آزمون Bayley دارای یک نسخه غربالگر نیز می‌باشد که فرم کوتاه شده نسخه تشخیصی این آزمون بوده و برای غربالگری اختلالات تکاملی استفاده می‌گردد. ترجمه فارسی و بررسی اعتبار و روایی آزمون غربالگر Bayley نیز توسط سلیمانی و همکاران صورت پذیرفته؛ که در آن تهیه نسخه فارسی آزمون غربالگر بیلی از روش‌های مشابه این مطالعه به منظور بررسی روایی و اعتبار آزمون غربالگر فارسی شده استفاده شد؛ و در نهایت روایی و اعتبار آن مورد تایید قرار گرفته است [۲۳].

از محدودیت‌های آزمون معیارهای تکاملی شیرخواران و نوپایان Bayley می‌توان به طولانی بودن زمان اجرای آزمون، نیاز به انجام آزمون توسط آزمونگران دوره دیده و در نتیجه هزینه نسبتاً بالای آن اشاره نمود.

در پایان، نتایج این مطالعه نشان می‌دهد که نسخه فارسی آزمون تشخیصی بیلی، از اعتبار و روایی قابل قبولی برخوردار می‌باشد که استفاده از آن را در کودکان ۱ تا ۴۲ ماهه فارسی زبان مقدور می‌سازد.

امید است استفاده از این آزمون خلاء وجود یک آزمون تشخیصی تکاملی به منظور ارزیابی دقیق‌تر کودکانی که در

نتایج مطالعه موجود با نتایج مطالعه Yu YT، که اعتبار آزمون Bayley را با روش آزمون - بازآزمون و روش اعتبار ارزیابی‌ها) خوب تا عالی گزارش نمود همسو می‌باشد [۱۸]. او روایی آزمون را از روش روایی هم‌زمان (Concurrent validity) و بررسی همبستگی این آزمون با نسخه دوم آن مورد تایید قرار داد. در پژوهش حاضر، روایی آزمون به روش‌های تحلیل عاملی و مقایسه میانگین نمرات صورت پذیرفته است.

Yu YT، در مطالعه خود نشان داد که نمرات به دست آمده از ویرایش سوم Bayley سطح تکامل کودکان تابوانی را بالاتر از ویرایش دوم Bayley ارزیابی می‌نماید. وی پیشنهاد نمود که نقاط برش ویرایش سوم Bayley باید بالاتر رود تا تاخیر تکاملی را نشان دهد [۱۸].

مطالعات دیگری نیز به این مورد اشاره نموده‌اند که ارزیابی سطح تکاملی کودک با استفاده از ویرایش سوم آزمون Bayley در مقایسه با ویرایش دوم آن ممکن است منجر به برآورد بالاتری از سطح تکاملی کودک و یا برآورد پایین‌تر در تشخیص اختلالات تکاملی گردد؛ و این امر بیش‌تر در حیطه شناختی آن آزمون ملاحظه می‌گردد [۲۰، ۱۹]؛ این مطلب می‌تواند موضوعی برای پژوهش‌های آتی محققین ایرانی باشد. روش دیگر بررسی روایی یک آزمون، استفاده از روایی همگرا و واگرا می‌باشد؛ که در آن دو آزمون به صورت هم‌زمان روی آزمودنی‌ها اجرا می‌شوند و نتایج آن‌ها با هم مقایسه می‌شود. Campbell SK، روایی آزمون Bayley را به روش بررسی روایی همگرا و واگرا بین این آزمون و Test of Infant Motor Performance (TIMP)، مورد بررسی قرار داد و همبستگی ضعیف تا متوسطی بین حیطه‌های مختلف Bayley و TIMP مشاهده نمود؛ وی نتیجه گرفت که TIMP برای ارزیابی زودرس شیرخواران مناسب‌تر است [۲۱]. با توجه به محدودیت نسخه‌های فارسی آزمون‌های تشخیصی تکاملی، در پژوهش حاضر از روش بررسی روایی همگرا (مقایسه هم‌زمان نتایج دو آزمون) به منظور تعیین روایی سازه

[10] Hunt SM, Bhopal R. Self report in clinical and epidemiological studies with non-English speakers: the challenge of language and culture. *J Epidemiol Commun Health* 2004; 58: 618-622.

[11] Gladstone M, Lancaster GA, Jones A, Maleta K, Mtitimila E, Ashorn P, et al. Can Western developmental screening tools be modified for use in a rural Malawian setting? *Arch Dis Child* 2008; 93: 23-29.

[12] Lipkin PH, Cartwright JD, Desch LW, Duby JC, Elias ER, Johnson CP, et al. Role of the medical home in family-centered early intervention services. *Pediatrics* 2007; 120: 1153-1158.

[13] Johnson S, Marlow N. Developmental screen or developmental testing? *Early Human Develop* 2006; 82: 173-183.

[14] Harris SR, Megens AM, Backman CL, Hayes VE. Stability of the bayley II scales of infant development in a sample of low-risk and high-risk infants. *Dev Med Child Neurol* 2005; 47: 820-823.

[15] Godamunne P, Liyanage C, Wimaladharmasooriya N, Pathmeswaran A, Wickremasinghe AR, Patterson C, et al. Comparison of performance of Sri Lankan and US children on cognitive and motor scales of the Bayley scales of infant development. *BMC Res notes* 2014; 7: 300.

[16] Deroma L, Bin M, Tognin V, Rosolen V, Valent F, Barbone F, et al. [Interrater reliability of the Bayley III test in the Italian Northern-Adriatic Cohort II]. *Epidemiol Prev* 2013; 37: 297-302.

[17] Zakaria S, Seok CB, Sombuling A, Ahmad MS, Hashmi SI. Reliability and validity for malay version of bayley scales of infant and toddler development-(bayley-III): preliminary study. *Int Proc Economic Dev Res* 2012; 40.

[18] Yu YT, Hsieh WS, Hsu CH, Chen LC, Lee WT, Chiu NC, 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: 3875-3883.

[19] 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: e794-e800.

[20] Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW. Underestimation of developmental delay by the new Bayley-III Scale. *Arch Pediatr Adolesc Med* 2010; 164: 352-356.

[21] Campbell SK, Zawacki L, Rankin KM, Yoder JC, Shapiro N, Li Z, et al. Concurrent validity of the TIMP and the Bayley III scales at 6 weeks corrected age. *Pediatric physical therapy: the official publication of the Section on Pediatrics of the American Physical Therapy Association*. 2013; 25: 395-401.

[22] Visser L, Ruiter SA, Van der Meulen BF, Ruijsenaars WA, Timmerman ME. Validity and suitability of the Bayley-III Low Motor/Vision version: A comparative study among young children with and without motor and/or visual impairments. *Res Dev Disabil* 2013; 34: 3736-3745.

[23] Soleimani F, Azari N, Kraskian-Mojembari A, Vameghi R, Shahshahani-Pour S, Sajedi F. Developing of the persian version of bayley scales of infant and toddlers development screening test and determine its validity and reliability. *Quart J Rehab* 2014; 14: 18-29.

آزمون‌های غربالگری مشکوک به تاخیر تکاملی شناخته شده‌اند، و نیز خلا وجود یک آزمون استاندارد طلائی به منظور مقایسه نتایج سایر آزمون‌های تکاملی با آن را برطرف نماید.

تشکر و قدردانی

از مرکز تحقیقات توان‌بخشی اعصاب اطفال و معاونت پژوهشی دانشگاه علوم بهزیستی و توان‌بخشی به دلیل حمایت مالی و معاونت‌های بهداشتی دانشگاه‌های علوم پزشکی و معاونت سلامت شهرداری تهران جهت استفاده از مراکز و از کلیه مراقبین کودکان شرکت‌کننده در طرح تقدیر و قدردانی می‌شود.

منابع

- [1] Boyle CA, Boulet S, Schieve LA, Cohen RA, Blumberg SJ, Yeargin-Allsopp M, et al. Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics* 2011; 127: 1034-1042.
- [2] Sajedi F, Doulabi MA, Vameghi R, Baghban AA, Mazaheri MA, Mahmodi Z, et al. Development of children in Iran: a systematic review and meta-analysis. *Global J Health Sci* 2015; 8: 145.
- [3] Mohammadi M, Sadollahi A, Ghorbani R. Prevalence of specific language impairment in 5 year-old children of an Iranian. *Koomesh* 2014; 15: 182-190.
- [4] Zare Delavar S, Bakhshi E, Soleimani F, Biglarian A. Application of penalized logistic regression for detecting risk factors interactions of children with cerebral Palsy. *Koomesh* 2015; 16: 159-165.
- [5] Glascoe FP. Early detection of developmental and behavioral problems. *Pediatr Rev* 2000; 21: 272-280.
- [6] Wu YT, Tsou KI, Hsu CH, Fang LJ, Yao G, Jeng SF. Brief report: Taiwanese infants' mental and motor development--6-24 months. *J Pediatr Psychol* 2008; 33: 102-108.
- [7] Abubakar A, Holding P, Van Baar A, Newton C, van de Vijver FJ. Monitoring psychomotor development in a resourcelimited setting: an evaluation of the Kilifi Developmental Inventory. *Ann Trop Paediatr* 2008; 28: 217-226.
- [8] Bayley N, Reuner G. Bayley scales of infant and toddler development: Bayley-III: Harcourt Assessment, Psych. Corporation; 2006.
- [9] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *SPINE* 2000; 25: 3186-3191.

Determination of validity and reliability of the Bayley scales in infant and toddler development

Nadia Azari (M.D)¹, Farin Soleimani (M.D)^{*1}, Roshanak Vameghi (M.D)¹, Firouzeh Sajedi (M.D)¹, Soheila Shahshahani (M.D)¹, Hossein Karimi (M.D)¹, Adis Kraskian (Ph.D)², Amin Shahrokhi (M.D)¹, Robab Teymouri (Ph.D)¹, Masoud Gharib (Ph.D Candidate)¹

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

2 – Dept. of counseling & guidance, Karaj Branch, Islamic Azad University (IAU), Karaj, Iran

3- Faculty of Paramedicine, Mazandaran University of Medical Sciences, Sari, Iran

(Received: 6 Mar 2016; Accepted: 27 Jul 2016)

Introduction: For diagnosing children who are suspected to have developmental delay through developmental screening tests, we need a valid and reliable diagnostic tool. The Bayley scales is a well-known diagnostic developmental assessment on cognitive, communication and motor domains. The aim of this study was validity and reliability determination of the Bayley test.

Materials and Methods: The method of this study was descriptive-analytic. The test was provided through translation- back translation and cultural adaptation. Content and face validity of tool was determined by experts' opinions. 260 children aged 1 to 42 months were recruited for developmental assessment by Bayley. Reliability of test was calculated through three methods; internal consistency, test-retest and inter-rater. Construct validity was calculated using factor analysis and comparison of the mean scores methods.

Results: Participants were 260 children 15 days to 42 months, including 134 boys (51.5%). Cronach's alpha coefficients were more than 0.76 for all domains. Pearson correlation coefficient in different domains, were at least 0.987 ($P < 0.001$) in test retest method; and 0.991 ($P < 0.001$) in inter-rater method. The face and content validity of test was approved by experts in child development. Construct validity of test was approved through factor analysis and statistically significant differences were observed between mean scores of different age groups, that confirms the validity of test.

Conclusion: These results indicated that the Bayley is a valid and reliable tool for developmental assessment in Persian children.

Keywords: Child Development, Infant, Reproducibility of Results, Neuropsychological Tests

* Corresponding author. Tel: +98 21 22180099

Soleimani_farin@yahoo.com