

# ویژگی‌های روان‌سنجی باورهای کارآمدی در تدریس ریاضی با استفاده از مدل اعتبار پاره‌ای تعمیم یافته

نورعلی فرخی<sup>۱</sup>

منصور حرفه دوست<sup>۲</sup>

هیوا محمد جانی<sup>۳</sup>

ثریا جباری<sup>۴</sup>

فرزانه رعایی<sup>۵</sup>

تاریخ پذیرش: ۹۳/۷/۵

تاریخ وصول: ۹۳/۱/۲۱

## چکیده

زمینه: مقیاس باورهای کارآمدی در تدریس ریاضیات توسط هونکر و اینوچ (۱۹۹۵) توسعه داده شده است. با این حال، نتایج پژوهش‌ها در باب ویژگی‌های روان‌سنجی این مقیاس متناقض است. هدف: پژوهش حاضر به منظور بررسی ویژگی‌های روان‌سنجی مقیاس باورهای کارآمدی در تدریس ریاضی با استفاده از مدل اعتبار پاره‌ای تعمیم یافته، یکی از مدل‌های نظریه سؤال-پاسخ برای داده‌های چند ارزشی، بر روی نمونه‌ای از معلمان اجرا شد. روش: تعداد ۳۳۲ نفر از معلمان دوره ابتدایی شهر تبریز به طور تصادفی انتخاب و به مقیاس باورهای کارآمدی معلم در تدریس ریاضیات (هونکر، اینوچ، ۱۹۹۵) پاسخ گفتند. یافته‌ها: ضریب آلفای کرونباخ برای دو خرده مقیاس انتظار پیامد، و باورهای کارآمدی به ترتیب برابر با ۰/۸۲ و ۰/۸۵ بود. نتایج تحلیل مؤلفه‌های اصلی و تحلیل عاملی تاییدی دال بر تک بعدی بودن هر دو خرده مقیاس بود. علی‌رغم شاخصه‌های آماری قابل قبول برای مقیاس، نتایج مدل اندازه‌گیری اعتبار پاره‌ای تعمیم یافته دال بر مشکلات عمده در سؤالات هر دو خرده

۱- استاد گروه سنجش و اندازه‌گیری دانشگاه علامه طباطبائی (farrokhinoorali@yahoo.com)

۲- دانشجوی دکتری سنجش و اندازه‌گیری دانشگاه علامه طباطبائی

۳- کارشناسی ارشد روانشناسی دانشگاه علوم پزشکی کردستان

۴- کارشناس ارشد پژوهش‌ات آموزشی دانشگاه شهید مدنی آذربایجان

۵- دانشجوی دکتری سنجش و اندازه‌گیری دانشگاه علامه طباطبائی

مقیاس بود. نتایج پژوهش نشان داد که ضرایب دشواری سؤالات همگی منفی، یعنی سؤالات آسان بودند؛ همچنین تابع آگاهی سؤالات دال بر این بود که برای طیف توانایی بالاتر از متوسط، سؤالات آگاهی بسیار پایینی دارند و تنها در طیف توانایی پایین، سؤالات آگاهی بخش بودند. بحث: نظر به اهمیت سنجش خود کارآمدی تدریس ریاضی ساخت ابزاری جدید و کارآمد با ویژگی‌های روان‌سنجی خوب توصیه می‌شود.

**واژگان کلیدی:** باورهای کارآمدی تدریس ریاضی، مدل اعتبار پاره‌ای تعمیم یافته، مدل‌های چندارزشی نظریه سؤال-پاسخ.

### مقدمه

یکی از مهمترین سازه‌هایی که در ادبیات پژوهش و در پژوهش‌های اخیر توجه خاصی به آن مبذول شده است، خود کارآمدی معلم است (برای مثال سوارز، دان و گيسن<sup>۱</sup>، ۲۰۰۶؛ سوارز، ۲۰۰۵؛ اوتلی، برایانت، موسلی<sup>۲</sup>، ۲۰۰۵؛ روس و بروس<sup>۳</sup>، ۲۰۰۷؛ کلاسن و چيو<sup>۴</sup>، ۲۰۱۰؛ وارد<sup>۵</sup>، ۲۰۰۹). براساس نظریه شناختی اجتماعی<sup>۶</sup>، خود کارآمدی معلم به صورت باورهای فردی معلم به توانایی فرد در برنامه‌ریزی، سازمان دهی، و انجام فعالیت‌هایی تعریف می‌شود که برای پیشبرد اهداف آموزشی ضرورت دارند (اسکالویک و اسکالویک<sup>۷</sup>، ۲۰۰۷). پژوهش‌های قبلی نشان داده‌اند که باورهای خود کارآمدی معلم در عملکرد و انگیزش آنها تأثیر بسزایی دارد (بندورا<sup>۸</sup>، ۱۹۹۷؛ چانان-موران، ولفولک و هوی<sup>۹</sup>، ۲۰۰۱؛ چانان-موران و ولفولک و هوی؛ ۱۹۹۸؛ ولفولک و هوی، ۲۰۰۴). در مقایسه با معلمانی که به کارآمدی خود شک می‌ورزند، معلمان خود کارآمد تمایل زیادی بر مشارکت در فعالیت‌های مدرسه، همکاران و دانش آموزان

1. Swars, Daane, & Giesen
2. Utley, Bryant, & Moseley
3. Ross, & Bruce
4. Klassen, R. M. & Chio
5. Ward
6. Bandura's social cognitive theory
7. Skaalvik
8. Bandura
9. Tschannen-Moran, Woolfolk Hoy, & Hoy

از خود نشان می‌دهند، و مدرسه را به عنوان سیستمی تلقی می‌کنند که قادر است هدف و رسالت آنها را به فرجام نیکی سوق دهد (کاپارا، باربرانلی، بورگوگنی، و استیکا، ۲۰۰۳). یافته‌های اخیر نشان داده‌اند که باورهای خودکارآمدی معلمان نقش اساسی در تعهد و رضایت شغلی معلمان دارد (کاپارا و همکاران، ۲۰۰۳). پژوهش‌ها نشان داده‌اند که معلمان دارای خودکارآمدی پایین مشکلات زیادی را در تدریس و سطوح بالایی از استرس‌های مربوط به شغل و سطوح پایین رضایت شغلی را تجربه می‌کنند (بیتورت<sup>۲</sup>، ۲۰۰۶؛ کلاسن<sup>۳</sup> و همکاران، ۲۰۰۹).

باورهای کارآمدی تلویحات مهمی در امر آموزش دارند. کارآمدی معلم سازه مهمی است که پیش‌بینی کننده واقعی تدریس اثر بخش می‌باشد (کلاسن و چیو، ۲۰۱۰). درک معلم از کارآمدی خود، قضاوتی در مورد ظرفیت او برای به دست آوردن پیامدهای دلخواه از یادگیری و به اشتیاق واداشتن دانش آموز، حتی در بین دانش آموزان بی انگیزه و دشوار، را فراهم می‌سازد. (چانان-موران و هوی، ۲۰۰۱). باورهای معلم زمانی که تثبیت می‌شوند، نسبت به تغییر مقاومت نشان می‌دهند (چانن-موران، ولفولک هوی، و هوی (۱۹۹۸). اما این باورها در معلمانی که هنوز تدریس را شروع نکرده‌اند تغییرپذیر هستند (هوی، و اسپيرو، ۲۰۰۵).

چون کارایی معلم وابسته به بافت (بندورا، ۱۹۸۶)، وابسته به حوزه تحصیلی خاص (چانان-موران و همکاران، ۱۹۹۸)، است. بنابراین فرض می‌شود سازه‌ای قدرتمند در پیش‌بینی توانایی آتی برای تدریس در حوزه خاص باشد. در نظریه شناختی-اجتماعی بندورا (۱۹۷۷) خودکارآمدی<sup>۴</sup>، سازه‌ای دو بعدی است که شامل انتظار پیامد<sup>۵</sup> و خودکارآمدی می‌شود. انتظار پیامد، به این باور برمی‌گردد که رفتار خاصی به پیامدهای مشخصی منجر می‌شود و

1. Caprara, , Barbaranelli, Borgogni, &Steca
2. Betoret
3. Klassen, Bong, Usher, Chong, Huan, Wong, Georgiou
4. self-efficiency
5. outcome expectancy

خودکارآمدی، اعتقاد راسخی است که یک فرد به طور موفقیت آمیزی برای رفتار مورد نیاز دارد. در زمینه تدریس، انتظار پیامد تدریس بر مبنای این باور است که تدریس اثربخش، پیامدهای یادگیری مثبتی را موجب می شود و خودکارآمدی معلم باور معلم برای اثربخش بودن تدریس را می رساند.

با توجه به اهمیت باورهای کارآمد در حوزه تدریس ریاضی و نظر به اینکه باورهای کارآمد در تدریس وابسته به حوزه بوده، وجود ابزاری برای سنجش باورهای کارآمد ریاضی ضرورت دارد. نگاهی به ادبیات پژوهش نشان می دهد که تنها یک مقیاس برای سنجش باورهای تدریس کارآمد ریاضی در معلمان دوره ابتدایی وجود دارد که توسط هوینکر و اینوچ<sup>۱</sup> (۱۹۹۵) توسعه داده شده است. در حال حاضر، این مقیاس تنها ابزار برای سنجش کارآمدی تدریس ریاضی است (وارد، ۲۰۰۹). اما نگاهی به ادبیات پژوهش نشانگر این واقعیت مهم است که مطالعات روان سنجی محدودی بر روی این مقیاس انجام شده است. از محدود مطالعات انجام شده در مورد ویژگی های روان سنجی این مقیاس مربوط به مطالعه اینوچ، اسمیت<sup>۲</sup> و هوینکر (۲۰۰۰) است، که به بررسی ضرایب تمیز، پایایی نمرات، و روایی سازه این ابزار پرداختند. این محققان با استفاده از تحلیل عاملی تاییدی دو عامل نسبتاً مستقل انتظارات از پیامد و خود کارآمدی را استخراج کردند، پایایی این دو خرده مقیاس به ترتیب برابر با ۰/۷۷ و ۰/۸۸ بود.

اما پژوهش انجام شده توسط اینوچ، اسمیت، هوینکر (۲۰۰۰) از چندین محدودیت اساسی برخوردار بود. مهمترین محدودیت این پژوهش این بود که آنها به منظور تأمین برآزش دادن مدل با داده ها از یک قیدی استفاده کرده بودند که واریانس خطای دو سؤال را به هم وصل می کردند. این امر موجب شد که یکی از مهمترین مفروضه های آمار، یعنی استقلال موضعی، مورد مخاطره قرار بگیرد. استقلال موضعی از این رو اهمیت دارد که بدون رعایت این پیش

- 
1. Huinker & Enochs
  2. Smith

فرض، پایایی و آگاهی آزمون احتمالاً به صورت جدی بیش برآورد می‌شود (واینر و تیسن<sup>۱</sup>، ۱۹۹۶)، و ضرایب تمیز سؤالات، دشواری سؤالات، و برآوردهای مربوط به توانایی شخص احتمالاً سوگیرانه و غلط باشند (تیسن، استینبرگ، و مونی<sup>۲</sup>، ۱۹۸۹).

در پاسخ به این محدودیت کیفیت‌بند، ناتسان، و ادی<sup>۳</sup> (۲۰۱۱) با استفاده از نظریه سؤال-پاسخ با رویکردی بیزی به بررسی ویژگی‌های روان‌سنجی سؤالات این ابزار با استفاده از نظریه سؤال و پاسخ پرداختند. یافته این محققان دال بر این بود که پایایی آزمون طبق آنچه که قبلاً متصور بود زیاد بالا نیست. مقیاس، واژه پردازی، و جایگاه سؤالات نیاز به تجدید نظر دارد. این محققان توصیه کردند که سؤالات دیگری برای اندازه‌گیری سطوح زیر متوسط کارآمدی باید ساخته شوند. با مرتب کردن سؤالات براساس دشواری، این محققان به توصیف ساختار باورهای کارآمدی تدریس ریاضی پرداختند و دلالت‌هایی برای معلمان تدریس ریاضی ارائه کردند.

به جز این دو مطالعه تاکنون پژوهش‌های بسیار معدودی در زمینه ویژگی‌های روان‌سنجی این مقیاس صورت گرفته است. از معدود مطالعاتی که در این حوزه صورت گرفته‌اند (برای مثال، رایانگک<sup>۴</sup>، ۲۰۱۰) عمدتاً معطوف به پایایی، روایی همگرا و روایی سازه این مقیاس بوده‌اند. محدودیت‌های چندی بر روش‌های اعتبارسنجی پرسش‌نامه‌ها از طریق نظریه کلاسیک وارد هستند. برای مثال، نظریه کلاسیک اندازه‌گیری برای کل مقیاس یک خطای اندازه‌گیری مفروض می‌دارد و کلیه محاسبات پایایی براساس این فرض ضعیف استوار است. این در حالی است که خطای اندازه‌گیری برای افراد مختلف با توجه به توانایی‌شان در آن سازه متفاوت است. از طرف دیگر پارامترهای برآورد شده براساس نظریه کلاسیک نامتغیر نبوده و از نمونه‌ای به نمونه‌ای دیگر احتمالاً نتایج کاملاً متفاوتی به بار می‌آورند. همچنین تایید ساختار عاملی و روایی همگرا یا واگرا به تنهایی نمی‌توانند تضمین‌کننده روایی سازه یک مقیاس باشند.

1. Wainer, & Thissen
2. Steinberg, & Mooney
3. Kieftenbeld, V., Natesan, P., & Eddy
4. Ryang

در پاسخ به این محدودیت‌ها، محققان در این پژوهش به بررسی ویژگی‌های روان‌سنجی مقیاس باورهای کارآمد در تدریس ریاضی با استفاده از نظریه جدید اندازه گیری، نظریه سؤال-پاسخ پرداخته‌اند. چون مقیاس اندازه گیری سؤالات این پژوهش از نوع لیکرتی بود، به همین جهت محققان از مدل امتیاز پاره‌ای تعمیم یافته که یکی از جدیدترین مدل‌های اندازه گیری IRT برای سؤالات چندگزینه‌ای هستند، به کار رفت. نتایج این پژوهش، می‌تواند علاوه بر بسط ادبیات پژوهش مفید بودن این ابزار را در بین نمونه‌ای از معلمان ایرانی نشان دهد.

با توجه به خلاصه پژوهشی که بدان اشاره شد. در این پژوهش محققان قصد دارند که با استفاده از مدل امتیاز پاره‌ای تعمیم یافته به بررسی ویژگی‌های روان‌سنجی مقیاس خود کارآمدی در تدریس ریاضیات بپردازند.

### روش پژوهش

روش پژوهش فعلی به لحاظ روش از نوع پژوهش‌های توصیفی-پیمایشی و از نظر هدف از نوع پژوهش توسعه‌ای است. داده‌های این پژوهش به صورت همزمان و از طریق مقیاس خودسنجی جمع آوری شده‌اند.

### نمونه پژوهش

نمونه این پژوهش را ۳۳۲ نفر از معلمان شهر تبریز تشکیل دادند که بین ۱ تا ۵ سال سابقه تحصیل داشتند. روش نمونه گیری به کار رفته در این پژوهش به صورت تصادفی خوشه‌ای بود و یکی از محققان به طور تصادفی با مراجعه به مدارس مقیاس پژوهش را بر روی معلمان مدارس ابتدایی اجرا کرد. تعداد ۱۷۵ نفر از این معلمان را معلمان زن تشکیل می‌داد. میانگین سن معلمان برابر با ۲۷/۶۷ (۳/۴۳) بود. ۲۱ درصد از معلمان دارای تحصیلات کاردانی، ۷۱ درصد از معلمان دارای تحصیلات کارشناسی، ۷ درصد کارشناسی ارشد و ۱ درصد نیز دارای تحصیلات دکتری یا دانشجوی دکتری بودند.

## ابزار

مقیاس باورهای کارآمدی معلم در تدریس ریاضیات توسط یکی از محققان به زبان فارسی ترجمه شد. سپس توسط یک متخصص زبان انگلیسی به انگلیسی برگردانده شد، و سپس توسط متخصصی دیگر به زبان فارسی ترجمه شد. سپس با همکاری ۳ متخصص روان‌شناسی تربیتی که دانش خوبی از زبان انگلیسی داشتند، سؤالات ویراستاری شدند. این مقیاس از ۲۱ سؤال (پیوست ۱) تشکیل یافته است که ۱۳ سؤال آن برای سنجش باورهای کارآمدی و ۸ سؤال آن برای سنجش انتظارات پیامد طراحی شده‌اند. هر سؤال از پنج گزینه تشکیل می‌شود که عبارتند از کاملاً مخالفم، مخالفم، نظری ندارم، موافقم، و کاملاً موافقم. ساختار عاملی این مقیاس در پژوهش‌های مختلف مورد تأیید قرار گرفته است (برای مثال رایانگ، ۲۰۱۱). همچنین پایایی این مقیاس با استفاده از آلفای کرونباخ توسط اینوچ و همکاران (۲۰۰۰) برای دو خرده مقیاس انتظارات پیامد و باورهای کارآمدی به ترتیب برابر با ۰/۷۷ و ۰/۸۸ بدست آمد. در پژوهش فعلی به منظور بررسی پایایی این مقیاس از آلفای کرونباخ استفاده شد که ۰/۷۸ برای خرده مقیاس انتظارات پیامد و ۰/۸۱ برای خرده مقیاس باورهای کارآمدی به دست آمد. با حذف سؤال ۷ از خرده مقیاس انتظار پیامد پایایی خرده مقیاس برابر با ۰/۸۲ و با حذف سؤال ۱۸ پایایی خرده مقیاس باورهای کارآمدی برابر با ۰/۸۵ به دست آمد که هر دو مقدار نشان دهنده پایایی خوب می‌باشند.

## روش تحلیل داده‌ها

برای بررسی و تحلیل ویژگی‌های روان‌سنجی این ابزار از نظریه سؤال-پاسخ برای داده‌های چندگزینه‌ای (مدل امتیاز پاره‌ای تعمیم یافته<sup>۱</sup>) استفاده شد. نظریه سؤال-پاسخ مبتنی بر اندازه گیری وابسته به مدل است که در صدد نشان دادن رابطه بین پاسخ‌های داده شده به سؤال و صفت مکنونی است که فرض می‌شود آن را مورد اندازه گیری قرار می‌دهد.

چندین مدل برای ارزیابی پاسخ‌های چندگزینه‌ای سؤالات مربوط به یک مقیاس یک بعدی طراحی شده‌اند. یکی از مدل‌های اندازه گیری که می‌تواند سؤالات چندگزینه‌ای را مورد ارزیابی قرار دهد مدل امتیاز پاره ای تعمیم یافته می‌باشد (GPCM، ۱۹۹۲). این مدل اطلاعات مختلفی در مورد سؤالات در مقایسه با مدل پاسخ مدرج می‌دهد. منحنی ویژگی‌های پاسخ یک سؤال در مدل پاسخ مدرج توسط معادله زیر مشخص می‌شود.

$$P_{ik} = \exp \left[ \sum_{i=1}^k a_i (\theta - \lambda_i + \tau_{ie}) \right] / \sum_{i=1}^{m_i} \exp \left[ \sum_{i=1}^c a_i (\theta - \lambda_i + \tau_{ik}) \right]$$

در این معادله،  $P_{ik}$  احتمال انتخاب طبقه  $k$  ام از طبقه  $m_j$  برای یک سؤال است،  $\theta$  سطح توانایی آزمودنی  $\tau$ ، مقدار پارامتر طبقه است،  $\lambda$  اندام مقدار پارامتر مکان سؤال (قابل مقایسه با آستانه در مدل پاسخ مدرج)، و  $a_i$  پارامتر شیب سؤال (قابل مقایسه با ضریب تمیز) می‌باشند. بنابراین، احتمال انتخاب طبقه خاص پاسخ حداقل به سطح صفت فردی و سه ویژگی مختلف سؤال، منظور مقادیر پارامتر طبقه، و پارامترهای شیب مکان سؤالات دارد. برای یک سؤال، تعداد کل پارامترهای مقوله برابر با تعداد گزینه‌های پاسخ منهای ۱ است. هر سؤال دارای یک شیب و یک پارامتر مکان می‌باشد. پارامتر  $\tau$  نشان دهنده دشواری مرحله در حرکت از یک گزینه پاسخ به دیگری است. برای مثال، اگر یک سؤال دارای ۴ گزینه پاسخ باشد، در این صورت سه گام خواهیم داشت، گام ۱ یعنی حرکت از گزینه پاسخ اول به گزینه دوم، گام دوم

1. Generalizes partial credit model



یعنی جابجایی از گزینه دوم به گزینه سوم، و گام سوم یعنی حرکت از گزینه سوم به گزینه چهارم. مقادیر مثبت بزرگتر نشانگر مراحل سخت بوده، در حالی که مقادیر منفی و پایین نشانگر گام‌های ساده می‌باشند. پارامتر شیب نشانگر چگونگی تغییر احتمال پاسخ در طبقات بالا به عنوان تابعی از افزایش در سطوح صفت است. در کل، مقادیر شیب بالا نشان می‌دهند که سؤال در تمیز سطوح مختلف صفت مکنون خوب عمل کرده است. دشواری یک سؤال توسط پارامتر لاندا مشخص می‌شود. مقادیر مثبت بزرگ نشان دهنده یک سؤال دشوار بوده و به معنای این است که آزمودنی‌های کمی در طبقات بالا به آن پاسخ داده‌اند. مقداری منفی دال بر سؤال آسان بوده یا دال بر این است که آزمودنی‌های کمی به طبقات پایین پاسخ داده‌اند. پارامتر مکان تابعی برای تغییر از پارامترهای طبقه به بالا و پایین در مقیاس صفت مکنون است. کلیه محاسبات IRT با استفاده از نرم افزار IRTPRO2.1 صورت گرفت.

### یافته‌ها

نظریه سؤال - پاسخ از چند مفروضه مهم تشکیل می‌شود. دو مفروضه مهم این نظریه تک بعدی بودن و استقلال موضعی می‌باشد. این دو مفروضه مهم IRT ارتباط بسیار نزدیکی باهم دارند. برای سنجش تک بعدی بودن از تحلیل عاملی تاییدی و برای سنجش استقلال موضعی از همبستگی بین باقیمانده‌ها استفاده شد.

طبق راهنمای این مقیاس این مقیاس از دو خرده مقیاس مجزا از هم انتظار پیامد و خود کارآمدی تشکیل یافته است. بنابراین نباید برای برآورد پارامترهای سؤالات از نظریه سؤال - پاسخ چند بعدی استفاده کرد. به عبارت دیگر، برای برآورد پارامترهای سؤالات باید از IRT تک بعدی استفاده کرد چرا که این دو ۲ سازه (خرده مقیاس‌های ابزار پرسش‌نامه) اساساً مجزا از هم هستند.

به منظور بررسی تک بعدی بودن هر یک از خرده مقیاس‌ها از تحلیل مؤلفه‌های اصلی<sup>۱</sup> و تحلیل عاملی تاییدی از روش حداقل مجذورات وزنی طبقه‌ای<sup>۲</sup> استفاده شد. نتایج تحلیل مؤلفه‌های اصلی برای سؤالات مربوط به خرده مقیاس انتظار پیامد با استفاده از روش چرخش واریماکس نشان داد که به جز سؤال ۷ که بار عاملی بسیار کمی بر روی عامل اول داشت (۰/۱۴) بقیه سؤالات وزن‌های عاملی بین ۰/۵۴ تا ۰/۸۱ بر روی عامل اول داشتند. بررسی مقادیر ویژه مؤلفه‌های استخراجی نشانگر این بود که اندازه مقدار ویژه مؤلفه اول در حدود ۵/۱۵ مقدار ویژه مؤلفه دوم بود. این اندازه به دست آمده نشانگر غالب بودن بعد اول است و از عوامل دیگر می‌توان صرف نظر کرد. برای اطمینان بیشتر از تک بعدی بودن خرده مقیاس انتظار پیامد از تحلیل عاملی تاییدی استفاده شد.

جدول ۱. آماره‌های مربوط به تحلیل عاملی اکتشافی و تاییدی به منظور بررسی تک بعدی بودن

خرده مقیاس	نسبت بار عاملی (اول به دوم)	خی دو	df	CFA	GFA	RMSEA
انتظار پیامد	۵/۱۵	۳۲/۱۴	۱۴	۰/۹۶	۰/۹۴	۰/۰۶۲
خود کارآمدی	۴/۴۸	۱۰۷/۱۳	۵۴	۰/۹۲	۰/۹۰	۰/۰۸۱

با توجه به نتایج به دست آمده از روش تحلیل عاملی اکتشافی سؤال ۷ به عنوان نشانگر این بعد وارد مدل نشد. نتیجه تحلیل دال بر شاخصه‌های برازش قابل قبول مدل با داده‌ها بود. همچنین بارهای عاملی سؤالات همگی از اندازه‌های قابل قبول برخوردار بودند (بین ۰/۵۵ تا ۰/۷۵).

برای بررسی تک بعدی بودن خرده مقیاس خود کارآمدی روال فوق الذکر تکرار شد. نسبت مقدار ویژه مؤلفه اول به دوم برابر با ۴/۴۸ بود. بررسی بارهای عاملی سؤالات نشانگر این بود که سؤال ۱۸ پرسش نامه از بار عاملی ۰/۰۸ برخوردار بود. در مرحله بعد این سؤال حذف

1. Principle component analysis
2. weighted least squares confirmatory

شد و نتایج تحلیل عاملی تاییدی نشانگر این بود که کلیه سؤالات بر روی یک عامل اصلی از وزن‌های قابل قبولی برخوردار هستند. همچنین شاخصه‌های برازش این مدل با داده‌ها دال بر نیکویی برازش مدل یک عاملی داشت.

به منظور بررسی استقلال موضعی از شاخص‌های وابستگی موضعی چن و تیسن (۱۹۹۷) استفاده شد. این شاخصه‌ها نشانگر این بودند که همبستگی بین باقیمانده‌ها در هیچ یک از متغیرهای دو بعد بالاتر از ۰/۱۰ نبود که نشانگر وجود استقلال موضعی است.

چون این دو مفروضه برای هر دو بعد ابزار سنجشی برقرار بودند. در مرحله بعد سؤالات مربوط به هر بعد به طور مجزا وارد تحلیل شد. همانطور که در بخش روش به آن اشاره شد، کلیه محاسبات براساس مدل ریاضی سؤال- پاسخ امتیاز پاره‌ای تعمیم یافته صورت گرفت. سؤال ۷ از مقیاس انتظارات پیامد و سؤال ۱۸ از مقیاس خودکارآمدی وارد مدل اندازه‌گیری نشد.

در جدول ۲ شاخص‌های دو برای نشان دادن برازش مدل اندازه‌گیری برای سؤالات نشان داده شده است. همانطور که در این جدول نشان داده شده است، هیچ یک از مقادیر  $\chi^2$  دو در درجه آزادی مربوطه معنی‌دار نیست که دال بر این است که مدل امتیاز پاره‌ای مدلی مناسب برای برازش داده‌ها می‌باشد.

جدول ۲. آمار برازش سؤال با مدل

خرده مقیاس	سؤال	$\chi^2$	درجه آزادی	احتمال
	۱	۲۳/۶۴	۱۶	۰/۰۹۷
	۴	۲۷/۶۷	۲۰	۰/۱۱۲
	۹	۲۵/۲۱	۱۸	۰/۱۱۸
انتظار پیامد	۱۰	۲۰/۴۰	۱۷	۰/۲۵۳
	۱۲	۲۵/۴۴	۱۶	۰/۰۶۲
	۱۳	۲۷/۸۴	۲۰	۰/۱۱۲
	۱۴	۳۵/۶۲	۲۷	۰/۱۲۳

۰/۲۱۳	۲۸	۳۲/۱۴	۲	
۰/۲۴۸	۲۶	۳۰/۰۳	۳	
۰/۱۹۲	۲۶	۳۲/۲۸	۵	
۰/۲۸۶	۲۹	۳۴/۶۷	۶	
۰/۲۳۲	۲۵	۲۸/۴۴	۸	
۰/۰۵۴	۲۳	۳۴/۸۷	۱۱	خود
۰/۲۴	۲۶	۳۰/۴۹	۱۵	کارآمدی
۰/۱۳۲	۲۰	۲۷/۳۱	۱۶	
۰/۳۴۵	۳۴	۳۸/۸۷	۱۷	
۰/۱۰۲	۱۸	۲۵/۹۳	۱۹	
۰/۲۱۰	۱۸	۲۲/۴۹	۲۰	
۰/۱۵۳	۲۸	۳۵/۴۸	۲۱	

مقادیر غیرمعنی دار برای همه سؤالات دال بر برآزش خوب داده با مدل دارد. لازم به ذکر است که سؤالات مربوط به دو خرده مقیاس در دو مرحله جداگانه وارد تحلیل شد و کلیه محاسبات به طور مجزا برای این دو خرده مقیاس انجام شد.

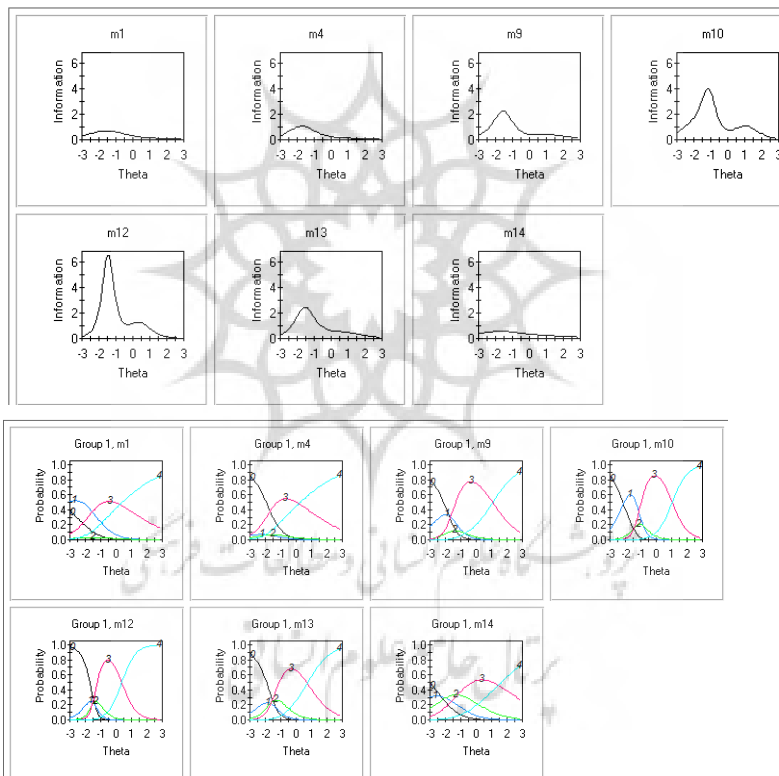
جدول ۳. پارامترهای سؤال برآورد شده از طریق مدل امتیاز پاره‌ای تعمیم یافته

خرده مقیاس	سؤال	شیب	پارامتر مکان	پارامترهای طبقه		
				گام اول	گام دوم	گام سوم
انتظار پیامد	۱	۰/۶۵	-۱/۵۵	۱/۸۵	-۳/۸۴	۳/۸۰
	۴	۰/۶۷	-۱/۲۲	۲/۲۲	۰/۵۸	۳/۰۴
	۹	۱/۲۳	-۱	۰/۷۸	-۰/۱۸	۱/۴۷
	۱۰	۲/۰۵	-۰/۸۳	۱/۲۹	-۰/۰۱	۰/۶۰
	۱۲	۲/۱۸	-۱	۰/۳۶	۰/۴۵	۰/۵۶
	۱۳	۱/۳۲	-۰/۹۷	۰/۳۵	۰/۷۰	۰/۶۵
	۱۴	۰/۷۰	-۱/۰۱	۱/۲۸	۰/۹۷	۰/۲۰
	۲	۰/۵۱	-۱/۲۳	-۰/۳۹	۰/۸۱	۰/۸۹
	خود کارآمدی					

۳	۰/۴۶	-۱/۳۷	-۰/۲۳	-۲/۰۸	۴/۹۱	-۲/۶۱
۵	۱	-۱/۳۹	۱/۸۹	۰/۵۰	-۰/۰۳	-۲/۳۷
۶	۰/۷۹	-۰/۸۷	۱/۶۸	۰/۵۶	۰/۰۰	-۲/۲۴
۸	۰/۶۲	-۱/۴۲	-۱/۲۲	۱/۴۹	۱/۱۴	-۱/۴۱
۱۱	۱/۴۱	-۰/۸۵	۱/۳۰	۰/۴۱	۰/۴۴	-۲/۱۴
۱۵	۰/۸۹	-۱/۰۳	۰/۵۹	۰/۷۹	۰/۸۲	-۲/۱۹
۱۶	۰/۷۱	-۱/۲۴	-۰/۷۳	-۰/۲۲	۲/۵۰	-۱/۵۵
۱۷	۰/۵۷	-۰/۹۲	۱/۴۳	-۰/۵۸	۰/۵۷	-۲/۴۲
۱۹	۱/۶۸	-۱/۲۷	۱/۶۶	-۰/۶۲	۰/۹۶	-۱/۹۹
۲۰	۱/۰۸	-۱/۴۸	۱/۱۱	-۰/۶۹	۰/۸۹	-۱/۳۱
۲۱	۰/۸۸	-۰/۵۱	۰/۰۱	-۰/۳۳	۱/۵۴	-۱/۲۲

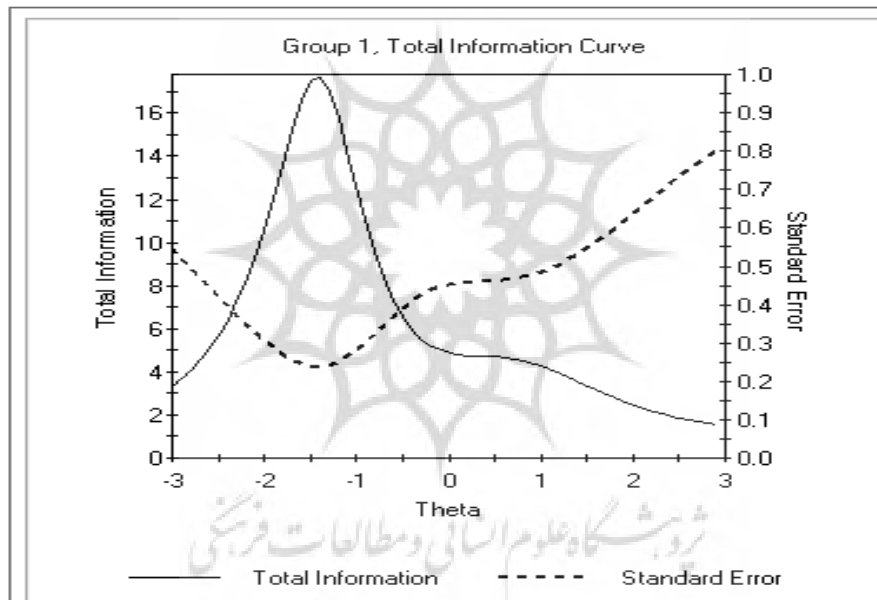
در جدول ۳ پارامترهای برآورد شده از طریق مدل امتیاز پاره‌ای به تفکیک برای هر دو خرده مقیاس نشان داده شده است. نگاهی به این جدول گویای این است که همه پارامترهای مکان برای همه سؤالات در هر دو خرده مقیاس دارای علامت منفی می‌باشند. این نتیجه نشان دهنده این نکته مهم می‌باشد که سؤالات این مقیاس از نظر آزمودنی‌ها آسان می‌باشد. به عبارت دیگر، ضریب دشواری این مقیاس پایین می‌باشد. پارامتر مکان شاخصی برای نشان دادن دشواری پاسخ به سؤالات می‌باشد. علامت مثبت نشانگر سؤالات سخت و علامت منفی نشانگر سؤالات ساده است. هر چند که ضرایب تمیز سؤالات مثبت می‌باشد ولی اندازه ضرایب تمیز اکثر سؤالات پایین است که این نتیجه دال بر این است که سؤالات به خوبی نمی‌توانند افراد با توانایی‌های مختلف را به خوبی از هم تفکیک کنند. نگاهی به پارامترهای طبقه نشانگر این است که اندازه گام چهارم یعنی جابجایی از گزینه موافقم به کاملاً موافقم دارای علامت منفی است که این نتیجه دال بر عدم تمایز گزینه‌های موافقم و کاملاً موافقم از نظر آزمودنی‌هاست. بنابراین، بهتر است که این دو گزینه را باهم ادغام کرده و یک گزینه پاسخ درست کرد. نگاهی به جدول ۳ گویای این است که علامت پارامتر طبقه برای گام سوم مثبت بود (به جز سؤال ۶)

که نشان دهنده این است که حرکت از پاسخ نظری ندارم به موافقم مستلزم توانایی (تتا) می‌باشد. بررسی روند نمرات برای پارامتر طبقه‌ای در گام اول و دوم برای سؤالات مختلف گویای روندهای مختلفی است. برای مثال سؤال اول دارای پارامتر طبقه‌ای مثبت برای گام اول ولی برای گام دوم علامت این پارامتر منفی و بسیار بزرگ بود. این نتیجه نشانگر این است که حرکت از گزینه ۱ تا ۲ مستلزم داشتن توانایی در متغیر مکنون بوده ولی حرکت از گزینه ۲ به ۳ نیاز کمی به توانایی داشت.



شکل ۱. منحنی ویژگی‌های سؤال و تابع آگاهی برای سؤالات مربوط به بعد انتظار پیامد

در شکل ۱ منحنی ویژگی‌های سؤالات (ICC) و تابع اطلاعات سؤالات انفرادی مربوط به بعد انتظار پیامد نشان داده شده است. نگاهی به تابع آگاهی سؤالات گویای این است که سؤال ۱ و ۱۴ هیچ اطلاعاتی در مورد توانایی آزمودنی‌ها ارائه نمی‌دهند. بقیه سؤالات هم آگاهی نسبتاً کمی در اطراف توانایی پایین ارائه می‌کنند. به عبارت دیگر هیچ یک از سؤالات در کل نمی‌توانند آگاهی را در طیف توانایی‌های بالا افزایش بدهند. به عبارت دیگر در طیف توانایی بالا ضریب آگاهی سؤال نزدیک صفر بود.



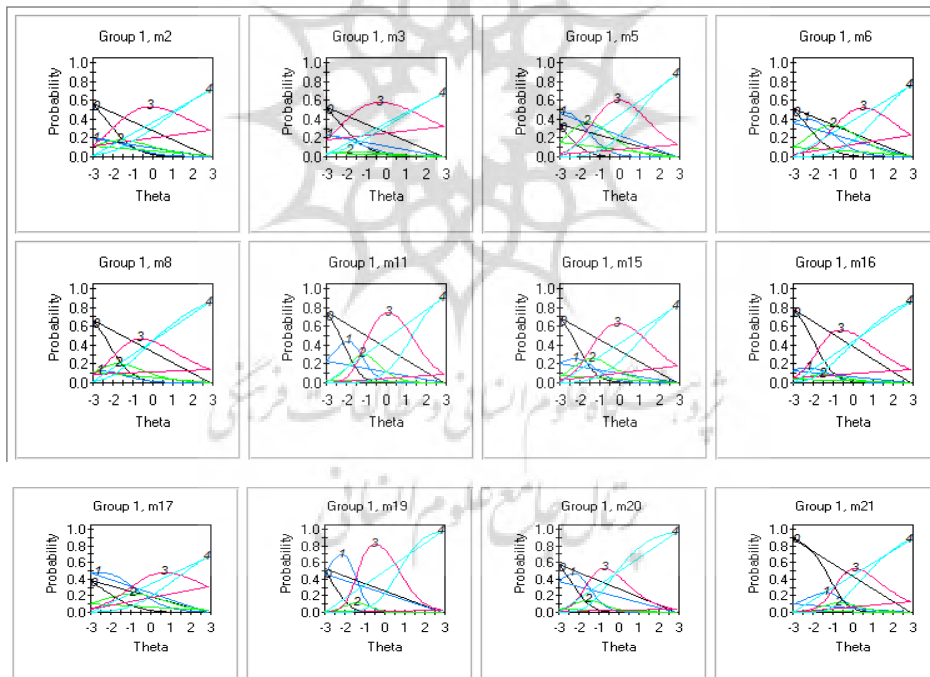
شکل ۲. تابع آگاهی آزمون برای بعد انتظار پیامد

در شکل ۲ منحنی تابع آگاهی آزمون برای بعد انتظار پیامد نشان داده شده است، نگاهی به این نمودار نشانگر این که نکته اوج تابع آگاهی در اطراف تنای  $1/5$  است. برای توانایی‌های بالا مقدار آگاهی آزمون بسیار پایین می‌آید.

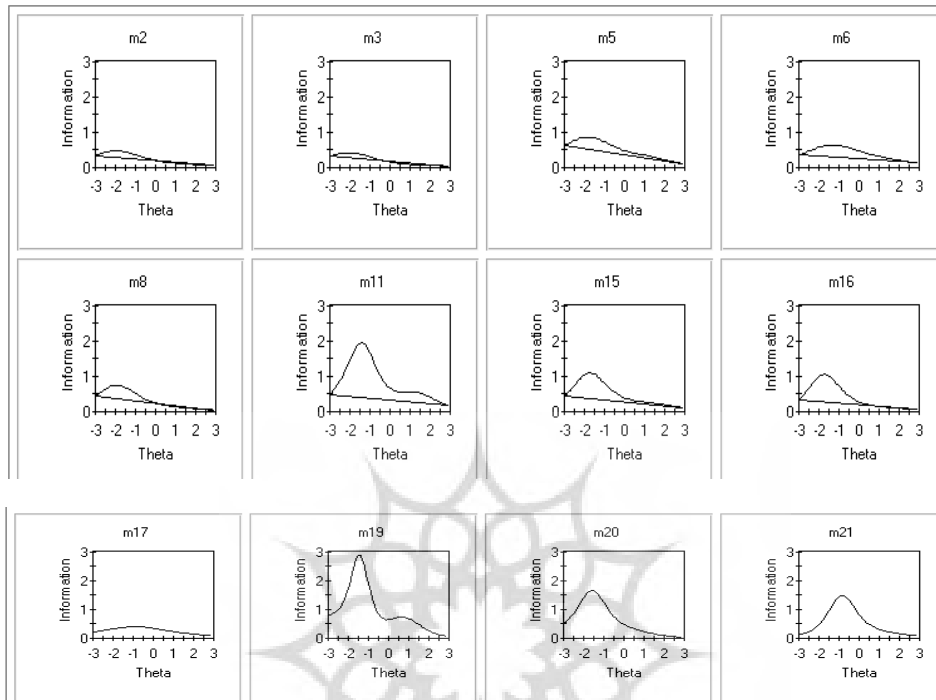
در شکل ۳ منحنی ویژگی‌های سؤال و توابع آگاهی برای سؤالات بعد کارآمدی تدریس نشان داده شده است.

با توجه به شکل ۳ مشخص است که تابع آگاهی سؤالات ۲، ۳، ۶، ۸ و ۱۷ دارای منحنی‌های کاملاً مسطح و با ارتفاع بسیار پایینی بودند که به هیچ وجه تأثیری در آگاهی نداشتند. سایر سؤالات هم همانند بعد انتظار پیامد قله تابع آگاهی برای توانایی پایین آزمودنی‌ها بود. به عبارت دیگر، برای افراد با توانایی بالا مقدار تابع آگاهی بسیار پایین بود.

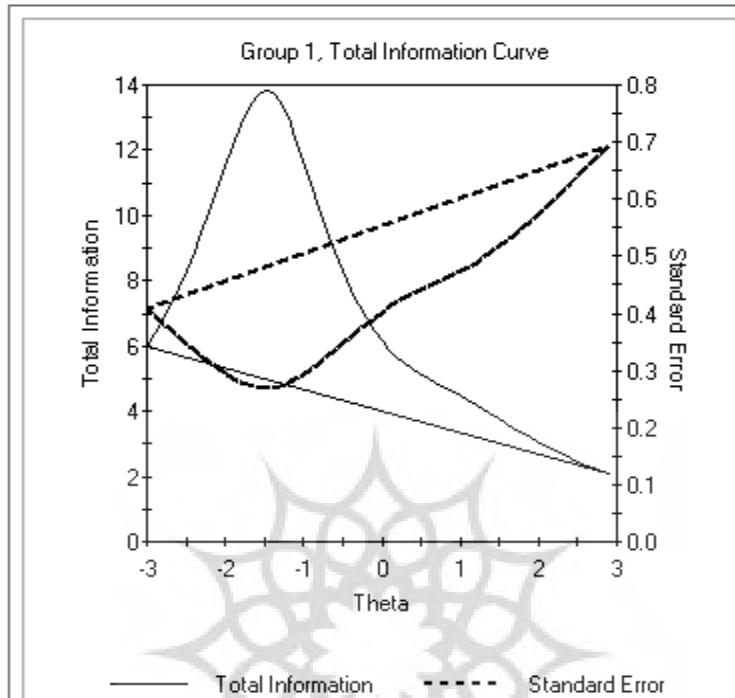
شکل ۴ نشان می‌دهد که تابع آگاهی آزمون برای بعد کارآمدی تدریس در نقطه تنای ۱/۵- دارای پیک است. برای توانایی‌های بالا مقدار آگاهی بسیار پایین بوده و خطای معیار دارای اندازه‌ای بالا می‌باشد.







شکل ۳. منحنی ویژگی‌های سؤال و تابع آگاهی برای سؤالات مربوط به بعد کارآمدی



شکل ۴. تابع آگاهی آزمون برای بعد کارآمدی تدریس

### بحث و نتیجه گیری

هدف این پژوهش بررسی ویژگی‌های روان‌سنجی مقیاس کارآمدی تدریس ریاضی معلمان دوره ابتدایی بود. نمونه‌ای متشکل از ۳۳۴ نفر از معلمان که سابقه تدریس آنها بین ۲ تا ۵ سال بود برای سنجش ویژگی‌های روان‌سنجی مقیاس انتخاب شدند. بررسی ویژگی‌های روان‌سنجی مقیاس با استفاده از شاخصه‌های کلاسیک اندازه‌گیری نشانگر ضریب همسانی بالا برای خرده مقیاس‌ها بود. همچنین تحلیل عاملی با هر دو روش تحلیل مؤلفه‌های اصلی و تحلیل عاملی تاییدی مؤید سازه دو بعدی این مقیاس بود. نتیجه این پژوهش با گزارشات طراحان این مقیاس همسو بود. در بخش اصلی این پژوهش محققان با استفاده از مدل اندازه‌گیری امتیاز پاره‌ای

تعمیم یافته به بررسی ویژگی‌های روان‌سنجی مقیاس مزبور پرداختند. نتایج پژوهش نشانگر این بود که سؤالات هر دو بعد این مقیاس دارای مشکلات متعددی در سنجش کارآمدی تدریس دارند. یکی از بزرگترین معایب سؤالات این بود که ضریب دشواری یا پارامتر مکان سؤالات همگی دارای علامت منفی بودند که دال بر این بود که پاسخ دادن به سؤالات نیاز به توانایی بالایی نداشت. به زبان ساده‌تر، این نتیجه دال بر این بود اکثر پاسخ‌های افراد به گزینه‌های موافقم و کاملاً موافقم بود. یعنی، طراحی سؤالات به گونه‌ای است که درج پاسخ‌های موافقم و کاملاً موافقم برای آزمودنی‌ها آسان می‌باشد. همچنین نتیجه این پژوهش نشان داد که خطای معیار اندازه‌گیری برای طیف مربوط به توانایی‌های بالا دارای مقداری بلا و به تبع آن آگاهی سؤالات بسیار پایین بود. به عبارتی دیگر این ابزار در سنجش خود کارآمدی تدریس معلمان در ریاضیات به هیچ وجه موفق عمل نمی‌کند. شاید تنها کارکرد این سؤالات در تمیز معلمان قوی از نظر خود کارآمدی با معلمان بسیار ضعیف بود. این در حالی است که یک پرسش‌نامه خوب باید دارای سؤالاتی باشد که بتواند تمام طیف توانایی را مورد هدف قرار دهد. یعنی سؤالات باید طوری شوند که بتوانند برای افراد با توانایی‌های مختلف اثر متمایزی در افتراق در سازه مکنون مورد نظر را داشته باشند. حال آنکه نظریه سؤال-پاسخ نشانگر این بود که سؤالات این قدرت افتراقی را ندارند. در کل نتایج این پژوهش گویای این بود که معیارهای ضریب پایایی، اعتبار عاملی به تنهایی نمی‌توانند مؤید مفید بودن سؤالات در عمل باشند، بلکه برای تأیید ساختار مفروض و مفید بودن سؤالات از نظریه سؤال-پاسخ استفاده کرد.

خود کارآمدی سازه بسیار مهم در روان‌شناسی تربیتی و آموزشی است که در عملکرد و انگیزش معلمان تأثیر بسزایی دارد (بندورا، ۱۹۹۷؛ چانان-موران، ولفولک و هوی، ۲۰۰۱؛ چانان-موران و ولفولک و هوی؛ ۱۹۹۸؛ ولفولک و هوی، ۲۰۰۴). معلمان خود کارآمد تمایل زیادی بر مشارکت در فعالیت‌های مدرسه، همکاران و دانش‌آموزان از خود نشان می‌دهند، و مدرسه را به عنوان سیستمی تلقی می‌کنند که قادر است هدف و رسالت آنها را به فرجام نیکی سوق دهد (کاپارا، باربرائلی، بورگوگنی، و استیکا، ۲۰۰۳). باورهای خود کارآمدی معلمان

نقش اساسی در تعهد و رضایت شغلی معلمان دارد (کاپارا و همکاران، ۲۰۰۳). پژوهش‌ها نشان داده‌اند که معلمان دارای خودکارآمدی پایین مشکلات زیادی را در تدریس و سطوح بالایی از استرس‌های مربوط به شغل و سطوح پایین رضایت شغلی را تجربه می‌کنند (بیتورت، ۲۰۰۶؛ کلاسن و همکاران، ۲۰۰۹). وجود ابزاری برای سنجش خودکارآمدی تدریس معلمان و به ویژه در درس ریاضی که چالش زیادی برای معلم می‌تواند به همراه داشته باشد. این در حالی است که در ادبیات پژوهش تنها یک ابزار برای سنجش این سازه مهم طراحی شده است. اما نتیجه این پژوهش نشانگر این بود که این ابزار فاقد ویژگی روان‌سنجی مکفی و قابل قبول برای سنجش خودکارآمدی تدریس معلمان بود. نتیجه این پژوهش گویای این است که پرسش نامه‌ای جدید برای سنجش این سازه باید توسعه داده شود. طراحی سؤالات این پرسش نامه باید به نحوی باشد که علاوه بر سنجش دو بعد اصلی (انتظارات پیامد و خودکارآمدی تدریس) این سؤالات باید پاسخ دهندگان با توانایی متفاوت را بتواند به خوبی تفکیک کرده و دشواری سؤالات معقول باشد. متأسفانه اکثر پرسش نامه‌های مربوط به سنجش خودکارآمدی به ویژه در مورد معلمان و اساتید به نحوی است که پاسخ دهندگان به خاطر مطلوب نشان دادن خود یا وانمود کردن خود به داشتن توانایی‌های بالا در درج پاسخ‌ها به انتخاب پاسخ‌های مطلوب اقدام می‌کنند. ساخت پرسش نامه‌هایی که بتواند پاسخ‌های واقعی فرد را مورد سنجش قرار بدهند چالشی است که بر پژوهش‌ها پرسش نامه به طور عام و بر پرسش نامه‌هایی نظیر خودکارآمدی به طور خاص وارد است. توصیه می‌شود محققان دیگر در طراحی سؤالات اقدام به انتخاب سؤالاتی بکنند که بتواند وانمود خوب آزمودنی‌ها را کنترل کند.

## منابع لاتین

- Baker, F. B. (2001). The basics of item response theory. ERIC clearinghouse on assessment and evaluation.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman and Company.
- Betoret, F. D. (2006). Stressors, self-efficacy, coping resources, and burnout among secondary school teachers in Spain. *Educational Psychology, 26*, 519° 539.
- Caprara, G. V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers job satisfaction. *Journal of Educational Psychology, 95*, 821° 832.
- construct. *Review of educational research, 62*, 307-332.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates Inc.
- Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the Mathematics Teaching Efficacy Beliefs Instrument. *School Science and Mathematics, 100*, 194-203.
- Huinker, D., & Enoch, L. G. (1995). *Mathematics Teaching Efficacy Beliefs Instrument (MTEBI)*. Milwaukee: University of Wisconsin, Center for Mathematics and Science Education Research.
- Kieftenbeld, V., Natesan, P., & Eddy, C. (2011). An Item Response Theory Analysis of the Mathematics Teaching Efficacy Beliefs Instrument. *29(5)*, 443° 454.
- Klassen, R. M. & Chio, M. M. (2010). Effects on Teachers Self-Efficacy and Job Satisfaction: Teacher Gender, Years of Experience, and Job Stress. *Journal of Educational Psychology, 102(3)*, 741° 756 .
- Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y., & Georgiou, T. (2009). Exploring the validity of the Teachers Self-Efficacy Scale in five countries. *Contemporary Educational Psychology, 34*, 67° 76.
- Pajares, F. (1992). Teachers beliefs and educational research: Cleaning up a messy.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of educational research, 66*, 543-578.
- Ross, J. A., & Bruce, C. D. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *The Journal of Educational Research, 101(1)*, 50-60.

- Ryang, D. (2010). Exploratory Analysis of Korean Elementary Preservice Teachers Mathematics Teaching Efficacy Beliefs. *International Electronic Journal of Mathematics Education*, 7(2), 45-61.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph*, 17.
- Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, 99, 611-625.
- Swars, S. L. (2005). Examining perceptions of mathematics teaching effectiveness among elementary preservice teachers with differing levels of mathematics teacher efficacy. *Journal of Instructional Psychology*, 32(2), 139-147.
- Swars, S. L., Daane, C. J., & Giesen, J. (2006). Mathematics anxiety and mathematics teacher efficacy: What is the relationship in elementary preservice teachers? *School Science and Mathematics*, 106(7), 306-315.
- Thissen, D., Steinberg, L., & Mooney, J. (1989). Trace lines for testlets: A use of multiple-categorical response models. *Journal of Educational Measurement*, 26, 247-260.
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, 68, 202-248.
- Utey, J., Bryant, R., & Moseley, C. (2005). Relationship between science and mathematics teaching efficacy of preservice elementary teachers. *School Science and Mathematics*, 105(2), 82-87.
- Wainer, H., & Thissen, D. (1996). How is reliability related to the quality of test scores? What is the effect of local dependence on reliability? *Educational Measurement: Issues and Practice*, 15, 22-29.
- Ward, E. K. (2009). *Latent transition analysis of pre-service teachers' efficacy in science and mathematics* (Unpublished doctoral dissertation), University of North Texas, Denton.
- Woolfolk, A. E., & Hoy, W. K. (2004). *What do teachers need to know about self-efficacy?* Paper presented at the annual meeting of the American educational research association, San Diego, CA, April 15, 2004. Session 52, 0700.