مقایسه توانمندی رویکرد بیزی مدل IRT چندسطحی و مدل کلاسیک چندسطحی: تحلیل داده‌های آزمون فیزیک تیمز پیشرفته (۲۰۰۸)

تاریخ دریافت: ۹۲/۴/۵ تاریخ پذیرش: ۹۲/۷/۶

جلیل یونسی[۱]
فرزاد اسکندری[۲]
علی دلاور[۳]
محمدرضا فلسفی نژاد[۴]
نورعلی فرخی[۵]

چکیده

زمینه: اعتبار تحلیل‌های چندسطحی با تأکید بر تفاوت‌های نظریه‌های اندازه‌گیری (رویکرد کلاسیک اندازه‌گیری و رویکرد نظریه جدید اندازه‌گیری (IRT)) به کمک داده‌های مختلف مورد پژوهشهای اخیر بوده است. **هدف:** هدف از انجام پژوهش حاضر بدست آوردن میزان تأثیر مدارس بر نمرات کسب شدهٔ دانش‌آموزان شرکت کننده در آزمون تیمز پیشرفته ۲۰۰۸ بود. **روش:** برای دستیابی به هدف اصلی پژوهش از داده‌های مربوط به اجرای آزمون فیزیک تیمز پیشرفته ۲۰۰۸ که روند آموزش فیزیک پیشرفته دانش‌آموزان سال آخر متوسطه (پیش‌دانشگاهی) را مورد ارزیابی قرار می‌دهد، استفاده شد. جامعه آماری و گروه نمونه پژوهش حاضر، جامعه آماری و گروه نمونه پایهٔ پیش دانشگاهی رشته ریاضی- فیزیک در آزمون فیزیک تیمز پیشرفته است که در سال تحصیلی ۱۳۸۷-۱۳۸۶ به اجرا در آمده است. حجم نمونه دانش‌آموزان ایرانی در این سنجش برابر ۲۵۵۶ نفر بوده است. **یافته‌ها:** نتایج تحلیل‌ها به طور کلی نشان داد که نخست تحلیل‌های چندسطحی نظریه سؤال- پاسخ (MLIRT) نسبت به تحلیل‌های چندسطحی نمرهٔ واقعی (MLTS) از قدرت بیشتری در آشکارسازی تفاوت بین مدارس برخوردار است. دوم آنکه، میزان تفاوت بین مدارس شرکت کننده در آزمون ریاضیات تیمز پیشرفته که در شاخص ICC (همبستگی درون رده‌ای) منعکس شده است در تحلیل‌های MLIRT بسیار زیاد است. **بحث و نتیجه‌گیری:** نخست آنکه، در نظر گرفتن خطای اندازه‌گیری تک تک سؤالات در تحلیل‌ها در چارچوب بیزی و با استفاده از نمونه‌گیری گیبس به طور چشمگیری توانمندی تحلیل‌های چندسطحی را بهبود می‌بخشد و افزایش معنادار نسبت واریانس تبیین شده را در پی خواهد داشت. دوم آنکه، تفاوت و تبعیض آموزشی بین مدارس زیاد است که این امر عمدتاً ناشی از متغیرهای در سطح مدرسه (نظیر متغیرهای مرتبط به معلم یا متغیرهای مرتبط

۱. استادیار دانشگاه علامه طباطبائی (نویسنده مسئول) jalilyounesi@gmial.com
۲. دانشیار دانشگاه علامه طباطبائی
۳. استاد دانشگاه علامه طباطبائی
۴. دانشیار دانشگاه علامه طباطبائی
۵. دانشیار دانشگاه علامه طباطبائی

به مدرسه) می‌باشد.
واژگان کلیدی: تحلیل چندسطحی، رویکرد بیزی، نظریه سؤال- پاسخ (IRT)، نظریه کلاسیک آزمون (CTT)، آزمون تیمز پیشرفته (TIMSS)، تحلیل حساسیت

مقدمه

یکی از مهم‌ترین دغدغه‌های اصلی تمامی زمامداران و فلاسفه برقراری عدالت در جامعه بوده است. پیرامون مفهوم عدالت در طول تاریخ نظریه‌های گوناگونی تبیین شده است. فیثاغورث عدالت را رعایت حقوق افراد مطابق استعدادهای طبیعی آنها می‌داند؛ ازنظر سقراط عدالت رعایت نوعی تناسب و هماهنگی است؛ از دیدگاه افلاطون عدالت نوعی تناسب و تعادل در امور است (جمشیدی، ۱۳۸۰). عدالت یعنی رعایت تساوی در زمینهٔ استحقاق‌های متساوی؛ عدالت اجتماعی عبارت است از ایجاد شرایط برای همه به‌طور یکسان و رفع موانع برای همه به‌طور یکسان (مطهری، ۱۳۵۲). با توجه به تعاریف گوناگون ارائه‌شده می‌توان دریافت که عدالت یک مفهوم چندوجهی است. یکی از مهمترین جنبه-های آن‌که به معرفت انسان برمی گردد، عدالت در امور آموزشی است. با توجه به مفهوم عدالت تعریف‌شده، در راستای عدالت آموزشی، همهٔ آحاد انسانی حق‌دارند از آموزش‌وپرورش یکسان برخوردار گردند. در جهان کنونی یکی از شاخص‌های پیشرفت اجتماعی، آموزش‌وپرورش و چگونگی بهره‌مندی از آن می‌باشد. بر همین مبنا کشورها به‌خصوص کشورهای درحال‌توسعه درصدد ایجاد فرصت‌های برابر دسترسی همگان به آموزش‌وپرورش هستند (عبدوس، ۱۳۸۱). برابری به مهیاسازی فرصت‌های برابر، متناسب با نیازهای ویژهٔ هر فرد اشاره دارد، زیرا افراد در دانش، مهارت، توانایی، انواع یادگیری و پیشینهٔ فرهنگی و نیازهای آموزشی‌شان باهم متفاوت‌اند. منظور از فرصت برابر، جلوگیری، حذف یا کاهش تبعیض بین افراد ازلحاظ جنسیت، نژاد، وضعیت جسمانی، سنی، زبانی، طبقهٔ اجتماعی می‌باشد (بنت، بوث و یندل[1]، ۲۰۰۱).

سیاست‌گذاران و برنامه‌ریزان آموزشی کشورها به دلایل متعددی ازجمله سنجش ارزیابی

1. Bennett, Both & Yeandle

بازده‌ها و عملکرد نظام آموزشی کشور، کشف و شناخت نقاط قوت و ضعف نظام آموزشی کشورشان در مقایسه با سایر کشورها و جستجوی راهکارهای عملی و علمی برای تقویت نظام آموزشی خود (ژائو و همکاران[1]، ۲۰۱۲؛ کریمی، ۱۳۸۷b) در چند دهه اخیر اقدام به برگزاری آزمون‌های معتبر بین‌المللی نظیر تیمز در دروسی همچون ریاضی، علوم و فیزیک کرده‌اند تا بدین طریق بتوانند هم به مقایسهٔ عملکرد خود با سایر کشورها بپردازند و هم با انجام مطالعات تکمیلی و تحلیل ثانوی داده‌های تیمز، عوامل مؤثر بر روند پیشرفت و پسرفت عملکرد دانش‌آموزان را با توجه به بافت و ساختار جامعه خود بهتر بشناسند و درنهایت بتوانند با تکیه‌بر یافته‌های پژوهشی جدید سیاست‌های آموزشی خود را تعدیل یا اصلاح کنند. به‌نوعی با استفاده از آزمون‌های بین‌المللی ازجمله تیمز، می‌توان این استنباط را در سطح بین‌المللی داشت که آیا بین دانش‌آموزان کشورهای گوناگون تفاوتی وجود دارد؛ در سطح ملی یکی از سؤالاتی که باید سیاست‌گذاران به آن حساس باشند آن است که آیا بین مدارس تفاوتی وجود دارد و این تفاوت به چه میزان است و آیا توانسته‌اند عدالت آموزشی را تا حدی رعایت کنند. با انجام این نوع تحلیل‌ها، می‌توان به این سؤال کلی‌تر پاسخ داد که تا چه اندازه سیاست‌های آموزشی سیاست‌گذاران توانسته است عدالت آموزشی رابین مدارس گوناگون در سطح کشور عملیاتی کند.

به‌منظور درک سهم مدارس در میزان موفقیت دانش‌آموزان باید از تکنیک‌های مدل‌سازی چندسطحی که دهه اخیر به‌طور روزافزون در دسترس محققان قرارگرفته است، استفاده شود. با به‌کارگیری مدل‌سازی چندسطحی می‌توان فهم بهتری از این موضوع داشت که آموزش رسمی چگونه تحت تأثیر مدارس، کلاس‌ها، معلمان و تغییراتی است که در طی زمان رخ می‌دهد. در سنجش‌های وسیع[2]، با در نظر گرفتن تعداد زیادی از متغیرهای قابل‌مشاهده و غیرقابل مشاهده باهم و تعامل در سطوح مختلف سیستم سنجش، نگرانی‌هایی دربارهٔ تعمیم و تفسیر نتایج سنجش وجود دارد. چالش سنجش‌های اجراشده

1. Zhao et al.
2. Large-scale assessment

در مقیاس وسیع جمعیتی عمدتاً مربوط می‌شود به ماهیت پیچیدهٔ داده‌ها، ساختار داده‌های جمع‌آوری‌شده، و مشکل در واحد تحلیل که نیازمند تکنیک‌های پیشرفته است. ازاین‌رو، به نظر می‌رسد تحلیل داده‌های آزمون تیمز که دارای ساختار آشیانه‌ای و چندسطحی هستند چه با رویکرد کلاسیک و چه با رویکرد جدید اندازه‌گیری امری ضروری و بسزاست. بنابراین مسئله اصلی موردبحث در این مقاله این بوده که آیا ویژگی‌های مدرسه، معلم و دانش‌آموز بر اساس رویکردهای کلاسیک اندازه‌گیری (متغیر وابسته یا توانایی دانش‌آموز از مجموع سادهٔ نمرهٔ سؤالات به دست می‌آید و نمرهٔ توانایی فرد بدون خطاست) و نظریه سؤال- پاسخ (نمرهٔ حاصل از اجرای تک‌تک سؤالات به‌عنوان متغیر وابسته در تحلیل‌های چندسطحی وارد شود) می‌تواند سهمی در برآورد نمرات توانایی دانش‌آموزان داشته باشند؟ و کدامیک از این دو رویکرد اندازه‌گیری در برآورد پارامترهای سؤال، توانایی و پارامترهای تصادفی از توانمندی بیشتری برخوردار است؟ مسئله اصلی پژوهش در قالب سؤالات زیر موردبررسی قرارگرفته است:

۱. آیا بر اساس رویکرد کلاسیک اندازه‌گیری (CTT) برآورد نمرات دانش‌آموزان تحت تأثیر مدرسه است؟

۲. آیا بر اساس رویکرد کلاسیک اندازه‌گیری (CTT) برآورد نمرات دانش‌آموزان تحت تأثیر ویژگی‌های دانش‌آموز و ویژگی‌های معلم و مدرسه است؟

۳. آیا بر اساس رویکرد IRT برآورد نمرات توانایی دانش‌آموزان تحت تأثیر مدرسه است؟

۴. آیا بر اساس IRT برآورد نمرات توانایی دانش‌آموزان تحت تأثیر ویژگی‌های دانش‌آموز و ویژگی‌های معلم و مدرسه است؟

روش پژوهش

به‌منظور دستیابی به اهداف مطالعه و تعیین توانمندی رویکرد بیزی نظریه سؤال- پاسخ چندسطحی از داده‌های واقعی مربوط به اجرای آزمون فیزیک تیمز پیشرفته ۲۰۰۸ استفاده‌شده است. این اطلاعات مبنای تصمیم‌گیری در مورد رتبه و وضعیت دانش‌آموزان

رشته ریاضی- فیزیک کشور ایران در درس فیزیک دورهٔ پیش‌دانشگاهی در بین سایر کشورهای شرکت‌های کننده در این آزمون بوده است. نظر به اینکه در این پژوهش محقق در پی مقایسه رویکردهای مختلف در تحلیل داده‌های با ساختار آشیانه‌ای، بررسی برازش مدل-های مختلف اندازه‌گیری، به دست آوردن برآوردهای باثبات‌تر پارامترها (اعم از پارامترهای سؤال و توانایی آزمودنی‌ها) و افزایش اعتبار تصمیم‌ها و دقت اندازه‌گیری خواهد بود، لذا روش این پژوهش به‌طور عام جزو پژوهش‌های توصیفی (غیرآزمایشی و غیر تاریخی) (سرمد، بازرگان، و حجازی، ۱۳۸۴) است. روش پژوهش حاضر از حیث نوع تجزیه‌وتحلیل‌های آماری و سنجشی جزو تحقیقات همبستگی (که خود زیرمجموعهٔ روش پژوهش توصیفی قرار می-گیرد) از نوع تحلیل ماتریس کوواریانس است.

جامعه آماری و گروه نمونه

به‌منظور دستیابی به اهداف پژوهش، از اطلاعات واقعی به‌دست‌آمده از اجرای آزمون تیمز پیشرفته (۲۰۰۸) استفاده‌شده است. بر این اساس، جامعه آماری و گروه نمونه پژوهش حاضر، جامعه آماری و گروه نمونه پایهٔ پیش‌دانشگاهی رشته ریاضی- فیزیک در آزمون فیزیک تیمز پیشرفته است که در سال تحصیلی ۱۳۸۷-۱۳۸۶ به اجرا درآمده است. حجم نمونه دانش‌آموزان ایرانی در این سنجش برابر ۲۵۵۶ نفر بوده است (گزارش تیمز پیشرفته، ۱۳۸۸).

ابزار اندازه‌گیری

ابزار اصلی اندازه‌گیری در این پژوهش، آزمون فیزیک تیمز پیشرفته ۲۰۰۸ است که خود جزو تازه‌ترین مطالعات IEA است که روند آموزش فیزیک پیشرفته دانش‌آموزان سال آخر متوسطه (پیش‌دانشگاهی) را مورد ارزیابی قرار می‌دهد و کشور ایران در این مطالعه نیز به همراه ده کشور دیگر شرکت نموده است (مرکز ملی مطالعات تیمز، ۱۳۸۸). آزمون فیزیک پیشرفته شامل ۷۰ سؤال با ۸۲ امتیاز است. هر سؤال چندگزینه‌ای یک امتیاز دارد. سؤالات پاسخ ساز عموماً یک یا دو امتیازی هستند ولی در پژوهش حاضر تمامی این سؤالات به‌صورت دو ارزشی نمره‌گذاری شده‌اند، بدین‌صورت که به پاسخ کاملاً صحیح نمرهٔ یک و به پاسخ‌های نیمه صحیح یا غلط نمرهٔ صفر اختصاص داده‌شده است.

یافته‌ها

سؤال اول پژوهش: آیا بر اساس رویکرد کلاسیک اندازه‌گیری (CTT) برآورد نمرات فیزیک دانش‌آموزان تحت تأثیر مدرسه است؟

به‌منظور پاسخگویی به این سؤال پژوهش در سؤالات آزمون فیزیک ابتدا تحلیل حساسیت انجام شد و چون نتایج پسین نسبت به تغییر پیشین‌ها نیرومندی لازم را نشان داد مدلی که دارای پیشین آگاهی‌بخش بود انتخاب شد و در ادامه شاخص‌های مربوط به تحلیل چندسطحی مدل صفر (با پیشین آگاهی‌بخش) به‌طور خلاصه در جدول زیر گزارش‌شده است.

جدول ۱. شاخص‌های آمار بیزی پارامترهای مربوط به تحلیل چندسطحی مدل صفر در آزمون فیزیک

	پارامتر	$\overline{X}$	SD	MCSE	MD	فاصله اطمینان ٪۹۵ بیزی	
آزمون فیزیک	γ_{00}	۴/۰۲۸۷	۶/۵۵۸۷	۰/۰۶۵۶	۰/۰۴۱۰	-۸/۸۶۶	۱۶/۹۳۹
	σ^2_{θ}	۰/۰۱۷۶	۰/۰۰۰۲	۰/۰۰۰	۰/۰۱۷۶	۰/۰۱۷۱	۰/۰۱۸۱
	τ^2_{00}	۰/۰۱۱۱	۰/۰۰۲۳	۰/۰۰۰	۰/۰۱۱۰	۰/۰۰۷۱	۰/۰۱۶۰
	DIC	۱۸۵۲۷/۳۰					
	ICC	۰/۳۸۶					

با توجه به شاخص ICC که برای آزمون فیزیک در جدول بالا ارائه‌شده است می‌توان استنباط کرد که در آزمون فیزیک تیمز پیشرفته ۳۸/۶ درصد از تغییرات واریانس نمرات فیزیک دانش‌آموزان به مدرسه‌ای برمی‌گردد که در آن تحصیل می‌کنند؛ یعنی ۳۸/۶ درصد از واریانس نمرات فیزیک دانش‌آموز از طریق مدارس ایشان قابل تبیین است. مقدار به‌دست‌آمده برخلاف مقدار حاصل در آزمون ریاضی پیشرفته نسبتاً قابل‌توجه است و نشان‌دهندهٔ آن است که در نظر گرفتن ساختار داده‌ها به‌صورت آشیانه‌ای ارزش عملی و معناداری دارد و بنابراین می‌توان نتیجه گرفت که مشاهدات مستقل نیستند. این نتیجه بدین معناست که بافت آموزشی درس فیزیک که دانش‌آموزان در آن مشغول به تحصیل هستند از مدرسه‌ای به مدرسه‌ای دیگر به نسبت بسیار متفاوت است و نمی‌توان از کنار این تفاوت به‌آسانی گذشت.

سؤال دوم: آیا بر اساس رویکرد کلاسیک اندازه‌گیری CTT برآورد نمرات فیزیک

دانش‌آموزان تحت تأثیر ویژگی‌های دانش‌آموز و ویژگی‌های معلم و مدرسه است؟

به‌منظور پاسخگویی به این سؤال پژوهش در آزمون فیزیک تیمز پیشرفته اثر کوواریت‌های سطح دانش‌آموز و سطح مدرسه به‌طور همزمان مدل‌سازی شده است. بدین منظور در ساخت مدل، ابتدا «متغیرهای سطح دانش‌آموز[۱]» برای تبیین تغییرات (واریانس) درون مدرسه ازلحاظ پیشرفت تحصیلی درس فیزیک انتخاب‌شده است. برای گزینش متغیرهای سطح یک به‌عنوان متغیرهای پیش‌بین (X_q)، از نظریه و تحقیقات پیشین استفاده‌شده است. با توجه به پرسشنامه‌های اجراشده در سطح دانش‌آموز متغیرهای جنسیت دانش‌آموز (پسر با کد صفر و دختر با کد ۱) و نگرش مثبت دانش‌آموز به فیزیک[۲]،"میزان زمانی که فرد در خانه به فیزیک اختصاص می‌دهد (به دقیقه)" و "فعالیت درسی در منزل (شامل حل تمرین‌ها، مرور درس داده‌شده، تست زدن به‌صورت فردی یا با دوستان)" به‌عنوان کوواریت انتخاب شد. البته لازم به ذکر است که نگرش دانش‌آموز به ریاضی از جمع نمرات چند سؤال نگرشی به‌دست‌آمده است. همچنین با توجه به پرسشنامه‌های اجراشده در سطح مدرسه «متغیرهای سطح مدرسه[۳]» شامل متغیرهای جنسیت معلم[۴]، سال‌های تدریس ریاضی معلم، رضایت شغلی معلم[۵]، ادراک معلم از برنامه درسی[۶] و انتظارات معلم[۷]، تأکید بر انجام تکلیف در منزل[۸]، تعامل معلم در مدرسه[۹] و منابع و امکانات مدرسه[۱۰] (شامل تجهیزات و وسایل لازم برای آزمایشگاه فیزیک، وجود کامپیوتر و نرم‌افزار برای آموزش فیزیک، وجود کتاب‌هایی در کتابخانه برای آموزش و تقویت فیزیک، امکانات صوتی- تصویری برای آموزش فیزیک)، به‌عنوان کوواریت انتخاب شد. بنابراین می‌توان مدل مذکور را در قالب معادله سطح اول و دوم به‌صورت زیر نوشت:

---

۱. Student-level variables

۲. Positive Affect toward Physics

3. School-level variables
4. Teacher gender
5. Teachers job satisfaction
6. Teachers understanding of the schools curricular goals
7. Teachers expectations
8. Teacher emphasis on Physics homework
9. Teacher interaction in school
10. School resources

سطح اول

$$\theta_{ij} = \beta_{0j} + \beta_{1j}Gender_{ij} + \beta_{2j}APH_{ij} + \beta_{3j}TDH_{ij} + \beta_{4j}HA_{ij} + e_{ij},$$

سطح دوم

$$\beta_{0j} = \gamma_{00} + \gamma_{01}T.Gender + \gamma_{02}T.T.Phy + \gamma_{03}T.JOB\ Sat + \gamma_{04}T.Und.Curr + \gamma_{05}T.Exp + \gamma_{06}T.Emphasis + \gamma_{07}T.Interaction + \gamma_{08}Resource + u_{0j},$$

در ادامه شاخص‌های مربوط به تحلیل چندسطحی مدل با کووریت (با پیشین آگاهی‌بخش) به‌طور خلاصه در جدول زیر گزارش‌شده است.

جدول ۲. شاخص‌های آمار بیزی مربوط به تحلیل HLM با کووریت دانش‌آموز و مدرسه در آزمون فیزیک

پارامتر	$\bar{X}$	SD	MCSE	MD	فاصله اطمینان ٪۹۵ بیزی	
σ^2_{θ}	۰/۱۰۵	۰/۰۰۲	۰/۰۰۰	۰/۱۰۵	۰/۱۰۲	۰/۱۰۸
τ^2_{00}	۰/۱۱۷	۰/۰۳۱	۰/۰۰۱	۰/۱۱۲	۰/۰۷۰	۰/۱۹۱
DIC	۱۴۲۶۶/۵					
ICC	۰/۵۲۷					
پارامتر	ضریب				فاصله اطمینان ٪۹۵ بیزی	
جنسیت	۰/۰۰۸				−۱/۹۲۶	۱/۹۴۵
نگرش به فیزیک	−۰/۲۹۲				−۱/۶۲۷	۱/۰۴۴
زمان در منزل برای فیزیک	۰/۰۰۰				−۰/۰۰۴	۰/۰۰۴
فعالیت در منزل	−۰/۱۴۹				−۱/۷۵۹	۱/۴۶۳
γ_{00}	۲/۴۱۹				۰/۵۴۷	۴/۲۶۹
γ_{01}	۱/۸۹۰				۰/۶۰۰	۳/۱۴۳
γ_{02}	۰/۰۸۷				۰/۰۴۵	۰/۱۳۰
γ_{03}	۰/۲۴۳				−۰/۳۲۷	۰/۸۰۹
γ_{04}	۱/۴۳۶				۰/۴۸۹	۲/۳۵۵
γ_{05}	۰/۰۴۷				−۰/۷۴۰	۰/۸۶۲
γ_{06}	۰/۵۶۱				−۰/۱۱۳	۱/۲۴۴
γ_{07}	۱/۱۷۴				۰/۱۲۳	۲/۲۱۶
γ_{08}	−۰/۰۱۱				−۰/۱۴۸	۰/۱۳۵

با توجه به نتایج جدول ۲، چنین استنباط می‌شود که در آزمون فیزیک در حدود ۵۳ درصد از واریانس نمرات فیزیک به‌دست‌آمده به ویژگی‌های دانش‌آموزان (جنسیت، نگرش وی به فیزیک، میزان زمانی که وی در خانه به فیزیک اختصاص می‌دهد و فعالیت درسی فیزیک در منزل) و ویژگی‌های مدرسه و معلم (جنسیت معلم، سابقه تدریس فیزیک معلم، رضایت شغلی معلم، ادراک معلم از برنامه درسی، انتظارات معلم، تأکید بر انجام تکلیف در منزل، تعامل معلم در مدرسه و منابع مدرسه) برمی‌گردد. از بین کوواریت‌های واردشده در مدل مذکور در آزمون فیزیک از متغیرهای سطح اول (دانش‌آموز) هیچ‌کدام از متغیرها و از متغیرهای سطح دوم، $_{01}$ (جنسیت معلم)، $_{02}$ (سابقه تدریس فیزیک معلم)، $_{04}$ (ادراک معلم از برنامه درسی) و $_{07}$ (تعامل معلم در مدرسه) معنادار (در سطح اطمینان بیزی ۹۵درصد) گزارش‌شده است.

سؤال سوم: آیا بر اساس رویکرد IRT برآورد نمرات توانایی فیزیک دانش‌آموزان تحت تأثیر مدرسه است؟

به منظور پاسخگویی به این سؤال پژوهش در آزمون فیزیک تیمز پیشرفته ابتدا تحلیل حساسیت انجام شد و چون نتایج پسین نسبت به تغییر پیشین‌ها نیرومندی لازم را نشان داد مدلی که دارای پیشین آگاهی‌بخش بود انتخاب شد و در ادامه شاخص‌های مربوط به تحلیل چندسطحی مدل صفر (با پیشین آگاهی‌بخش) به‌طور خلاصه در جدول زیر گزارش‌شده است. در ادامه جداول مربوط به شاخص DIC، پارامترهای آلفا و بتا بیزی، و پارامترهای واریانس خطا و اثر مدرسه در جداول جداگانه ارائه‌شده است.

جدول ۳. شاخص DIC به‌منظور برازش مناسب‌ترین مدل‌های صفرِ MLIRT یک، دو و سه پارامتری آزمون فیزیک

مدل‌های صفر	Dbar	Dhat	pD	DIC
یک پارامتری	۶۱۹۹۰/۴	۶۰۰۷۵/۹	۱۹۱۴/۵۴	۶۳۹۰۵
دو پارامتری	۶۰۹۵۳/۳	۵۸۹۲۰/۸	۲۰۳۲/۵	۶۲۹۸۵/۸
سه پارامتری	۸۰۴۷۶	۸۰۳۴۷/۳	۱۲۸/۶۳۹	۸۰۶۰۴/۶

با توجه به نتایج جدول ۳، شاخص DIC برای مدل صفر دو پارامتری (۶۲۹۸۵/۸) نسبت دو مدل یک پارامتری (۶۳۹۰۵/۰) و سه پارامتری (۸۰۶۰۴/۶)، میزان کمتری را گزارش کرده است که نشان‌دهندهٔ برازش مناسب‌تر مدل مذکور با داده‌ها بوده است. به‌عبارت‌دیگر مدل صفر دو پارامتری MLIRT به‌منظور برآورد دقیق پارامترهای سؤال (آلفا و بتا) و نیز توانایی افراد (θ) مناسب گزارش‌شده و مورداستفاده قرار گرفت. به‌منظور تبیین میزان اثر مدرسه در واریانس نمرات توانایی فیزیک دانش‌آموزان بر اساس IRT از شاخص ICC بهره برده شده است؛ شاخص مذکور به همراه شاخص‌های آمار بیزی پارامترهای واریانس نمره خطا (σ_θ^2) و واریانس مدرسه (τ_{00}^2) در جدول زیر ارائه‌شده است:

جدول ۴. شاخص‌های آمار بیزی پارامترهای واریانس نمره خطا و واریانس مدرسه در آزمون فیزیک

پارامتر	$\bar{X}$	SD	MCSE	MD	فاصله اطمینان ٪۹۵ بیزی	
σ_θ^2	۱/۴۸۳	۰/۲۳۳	۰/۰۲۹	۱/۴۹۵	۰/۹۹۹	۱/۹۲۴
τ_{00}^2	۲/۷۵۵	۰/۷۴۶	۰/۰۵۲	۲/۶۷۸	۱/۵۴۶	۴/۴۲۲
ICC	۰/۶۵					
پارامتر	ضریب				فاصله اطمینان ٪۹۵ بیزی	
γ_{00}	-۰/۱۷۰				-۱/۰۶۲	۰/۷۱۸

با توجه به نتایج جدول ۴، چنین استنباط می‌شود که بر اساس IRT در آزمون فیزیک، حدود ۶۵ درصد از واریانس نمرات توانایی فیزیک به‌دست‌آمده به مدرسه‌ای برمی‌گردد که دانش‌آموز در آن مشغول به تحصیل بوده است.

سؤال چهارم: آیا بر اساس IRT برآورد نمرات توانایی فیزیک دانش آموزان تحت تأثیر ویژگی‌های دانش‌آموز و ویژگی‌های معلم و مدرسه است؟

به‌منظور پاسخگویی به همین سؤال پژوهش در آزمون فیزیک تیمز پیشرفته ابتدا ارزیابی نیرومندی از طریق تحلیل حساسیت انجام شد و چون نتایج پسین نسبت به تغییر پیشین‌ها نیرومندی لازم را نشان داد مدلی که دارای پیشین آگاهی‌بخش بود انتخاب شد. در ادامه جداول مربوط به شاخص DIC به‌منظور میزان برازش مدل‌های یک، دو و سه پارامتری ،

پارامترهای آلفا و بتا بیزی، و پارامترهـای واریـانس خطـا و اثـر مدرسـه مربـوط بـه آزمـون فیزیک در جداول جداگانه ارائه‌شده است.

جدول ۵. شاخص‌های DIC مدل‌های MLIRT یک، دو و سه پارامتری به همراه کووریت (آزمون فیزیک)

مدل‌های صفر	Dbar	Dhat	pD	DIC
یک پارامتری	۶۱۹۹۸/۵	۶۰۳۱۵/۶	۱۶۸۳	۶۳۶۸۱/۵
دو پارامتری	۶۰۹۶۸/۳	۵۹۱۶۷/۲	۱۸۰۱/۱	۶۲۷۶۹/۴
سه پارامتری	۸۰۲۴۶/۶	۸۰۰۸۲/۹	۱۶۳/۷	۸۰۴۱۰/۳

با توجه به نتایج جدول ۵، شاخص DIC برای مدل صفر دو پارامتری (۶۲۷۶۹/۴) نسبت بـه دو مدل یک پارامتری (۶۳۶۸۱/۵) و سه پـارامتری (۸۰۴۱۰/۳)، میـزان کمتـری را گـزارش کرده است که نشان‌دهندهٔ برازش مناسب‌تر مدل مذکور با داده‌ها بوده است. به‌عبارت‌دیگر مدل MLIRT دو پارامتری به همـراه کووریـت‌هـا (ویژگـی‌هـای دانـش‌آمـوز و مدرسـه و معلم) به‌منظور برآورد دقیق پارامترهای سؤال (آلفا و بتـا) و نیـز توانـایی افـراد (θ) مناسـب گزارش‌شده و مورداستفاده قرار گرفت. در جدول زیر شاخص‌های آماری بیزی مربـوط بـه پارامترهای متغیرهای سطح اول و متغیرهای سطح دوم گزارش‌شده است:

جدول ۶. اثرات متغیرهای سطح دوم و معناداری آنها (آزمون فیزیک)

پارامتر	$\bar{X}$	SD	MCSE	MD	فاصله اطمینان ٪۹۵ بیزی	
σ^2_{θ}	۰/۶۳	۰/۱۲	۰/۰۱	۰/۶۲	۰/۴۰	۰/۸۷
τ^2_{00}	۱/۴۰	۰/۳۳	۰/۰۲	۱/۳۶	۰/۹۰	۲/۱۸
ICC	۰/۶۸۹					

پارامتر		ضریب	فاصله اطمینان ٪۹۵ بیزی	
γ_{00}	عرض از مبدأ	۰/۲۷	−۰/۹۶	۱/۴۲
γ_{01}	جنسیت معلم	۰/۶۲	۰/۱۲	۱/۱۲
γ_{02}	سابقه تدریس	۰/۰۱۴	۰/۰۰۱	۰/۰۲۷
γ_{03}	رضایت شغلی	−۰/۰۴	−۰/۲۱	۰/۱۳
γ_{04}	ادراک معلم از برنامه	۰/۳۷	۰/۰۶	۰/۷۰
γ_{05}	انتظارات معلم	−۰/۲۶	−۰/۵۴	۰/۰۱

پارامتر		$\bar{X}$	SD	MCSE	MD	فاصله اطمینان ٪۹۵ بیزی	
γ_{06}	تأکید بر تکلیف منزل	-۰/۰۶				-۰/۲۸	۰/۱۸
γ_{07}	تعامل معلم در مدرسه	-۰/۱۱				-۰/۴۶	۰/۲۶
γ_{08}	منابع مدرسه	۰/۰۶				۰/۰۲	۰/۱۰
Gender	جنسیت دانش‌آموز	-۰/۰۰۴				-۱/۷۵۴	۱/۷۴۸
PHM	نگرش به فیزیک	۰/۲۲۷				۰/۱۰۷	۰/۵۵۹
TDH	زمان در منزل برای فیزیک	۰/۰۰۱				-۱/۹۴۲	۱/۹۴۴
AH	فعالیت در منزل	۰/۰۰۶				-۱/۹۶۶	۱/۹۵۴

با توجه به نتایج جدول ۶، چنین استنباط می‌شود که در آزمون فیزیک ۶۹ درصد از واریانس نمرات توانایی فیزیک به‌دست‌آمده به ویژگی‌های دانش‌آموزان (جنسیت دانش‌آموز، نگرش دانش‌آموز به فیزیک، میزان زمانی که وی در خانه به فیزیک اختصاص می‌دهد و فعالیت درسی فیزیک در منزل) و ویژگی‌های مدرسه و معلم (جنسیت معلم، سابقه تدریس فیزیک معلم، رضایت شغلی معلم، ادراک معلم از برنامه درسی، انتظارات معلم، تأکید بر انجام تکلیف در منزل، تعامل معلم در مدرسه و منابع مدرسه) برمی‌گردد. از بین کوواریت‌های واردشده در مدل مذکور γ_{01} (جنسیت معلم)، γ_{02} (سابقه تدریس فیزیک معلم)، γ_{04} (ادراک معلم از برنامه درسی) و γ_{08} (منابع و امکانات مدرسه برای درس فیزیک) و نگرش دانش‌آموز به فیزیک معنادار (در سطح اطمینان بیزی ۹۵درصد) گزارش‌شده است.

بحث و نتیجه‌گیری

آنچه در این پژوهش مدّ نظر بوده، بدست آوردن میزان تأثیر مدارس بر نمرات کسب‌شدهٔ دانش آموزان شرکت‌کننده در آزمون تیمز پیشرفته ۲۰۰۸ بوده است. دراین‌بین سعی شد اعتبار تحلیل‌های چندسطحی با نگاهی به نظریه‌های اندازه‌گیری مورد کاوش بیشتر قرار گیرد. به همین جهت تحلیل‌های چندسطحی هم با رویکرد کلاسیک اندازه‌گیری و هم با رویکرد نظریه جدید اندازه‌گیری (IRT) انجام شدند. در مطالعه حاضر مزایای آماری

فرمول‌بندی چندسطحی و برآورد بیزی مدل‌های سؤال- پاسخ به‌ویژه مدل دو پارامتری چندسطحی (2PL MLIRT) موردتوجه و بحث قرار گرفت. در جدول زیر، برآوردهای پارامتر مدل صفر (بدون متغیرهای اکتشافی) و مدل همراه با کوواریت تحت مدل MLIRT و مدل چندسطحی خطی (ML) برای آزمون فیزیک تیمز پیشرفته ارائه‌شده است.

جدول ۷. مقایسه برآورد پارامترهای مدل صفر و مدل همراه با کوواریت تحلیل‌های ML و MLIRT

آزمون	شاخص	ML		MLIRT	
		مدل صفر	مدل با کوواریت دانش‌آموز، مدرسه و معلم	مدل صفر	مدل با کوواریت دانش‌آموز، مدرسه و معلم
فیزیک	σ_{θ}^{2}	۴/۰۲۸۷	۰/۱۰۵	۱/۴۸۳	۰/۶۳
	τ_{00}^{2}	۰/۰۱۷۶	۰/۱۱۷	۲/۷۵۵	۱/۴۰
	ICC	۰/۳۹	۰/۵۳	۰/۶۵	۰/۶۹
	DIC	۱۸۵۲۷/۳۰	۱۴۲۶۶/۵	۶۲۹۸۵/۸	۶۲۷۶۹/۴

بر اساس نتایج جدول ۷ می‌توان مشاهده کرد که اولاً در آزمون فیزیک میزان واریانس تبیین شده اثر مدارس در مدل‌های چندسطحی نظریه سؤال- پاسخ (MLIRT) نسبت به مدل‌های چندسطحی نمرهٔ واقعی (MLTS) بیشتر است و مدل‌های MLIRT توانسته است تفاوت‌های بین میانگین‌های مدارس را بیشتر استخراج کند؛ آنچه از این نتیجه برمی‌آید این است که برخورد با سازهٔ پیشرفت تحصیلی به‌عنوان متغیری مکنون و در نظر گرفتن خطای اندازه‌گیری تک‌تک سؤالات در تحلیل‌ها در چارچوب بیزی و با استفاده از نمونه‌گیری گیبس به‌طور چشمگیری توانمندی تحلیل‌های چندسطحی را بهبود می‌بخشد و افزایش معنادار نسبت واریانس تبیین شده را در پی خواهد داشت. ثانیاً، زمانی که کوواریت به مدل‌های ML صفر و یا MLIRT صفر افزوده می‌شود، میزان واریانس تبیین شده اثر مدارس نیز افزایش می‌یابد. به‌گونه‌ای که مدل MLIRT با کوواریت‌های سطح دانش‌آموز و مدرسه نسبت به سایر مدل‌های موردبحث بیشترین قدرت آشکارسازی اثر مدارس را دارد.

دوم آنکه، میزان تفاوت بین مدارس شرکت‌کننده در آزمون ریاضیات تیمز پیشرفته که در

شاخص ICC (همبستگی درون رده‌ای) منعکس‌شده است در تحلیل‌های MLIRT بسیار زیاد است و این امر به‌طور تلویحی نشان می‌دهد که تفاوت و تبعیض آموزشی بین مدارس زیاد است که این امر عمدتاً ناشی از متغیرهای در سطح مدرسه (نظیر متغیرهای مرتبط به معلم یا متغیرهای مرتبط به مدرسه) می‌باشد. این نتیجه به‌ویژه برای آموزش‌وپرورش ایران که دارای ساختار متمرکز است و کل تصمیم‌های آموزشی از بالا هدایت می‌شود چندان امیدوارکننده و مطلوب نیست. میزان تفاوت بالای بین مدارس گویای پیامی روشن برای متصدیان و تصمیم-گیران حوزهٔ آموزش‌وپرورش و سنجش کشور است و آن اینکه مقایسه دانش‌آموزان و اتخاذ تصمیم‌های سرنوشت‌ساز بر اساس آزمون‌های متداول (و حتی آزمون‌های باکیفیت بالا نظیر آزمون‌های تیمز) در شرایط فعلی چندان منصفانه نمی‌نماید و حداقل تا رفع بخشی از تبعیض-های آموزشی و برقراری نسبی عدالت آموزشی، استفاده از کنکور معقول‌ترین و منصفانه‌ترین راهکار است و اجرایی شدن قانون حذف کنکور و جایگزینی سوابق تحصیلی دانش‌آموزان در مدرسه بر اساس برگزاری آزمون‌های استاندارد نیازمند برنامه‌ریزی دقیق به‌منظور برقراری عدالت آموزشی در مدارس کل کشور است. در کل نتایج این پژوهش حاکی از آن است که در آزمونی استاندارد همچون تیمز، نمرات دانش‌آموزان به مدارس محل تحصیل آنها ارتباط قوی دارد، چه برسد به آزمون‌های غیراستاندارد نهایی پایان‌ترم که به‌جز مسائل امنیتی برگزاری، خطای مصححان نیز در آن دخیل است.

همچنین از بین عوامل سطح مدرسه و معلم در درس ریاضی متغیرهای منابع مدرسه، سابقه تدریس معلم و انتظارات معلم و در درس فیزیک متغیرهای منابع مدرسه، سابقه تدریس معلم، جنسیت معلم و ادراک معلم از برنامه درسی معنادارترین متغیرها از بین متغیرهای کمکی به دست آمدند. مروری بر پیشینه پیشرفت تحصیلی نشان می‌دهد که نتایج این مطالعه همسو با نتایج برخی از پژوهشگران بوده و حمایت‌ها و انتظارات آموزشی معلم (مایا[1]، ۲۰۰۲)، منابع و امکانات مدرسه برای آموزش ریاضی[2] (رامیرز، ۲۰۰۶؛ ساتولین

---

۱. Maya

۲. School resources for mathematics instruction

جی، ۱۹۹۶)، مکان مدرسه[۱] (چیپت، ۲۰۰۸؛ هاوی، ۲۰۰۶؛ رامیرز، ۲۰۰۶)، سابقه تدریس معلم (سائولین جی[۲]، ۱۹۹۶)، جنسیت معلم (کال و همکاران[۳]، ۲۰۰۰)، و ادراک معلم از برنامه درسی ارتباط مثبت و معناداری با پیشرفت تحصیلی ریاضی دانش‌آموزان دارد که ازاین‌بین مهمترین متغیر مؤثر بر تفاوت بین مدارس متغیر انتظارات آموزشی معلم است. مرور پیشینه گویای آن است که هرچند مدارس و متغیرهای مربوط به آن نمی‌توانند پیشرفت تحصیلی دانش‌آموزان را به‌طور کامل تبیین کنند و مسئول کامل تفاوت‌هایی که بین گروه‌های مختلف مشاهده می‌شود نیستند، ولی کیفیت مدرسه و ماهیت محیط یادگیری که فراهم می‌کنند واقعاً تفاوت ایجاد می‌کند (چو[۴]، ۲۰۱۰) که مهمترین مؤلفه معنادار محیط یادگیری فعال، معلم، ویژگی‌های مرتبط با معلم و نوع تعامل یاددهی-یادگیری بین معلم و دانش‌آموز است. روشن است که نوع نگاه معلم به فرایند تدریس در انگیزش دانش‌آموز و به دنبال آن پیشرفت تحصیلی وی نقش بسزایی دارد. آگاهی از این موضوع که در تحلیل چندسطحی پژوهش حاضر انتظارات معلم متغیری معنادار بوده و اینکه اثر انتظار تأثیرات تعیین‌کننده‌ای را به همراه دارد، می‌تواند در انگیزش و اعتمادبه‌نفس دانش‌آموزان و به دنبال آن کاهش تفاوت بین مدارس نقش عمده‌ای داشته باشد.

یکی از متغیرهای مهم و مؤثر بر پیشرفت تحصیلی در درس علوم فیزیکی و فیزیک، که در پرسشنامه‌های آزمون تیمز پیشرفته نیز مورد تأکید بوده، منابع و امکانات مدرسه[۵] (شامل تجهیزات و وسایل لازم برای آزمایشگاه فیزیک، وجود کامپیوتر و نرم‌افزار برای آموزش فیزیک، وجود کتاب‌هایی در کتابخانه برای آموزش و تقویت فیزیک، امکانات صوتی-تصویری برای آموزش فیزیک) است. وجود منابع و امکانات در یادگیری بهتر درس

۱. Location of the school

۲. Saulinggi

۳. Kahle et al.

۴. Cho

۵. School resources

فیزیک و علوم از اهمیت ویژه‌ای برخوردار است، چون هر چه بیشتر بتوان از آزمایشگاه فیزیک، کامپیوتر و وسایل صوتی- تصویری برای آموزش استفاده کرد امکان ملموس کردن مطالب و واقعیت‌های قابل‌مشاهده جهان هستی برای دانش‌آموزان بیشتر است (تیت، ۲۰۰۱؛ روگان و گریسون[۱]، ۲۰۰۳؛ روگان و آلدوس[۲]، ۲۰۰۵؛ جورج و کاپلان، ۱۹۹۸). این امر به‌ویژه برای دانش‌آموزان محروم یا دانش‌آموزان با وضعیت اقتصادی- اجتماعی پایین از اهمیت بیشتری برخوردار است (برکام و همکاران[۳]، ۱۹۹۷). البته مسئله مهمی که باید به آن اشاره کرد آن است که وجود منابع و امکانات به‌خودی‌خود استفاده از آنها را تضمین نمی‌کند، چه‌بسا مدرسه‌ای دارای آزمایشگاهی مجهز باشد ولی معلم نتواند به‌خوبی از آن استفاده کند یا کامپیوتر در مدرسه وجود داشته باشد ولی معلم نتواند از کامپیوتر برای آموزش و اهداف آموزشی بهره بگیرد. استارک و گری[۴] (۱۹۹۹) در مطالعه خود بر این موضوع تأکید کرده‌اند که در گزارش‌های تیمز علیرغم وجود منابع و امکانات خوب در برخی مدارس میزان استفاده از آنها چندان بالا گزارش نشده است.

مسئله دیگری که در رابطه باوجود منابع و امکانات به‌ویژه برای درس فیزیک اهمیت دارد این است که هرچه قدر این منابع کمتر باشد و دانش‌آموزان به آزمایشگاه و تجهیزات مرتبط با یادگیری درس فیزیک کمتر دسترسی داشته باشند این امر آنها را به حفظ کردن مطالب بیشتر سوق می‌دهد و مهارت حل مسئله در دانش آموزان تقویت نخواهد شد (بلک و همکاران[۵]، ۱۹۹۸). میر و همکاران[۶] (۲۰۰۰) نیز گزارش کرده‌اند که کمبود منابع و مواد آموزشی تأثیرات منفی چشمگیری بر آموزش و یادگیری درس علوم و فیزیک دارد.

---

۱. Rogan & Grayson

۲. Rogan & Aldous

۳. Burkam et al.

۴. Stark & Gray

۵. Black et al.

۶. Mayer et al.

همچنین بر اساس نظر کلمن[۱] (۱۹۷۵) آموزش علوم و فیزیک کمتر تحت تأثیر متغیرهای خانوادگی و بیشتر تحت تأثیر متغیرهای مربوط به مدرسه است. یعنی فرصت‌های یادگیری فیزیک و علوم فیزیکی بیشتر در مدارس اتفاق می‌افتد و هر چه منابع و امکانات مدرسه برای درس علوم فیزیکی و فیزیک کمتر باشد پیشرفت دانش‌آموزان در علوم فیزیکی و فیزیک دانش‌آموزان کمتر است (کامبر و کیوس[۲]، ۱۹۷۳). طبق نظر ردی[۳] (۲۰۰۵b)؛ تیت[۴] (۲۰۰۱)؛ همیلتون و همکاران[۵] (۱۹۹۵) مدارس تنها منبع غنی‌ای هستند که فرصت یادگیری علوم و فیزیک را در اختیار دانش‌آموزان می‌گذارند.

بنابر گزارش‌های مرکز ملی مطالعات تیمز و پرلز (۱۳۸۹)، میانگین عملکرد دانش‌آموزان ایران در درس علوم فیزیکی پایه چهارم (۴۵۴) و فیزیک پایه هشتم (۴۷۰) و فیزیک تیمز پیشرفته (۴۶۰) به‌طور معنادار پایین‌تر از میانگین بین‌المللی (۵۰۰) است. جایگاه کشور ایران در فیزیک تیمز پیشرفته ۲۰۰۸ از میان ۹ کشور (با میانگین عملکرد ۵۰۰) در رتبه هفتم (با میانگین عملکرد ۴۶۰) می‌باشد که با میانگین کشورهای شرکت‌کننده تفاوت معناداری دارد. متوسط عملکرد دانش‌آموزان رشته ریاضی- فیزیک ایران در درس ریاضیات برابر ۴۹۷ (رتبه چهارم از بین ۱۱ کشور)، اما در درس فیزیک برابر با ۴۶۰ (رتبه هفتم از بین ۱۱ کشور) می‌باشد و اختلاف بین این دو درس ۳۷ نمره و معنادار گزارش‌شده است. مرور پیشینه تجربی (ردی، ۲۰۰۵b؛ تیت، ۲۰۰۱؛ همیلتون و همکاران، ۱۹۹۵) نشان می‌دهد که در یادگیری درس فیزیک مدارس و امکانات آنها نقش عمده‌ای ایفا می‌کنند و بنابراین یکی از دلایل مهم ضعف دانش‌آموزان در درس فیزیک نسبت به ریاضی به مسئله وجود و به‌کارگیری منابع و امکانات برای یادگیری درس فیزیک در مدارس به می‌گردد. میزان بالای تفاوت بین مدارس در تحلیل‌های چندسطحی پژوهش حاضر نیز مؤید این مطلب

---

۱. Coleman

۲. Comber & Keeves

۳. Reddy

۴. Tate

۵. Hamilton et al.

است.

با توجه به آنکه پژوهشگر در روند جمع‌آوری داده‌ها هیچ نقشی نداشته و به بیانی دیگر تحلیل‌های موردنظر بر روی‌داده‌های آزمون تیمز پیشرفته ۲۰۰۸ انجام‌شده است، تعمیم نتایج به رشته‌های دیگر و یا مقاطع تحصیلی ابتدایی و راهنمایی باید بااحتیاط صورت‌پذیرد. در این پژوهش، مدل‌های MLIRT ارائه‌شده بر اساس یک رابطهٔ غیرخطی بین متغیرهای پاسخ و متغیرهای اکتشافی (کووریت‌ها) بود و در آن فرض بر این بود که متغیرهای اکتشافی بدون خطا اندازه‌گیری می‌شوند. اغلب امکان اندازه‌گیری تمامی متغیرهای اکتشافی مرتبط به‌طور دقیق وجود ندارد. سنجش خطاهای اندازه‌گیری مهم است، چون وجود خطاها در اندازه‌گیری متغیرهای اکتشافی می‌تواند رابطهٔ بین متغیرهای پاسخ و اکتشافی را دچار سوگیری کند. بر این اساس پیشنهاد می‌شود که تمامی سؤالات و گویه‌های متغیرهای اکتشافی (کووریت‌ها) نیز به‌صورت IRT (و نه به‌صورت یک نمرهٔ کل) وارد تحلیل شوند و بسته به نوع متغیر با کووریت‌ها به‌عنوان متغیرهای مکنون (و نه آشکار) برخورد شود و نتایج حاصل از این دو رویکرد مقایسه شود.

مقادیر محتمل[1] یا ارزش‌های پذیرفتنی یکی از تکنیک‌های مورداستفاده در مقیاس‌سازی مطالعات در مقیاس جمعیتی بزرگ است که ویژگی‌های آماری دقیقی نسبت به نمرات مشاهده‌شده برای مقایسه‌های زیرگروه‌ها دارد و می‌تواند به تشکیل توزیع واقعی شایستگی کمک کند (ون داویر، گنزالس و میسلوی، ۲۰۰۹) و مناسب توصیف عملکرد جامعه آماری است. پیشنهاد می‌شود کارایی این روش نسبت به سایر روش‌های برآورد (همچون EAP، بیشینه احتمالات وارم[2] (WML) و MLE) چه با روش‌های شبیه‌سازی و چه با داده‌های واقعی مورد مقایسه قرار گیرد.

1. Plausible values
2. Warm maximum likelihood (WML)

منابع

جمشیدی، محمدحسین (۱۳۸۰). نظریه عدالت از دیدگاه فارابی، امام خمینی و شهید صدر. تهران: پژوهشکده امام خمینی (ره) و انقلاب اسلامی.

سرمد، زهره، بازرگان، عباس، حجازی، الهه. (۱۳۸۴). روش‌های تحقیق در علوم رفتاری. تهران. نشر آگاه.

عبدوس، میترا. (۱۳۸۱). بررسی عوامل مؤثر در فرصت‌های آموزشی نابرابر بین دختران و پسران. خلاصه مقالات مهندسی اصلاحات در آموزش‌وپرورش. تهران: پژوهشکده تعلیم و تربیت.

کریمی، عبدالعظیم (۱۳۸۸). مجموعه سؤال‌های علوم و ریاضیات تیمز TIMSS (پایه سوم راهنمایی). تهران: پژوهشگاه مطالعات آموزش‌وپرورش.

کریمی، عبدالعظیم (۱۳۸۹). مجموعه سؤال‌های قابل‌انتشار فیزیک و ریاضیات تیمز پیشرفته ۲۰۰۸. تهران: مرکز ملی بین‌المللی تیمز و پرلز. پژوهشگاه مطالعات آموزش‌وپرورش.

کریمی، عبدالعظیم (۱۳۸۷b). تیمز: گزارش اجمالی مهمترین یافته‌های ملی و بین‌المللی تیمز ۲۰۰۷ در مقایسه با تیمز ۹۵، ۹۹ و ۲۰۰۳. تهران: پژوهشگاه مطالعات آموزش‌وپرورش.

مطهری، مرتضی. (۱۳۵۲). عدل الهی. تهران: صدرا.

Bennett, C. , Both, C. & Yeadle, S. (2001). Mainstreaming equality in the committees of the Scottish parliament. University of strathclyde.

Black, T. R., Atwaru-Okello, D., Kiwanuka, J., Serwadda, D., Birabi, O., Malinga, F., Biumigishu, A., & Rodd, A. (1998). Science education in Uganda: progress and possibilities. International Journal of Science Education, 20(2), 239-252.

Burkam, D. T., Lee, V. E., & Smerdon, B. A. (1997). Gender and science learning early in high school: subject matter and laboratory experiences. American Educational Research Journal, 34(2), 297-331.

Chepete, P. (2008). Modeling of the Factors Affecting Mathematical Achievement of Form 1 Students in Botswana based on the 2003 Trends in International Mathematics and Science Study. Unpublished Doctor of Philosophy, Indiana University.

Cho, M. O. (2010). A comparison of the effectiveness of science education in Korea and South Africa: A multilevel analysis of TIMSS 2003 data. PhD dissertation, University of Peretoria, Faculty of Education.

Coleman, J. S. (1975). Methods and results in the IEA studies of effects of school on learning. Review of Educational Research, 45(3), 355-386.

Comber, L. C., Keeves, J. P. (1973). Science education in nineteen countries. New York: John Wiley & Sons.

George, R., & Kaplan, M. D. (1998). A structural model of parent and teacher influences on science attitudes of eighth grades: Evidence from NELS: 88. Science Education, 82(1), 93-109.

Hamilton, L. S., Nussbaum, E. M., Kupermintz, H., Kerkhoven, J. I. M., & Snow, R. E. (1995). Enhancing the validity and usefulness of large-scale educational assessments: NELS: 88 science achievement. American Educational Research Journal, 32(3), 555-581.

Howie, S. (2006). Multi-level Factors Affecting the Performance of South African Pupils in Mathematics. In S. Howie & T. Plomp (Eds.), *Contexts of Learning Mathematics and Science(pp.157-176):* Routledge.

Kahle, J. B., Meece, J., & Scantlebury, K. (2000). Urban African-American middle school science students: does standards-based teaching make a difference? Journal of Research in Science Teaching, 37 (9), 1019-1041.

Maya, Cynthia. (2002). Factors affecting the achievement motivation of high school student in Maine. University of Southern Maine.

Mayer, D. P., Mullens, J. E., Moore, M. T., & Ralph, J. (2000), Monitoring school quality: an indicators report. Washington, D. C.: National Center for Education Statistics. Retrived March 14, 2007 from http://nces.ed.gov/pubs2001/2001030.pdf

Ramírez, M. J. (2006). Factors Related to Mathematics Achievement in Chile. In S. Howie & T. Plomp (Eds.), *Contexts of Learning Mathematics and Science*: Routledge.

Reddy, V. (2005b). State of mathematics and science education: schools are not equal. *Perspectives in Education*, 23(3), 125.

Rogan, J. M., & Aldous, C. (2005). Relationships between the constructs of a theory of curriculum implementation. *Journal of Research in Science Teaching*, 42(3), 313-336.

Rogan, J. M., & Grayson, D. (2003). Towards a theory of curriculum implementation with particular reference to science education in developing countries. *International Journal of Science Education*, 25(11), 1171-1204.

Saulinggi, Simon. (1996). The Relationships Between school and out of

school factors and mathematics Achievement of students in Indonesian primary schools. Source: DAI ˚ A 56/ 09.

Stark, R., & Gray, D. (1999). Gender preferences in learning science. *International Journal of Science Education*, 21(6), 633-643.

Tate, W. (2001). Science education as a civil right: urban schools and opportunity-to-learn considerations. Journal of Research in Science Teaching, 38(9), 1015-1028.

Von Davier, M., Gonzalez, E., & Mislevy, R. J. (2009). What is plausible values and why are they useful? *Issues and Methodologies in Large-Scale Assessments* (Vol. 2, pp. 9- 36). Hamburg: IEA-ETS Research Institute.

Zhao, N.N., Valcke, M., Desoete, A., Verhaeghe, J. (2012). The quadratic relationship between socioeconomic status and learning performance in China by multilevel analysis: Implications for policies to foster education equity. International Journal of Educational Development. 32 . 412˚422.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی