

مقاله پژوهشی اصیل

## ساخت و رواسازی آزمون تشخیصی حساب‌نارسانی برای کودکان پایه سوم دبستان

**حجت‌الله فراهانی<sup>۱</sup>**

گروه روانشناسی، دانشگاه اصفهان

**دکتر مختار ملک‌پور**

گروه روانشناسی، دانشگاه اصفهان

**دکتر شعله امیری**

گروه روانشناسی، دانشگاه اصفهان

**دکتر حسین مولوی**

گروه روانشناسی، دانشگاه اصفهان

**دکتر مهرداد کلانتری**

گروه روانشناسی، دانشگاه اصفهان

**هدف:** هدف این پژوهش ساخت و رواسازی آزمون تشخیصی حساب‌نارسانی پایه سوم دبستان شهر اصفهان بود. **روش:** در این پژوهش پیمایشی، کتاب ریاضی پایه سوم دبستان تحلیل محتوا شد. ضریب کریپندورف مبنی بر توافق تحلیل محتوا ۰/۹۱ به دست آمد. پس از آن بر اساس مقوله‌های استخراج‌شده و اهداف رفتاری هر مقوله و با توجه به نظریه شناختی بلوم و حجم هر مقوله در کتاب ریاضی پایه سوم تعداد زیادی سؤال طرح و به‌طور مقدماتی اجرا گردید. سپس فرم نهایی روی یک نمونه ۲۰۰ نفری از دانش‌آموزان پایه سوم دبستان (۱۵۵ پسر و ۱۴۵ دختر) که در سال تحصیلی ۸۶-۱۳۸۵ به روش نمونه‌برداری خوشه‌ای چندمرحله‌ای از نواحی پنج‌گانه آموزشی شهر اصفهان انتخاب شده بودند، اجرا و پارامترهای سؤال آزمون بر اساس نظریه کلاسیک آزمون (CTT) و نظریه سؤال - پاسخ (IRT) محاسبه گردید. **یافته‌ها:** نتایج نشان داد که این آزمون دارای روایی مطلوب و نیز پایایی مناسب ( $\alpha = 0/92$ )، ضریب بازآزمایی ( $0/87$ ) است و حساسیت و ویژگی (دقت) آن به ترتیب ۰/۹۸ و ۰/۹۶ به دست آمد. پارامتر تشخیصی  $a(g)$  و دشواری  $b(g)$  بر اساس نظریه سؤال - پاسخ نیز نشان داد که همه سؤال‌ها دارای توان تشخیص مناسب (بالای ۰/۵) و دشواری مطلوب‌اند و با مدل دو پارامتری برازش کامل دارند ( $p > 0/05$ )، منحنی تابع آگاهی آزمون نیز نشان داد که مناسب‌ترین کاربرد آن در سطوح توانایی ریاضی  $R = -0/6$  تا  $R = 0/6$  است. **نتیجه‌گیری:** این آزمون دارای ویژگی‌های روان‌سنجی مناسب برای تشخیص کودکان حساب‌نارسانی پایه سوم می‌باشد.

کلید واژه‌ها: آزمون تشخیصی، حساب‌نارسانی، نظریه کلاسیک آزمون (CTT)، نظریه سؤال - پاسخ (IRT)

**مقدمه**

ریاضی یکی از درس‌هایی است که با وجود توجه زیادی که در مدارس به آن می‌شود، یادگیری آن برای بسیاری از دانش‌آموزان مشکل است. این در حالی است که معلمان می‌توانند با ارزشیابی درست توانایی‌های ریاضی دانش‌آموزان، بازده تدریس خود و در نتیجه موفقیت تحصیلی و پیشرفت ریاضی دانش‌آموزان را افزایش دهند. زیرا شناسایی نقاط ضعف و نارسانی‌هایی ریاضی

کودکان سبب انطباق مواد درسی و روش آموزش با سطح یادگیری دانش‌آموز می‌گردد. از طریق سنجش و اندازه‌گیری روایی و پایایی می‌توان برنامه آموزشی درست و جامعی تدوین کرد که به بهترین روش درک مفاهیم نظری ریاضی را با کاربرد آن در زندگی روزمره همراه سازد.

دشواری در فهم روابط عددی برای برخی از کودکان از همان سال‌های نخست زندگی که سایر کودکان در شمارش، درک واحد، جورکردن و مقایسه تجاری به دست می‌آورند، آغاز می‌شود که می‌تواند کارکردی پایین‌تر از حدانتظار را در آزمون‌های پیشرفت

۱- نشانی تماس: اصفهان، خیابان هزارچریب، گروه روانشناسی، دانشگاه اصفهان.

Email: icphaf@yahoo.com

هنجاربایی گردید (محمد اسماعیل و هومن، ۱۳۸۱). آزمون دیگری که برای سنجش شایستگی ریاضی طراحی شده، آزمون شایستگی اوترخت برای ریاضیات کودکان است، که شامل هشت مقیاس و در مجموع ۱۲۰ سوال است (کرپرال<sup>۱۴</sup>، ۱۹۹۵؛ اندوریخت<sup>۱۵</sup>، ۱۹۹۲، به نقل از عریضی، کاووسیان و کدیور، ۱۳۸۳). این آزمون را عریضی و همکاران (۱۳۸۳) پایایی سنجی و رواسازی کرده‌اند. در پژوهش دیگری چین آوه، عرب‌زاده و منشی‌زاده (۱۳۸۳) به ساخت و هنجاربایی آزمون تشخیصی ریاضی برای دانش‌آموزان ۶ تا ۱۳ ساله شهر شیراز پرداختند. آنچه در این پژوهش مورد توجه قرار گرفت، این بود که با توجه به تنوع این اختلال و یک‌دست نبودن ویژگی‌های آن (احدی و کاکاوند، ۱۳۸۲)، به نظر می‌رسد استفاده از یک آزمون استاندارد ملی پژوهشگران را به نتیجه مطلوبی نرساند و بر روی محتوایی این آزمون‌ها تأکید خاص نشده است. از سوی دیگر بنا به نظر باترورث<sup>۱۶</sup> (۲۰۰۳، ۲۰۰۵)، جوردن و مونتانی<sup>۱۷</sup> (۱۹۹۷)، لندرل<sup>۱۸</sup>، بوان<sup>۱۹</sup> و باترورث (۲۰۰۴) و آشکرافت<sup>۲۰</sup>، یاماشیتا<sup>۲۱</sup> و آرام<sup>۲۲</sup> (۱۹۹۲) اگر در مقوله شمارش زمان لحاظ نگردد، کودکان حساب‌نارسا نیز می‌توانند مانند سایر کودکان عادی به سؤال‌ها پاسخ درست بدهند، زیرا آنها از راهکارهای نادرست یعنی شمارش با انگشتان استفاده کرده و در مدت زمان طولانی‌تر پاسخ درست را ایجاد می‌کنند. بنابراین در این پژوهش، خرده‌آزمون شمارش زمان‌مند گردید و از سوی دیگر برای دستیابی به پارامترهای دقیق سؤال (ضرایب دشواری و تشخیصی)، جامعه آماری آزمودنی‌ها (نتا) و تعیین منحنی آگاهی آزمون علاوه بر نظریه کلاسیک آزمون<sup>۲۳</sup> (CTT) از نظریه سؤال-پاسخ<sup>۲۴</sup> (IRT) نیز استفاده شد. در ایران بر اساس نظریه سؤال-پاسخ (IRT) در زمینه ریاضیات هیچ آزمونی ساخته

تحصیلی دوران دبستان برای آنان پیش‌بینی کند (رمضانی، ۱۳۸۰). حساب‌نارسایی که به عنوان یک اختلال در سومین نسخه دستامه تشخیصی و آماری اختلالات روانی<sup>۱</sup> (DSM-III) (یعنی از سال ۱۹۸۰ به بعد) مطرح گردید، عبارت است از ناتوانی در انجام مهارت‌های حساب با توجه به ظرفیت هوش و سطح آموزشی مورد انتظار از کودک که این مهارت‌ها می‌بایست به کمک آزمون‌های میزان‌شده فردی اندازه‌گیری شده باشد. بر اساس ویراست چهارم اصلاح‌شده دستامه تشخیصی و آماری اختلالات روانی<sup>۲</sup> (DSM-IV-TR) این کودکان در چهار گروه از مهارت‌های زبانی، ادراکی، ریاضی و توجهی مرتبط با ریاضیات مشکل دارند.

با وجود اینکه این اختلال مشکل جدیدی تلقی نمی‌شود، اما ماهیت آن تا حد زیادی مبهم مانده است (فوکس<sup>۳</sup> و فوکس<sup>۴</sup>، ۲۰۰۵). به همین دلیل پژوهشگران از این اختلال تقسیم‌بندی‌های متعددی ارائه داده‌اند (جوردن<sup>۵</sup>، هانیک<sup>۶</sup> و کاپلان<sup>۷</sup>، ۲۰۰۳؛ روسلی<sup>۸</sup> و آردیلا<sup>۹</sup>، ۱۹۸۹). فوکس و فوکس (۲۰۰۵) همه‌گیری اختلال ریاضی را در دبستان پنج درصد، رمضانی (۱۳۸۲) و حمید (۱۳۸۵) در شهر تهران به ترتیب ۱/۶۳ و ۳/۶ درصد برآورد کرده‌اند.

در سبب‌شناسی این اختلال فرضیه‌هایی مطرح شده که در مجموع می‌توان به ترکیبی از تأثیرات محیط و ژنتیک اشاره کرد (روسولی<sup>۸</sup>، نونل<sup>۱۰</sup>، ۲۰۰۷). بنا به تأکید انجمن روان‌پزشکی آمریکا تشخیص قطعی فقط پس از اجرای آزمون استاندارد شده انفرادی در ریاضی و کسب نمره‌ای بسیار پایین‌تر از سطح مورد انتظار (بر حسب میزان تحصیلات و ظرفیت هوشی کودک) امکان‌پذیر است (کاپلان و سادوک<sup>۱۱</sup>، ۱۳۸۳). بنابراین تشخیص صحیح و دقیق آن نیازمند اجرای یک آزمون تشخیصی استاندارد می‌باشد. برای سنجش این اختلال روش‌های رسمی و غیررسمی متعددی طراحی شده است که به عنوان نمونه می‌توان به آزمون ریاضی کی‌مت<sup>۱۲</sup> اشاره نمود که کنولی<sup>۱۳</sup> (۱۹۸۸)، به نقل از رمضانی (۱۳۸۲) آن را در سال ۱۹۸۸ هنجاربایی کرد. این آزمون که به منظور تعیین نقاط قوت و ضعف دانش‌آموزان در حوزه‌های مختلف ریاضی به کار می‌رود، شامل ۱۳ خرده‌آزمون و سه بخش مفاهیم اساسی، عملیات و کاربردها است که از اهمیت آموزشی تقریباً یکسانی برخوردارند. این آزمون در ایران برای دانش‌آموزان

- 1- Diagnostic and Statistical Manual of Mental Disorders-Third Edition
- 2- Diagnostic and Statistical Manual of Mental Disorders-Text Revision
- 3- Fuchs
- 4- Jordan
- 5- Hanich
- 6- Kaplan
- 7- Rosselli
- 8- Ardila
- 9- Rrusselle
- 10- Noel
- 11- Sadoek
- 12- Key Math
- 13- Conolly
- 14- Korporal
- 15- Andoricht
- 16- Butterworth
- 17- Montani
- 18- Landrel
- 19- Bevan
- 20- Ashcraft
- 21- Yamashita
- 22- Aram
- 23- Classic Test Theory
- 24- Item-Response Theory

نشده و این اولین پژوهش در این راستا است. هدف این رویکرد آن است که علاوه بر برآورد دقیق پارامترهای آزمون و سؤال، نارسانی‌های نظریه کلاسیک آزمون، از جمله وابستگی پارامتر سؤال به نمونه خاص را از میان بردارد (هومن، ۱۳۷۱). بنابراین، مطالعه حاضر به منظور طراحی آزمون تشخیصی حساب‌نارسانی برای کودکان پایه سوم دبستان شهر اصفهان و بررسی ویژگی‌های روان‌سنجی و هنجاریابی آن برای این کودکان انجام شد.

## روش

### جامعه، نمونه آماری و روش نمونه‌گیری

جامعه آماری این پژوهش کلیه کودکان دختر و پسر دبستانی پایه سوم شهر اصفهان بودند که در سال تحصیلی ۸۶-۱۳۸۵ در مدارس ابتدایی این شهر تحصیل می‌کردند (جدول ۱). از این تعداد با استفاده از فرمول موزر<sup>۱</sup> و کالتون<sup>۲</sup> و با در نظر گرفتن دو درصد برآورد نسبت همه‌گیرشناسی این اختلال در جامعه و یک درصد خطای مورد پذیرش، ۳۰۰ دانش‌آموز (۱۵۵ پسر و ۱۴۵ دختر) از نواحی پنج‌گانه به‌عنوان نمونه انتخاب شدند (رمضانی، ۱۳۸۲).

در جدول ۱ تعداد افراد جامعه مورد بررسی به تفکیک ناحیه آموزشی و جنسیت آمده است. بدین صورت که با مراجعه به اداره کل آموزش و پرورش اصفهان فهرست مدارس ابتدایی هر پنج ناحیه تهیه شد و سپس با توجه به جمعیت دانش‌آموزان پایه سوم، از هر ناحیه چند مدرسه و از میان آنها یک یا چند کلاس به‌صورت تصادفی انتخاب گردیدند.

### ابزار پژوهش و مراحل ساخت آن

برای تهیه پرسش‌های این آزمون ابتدا کتاب ریاضی پایه سوم به‌وسیله تعدادی از اساتید روان‌شناسی و علوم تربیتی، متخصصان

روان‌سنجی، پژوهشگران و سرگروه‌های آموزش ریاضی پایه سوم تحلیل محتوا شد و ضریب توافق کربیندورف<sup>۳</sup> ۰/۹۱ به‌دست آمد که نشان دهنده توافق افراد در تحلیل محتوای این کتاب است. سپس عناوین مجدداً به تأیید سرگروه‌های آموزشی رسید و سپس سؤال‌ها با توجه به نظریه شناختی بلوم<sup>۴</sup> (۱۹۶۵) که همسو با نظریه گانیه<sup>۵</sup> (۱۹۷۰)، به نقل از چین آوه و همکاران، (۱۳۸۳) است، طرح گردید. تعداد سؤال‌های هر مقوله، با توجه به حجم آن در کتاب ریاضی تعیین شد. این مقولات شامل نه حیطه اعداد، جمع و تفریق، ارزش مکانی، اندازه‌گیری، ضرب و تقسیم، هندسه، کسر و اعشار، نسبت و درصد، تخمین و حل مسأله بود. سپس این سؤال‌ها مورد بررسی مقدماتی قرار گرفتند و بعد روی یکصد دانش‌آموز (۵۰ دختر و ۵۰ پسر) اجرا شدند. پرسش‌ها به‌صورت انفرادی و تشریحی اجرا گردید و برای هر پاسخ درست نمره یک و برای هر پاسخ غلط نمره صفر منظور شد. نتایج تحلیل و پرسش‌ها از نظر سطح دشواری (p)، ضریب تشخیص (d)، همسانی درونی ( $\alpha$ ) و روایی، همزمان با آزمون کمی‌ت ریاضی ایران (محمد اسماعیل و هومن، ۱۳۸۱) و نمره حساب آزمون و وکسلر کودکان (وکسلر، ۱۳۷۳) محاسبه گردید. سپس سؤال‌های دارای ضریب بالا ( $\alpha$ ) بالای ۰/۸ و ضرایب تشخیص معنادار (بالای ۰/۴) انتخاب شدند. برای کمک به استقلال موضعی سؤال‌ها (شرط لازم برای تحلیل سؤال - پاسخ) سؤال‌ها بر اساس درجه دشواری به‌طور تصادفی در هر خرده‌آزمون قرار داده شدند.

در پایان ۵۰ سؤال باقی ماند و از آنجا که ریاضی یک مجموعه به‌هم پیوسته تلقی می‌شود، آزمون تشخیصی حساب‌نارسانی پایه دوم (۳۰ سؤال) نیز که بر اساس اصولی مشابه این آزمون ساخته شده بود،\* به سؤال‌ها اضافه و در مجموع ۸۰ سؤال به‌عنوان آزمون تشخیصی حساب‌نارسانی پایه پنجم اجرا گردید. برای آگاهی از وضعیت هوشی و ادراکی - دیداری شرکت‌کنندگان علاوه بر اجرای آزمون هوش کودکان

جدول ۱- تعداد افراد جامعه مورد بررسی به تفکیک ناحیه آموزشی و جنسیت

	ناحیه ۱	ناحیه ۲	ناحیه ۳	ناحیه ۴	ناحیه ۵
پسر	۹۹۳	۱۵۵۸	۲۷۵۳	۳۶۰۰	۳۳۰۸
دختر	۱۰۰۵	۱۵۳۳	۲۵۲۸	۳۳۲۱	۳۱۰۳

1- Mozer  
3- Krippendorff  
5- Ganie

2- Kalton  
4- Bloom

\* این مقاله قسمتی از پایان‌نامه دکترای روان‌شناسی با تأکید بر روان‌سنجی بود که ضمن آن برای هر پنج پایه دبستان، یک آزمون تشخیصی حساب‌نارسانی ساخت و رواسازی شد.

دو نیمه‌سازی محاسبه و نقاط برش، حساسیت و ویژگی (دقت) سؤال‌ها از طریق تحلیل تمیزی<sup>۳</sup> تعیین شد. منظور از حساسیت یک آزمون امکان تشخیص افراد حساب‌نارسانا از افراد عادی و منظور از ویژگی (دقت) یک آزمون، توان یک آزمون در تشخیص موارد سالم است. پارامترهای سؤال در این پژوهش براساس مدل دو پارامتری سؤال- پاسخ (IRT) برآورد شد. ضرایب تشخیص سؤال‌ها  $a(g)$ ، پارامتر دشواری  $b(g)$  و عامل حدس  $c(g)$  بود که به دلیل ماهیت تشریحی این آزمون عامل حدس در پاسخگویی تأثیر نداشت و صفر در نظر گرفته شد. بنابراین از مدل دو پارامتری استفاده شد. در این مدل، پارامتر تشخیص سؤال به قدرت سؤال‌های یک آزمون در تشخیص جداسازی افراد قوی و ضعیف اشاره دارد (آلن<sup>۴</sup> و ین<sup>۵</sup>، ۱۳۷۴). معمولاً دامنه  $a(g)$  در عمل بین ۰/۵ تا ۲/۵ تغییر می‌کند که مقادیر کمتر از ۰/۵ نشان‌دهنده توان تشخیص پایین سؤال‌ها و مقادیر بزرگتر نیز (که به ندرت به دست می‌آید) (هومن، ۱۳۷۱)، پارامتر دشواری سؤال  $b(g)$  است که معمولاً بین ۲/۵- تا ۲/۵+ مقیاس‌پردازی می‌شود. پارامتر دشواری سؤال نشان‌دهنده نقطه عطف منحنی شاخص سؤال است که ۵۰ درصد آزمودنی‌هایی که در آن سطح قرار دارند جواب سؤال را می‌دانند یا می‌توانند جواب آن را پیدا کنند (ستاری، ۱۳۸۲).

برای تعیین برازش سؤال‌ها با مدل دو پارامتری سؤال - پاسخ مقادیر مجذور کای محاسبه و سطح معناداری آنها تعیین شد. بررسی داده‌ها نشان داد که برای احراز دو پیش فرض تحلیل IRT یعنی استقلال موضعی<sup>۶</sup> و تک‌بعدی بودن<sup>۷</sup> آزمون، بیش از ۲۰ درصد واریانس آزمون در یک عامل قرار می‌گیرد که این نشان‌دهنده برقراری فرض تک‌بعدی بودن است که اگر رعایت شود استقلال نیز قدری رعایت می‌شود (باقی، ۱۹۹۵، به نقل از ستاری، ۱۳۸۲). برای کمک به برقراری این فرض سؤال‌هایی با درجه دشواری‌های تصادفی قرار داده شدند. در تمام مراحل از نرم‌افزارهای SPSS، Testfact و Bilog استفاده شد.

(وکسلر، ۱۳۷۳) و آزمون ادراکی- دیداری آندره‌ری (میرهاشمی، ۱۳۷۱) که از روایی و پایایی مطلوبی برخوردار بودند، نمرات روخوانی فارسی، انشا و دیکته نیز از آنها پرسیده شد. همچنین طی یک مصاحبه بالینی از آنها خواسته شد متن مشخصی را بخوانند و آن را برای آزمونگر بازگو کنند و در مواردی که شرکت‌کننده دچار مشکل ادراکی، بالینی یا هوشی بود از نمونه خارج و به جای آن فرد دیگری به‌طور تصادفی انتخاب و جایگزین می‌شد. برای تعیین زمان مطلوب در خرده‌آزمون شمارش، مدت زمانی که ۷۰ درصد شرکت‌کنندگان (آناستازی<sup>۱</sup>، ۱۳۷۳؛ ترندایک<sup>۲</sup>، ۱۳۷۵؛ سیف، ۱۳۷۵) قادر به پاسخ‌گویی سؤال‌ها بودند، به‌عنوان زمان مطلوب در نظر گرفته شد. برای تعیین نمونه دارای حساب‌نارسایی بر اساس روش غیر آزمون، به‌منظور تعیین نقطه برش، حساسیت و ویژگی (دقت) و روایی تشخیصی سؤال‌های آزمون تشخیصی نهایی، از آموزگار پایه سوم کلاس مربوطه خواسته شد بر مبنای شناخت فردی، دانش آموزشی را که فقط در درس ریاضی با مشکل روبه‌رو بودند، معرفی کند. نمره کمتر از ۱۲ (۱۰ درصد پایین توزیع) به‌عنوان اولین ضابطه تعیین گردید و سپس از آموزگار خواسته شد که به فهرست واری تشخیصی DSM-IV-TR برای هر مورد پاسخ دهد و پس از پاسخ‌دهی، ماده‌هایی که با بلی پاسخ داده می‌شدند جمع می‌شدند و در صورتی که آزمودنی حداقل چهار ماده (از ۱۱ ماده فهرست واری تشخیصی) را با «بلی» پاسخ داده بود به‌عنوان حساب‌نارسایی تلقی می‌شد. بدین ترتیب ۲۵ نفر تعیین شدند.

این آزمون به‌وسیله پژوهشگر و دانشجویان کارشناسی ارشد و کارشناسی که در مورد ایجاد رابطه و برقراری ارتباط مؤثر و همدلانه با کودک و اجرای آزمون تجربه داشتند و آموزش کامل دیده بودند، در هفته‌های پایانی سال تحصیلی ۸۶-۱۳۸۵ اجرا گردید.

پارامترهای سؤال در این پژوهش در نظریه کلامیک آزمون (CTT)، ضریب دشواری (یعنی نسبت افرادی که به سؤال‌ها پاسخ درست دادند، و ضریب تشخیص (روایی) سؤال بود که از طریق همبستگی دو رشته‌ای نقطه‌ای سؤال با نمره کل به‌دست آمد. پایایی از طریق آلفای کرونباخ، بازآزمایی و

1- Anastasia  
3- discrimination analysis  
5- Yen  
7- unidimensionality

2- Thorndike  
4- Allen  
6- local independence

## یافته‌ها

که حاکی از روایی همزمان مطلوب خرده‌آزمون‌ها و نمره کل این آزمون است.

همچنین همبستگی بین آزمون تشخيصي حساب‌نارسايي پایه سوم و خرده‌آزمون حساب در آزمون هوش ویسک-آرنیز معنادار بود که نشان‌دهنده روایی همزمان و مطلوب این آزمون است ( $r=0/79$ ,  $p<0/001$ ).

برای تعیین روایی تشخيصي سؤال‌های آزمون تشخيصي حساب‌نارسايي، این آزمون روی افراد عادی و گروه حساب‌نارسا که بر اساس فهرست واریسی DSM-IV-TR و با نمره کمتر از ۱۲ در ریاضی انتخاب شده بودند، اجرا و تحلیل شد. تحلیل MANOVA سؤال‌ها نشان داد که براساس همه سؤال‌ها می‌توان افراد حساب‌نارسا و عادی را به‌طور معناداری ( $p<0/05$ ) تفکیک کرد. اگر آزمونی این ویژگی را داشته باشد، دارای روایی سازه است (سرمد، بازرگان و حجازی، ۱۳۷۶).

**پایایی آزمون تشخيصي حساب‌نارسايي:** برای تعیین پایایی این آزمون از سه روش آلفای کرونباخ، دونیمه‌سازی (همسانی درونی) و بازآزمایی (ثبات) استفاده شد. ضرایب آلفای کرونباخ در خرده‌آزمون‌ها در محدوده ۰/۷۴-۰/۹۲ و در کل ۰/۹۳، در روش دونیمه‌سازی در محدوده ۰/۷۱-۰/۹۱ و در کل ۰/۸۹ و در بازآزمایی در محدوده ۰/۸۲-۰/۹۲ و در کل ۰/۹۲ به‌دست آمد. برای تعیین ضریب بازآزمایی، شرکت‌کنندگان پس از چهار هفته مجدداً آزمون شدند.

**نقطه برش، حساسیت و ویژگی (دقت) آزمون تشخيصي حساب‌نارسايي:** نقطه برش در خرده‌آزمون‌ها در محدوده ۱۰-۲ و در کل ۵۱، حساسیت خرده‌آزمون‌ها در محدوده ۹۸-۸۸ درصد و در کل ۹۸ درصد و ویژگی (دقت) خرده‌آزمون‌ها در محدوده ۹۶-۸۴ درصد و در کل ۹۶ درصد به‌دست آمد.

حساسیت و ویژگی (دقت) آزمون در حد بسیار مطلوبی بود، یعنی با این آزمون می‌توان ۹۹ درصد موارد حساب‌نارسا و ۹۶ درصد موارد سالم را به درستی تشخیص داد.

**ضرایب تشخيصي و دشواری سؤال‌های آزمون براساس نظریه کلاسیک آزمون (CTT) و سؤال - پاسخ (IRT):** ضرایب تشخيصي و دشواری سؤال‌های آزمون براساس نظریه کلاسیک آزمون (CTT) و سؤال - پاسخ (IRT) در جدول ۳ آورده شده است.

**روایی صوری و محتوایی آزمون:** برای تعیین روایی صوری و محتوایی این آزمون مراحل زیر انجام شد:

الف) کتاب ریاضی پایه سوم دبستان تحلیل محتوا و مقوله‌های مختلف استخراج گردید.

ب) ضریب توافق کریپندورف مبنی بر توافق تحلیل محتوا محاسبه شد که ۰/۹۱ به‌دست آمد.

ج) اهداف آموزشی و رفتاری بر اساس مقوله‌های استخراج‌شده تعریف و تدوین شدند.

د) بر اساس بند «ج» مهارت‌هایی که نشان‌دهنده محتوای مورد نظر بودند، تعیین شدند.

ه) از میان سؤال‌های امتحان‌های ریاضی دو نوبت (صبح و عصر) پایه سوم دبستان‌های نواحی پنج‌گانه آموزش و پرورش شهر اصفهان خزانه‌ای از سؤال‌های پنج سال اخیر تهیه و بررسی شد.

و) بر اساس مراحل «الف» تا «ه» مجموعه‌ای از سؤال‌ها طراحی و تنظیم شدند.

ز) سؤال‌های نهایی به‌وسیله اساتید گروه علوم تربیتی و روان‌شناسی، متخصصان روان‌سنجی، اساتید راهنما و مشاور و سرگروه‌های آموزش ریاضی پایه سوم دبستان مورد بازنگری نهایی قرار گرفت و نسبت روایی محتوایی (CVR<sub>i</sub>) بر اساس توافق در تعیین روایی محتوایی سؤال‌ها محاسبه و دامنه‌ای از ۰/۷۹-۰/۸۲ به‌دست آمد که معنادار و حاکی از روایی محتوایی بالا و مطلوب این آزمون بود.

ح) پس از اجرای مقدماتی، سؤال‌های مبهم (از نظر دانش‌آموزان) شناسایی، مجدداً عبارت‌بندی و در صورت لزوم در محتوای آنها تغییر داده شد.

**روایی آزمون تشخيصي حساب‌نارسايي پایه پنجم:** برای تعیین روایی همزمان این آزمون از خرده‌آزمون‌های مرتبط در آزمون ریاضی ایران کی‌مت (محمد اسماعیل و هومن، ۱۳۸۱) و خرده‌آزمون حساب در آزمون هوش ویسک-آر (وکسلر، ۱۳۷۳) استفاده شد. نتایج نشان داد که همبستگی بین همه خرده‌آزمون‌ها و نمره کل این آزمون و خرده‌آزمون‌های مرتبط در آزمون ریاضی ایران کی‌مت در محدوده ۰/۸۷-۰/۸۹ و معنادار بودند ( $p<0/001$ ).

جدول ۲- ضرایب تشخیصی و دشواری سؤالات، مقادیر کای دو و سطح معناداری

ضریب							ضریب						
df	P	$\chi^2$	b(g)	a(g)	$r_{pbis}$	pi	df	P	$\chi^2$	b(g)	a(g)	$r_{pbis}$	pi
8	0.028	11.71	1.5219	1.843	0.75	0.67	سوال ۴۱	8	0.011	13.71	1.843	0.76	0.57
8	0.051	9.31	1.471	1.847	0.64	0.55	سوال ۴۲	8	0.011	13.71	1.843	0.76	0.57
8	0.030	11.8	1.5026	1.857	0.67	0.63	سوال ۴۳	8	0.028	11.8	1.843	0.76	0.57
8	0.081	10.37	1.495	1.86	0.65	0.58	سوال ۴۴	8	0.037	11.87	1.843	0.76	0.57
9	0.014	13.5	1.516	1.87	0.72	0.63	سوال ۴۵	8	0.010	13.87	1.87	0.75	0.51
9	0.011	13.37	1.5285	1.882	0.69	0.69	سوال ۴۶	8	0.010	14.31	1.87	0.76	0.53
9	0.014	13.5	1.525	1.891	0.68	0.59	سوال ۴۷	8	0.081	11.51	1.88	0.77	0.52
9	0.081	10.8	1.5887	1.909	0.67	0.61	سوال ۴۸	8	0.081	11.7	1.891	0.78	0.53
9	0.052	11.31	1.5711	1.918	0.67	0.62	سوال ۴۹	9	0.048	8.71	1.911	0.81	0.56
9	0.037	11.81	1.5669	1.927	0.67	0.63	سوال ۵۰	9	0.019	12.31	1.911	0.79	0.57
9	0.015	13.7	1.508	1.937	0.81	0.60	سوال ۵۱	8	0.041	11.71	1.920	0.81	0.56
9	0.050	11.8	1.5356	1.945	0.65	0.67	سوال ۵۲	8	0.032	11.9	1.931	0.717	0.65
9	0.011	13.8	1.5422	1.953	0.67	0.59	سوال ۵۳	8	0.031	13.1	1.935	0.78	0.66
9	0.017	13.1	1.5599	1.961	0.67	0.67	سوال ۵۴	9	0.018	13.8	1.949	0.76	0.57
7	0.014	8.87	1.462	1.975	0.67	0.61	سوال ۵۵	9	0.038	11.7	1.959	0.64	0.61
8	0.055	11.7	1.509	1.983	0.64	0.62	سوال ۵۶	8	0.031	11.92	1.969	0.65	0.69
8	0.037	11.6	1.508	1.991	0.67	0.68	سوال ۵۷	8	0.053	11.7	1.975	0.68	0.62
9	0.017	13.1	1.529	1.999	0.66	0.68	سوال ۵۸	8	0.038	11.78	1.979	0.67	0.63
9	0.031	11.1	1.501	2.001	0.65	0.62	سوال ۵۹	8	0.011	13.8	1.980	1.982	0.66
9	0.054	11.6	1.509	2.009	0.69	0.68	سوال ۶۰	9	0.038	11.93	1.988	0.650	0.62
9	0.031	11.7	1.511	2.018	0.67	0.61	سوال ۶۱	9	0.061	11.03	1.991	0.657	0.61
9	0.011	13.87	1.522	2.028	0.66	0.60	سوال ۶۲	8	0.031	11.81	1.992	0.658	0.64
7	0.038	9.5	1.492	2.035	0.64	0.62	سوال ۶۳	9	0.038	11.83	1.999	0.683	0.68
7	0.043	13.51	1.505	2.044	0.68	0.67	سوال ۶۴	8	0.052	8.71	2.0	0.64	0.67
7	0.031	11.81	1.497	2.051	0.61	0.68	سوال ۶۵	8	0.061	9.32	2.002	0.67	0.68
7	0.015	13.8	1.509	2.06	0.67	0.67	سوال ۶۶	7	0.031	11.87	2.004	0.68	0.63
9	0.031	11.7	1.508	2.069	0.64	0.61	سوال ۶۷	9	0.053	11.7	2.005	0.68	0.67
9	0.031	11.7	1.508	2.078	0.64	0.61	سوال ۶۸	9	0.011	12.4	2.008	0.82	0.61
7	0.037	11.7	1.5285	2.087	0.65	0.66	سوال ۶۹	9	0.031	11.3	2.012	0.67	0.68
7	0.037	11.7	1.5285	2.096	0.68	0.66	سوال ۷۰	8	0.087	10.2	2.014	0.68	0.65
7	0.037	11.7	1.5285	2.105	0.61	0.67	سوال ۷۱	7	0.037	11.77	2.016	0.63	0.62
7	0.015	13.8	1.5269	2.114	0.65	0.61	سوال ۷۲	7	0.010	11.7	2.019	0.60	0.65
9	0.031	11.7	1.5269	2.123	0.64	0.62	سوال ۷۳	7	0.057	11.2	2.027	0.67	0.65
7	0.037	11.7	1.5269	2.132	0.64	0.62	سوال ۷۴	7	0.062	11.82	2.031	0.60	0.63
7	0.037	11.7	1.5269	2.141	0.64	0.62	سوال ۷۵	7	0.062	11.82	2.031	0.60	0.63
9	0.011	13.87	1.5269	2.150	0.64	0.62	سوال ۷۶	7	0.062	11.82	2.031	0.60	0.63
8	0.011	13.87	1.5269	2.159	0.64	0.62	سوال ۷۷	8	0.052	11.02	2.035	0.63	0.64
8	0.041	11.87	1.5269	2.168	0.64	0.62	سوال ۷۸	7	0.065	8.8	2.04	0.67	0.68
8	0.041	11.87	1.5269	2.177	0.64	0.62	سوال ۷۹	7	0.062	11.02	2.04	0.64	0.68
8	0.039	11.87	1.5269	2.186	0.64	0.62	سوال ۸۰	9	0.041	11.02	2.04	0.62	0.64

PI: ضریب دشواری هر سؤال در نظریه کلاسیک آزمون (CTT)؛ همبستگی دو رشته‌ای نقطه‌ای سوال با نمره کل (ضریب تشخیص)؛ در نظریه کلاسیک؛ a(g): ضریب تشخیص هر سؤال در نظریه سؤال پاسخ (ITR)؛ b(g): ضریب دشواری در نظریه سؤال پاسخ (ITR)؛ P: سطح معناداری؛ df: درجه آزادی

جدول ۳- رتبه‌های درصدی نمرات در هر یک از خرده‌آزمون‌های آزمون تشخيصي حساب‌نارسايي و کل

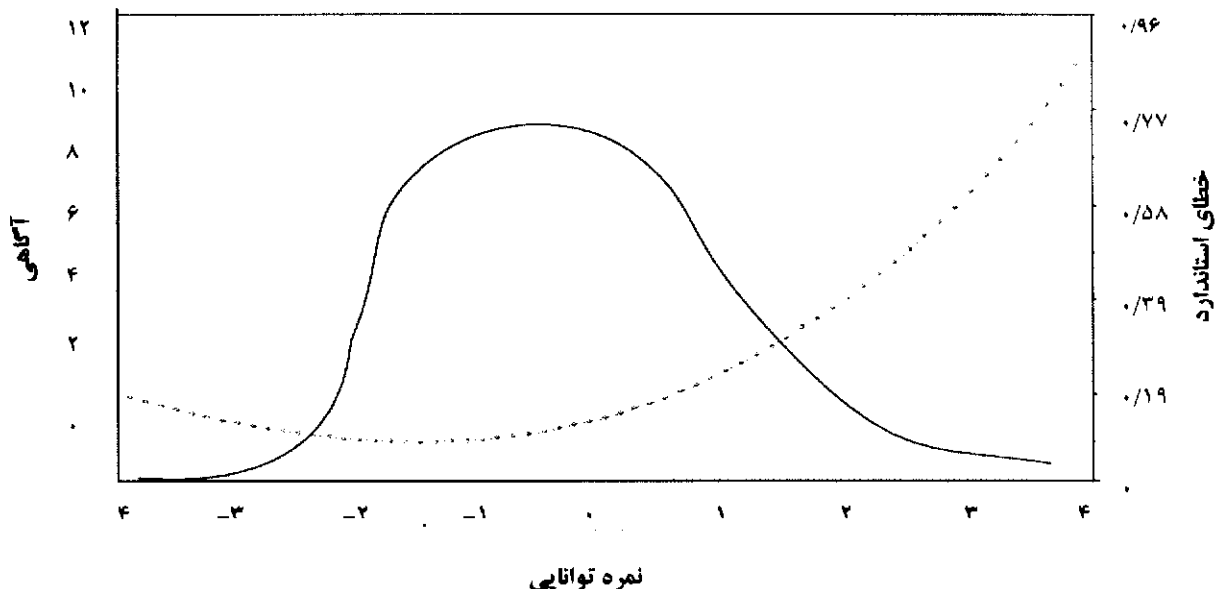
رتبه درصدی	خرده‌آزمون‌ها						
	شمارش	هندسه	جمع و تفریق	ضرب و تقسیم	کسر و اعشار	اندازه‌گیری	حل مسأله
۵	۸	۲	۸	۷	۲	۸	۷
۲۵	۱۰	۲	۹	۸	۳	۹	۸
۵۰	۱۲	۳	۱۱	۱۱	۴	۱۲	۹
۷۵	۱۴	۳	۱۳	۱۳	۴	۱۳	۱۱
۹۵	۱۵	۴	۱۵	۱۴	۵	۱۵	۱۲

رتبه‌های درصدی نمرات در هر یک از خرده‌آزمون‌های آزمون تشخيصي حساب‌نارسايي: همان گونه که در جدول ۴ ملاحظه می‌گردد، نمرات پنج درصد افراد در خرده‌آزمون شمارش کمتر از هشت، در خرده‌آزمون هندسه کمتر از دو، در خرده‌آزمون جمع و تفریق کمتر از هشت، در خرده‌آزمون ضرب و تقسیم کمتر از هفت، در خرده‌آزمون کسر و اعشار کمتر از دو، در خرده‌آزمون اندازه‌گیری کمتر از هشت، در خرده‌آزمون حل مسأله کمتر از هفت و در کل کمتر از ۴۷ است.

**منحنی تابع آگاهی آزمون (TIF):** همان گونه که در شکل ۱ ملاحظه می‌شود، بیشترین میزان آگاهی‌دهندگی این آزمون در فاصله‌ی تنای ۰/۶- تا ۰/۶+ قرار گرفته که نشان‌دهنده‌ی آن است که این آزمون برای افرادی که دارای این سطوح توانایی هستند، بیشترین و مناسب‌ترین کاربرد را دارد. میزان آگاهی‌دهندگی این

همان گونه که در جدول ۳ ملاحظه می‌گردد، ضرایب براساس نظریه کلاسیک ( $F_{pbis}$ ) بین ۰/۴ تا ۰/۸۵ نوسان دارد که نشان می‌دهد هیچ یک از ضرایب، منفی و پایین نیست. همچنین ضرایب دشواری سؤال‌ها ( $P_i$ ) بین ۰/۲۸ تا ۰/۸۸، ضرایب تشخيص  $a(g)$  براساس نظریه سؤال- پاسخ (IRT) بین ۲/۵ تا ۰/۵۲۳ با میانگین ۱/۱۹۸ و انحراف معیار ۰/۱۶۰ است. هیچ یک از ضرایب کمتر از ۰/۵ نیست. ضرایب دشواری  $b(g)$  نیز بین ۳/۰۰۰ تا ۰/۹- با میانگین ۱/۷۲ و انحراف معیار ۰/۷۱۵ است. سطح معناداری مقادیر کای دو سؤال‌ها نیز نشان می‌دهد که همگی با مدل دو پارامتری (2PLM) برازش دارند ( $p > ۰/۰۵$ ).

بررسی پارامتر توانایی ( $\theta$ ) آزمودنی‌ها در ریاضیات نیز نشان داد که آزمودنی‌ها دارای پارامتر توانایی ( $\theta$ ) بین ۲/۱۳۹- تا ۲/۹۵ با میانگین ۱/۴۳۵ و انحراف معیار ۰/۰۲۱ هستند.



شکل ۱- منحنی تابع آگاهی آزمون (TIF) آزمون تشخيصي حساب‌نارسايي پایه سوم

رواسازی آزمون تشخیصی دقیق برای سنجش حساب‌نارسانی پایه سوم به‌عنوان هدف این آزمون در نظر گرفته شد.

بررسی همه‌جانبه، تأکید بر تحلیل و مقوله‌بندی محتوای کتاب ریاضی پایه سوم، طراحی سؤال‌های مختلف و نظرخواهی از اساتید و صاحب‌نظران و سرگروه‌های آموزشی ریاضی پایه سوم، محاسبه و شاخص‌های کمی، توافق در تحلیل محتوا و تعیین رویی صوری و محتوایی این آزمون، همگی مطلوب‌بودن رویی آزمون را نشان می‌دهند. همچنین همبستگی خرده‌آزمون‌ها و نمره کل این آزمون با آزمون ریاضیات ایران کی‌مت و خرده‌آزمون حساب در آزمون هوش ویسک-آر (وکسلر، ۱۳۷۳) هم حاکی از رویی هم‌زمان و هم‌نشان‌دهنده رویی همگرای (سازه) این آزمون است که امکان تفکیک معنادار افراد حساب‌نارسا و عادی را براساس سؤال‌های این آزمون فراهم می‌آورد. وجود نقاط برش گواه دیگری برای رویی سازه این آزمون است، زیرا اگر آزمون‌دنی بتواند بین دو گروه تمایز قائل شود سازه دارای رویی است (سرمد و همکاران، ۱۳۷۶). بررسی پایایی خرده‌آزمون‌های این آزمون نشان داد که همسانی درونی این آزمون از طریق آلفای کرونباخ، دونیمه‌سازی و ثبات خرده‌آزمون‌ها، و پایایی نمره کل آن از طریق بازآزمایی مطلوب و قابل قبول است. مطابق با اعتقاد نانالی و برنشتاین (۱۹۹۴)، به نقل از مستامورنن، (۱۳۸۵) ضریب پایایی بالای ۰/۷ مطلوب تلقی می‌شود، بنابراین ضرایب پایایی خرده‌آزمون‌ها و کل این آزمون مطلوب و پذیرفتنی است.

حساسیت این آزمون نشان داد که در ۹۸ درصد موارد با این آزمون می‌توان افراد حساب‌نارسا را از سالم تفکیک کرد. بررسی ویژگی (دقت) این آزمون نیز نشان داد که با استفاده از این آزمون می‌توان در ۹۶ درصد موارد، افراد سالم را به درستی تشخیص داد. بررسی ضرایب همبستگی دو رشته‌ای نقطه‌ای سؤال‌های آزمون (ضرایب تشخیص) در نظریه کلاسیک نشان داد که هیچ یک از این ضرایب منفی و کمتر از ۰/۲ نیستند. حداقل ضرایب تشخیصی در نظریه کلاسیک آزمون (CTT) ۰/۲ می‌باشد (مستامورنن، ۱۳۸۵). از سوی دیگر ضرایب دشواری ( $P_i$ ) در این آزمون نشان داد که ضرایب دشواری سؤال‌ها در محدوده ۰/۲۸ تا ۰/۸۸ است، بنابراین سؤال ۷۵ این آزمون ساده‌ترین و سؤال‌های ۳۰ و ۵۹ دشوارترین

آزمون در سطوح توانایی بسیار پایین (-۴) و بسیار بالا (+۴) بسیار اندک است. میزان آگاهی‌دهندگی این آزمون در این سطوح به کمتر از ۲ واحد می‌رسد.

**مدت زمان لازم برای اجرای خرده‌آزمون شمارش در آزمون تشخیصی:** زمان لازم برای اجرای خرده‌آزمون شمارش در این آزمون، مدت زمانی در نظر گرفته شد که ۷۰ درصد آزمون‌دنی‌ها در آن دوره زمانی توانستند به همه سؤال‌ها پاسخ بدهند. این زمان در خرده‌آزمون شمارش ۶ دقیقه و ۵۰ ثانیه به‌دست آمد.

## بحث

به نظر استیونس (۱۹۸۸)، به نقل از چین آوه و همکاران، (۱۳۸۳) تعداد روش‌های ارزشیابی که طراحی آنها به گونه‌ای باشد که مستقیماً به آموزش‌های جبرانی منجر شود اندک است، بنابراین هدف اصلی این پژوهش ساخت و رواسازی آزمون دقیق با حداکثر قدرت تشخیص موارد حساب‌نارسا و عادی بود. بررسی آزمون‌های مختلفی که در حیطه ریاضیات تدوین شده‌اند نشان می‌دهد که در هر یک از این آزمون‌ها اگرچه نقاط مثبت هم وجود دارد، اما ضعف‌های جدی نیز به چشم می‌خورد. مثلاً در آزمون تشخیصی ریاضی فقط به ارزشیابی حل مسأله پرداخته شده (روچ و همکاران، ۱۹۲۱)، به نقل از چین آوه و همکاران، (۱۳۸۳) و یا پرسشنامه ریاضی پیش‌دبستانی کرانر (۱۹۷۴) برای ارزشیابی توانایی کاربرد مفاهیم کمی کودکان تهیه شده است. از سوی دیگر برخی از این آزمون‌ها فقط برای پایه‌های خاصی تهیه شده و در آنها به پیوستاری‌بودن مفاهیم و مهارت‌های ریاضی توجهی نشده است؛ مثلاً در آزمون تشخیصی ریاضی کالیفرنیا مخصوص پایه‌های سوم تا هشتم، پرسشنامه ریاضی پیش‌دبستانی کرانر برای کودکان پیش‌دبستانی و پایه اول دبستان (به نقل از باس و وگن، ۲۰۰۰)، و نیز در برخی ابزارهای رویی محتوایی که صرفاً هنجاریابی شده‌اند محتوای دروس ریاضی پایه‌های مختلف و حجم این مطالب در کتاب‌های ریاضی ایرانی کمتر مورد توجه قرار گرفته است. همچنین هیچ یک از این آزمون‌ها زمان‌مند نبوده و در نتیجه قادر به تفکیک افراد حساب‌نارسا و عادی نیستند (باترورث، ۲۰۰۳، لندل و همکاران، ۲۰۰۴؛ جوردن و مونیتانی، ۱۹۹۷؛ آشکرافت و همکاران، ۱۹۹۲). بنابراین ساخت و



(۲۰۰۳) هماهنگ است. براساس یافته‌های این پژوهشگران اگر در آزمون‌های تشخیصی ریاضی (خرده‌آزمون‌های شمارش) زمان لحاظ نشود، کودکان حساب‌نارسان نیز مانند کودکان عادی عمل خواهند کرد. در نتیجه زمانی این آزمون‌های تشخیصی دارای حساسیت کافی خواهند بود که زمانمند باشند.

بررسی منحنی آگاهی آزمون نشان داد که بیشترین میزان آگاهی‌دهندگی آن در فاصله سطح توانایی  $+0/6$  تا  $-0/6$  است که نشان‌دهنده آن است که این آزمون برای افرادی که در این سطح توانایی قرار دارند، بهترین و بیشترین کاربرد را خواهد داشت. میزان آگاهی‌دهی این آزمون در این سطوح توانایی حدود ۹ و پایین‌ترین میزان در سطوح توانایی بسیار پایین ۳- و بسیار بالا ۴+ است که در این سطوح این میزان به کمتر از یک واحد می‌رسد. با توجه به متوسط تابع آگاهی این آزمون که به وسیله رایانه  $4/87$  به دست آمده، منحنی این آزمون یک تابع آگاهی متوسط به بالا را نشان می‌دهد.

بنابراین این آزمون دارای قدرت لازم جهت سنجش متغیر زیربنایی (خصیصه مکتون ریاضی) است و بر اساس آن می‌توان به ارزشیابی نسبتاً دقیق وضعیت پیشرفت ریاضی و تفکیک افراد دارای ناتوانی ریاضی از افرادی عادی پرداخت. موارد استفاده از آن با کاربرد آزمون ایرانی کمی‌ت ریاضیات (محمداسماعیل و هومن، ۱۳۸۱) و آزمون پیشرفت ریاضی (چین آوه و همکاران، ۱۳۸۳) مشابه است. از این آزمون می‌توان در تعیین نقاط ضعف و قوت دانش آموزان پایه پنجم در حوزه‌های مختلف ریاضی و نیز تفکیک دقیق افراد حساب‌نارسان و عادی سود جست. به پژوهشگران نیز پیشنهاد می‌شود که برای تشخیص دیگر مشکلات یادگیری کودکان، مانند اختلال در دیکته‌نویسی و خواندن، با استفاده از نظریه جدید روان‌سنجی سؤال - پاسخ (IRT) آزمون‌هایی طراحی و ویژگی‌های دقیق روان‌سنجی آنها را تعیین کنند.

بودند. بهترین ضریب دشواری آزمون‌های تشریحی، ضریب دشواری نزدیک به  $0/5$  می‌باشند (عباس‌زادگان و فتوت، ۱۳۸۴؛ شریفی و نجفی‌زند، ۱۳۷۸). بررسی ضرایب دشواری این سؤال‌ها نشان داد که بیشتر آنها دارای ضریب دشواری بهینه یعنی نزدیک به  $0/5$  هستند.

برای تفسیر ضرایب دشواری و تشخیصی سؤال‌ها در نظریه کلاسیک، باید این دو در ارتباط با یکدیگر در نظر گرفته شوند؛ سؤال‌هایی که ضرایب تشخیصی بالا و دشواری متوسط (در سؤال‌های تشریحی  $0/5$ ) دارند، در تفکیک افراد قوی از ضعیف مطلوب تلقی می‌شوند (متسامورزن، ۱۳۸۵)، بنابراین همه سؤال‌ها براساس نظریه کلاسیک دارای ضرایب تشخیصی مطلوب و ضرایب دشواری نزدیک به  $0/5$  هستند و در نتیجه می‌توان بر اساس آنها افراد دارای توانمندی ریاضی بالا و پایین را تفکیک کرد.

بررسی ضرایب تشخیص  $a(g)$  و دشواری  $b(g)$  سؤال‌ها براساس نظریه سؤال- پاسخ (IRT) نشان داد که سؤال‌های این آزمون دارای ضرایب تشخیصی  $a(g)$  مطلوب (بالای  $0/5$ ) هستند و براساس آنها می‌توان افراد ضعیف و قوی را تفکیک کرد. همچنین میانگین کل ضرایب تشخیصی نیز نشان می‌دهد که در مجموع قدرت تمیز آزمون نیز مطلوب است. ضرایب دشواری آزمون نیز نشان داد که سؤال‌ها در طیف وسیعی از بسیار ساده تا بسیار دشوار در نوسان‌اند ( $-0/09$  تا  $+3$ ) و میانگین ضرایب دشواری نیز نشان می‌دهد که متوسط دشواری این آزمون بیشتر از متوسط توانمندی ریاضی افراد است. به دلیل ماهیت تشریحی سؤال‌ها عامل شانس صفر بود ( $c(g)=0$ ) و احتمال پاسخ‌دهی شانس وجود نداشت. بررسی سطح معنادار مقادیر  $\chi^2$  دو نیز نشان داد که همه سؤال‌ها با مدل دو آماره‌ای برازش کامل دارند.

در این آزمون زمان لازم برای خرده‌آزمون شمارش شش دقیقه و  $50$  ثانیه به دست آمد که مدت زمانی بود که  $70$  درصد آزمودنی‌ها در آن فاصله زمانی توانستند به سؤال‌های این خرده‌آزمون پاسخ دهند. این نکته با یافته‌های جوردن و مونتسیانی (۱۹۹۷)، لندرل و همکاران (۲۰۰۴) و باترورث

## منابع

- احدی، ح. و کاوند، ع. ر. (۱۳۸۲). *اختلال‌های یادگیری (از نظریه تا عمل)*. تهران: انتشارات ارسباران.
- آلن، م. جی. و یو. و. ام. (۱۳۷۴). *مقدمه‌ای بر نظریه‌های اندازه‌گیری (روان‌سنجی)* (ترجمه ع. دلاور). تهران: انتشارات سمت.
- آناستازی، ا. (۱۳۷۳). *روان‌آزمایی* (ترجمه م. ن. براهنی). تهران: انتشارات دانشگاه تهران.
- چین آوه، م.، عرب‌زاده، ز.، و منشی‌زاده، ف. (۱۳۸۳). ساخت و هنجاریابی آزمون تشخیص ریاضی برای دانش‌آموزان ۶ تا ۱۳ ساله شهر شیراز. *مجله دانش و پژوهش در علوم تربیتی، دانشگاه آزاد واحد اسلامی خوارسگان اصفهان*، ۳، ۳۲-۱۵.
- حمید، ن. (۱۳۸۵). بررسی اختلال یادگیری ریاضی در دانش‌آموزان دختر و پسر دوره ابتدایی ناحیه یک شهر تهران و اثر آموزش کاربردی، تقویت ژنتونی و آرامش عضلانی در کاهش اختلال یادگیری ریاضی در آنان. *مجله علوم تربیتی و روان‌شناسی دانشگاه شهید چمران اهواز*، ۴، ۱۳۶-۱۱۹.
- رضایانی، م. (۱۳۸۰). بررسی میزان شیوع حساب‌نارسایی در دانش‌آموزان پایه‌های چهارم و پنجم مدارس ابتدایی شهر اصفهان. تهران: پژوهشکده کودکان استثنایی.
- رضایانی، م. (۱۳۸۲). تحلیل خطاهای محاسباتی دانش‌آموزان حساب‌نارسانا. *مجله پژوهش در حیطه کودکان استثنایی*، ۴، ۲۴۴-۲۲۳.
- ژندایک، ا. (۱۳۷۵). *روان‌سنجی کاربردی* (ترجمه ح. ع. هومن). تهران: انتشارات دانشگاه تهران.
- ستاری، ب. (۱۳۸۲). *روان‌سنجی پیشرفته کاربردی*. مشهد: انتشارات آستان قدس رضوی.
- سرمه، ز.، بازرگان، ع.، و حجازی، ا. (۱۳۷۶). *روش‌های تحقیق در علوم رفتاری*. تهران: انتشارات آگاه.
- سیف، ع. ا. (۱۳۷۵). *روش‌های اندازه‌گیری و ارزشیابی آموزشی*. تهران: نشر دوران.
- شریفی، ح. پ.، و نجفی‌زند، ج. (۱۳۷۸). *اصول روان‌سنجی و روان‌آزمایی*. تهران: انتشارات رشد.
- عباس‌زادگان، م. ر.، و فوت، ا. ر. (۱۳۸۴). کاربرد پایایی و روایی در پژوهش. تهران: انتشارات کیهان.
- عریضی، ح. ر.، کاووسیان، ح.، و کدبورد، ب. (۱۳۸۳). پایایی سنجی و اعتباریابی مقیاس شایستگی ریاضی کودکان اوت‌ترخت در دانش‌آموزان دختر و پسر پیش‌دبستانی و دبستانی شهر اصفهان. *فصلنامه نوآوری‌های آموزشی*، ۹، ۶۵-۴۹.
- کاپلان، ه. ا.، و سادوک، ب. جی. (۱۳۸۳). *خلاصه روان‌پزشکی* (ترجمه ح. رفیعی). تهران: انتشارات ارجمند.
- محمد اسماعیل، ا.، و هومن، ح. ع. (۱۳۸۱). انطباق و هنجاریابی آزمون ریاضیات ایران کمی‌مت. *مجله پژوهش در حیطه کودکان استثنایی*، ۴، ۳۳۲-۳۲۳.
- مستامورتن، ح. (۱۳۸۵). مبانی نظری آزمون و آزمون‌سازی (ترجمه ش. کامکار و ا. اسرایی). تهران: انتشارات بهینه.
- میرهاشمی، م. (۱۳۷۱). هنجاریابی آزمون تصاویر در هم‌اندازه ری در دانش‌آموزان ۷ تا ۱۵ سال و بزرگسالان شهر تهران. پایان‌نامه کارشناسی ارشد دانشگاه آزاد کرج.
- هومن، ح. ع. (۱۳۷۱). *اندازه‌گیری‌های روانی و تربیتی (فصل تهیه تست)*. تهران: انتشارات درس.
- و کسلر، د. (۱۳۷۳). مقیاس تجدید نظر شده هوش و کسلر برای کودکان، انطباق و هنجاریابی (ترجمه س. شهیم). شیراز: انتشارات دانشگاه شیراز.
- Ashcraft, M. H., Yamashita, T. S., & Aram, D. M. (1992). Mathematics performance in left and right brain-lesioned children and adolescents. *Brain and Cognition*, 19, 208-252.
- Bloom, B. S. (1965). *Taxonomy of educational objectives; the classification of educational goals, by a committee of college and university examiners*. New York: Longmans, Green.
- Butterworth, B. (2003). *Dyscalculia screener. Highlighting pupils with specific learning difficulties in maths*. London, UK: Nelson Publishing Company.
- Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 46(1), 3-18.
- Fuchs, D., & Fuchs, L. D. (2005). Enhancing mathematical problem solving for student with disabilities. *Journal of Special Education*, 39, 45-57.
- Jordan, N. C., & Montani, T. O. (1997). Cognitive arithmetic and problem solving: A comparison of children with specific and general mathematics difficulties. *Journal of Learning Disabilities*, 30, 624-634.
- Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development*, 74, 834- 850.
- Korporal, A. (1995). *The utrecht early mathematical competences scales. Analysis from classical and modern test theory*. Utrecht University, The Netherlands.
- Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8-9-year-old student. *Cognition*, 93, 99-125.
- Rosselli, M., & Ardila, A. (1989). Calculation deficits in patients with right and left hemisphere damage. *Neuropsychologia*, 27, 607-617.
- Rousselle, L., & Noel, M-P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, 102, 361-365.