

## ساخت و اعتباریابی آزمون تشخیصی حساب نارسایی برای کودکان پایه پنجم دبستان

## Development and validation of a diagnostic test for dyscalculia in fifth grade primary students

Hojjat A. Farahani, Ph.D.

Mokhtar Malekpour, Ph.D.

Shole Amiri, Ph.D.

Hossein Molavi, Ph.D.

Mehrdad Kalantary, Ph.D.

دکتر حجت‌اله فراهانی \*

دکتر مختار ملک‌پور \*\*

دکتر شعله امیری \*\*

دکتر حسین مولوی \*\*

دکتر مهرداد کلانتری \*\*

## Abstract

The main purpose of this study was to develop and validate a diagnostic test for dyscalculia in fifth grade in primary schools of Isfahan. For this purpose, content analysis was conducted on the content of fifth grade math textbook. Keripendorf's coefficient of 0.88 was found based on consistency of content analysis. Based on Bloom's (1926) Cognitive Theory, 150 questions were designed and test via a pilot study. The test with revised questions were administered to 300 students (152 males, 148 females) who were selected using multi-level cluster sampling method from five districts of Isfahan. Questions and test parameters were assessed based on the classic test and item-response theory. Results showed that the test had desirable validity and reliability ( $\alpha = 0.92$ , test-retest coefficient = 0.94). Specificity and sensitivity of 0.99 and 0.96 were obtained for this test, respectively. Based on item-response theory, discrimination a (g) and difficulty b (g) Parameters showed that the questions had desirable discrimination power and acceptable difficulty. Furthermore, the questions had complete fitness with two-parameter model ( $p > 0.05$ ). TIF curve demonstrated that the test was suitable for students with math ability ( $\theta$ ) between +2 and -0.2.

**Keywords:** diagnostic test, dyscalculia, classic test, item-response theory

## چکیده

هدف از این پژوهش ساخت و اعتباریابی آزمون تشخیصی حساب نارسایی برای دانش‌آموزان پایه پنجم دبستان شهر اصفهان بود. نمونه مورد نظر در این پژوهش ۳۰۰ نفر از دانش‌آموزان پایه پنجم دبستان (۱۵۲ پسر و ۱۴۸ دختر) بودند که به روش نمونه‌گیری خوشه‌ای چند مرحله‌ای از نواحی پنج‌گانه آموزشی شهر اصفهان در سال تحصیلی ۸۶-۸۵ انتخاب شدند. در این پژوهش، کتاب ریاضی پایه پنجم دبستان تحلیل محتوا شد. ضریب کریپندورف مبنی بر توافق تحلیل محتوا ۰/۸۴ محاسبه شد. بر اساس مقوله‌های استخراج شده و هدف‌های رفتاری هر مقوله و با توجه به نظریه شناختی بلوم و حجم هر مقوله در کتاب درسی ریاضی ۱۵۰ سؤال طرح و به صورت مقدماتی اجرا گردید. سپس فرم نهایی اجرا و پارامترهای سؤالات آزمون بر اساس نظریه کلاسیک آزمون و نظریه سؤال - پاسخ محاسبه گردید. نتایج نشان داد که این آزمون دارای روایی مطلوب و نیز پایایی مناسب ( $\alpha = 0.92$ ) و ضریب بازآزمایی ( $0.94 = 0.96$ ) بوده و حساسیت و دقت این آزمون به ترتیب ۹۹ و ۹۶ درصد به دست آمد. پارامتر تشخیصی a g و دشواری b g بر اساس نظریه سؤال - پاسخ نیز نشان داد که همه سؤالات دارای قدرت تشخیص مناسب (بالای ۰/۵) و دشواری مطلوب بوده و با مدل ۲ پارامتری برازش کامل دارند ( $P > 0.05$ ). منحنی تابع آگاهی آزمون نیز نشان داد که مناسب‌ترین کاربرد این آزمون در سطوح توانایی ریاضی ( $\theta$ ) ۰/۲- تا ۲+ است.

**واژه‌های کلیدی:** آزمون تشخیصی، حساب نارسایی، نظریه کلاسیک آزمون، نظریه سؤال - پاسخ

حد زیادی مبهم مانده است (فاچس و فاجس<sup>۶</sup>، ۲۰۰۵). به همین دلیل، پژوهشگران مختلف تقسیم‌بندی‌های متعدد برای این اختلال انجام داده‌اند. از جمله این تقسیم‌بندی‌ها، می‌توان دسته‌بندی جوردن<sup>۷</sup> و همکاران (۲۰۰۳) و روسلی و آردیلا<sup>۸</sup> (۱۹۸۹)، را نام برد. اما تقسیم‌بندی جامع این اختلال را گیری<sup>۹</sup> (۲۰۰۴) بر اساس مرور مبانی شناختی، عصب‌شناختی و ژنتیک انجام داد و این اختلال را به زیر دسته روندی<sup>۱۰</sup>، حافظه معنایی<sup>۱۱</sup> و نوع خوانش پریشی واجی<sup>۱۲</sup> تقسیم کرد. وگن و باس (۲۰۰۰) درصد دانش‌آموزان دارای اختلال یادگیری را ۱۵ تا ۲۵ درصد اعلام می‌کنند (چین‌آوه، عرب‌زاده و منشی‌زاده، ۱۳۸۳). فاچس و فاجس (۲۰۰۵) همه‌گیری اختلال ریاضی در دبستان را ۵ درصد اعلام کردند. البته، مشخص نیست که چه میزان از این همه‌گیرشناسی ناشی از محدودیت‌های آموزشی است (کاپلن و سادوک، ۱۳۸۳). در مورد همه‌گیری‌شناسی این اختلال در ایران نیز پژوهش‌هایی انجام شده است. در پژوهش رضانی (۱۳۸۳) میزان همه‌گیری‌شناسی این اختلال در شهر تهران ۱/۴۳ و در پژوهش حمید (۱۳۸۵) میزان آن ۳/۶ درصد برآورد شده است. نسبت این اختلال در دو جنس تحت بررسی است (گیری، ۲۰۰۴). این اختلال معمولاً از پایان کلاس اول قابل شناسایی است اما اگر کودک دارای هوش بالایی باشد، ممکن است اختلال در ریاضیات تا کلاس پنجم و بالاتر آشکار نگردد (احدی و کاکاوند، ۱۳۸۱). در سبب‌شناسی این اختلال فرضیه‌هایی مطرح شده که در مجموع می‌توان در تبیین علت آن به ترکیبی از تأثیرات محیطی و ژنتیک اشاره کرد (راسلی، نوئل<sup>۱۳</sup>، ۲۰۰۷) بنا به تأکید انجمن روان‌پزشکی آمریکا (۲۰۰۱) تشخیص قطعی فقط پس از اجرای آزمون استاندارد شده انفرادی در حساب و کسب نمره‌ای بسیار پایین‌تر از سطح مورد انتظار بر حسب میزان تحصیلات و ظرفیت هوشی کودک امکان‌پذیر است (کاپلن و سادوک، ۱۳۸۳). بنابراین، برای تشخیص صحیح و دقیق وجود یک آزمون تشخیص استاندارد ضرورت می‌یابد. برای سنجش این اختلال، روش‌های رسمی و غیررسمی متعددی طراحی شده است. از جمله می‌توان به آزمون ریاضی کی‌مت<sup>۱۴</sup> اشاره کرد که

سالانه درصد بسیاری از دانش‌آموزان به علت مشکلات و ناتوانی‌های یادگیری و از جمله ناتوانی در ریاضیات مردود می‌شوند. در بسیاری از این مواقع علت مردودی مبتلا بودن کودک به اختلال‌های یادگیری<sup>۱</sup> است، یکی از این اختلال‌ها، حساب نارسایی<sup>۲</sup> یا ناتوانی‌های ریاضی<sup>۳</sup> است. این کودکان با وجود برخورداری از هوش متوسط و بالاتر، در شرایط تقریباً یکسان آموزشی نسبت به دانش‌آموزان دیگر، عملکرد تحصیلی ضعیف‌تری دارند و علی‌رغم قرار داشتن در محیط آموزشی مناسب و نیز فقدان ضایعات زیست‌شناختی بارز و عدم مشکلات اجتماعی حاد، قادر به یادگیری در زمینه‌های خاص (خواندن، نوشتن و ریاضی) نیستند (تبریزی، ۱۳۷۹).

در مورد برخی از کودکان دشواری در فهم روابط عددی از همان سال‌های نخست زندگی آغاز می‌شود، سال‌هایی که کودکان دیگر تجربه‌هایی را در شمارش، درک واحد، جور کردن و مقایسه به‌دست می‌آورند و می‌توان برای آنها در دوره دبستان کارکردی پایین‌تر از حد انتظار در آزمون‌های پیشرفت تحصیلی متصور شد (رضانی، ۱۳۸۰). حساب نارسایی به عنوان یک اختلال در گستره مجموعه‌های تشخیصی و آماری انجمن روان پزشکی آمریکا تنها از سومین مجموعه (DSM-III) یعنی از سال ۱۹۸۰ به بعد مطرح گردید که عبارت است از ناتوانی در انجام مهارت‌های حساب با توجه به ظرفیت هوش و سطح آموزشی مورد انتظار کودک که این مهارت‌ها به کمک آزمون‌های استاندارد شده فردی اندازه‌گیری شده باشند. بر اساس DSM-IV-TR<sup>۴</sup> (به نقل از کاپلن و سادوک<sup>۵</sup>، ۱۳۸۳) این کودکان در چهار گروه از مهارت‌ها دچار مشکل‌اند:

با وجود اینکه این اختلال مشکل جدیدی نیست، ماهیت آن تا

6. Fuchs, L. S. & Fuchs, D.

7. Jordan, N. C.

8. Rosselli, M., & Ardila, A.

9. Geary, D. C.

10. procedural

11. semantic memory

12. phonetic dyscalculia

13. Rousselle, L., & Noel, M. P.

14. Key Math

1. learning disorders

2. dyscalculia

3. mathematic disabilities

4. Diagnostic and Statistical manual of mental disorders (Text - Revised)

5. Kaplan, H., & Sadock, B. J.

افراد غیرمتخصص نیز قابل اجرا می‌باشند. اما نظریه‌ی سؤال - پاسخ<sup>۹</sup> که مبتنی بر نظریه‌ی صفت مکنون<sup>۱۰</sup> است مبتنی بر روش‌های نسبتاً پیچیده‌ی تحلیل آزمون بوده و مستلزم کاربرد رایانه است. نظریه سؤال - پاسخ در صدد رفع نقایص نظریه‌ی کلاسیک آزمون است از جمله این نقایص وابستگی پارامترهای سؤال و آزمون به ویژگی‌های نمونه است. در نظریه‌ی سؤال - پاسخ به جای تأکید بر نمره‌ی کل آزمون بر نمره‌ی حاصل از تک تک سؤالات آزمون تأکید می‌شود. بر اساس این نظریه می‌توان احتمال پاسخ درست به یک سؤال آزمون را به عنوان تابعی از «توانایی» آزمون شونده و همچنین برخی ویژگی‌های سؤال را پیش‌بینی کرد. این نظریه از طریق چندین مدل قابل بررسی است (سیف، ۱۳۸۷). بنابراین یکی دیگر از اهداف این پژوهش برآورد پارامترهای دقیق سؤال (ضرایب دشواری و تشخیصی)، پارامتر توانایی آزمودنی‌ها (تتا) و تعیین منحنی آگاهی آزمون علاوه بر نظریه کلاسیک آزمون (CTT) از طریق نظریه سؤال - پاسخ (IRT) بود.

## روش

**جامعه، نمونه و روش اجرای پژوهش:** در این پژوهش جامعه آماری عبارت بود از کلیه دانش‌آموزان دختر و پسر دبستانی شهر اصفهان که در سال تحصیلی ۸۶-۸۵ در مدارس ابتدایی این شهر مشغول به تحصیل بودند، (جدول ۱) از این تعداد با استفاده از فرمول موزو کالتون<sup>۱۱</sup> و با در نظر گرفتن ۲ درصد به عنوان برآورد نسبت همه‌گیرشناسی این اختلال در جامعه تهران (رضانی، ۱۳۸۱) و خطای مورد پذیرش ۱ درصد، ۳۰۲ به عنوان نمونه (۱۵۴ پسر و ۱۴۸ دختر) از نواحی ۵ گانه شهر اصفهان انتخاب شدند. که ۲ نفر به دلیل مخدوش بودن آزمون از نمونه خارج و ۳۰۰ نفر مورد تحلیل قرار گرفتند. هرچند مدل ۲ پارامتری پاسخ - سؤال نیازمند حجم نمونه زیاد است اما می‌توان آن را با حجم حداقل ۳۰۰ نفر نیز انجام داد (سلیمی‌زاده، ۱۳۸۲؛ ارتباط شخصی). در جدول (۱)، تعداد افراد جامعه مورد بررسی به تفکیک ناحیه آموزشی و جنسیت آمده است.

شیوه نمونه‌گیری در این پژوهش به این صورت بود که با مراجعه به اداره کل آموزش و پرورش اصفهان لیست مدارس ابتدایی هر پنج ناحیه تهیه گردید و سپس با توجه به جمعیت

در سال ۱۹۸۸ توسط کنولی<sup>۱</sup> هنجاریابی شده است. این آزمون که شامل ۱۳ زیرمقیاس است، به منظور تعیین نقاط قوت و ضعف دانش‌آموزان در حوزه‌های مختلف ریاضی به کار می‌رود و در سه بخش مفاهیم اساسی، عملیات و کاربردها تدوین شده است که از اهمیت آموزشی تقریباً یکسانی برخوردارند. این آزمون برای دانش‌آموزان ۶/۵ تا ۱۲/۸ در ایران هنجاریابی شده است (محمد اسماعیل و هومن، ۱۳۸۱). آزمون دیگری که برای سنجش شایستگی ریاضی طراحی شده، آزمون شایستگی ریاضیات کودکان اوترخت (کراپرال<sup>۲</sup>، ۱۹۹۵ و اندوریکت<sup>۳</sup>، ۱۹۹۲) است که شامل ۸ مقیاس و در مجموع ۱۲۰ سؤال است. این آزمون توسط عریضی و همکاران (۱۳۸۳) اعتباریابی شد. در پژوهش دیگری، چین‌آوه و همکاران (۱۳۸۳) به ساخت و هنجاریابی آزمون پیشرفت ریاضی برای دانش‌آموزان ۶ تا ۱۳ ساله شهر شیراز پرداختند. در پژوهش دیگری رضانی (۱۳۸۲) به تحلیل خطاهای محاسباتی دانش‌آموزان مبتلا به حساب نارسایی پرداخت و در پژوهش دادستان و همکاران (۱۳۸۱) به بررسی تحولی این کودکان پرداخته شده است.

هدف پژوهش حاضر این بود که با توجه به تنوع این اختلال و یک‌دست نبودن ویژگی‌های آن (احدی و کاکاوند، ۱۳۸۱)، آزمونی طراحی شود که روایی محتوایی بیشتری داشته و برای کودکان اصفهانی دارای حداکثر کاربرد باشد. هدف دیگر این پژوهش ساخت آزمونی زمانمند بود زیرا بنا به نظر محققان (اشکرفت، یاماشیتا و آرام<sup>۴</sup>، ۱۹۹۹؛ باترورث<sup>۵</sup>، ۲۰۰۵، ۲۰۰۳؛ جوردن و مونتانی<sup>۶</sup>، ۱۹۹۷؛ لندرل، بیوان و باترورث<sup>۷</sup>، ۲۰۰۴) اگر آزمون تشخیصی ریاضیات در مقوله شمارش زمان لحاظ نگردد، کودکان حساب نارسا نیز می‌توانند مانند سایر کودکان عادی به سؤالات پاسخ درست بدهند زیرا آنها از راهکارهای معیوب یعنی شمارش با انگشتان استفاده کرده و یک مدت زمان طولانی‌تر پاسخ درست را ایجاد می‌کنند. از سوی دیگر آزمون‌های موجود بر اساس نظریه‌ی کلاسیک<sup>۸</sup> مورد تجزیه و تحلیل قرار گرفته‌اند. نظریه کلاسیک آزمون با روش‌های متداول و ساده سر و کار دارند. این روش‌ها روش‌های نسبتاً سراسری هستند که توسط

1. Connolly, A. J.
2. Korporal, A.
3. Endorecht, B.
4. Ashcraft, M. H., Yamashita, T. S., & Aram, D. M.
5. Butterworth, B.
6. Jordan, N. C., & Montani, T. O.
7. Landrel, K., Bevan, A., & Butterworth, B.
8. classic test theory

9. item-response theory

10. latent trait

11. Mozer & Kalton

جدول ۱- تعداد افراد جامعه و نمونه به تفکیک ناحیه آموزشی و جنس

ناحیه	پسر جامعه (نمونه)	دختر جامعه (نمونه)	کل جامعه (نمونه)
۱	۹۲۵ (۱۳)	۸۷۰ (۱۱)	۱۸۰۵ (۲۴)
۲	۱۴۷۰ (۲۲)	۱۴۱۸ (۱۹)	۲۸۸۸ (۴۱)
۳	۲۶۸۶ (۳۷)	۲۴۵۰ (۳۲)	۵۱۳۶ (۶۹)
۴	۳۳۹۷ (۴۵)	۳۱۸۲ (۴۲)	۶۵۷۹ (۸۷)
۵	۳۱۶۴ (۴۲)	۲۹۹۱ (۳۹)	۶۱۵۵ (۸۱)

### ابزار سنجش

برای تهیه سؤالات این آزمون ابتدا کتاب ریاضی پایه پنجم توسط ۵ نفر از اساتید روان‌شناسی و علوم تربیتی و متخصصان روان‌سنجی و ۳ نفر از سرگروه‌های آموزشی ریاضی پایه پنجم دبستان مورد تحلیل محتوا قرار گرفت و ضریب توافق کریپندورف<sup>۱</sup> ۰/۸۴ به دست آمد که نشان‌دهنده توافق افراد در تحلیل محتوای این کتاب بود. سپس این مقوله‌ها مجدداً به تأیید سرگروه‌های آموزشی رسید و در نهایت با توجه به نظریه شناختی بلوم (۱۹۶۵) که همسو با نظریه گانیه (۱۹۷۰) است سؤالات طرح گردید. تعداد سؤالات در هر مقوله با توجه به حجم آن مقوله در کتاب ریاضی تعیین شد. این مقوله‌ها شامل ۹ حیطه بودند که عبارتند از: اعداد، جمع و تفریق، ارزش مکانی، اندازه‌گیری، ضرب و تقسیم، هندسه، کسر و اعشار، نسبت و درصد، تخمین و حل مسأله سپس این سؤالات به صورت مقدماتی بر روی ۱۰۰ نفر دانش‌آموز پسر و دختر (۵۰ دختر و ۵۰ پسر) پایه پنجم دبستان اجرا شد. سؤالات به صورت انفرادی و تشریحی اجرا گردید و برای پاسخ درست نمره ۱ و پاسخ غلط نمره صفر منظور شد. نتایج مورد تحلیل قرار گرفت و سؤالات از نظر سطح دشواری (p) ضریب تشخیص (D) و همسانی درونی (α) بر اساس نظریه کلاسیک آزمون مورد بررسی قرار گرفت و روایی همزمان با آزمون ریاضی ایران کی‌مت (محمد اسماعیل و هومن، ۱۳۸۱) و نمره حساب آزمون وکسلر کودکان (شهیم، ۱۳۷۹) محاسبه گردید. سؤالات دارای ضریب (α) بالا (بالای ۰/۸) و ضرایب تشخیص معنادار (بالای ۰/۴) برای فرم نهایی انتخاب شدند، زیرا ضرایب تشخیص ۰/۴ به بالا مطلوب تلقی

می‌شوند (متسامورنن، ۱۳۸۵). سپس برای کمک به استقلال موضوعی سؤالات (شرط لازم برای تحلیل سؤال - پاسخ) سؤالات در هر زیرمقیاس بر اساس درجه دشواری به طور تصادفی قرار داده شدند. در نهایت ۶۳ سؤال باقی ماند و از آنجا که ریاضی یک امر پیوسته تلقی می‌شود، آزمون تشخیص حساب نارسایی پایه چهارم (۵۲ سؤال) نیز که بر اساس اصولی مشابه این آزمون ساخته شده بود به سؤالات این آزمون اضافه و در مجموع ۱۱۵ سؤال به عنوان آزمون تشخیص حساب نارسایی پایه پنجم اجرا گردید. برای آگاهی از وضعیت هوشی شرکت‌کنندگان، از آزمون هوش وکسلر کودکان (شهیم، ۱۳۷۹) و برای آگاهی از وضعیت ادراکی - دیداری شرکت‌کنندگان از آزمون ادراکی - دیداری آندره‌ری (میرهاشمی، ۱۳۷۶) که هر دو آزمون از روایی و پایایی مطلوبی برخوردارند، استفاده شد. همچنین با شرکت‌کنندگان مصاحبه بالینی صورت گرفت به این صورت که از آنها خواسته شد متن مشخصی را بخوانند و آن را برای آزمونگر بازگو کنند در مواردی که شرکت‌کننده دچار مشکل ادراکی، بالینی یا هوشی بود از نمونه خارج و به جای آن فرد دیگری به طور تصادفی انتخاب شد. برای تعیین زمان مطلوب در زیرمقیاس شمارش مدت زمانی که ۷۰ درصد شرکت‌کنندگان (آناستازی، ۱۳۷۳ و ثرندایک، ۱۳۶۹؛ سیف، ۱۳۷۸) توانستند به سؤالات آن پاسخ بدهند به عنوان زمان مطلوب در نظر گرفته شد. جهت تعیین نقطه برش، حساسیت، دقت و روایی تشخیصی سؤالات آزمون تشخیصی نهایی، یک نمونه ۲۵ نفری از کودکان دارای حساب نارسایی انتخاب شدند. به این منظور ابتدا از آموزگار پایه پنجم در کلاس مربوطه خواسته شد بر مبنای شناخت خود، دانش‌آموزانی را که منحصراً در درس ریاضی با مشکل روبه‌رو هستند معرفی کند، نمره کمتر از ۱۲ به عنوان ضابطه نخست در

1. Krippendorff, C.

نظر گرفته شد. سپس از آموزگار خواسته شد که تا در هر مورد فهرست واریتی تشخیصی DSMIV-TR را پاسخ دهند. پس از پاسخ‌دهی، نمره‌های حاصل از ماده‌هایی که با بلی «پاسخ» داده شده بود جمع و در صورتی که در مورد آزمودنی حداقل ۴ ماده (از ۱۱ ماده فهرست واریتی تشخیصی) با پاسخ «بلی» پاسخ داده شده بود آن آزمودنی به عنوان حساب نارسا در نظر گرفته شد. این آزمون توسط پژوهشگر و دانشجویان کارشناسی ارشد و کارشناسی که در مورد ایجاد رابطه و برقراری ارتباط مؤثر و همدلانه با کودک و اجرای آزمون داری تجربه بوده و آموزش کامل داده شده بودند در هفته‌های پایانی سال تحصیلی ۸۶-۸۵ به صورت انفرادی و تشریحی اجرا گردید و به سؤالات درست نمره (۱) و به سؤالات غلط نمره (۰) داده شد.

پارامترهای سؤال در این پژوهش در نظریه کلاسیک آزمون ضریب دشواری یعنی نسبت افرادی که به سؤالات پاسخ درست دادند و ضریب تشخیص سؤال بود که از طریق همبستگی دو رشته‌ای نقطه‌ای سؤال با نمره کل به دست آمد. پایایی از طریق آلفای کرونباخ و بازآزمایی محاسبه گردید. نقاط برش، حساسیت و دقت سؤالات از طریق تحلیل تفکیکی<sup>۱</sup> تعیین شد. منظور از حساسیت یک آزمون امکان تشخیص افراد حساب نارسا از افراد عادی و منظور از دقت یک آزمون، توان یک آزمون در تشخیص موارد سالم است. پارامترهای سؤال در این پژوهش بر اساس مدل دو پارامتری سؤال - پاسخ برآورد شد. ضرایب تشخیص سؤالات  $a(g)$  و پارامتر دشواری  $b(g)$  و عامل حدس  $c(g)$  است که به دلیل ماهیت تشریحی این آزمون عامل حدس در پاسخگویی سؤالات تأثیر نداشت و صفر در نظر گرفته شد. بنابراین، از مدل ۲ پارامتری استفاده شد. در این مدل پارامتر تشخیص سؤال به قدرت سؤالات در تشخیص و جداسازی افراد قوی و ضعیف در یک آزمون اشاره دارد (آلن و وین، ۱۳۷۴). معمولاً دامنه  $a(g)$  بین ۰/۵ تا ۲/۵ تغییر می‌کند. مقادیر کمتر از ۰/۵ نشان‌دهنده قدرت تشخیص پایین سؤالات است و مقادیر بزرگ‌تر نیز به ندرت به دست می‌آید (هومن، ۱۳۷۱). پارامتر دشواری سؤال معمولاً بین ۲/۵- تا ۲/۵+ مقیاس‌پردازی می‌شود و نشان‌دهنده نقطه عطف منحنی شاخص سؤال است که ۵۰٪ آزمون شونده‌گانی که در آن سطح قرار دارند جواب سؤال را می‌دانند یا می‌توانند جواب آن را بیابند (ستاری، ۱۳۸۳).

برای تعیین برازش سؤالات با مدل ۲ پارامتری سؤال - پاسخ مقادیر خی‌دو محاسبه و سطح معناداری آنها تعیین شد. برای

اطمینان از صحت مدل ۲ پارامتری در تحلیل سؤالات، سؤالات با مدل تک پارامتری و ۲ پارامتری بررسی شدند. سپس تفاضل مقدار  $2\text{-Log Likelihood}$  دو مدل، باخی - دو بحرانی با درجه آزادی  $n-1$  مقایسه شد. چون این مقدار معنادار نبود بنابراین مدل ۲ پارامتری به عنوان مدل با بهترین برازندگی انتخاب شد. لازم به ذکر است که برای احراز دو پیش فرض تحلیل نظریه سؤال - پاسخ یعنی استقلال موضعی<sup>۲</sup> و تک بعدی بودن<sup>۳</sup> آزمون، بررسی داده‌ها نشان داد که بیش از ۲۰ درصد واریانس آزمون در یک عامل قرار می‌گیرد که نشان‌دهنده برقراری مفروضه تک بعدی بودن است (باقی، ۱۹۹۹) و اگر ملاک تک بعدی بودن رعایت شود استقلال نیز تا حدی رعایت می‌شود (سلیمی‌زاده، ۱۳۸۲، ارتباط شخصی). با این حال برای کمک به برقراری این مفروضه سؤالات با درجه دشواری‌های تصادفی قرار داده شدند.

## یافته‌ها

### روایی صوری و محتوایی این آزمون چگونه است؟

برای تعیین روایی صوری و محتوایی این آزمون گام‌های زیر برداشته شد:

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

ب) ضریب توافق کرپیندورف مبنی بر توافق تحلیل محتوا محاسبه گردید که این ضریب ۰/۸۴ به دست آمد.

ج) هدف‌های آموزشی و رفتاری بر اساس مقوله‌های استخراج شده تعریف و تدوین شد.

د) بر اساس بند «ج» مهارت‌هایی که نشان‌دهنده محتوای مورد نظر بود تعیین شد.

ه) از سؤالات امتحان ریاضی پایه پنجم در نوبت‌های دو گانه در دبستان‌های نواحی ۵ گانه آموزش شهر اصفهان خزانه‌ای از سؤالات پنج سال اخیر تهیه و مورد بررسی قرار گرفتند.

و) بر اساس گام‌های «الف» تا «ه» مجموعه‌ای از سؤالات طراحی و تنظیم شد.

ز) سؤالات نهایی توسط اساتید گروه علوم تربیتی و روان‌شناسی، متخصصان روان‌سنجی، و سرگروه‌های آموزش ریاضی پایه پنجم دبستان مورد بازنگری نهایی قرار گرفت و نسبت روایی محتوایی  $(CVR_i)$  مبنی بر توافق در تعیین روایی محتوایی سؤالات محاسبه و در دامنه‌ای از ۰/۸۵ - ۰/۸ به دست آمد که معنادار

2. local independence  
3. unidimensionality

1. discrimination analysis

بوده و حاکی از روایی محتوایی بالا و مطلوب این آزمون بود. ح- سؤالات اجرای مقدماتی شدند و سؤالات مبهم برای دانش‌آموزان شناسایی و عبارت‌بندی و در صورت لزوم محتوای آن تغییر داده شد.

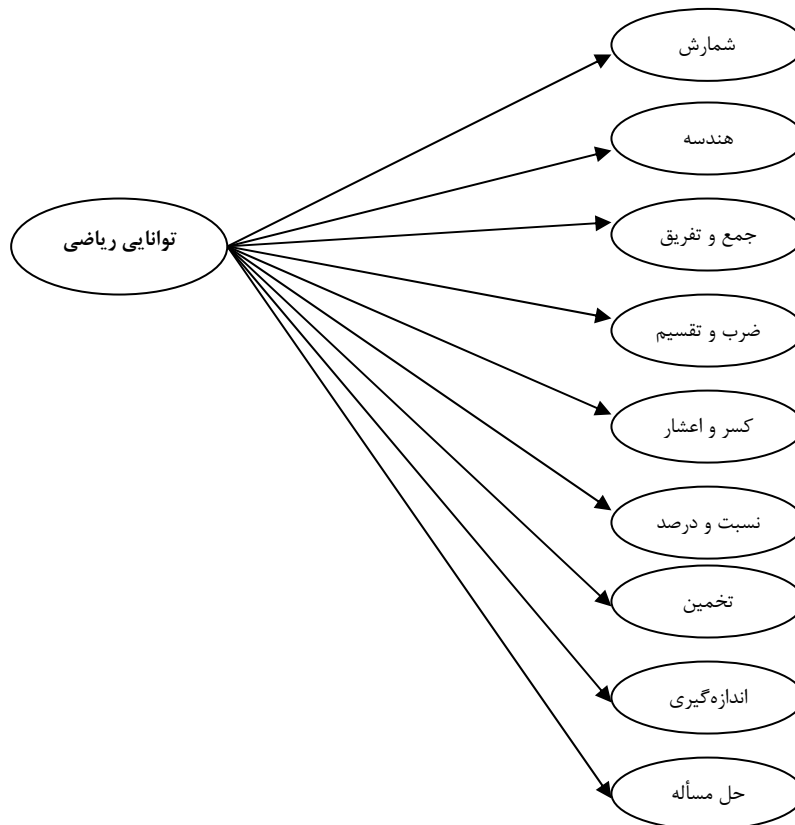
### روایی آزمون تشخیصی حساب نارسایی پایه پنجم چگونه است؟

روایی ملاکی همزمان آزمون: برای تعیین روایی همزمان این آزمون از خرده آزمون‌های مرتبط در آزمون ریاضی ایران کی‌مت (محمد اسماعیل و هومن، ۱۳۸۱) و زیرمقیاس حساب در آزمون هوش ویسک - آر (شهیم، ۱۳۷۳) استفاده شد. نتایج نشان داد که رابطه بین خرده‌آزمون شمارش در این ابزار و خرده آزمون شمارش در آزمون ریاضی ایران کی‌مت ( $P=0/001$  و  $r=0/68$ )، رابطه بین خرده آزمون هندسه در این ابزار و خرده آزمون هندسه در آزمون ریاضی کی‌مت ( $P=0/001$  و  $r=0/67$ )، رابطه بین خرده آزمون جمع و تفریق در این ابزار و خرده آزمون جمع و تفریق در آزمون ریاضی کی‌مت ( $P=0/001$  و  $r=0/65$ )، رابطه بین خرده آزمون ضرب و تقسیم در این ابزار و خرده آزمون ضرب

و تقسیم در آزمون ریاضی کی‌مت ( $r=0/70$  و  $P=0/001$ )، رابطه بین خرده آزمون کسر و اعشار در این ابزار و خرده آزمون اعداد گویا در آزمون ریاضی کی‌مت ( $r=0/64$  و  $P=0/001$ )، رابطه بین خرده آزمون نسبت و درصد در این ابزار و خرده آزمون اعداد گویا در آزمون ریاضی کی‌مت ( $r=0/65$  و  $P=0/001$ )، رابطه بین خرده آزمون تخمین در این ابزار و خرده آزمون تخمین در آزمون ریاضی کی‌مت ( $r=0/71$  و  $P=0/001$ )، رابطه بین خرده آزمون اندازه‌گیری در این ابزار و خرده آزمون اندازه‌گیری در آزمون ریاضی کی‌مت ( $r=0/63$  و  $P=0/001$ )، رابطه بین خرده آزمون حل مسأله در این ابزار و خرده آزمون حل مسأله در آزمون ریاضی کی‌مت ( $r=0/68$  و  $P=0/001$ ) و رابطه بین نمره کل این ابزار و نمره کل آزمون ریاضی کی‌مت ( $r=0/73$  و  $P=0/001$ ) به‌دست آمد که همگی مثبت و معنادار بوده و حاکی از روایی همزمان مطلوب این ابزار است. همچنین همبستگی بین آزمون تشخیصی حساب نارسایی پایه پنجم و زیرمقیاس حساب در آزمون هوش ویسک - آر معنادار بوده که نشان‌دهنده روایی همزمان و مطلوب این آزمون است ( $r=0/8$  و  $P=0/001$ ).

جدول ۲ - شاخص‌های برازندگی مدل ۹ عاملی آزمون تشخیصی حساب نارسایی

RMSEA	CFI	AGFI	GFI	p	df	$\chi^2$
۰/۰۰۲	۰/۹۷	۰/۸۶	۰/۹۴	۰/۰۷۲	۲۷	۲۳/۷۳



**روایی تشخیصی آزمون:** برای تعیین روایی تشخیصی سؤالات این آزمون بر روی یک گروه ۲۵ نفری از کودکان حساب نارسا و عادی اجرا گردید. تحلیل MANOVA سؤالات نشان داد که می‌توان براساس همگی سؤالات افراد حساب نارسا و عادی را به‌طور معناداری ( $P < 0/05$ ) تفکیک کرد.

**روایی سازه آزمون:** برای بررسی روایی‌سازه با استفاده از نرم‌افزار لیزرل بر روی نمرات حاصل از نمونه ۳۰۰ نفری تحلیل عامل تأییدی انجام شد. نتایج نشان داد که این آزمون دارای روایی سازه مطلوب است و مدل نظری با داده‌های حاصل برازش کامل دارد. جدول (۲) شاخص‌های برازندگی مدل ۹ عاملی آزمون تشخیصی حساب نارسائی را نشان می‌دهد.

**پایایی آزمون تشخیص حساب نارسایی چگونه است؟** برای تعیین پایایی این آزمون از دو روش آلفای کرونباخ (همسانی درونی) و بازآزمایی (ثبات) آزمون استفاده شد. ضرایب آلفای کرونباخ در خرده آزمون‌ها در محدوده ۰/۷۵-۰/۸۷ و در کل ۰/۹۲ و در بازآزمایی در خرده آزمون‌ها در محدوده ۰/۹۴-۰/۹۰ و در کل ۰/۹۴ به‌دست آمد، برای تعیین ضریب بازآزمایی شرکت کنندگان پس از ۴ هفته مجدداً آزمون شدند.

**نقطه برش، حساسیت و دقت آزمون تشخیصی حساب نارسایی چگونه است؟** نقطه برش، حساسیت و دقت هر یک از

خرده آزمون‌های آزمون تشخیص حساب نارسایی و کل نشان داد که نقطه برش در خرده آزمون‌های این آزمون در محدوده ۱۵-۲ و در کل ۶۳ و حساسیت خرده آزمون‌ها در محدوده ۹۶-۸۰ درصد و در کل ۹۹ درصد و دقت خرده آزمون‌ها در محدوده ۹۶-۸۶ و در کل ۹۶ درصد به دست آمد. حساسیت و ویژگی (دقت) آزمون در حد بسیار مطلوبی است یعنی با این آزمون می‌توان در ۹۹ درصد موارد حساب نارسا و در ۹۶ درصد موارد سالم را به درستی تشخیص داد. ضرایب بر اساس نظریه کلاسیک ( $r_{pbis}$ ) بین ۰/۴ تا ۰/۸ در نوسان بوده که نشان می‌دهد هیچ یک از ضرایب منفی و پایین نیست. همچنین ضرایب دشواری سؤالات ( $P_i$ ) بین ۰/۸۹ تا ۰/۲۴ ضرایب تشخیص ( $g$ ) بر اساس نظریه سؤال - پاسخ (IRT) در دامنه‌ای از ۲/۵ تا ۰/۵۰۱ با میانگین ۱/۲۵۱ و انحراف استاندارد ۰/۵۵۴ بوده هیچ یک از ضرایب کمتر از ۰/۵ نیست، ضرایب دشواری ( $g$ ) نیز بین ۳ تا ۲/۳۴- با میانگین ۱/۳۱۴ و انحراف استاندارد ۱/۱۴ است. سطح معناداری مقادیر  $\chi^2$  دو سؤالات نیز نشان می‌دهد که همگی با مدل ۲ پارامتری (2PLM) برازش دارند ( $P > 0/05$ ) (جدول ۳ و ۴). همچنین بررسی پارامتر توانایی ( $\theta$ ) آزمودنی‌ها در ریاضیات نشان داد که آزمودنی‌ها دارای پارامتر توانایی ( $\theta$ ) بین ۲/۷۵- تا ۲/۱۴ با میانگین ۱/۰۹ و انحراف استاندارد ۰/۹۳ هستند.

جدول ۳- خلاصه نتایج تحلیل سؤالات آزمون بر اساس نظریه کلاسیک آزمون

مشخصه	ضرایب تشخیصی ( $r_{pbis}$ )	ضرایب دشواری ( $P_i$ )
حداقل	۰/۴	۰/۲۴
حداکثر	۰/۸	۰/۸۹

جدول ۴- خلاصه نتایج مدل ۲ پارامتری سؤالات آزمون بر اساس نظریه سؤال - پاسخ

مشخصه	ضرایب تشخیصی ( $g$ )	دشواری ( $b$ )
حداقل	۰/۵۰۱	-۲/۳۴
حداکثر	۲/۵	۳
میانگین	۱/۲۵۱	۱/۳۱۴
انحراف معیار	۰/۵۵۴	۱/۱۴

شمارش نمره کمتر از ۴، در زیرمقیاس هندسه نمره کمتر از ۹، در زیرمقیاس جمع و تفریق نمره کمتر از ۵ در زیرمقیاس ضرب و تقسیم نمره کمتر از ۹ در زیرمقیاس کسر و اعشار نمره کمتر از ۱۲ در زیرمقیاس حل مسأله نمره کمتر از ۹ و در کل نمره کمتر از ۵۹ به دست آورند.

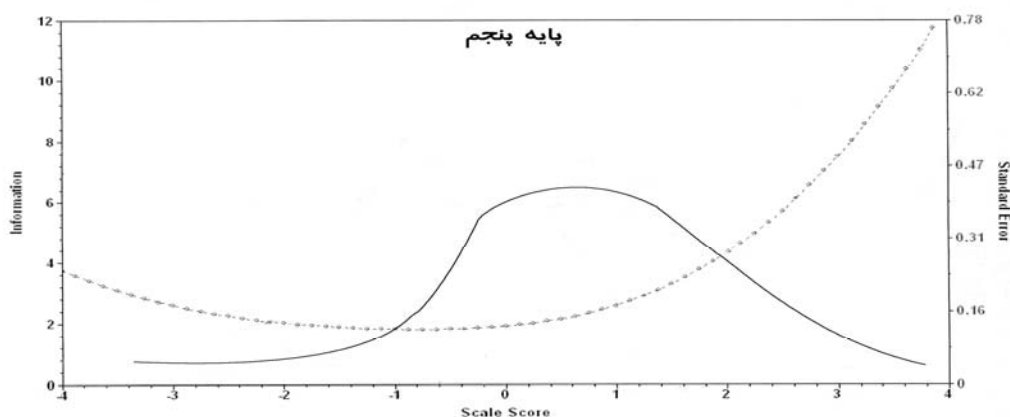
**رتبه‌های درصدی نمره‌های در هر یک از زیرمقیاس‌های آزمون تشخیص حساب نارسایی چگونه است؟** رتبه‌های درصدی نمره‌های در هر یک از خرده آزمون‌های آزمون تشخیصی حساب نارسایی در جدول (۵) آمده است. همان‌گونه که در جدول (۵) ملاحظه می‌گردد، ۵ درصد افراد در زیرمقیاس

جدول ۵- رتبه‌های درصدی نمره‌های در هر یک از زیر مقیاس‌های آزمون تشخیصی

حساب نارسایی و کل					زیر مقیاس
رتبه درصدی					
۹۵	۷۵	۵۰	۲۵	۵	
۹۹	۸۸	۸۰	۷۱	۵۹	کل
۱۸	۱۶	۱۴	۱۱	۹	حل مسأله
۹	۸	۷	۶	۴	اندازه‌گیری
۵	۴	۴	۳	۲	تخمین
۶	۵	۴	۳	۲	نسبت و درصد
۲۲	۱۹	۱۷	۱۵	۱۳	کسر و اعشار
۱۸	۱۵	۱۳	۱۱	۹	ضرب و تقسیم
۹	۸	۷	۶	۵	جمع و تفریق
۱۷	۱۴	۱۱	۱۰	۹	هندسه
۸	۷	۶	۵	۴	شمارش

هستند بهترین و مناسب‌ترین کاربرد را خواهد داشت. میزان آگاهی‌دهندگی این آزمون در سطوح توانایی بسیار پایین (۴-) و بسیار بالا (۴+) است. در این سطوح میزان آگاهی‌دهندگی این آزمون به کمتر از ۲ واحد می‌رسد.

منحنی تابع آگاهی آزمون (TIF) چگونه است؟ نمودار ۱ تابع آگاهی آزمون (TIF) نشان می‌دهد. همان‌گونه که در (شکل ۱) ملاحظه می‌گردد بیشترین میزان آگاهی‌دهندگی این آزمون در فاصلهٔ تتای ۰/۲- تا ۲+ قرار گرفته است که نشان‌دهندهٔ آن است که این آزمون برای افرادی که دارای این سطوح توانایی



نمودار ۱- منحنی تابع آگاهی آزمون (TIF) آزمون تشخیص حساب نارسایی پایه پنجم

مختلف و نظرخواهی از اساتید و صاحب‌نظران و سرگروه‌های آموزشی ریاضی پایه پنجم، محاسبه شاخص‌های کمی توافق در تحلیل محتوا، روایی صوری و محتوایی مطلوبی برای این آزمون احراز شد. همچنین همبستگی خرده آزمون‌ها و نمره کل این آزمون با آزمون ریاضیات ایران کمی مت و زیرمقیاس حساب در آزمون هوش ویسک - آر (شهیم، ۱۳۷۳) حاکی از روایی همزمان و برازش مدل ۹ عاملی در تحلیل عامل تأییدی حاکی از روایی سازه مطلوب آزمون بود. امکان تفکیک معنادار افراد حساب نارسا و عادی بر اساس سؤالات این آزمون با توجه به نقاط برش از یکدیگر دال بر روایی تفکیکی مطلوب این آزمون است. بررسی

مدت زمان لازم برای اجرای زیرمقیاس شمارش در این آزمون تشخیص چقدر است؟ برای تعیین مدت زمان لازم برای اجرای زیرمقیاس شمارش در این آزمون مدت زمانی که ۷۰ درصد آزمونی‌ها در آن دورهٔ زمانی توانستند سؤالات را به اتمام برسانند به عنوان زمان لازم در نظر گرفته شد. این زمان در زیرمقیاس شمارش ۲ دقیقه و ۴۵ ثانیه به دست آمد.

### بحث

یافته‌ها نشان داد که با بررسی همه جانبه، تأکید بر تحلیل و مقوله‌بندی محتوای کتاب ریاضی پایه پنجم، طراحی سؤالات



پایایی خرده آزمون‌های این آزمون نشان داد که همسانی درونی این آزمون از طریق آلفای کرونباخ و ثبات خرده آزمون‌ها و نمره کل این آزمون از طریق بازآزمایی مطلوب و قابل قبول است. مطابق با اعتقاد نانالی و برنشتاین (۱۹۹۴) ضریب پایایی بالای ۰/۷ مطلوب تلقی می‌شود بنابراین ضرایب پایایی خرده آزمون‌ها و کل این آزمون مطلوب و قابل قبول است.

حساسیت این آزمون نشان داد که می‌توان ۹۹ درصد موارد حساب نارسا را بر اساس این آزمون درستی تشخیص داد. بررسی و دقت این آزمون 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 سؤالات آزمون (ضرایب تشخیص) در نظریه کلاسیک نشان داد که هیچ یک از این ضرایب منفی و کمتر از ۰/۲ نبودند، حداقل ضرایب تشخیص در نظریه کلاسیک آزمون ۰/۲ است (متسامورنن، ۱۳۸۵). از سوی دیگر ضرایب دشواری ( $P_i$ ) در این آزمون نشان داد که ضرایب دشواری سؤالات در محدوده ۰/۲۴ تا ۰/۸۱ بوده. بنابراین ساده‌ترین سؤال این آزمون سؤال ۷۴ و دشوارترین آن سؤال ۶۵ بود. بهترین ضریب دشواری آزمون‌های تشریحی، ضرایب دشواری هستند که به ۰/۵ نزدیک باشند (شریفی و نجفی‌زند، ۱۳۷۸). بررسی ضرایب دشواری این سؤالات نشان داد که بیشتر آنها دارای ضریب دشواری بهینه یعنی نزدیک به ۰/۵ هستند.

در تفسیر ضرایب دشواری و تشخیص سؤالات در نظریه کلاسیک باید این دو در ارتباط با یکدیگر در نظر گرفته شوند. سؤالاتی که ضرایب تشخیصی بالا و دشواری متوسط (در سؤالات تشریحی ۰/۵) دارند، سؤالات مطلوبی در تفکیک افراد قوی از ضعیف تلقی می‌شوند (متسامورنن، ۱۳۸۵). بنابراین، همگی سؤالات بر اساس نظریه کلاسیک دارای ضرایب تشخیصی مطلوب و ضرایب دشواری نزدیک به ۰/۵ بوده و در نتیجه می‌توان بر اساس آنها افراد دارای توانمندی ریاضی بالا و پایین را تفکیک نمود.

بررسی ضرایب تشخیص ( $g$ ) و دشواری ( $g$ ) سؤالات این آزمون بر اساس نظریه سؤال - پاسخ نشان داد که سؤالات این آزمون دارای ضرایب تشخیص ( $g$ ) مطلوب (بالای ۰/۵) بوده و می‌توان بر اساس آنها افراد ضعیف و قوی را تفکیک نمود. همچنین میانگین کل ضرایب تشخیصی نیز نشان‌دهنده این است که در مجموع قدرت تمیز آزمون نیز مطلوب است. ضرایب دشواری این آزمون نیز نشان داد که سؤالات در طیف وسیعی از بسیار ساده تا بسیار دشوار در نوسان بوده (۲/۳۲ - تا ۳+) و

میانگین ضرایب دشواری نیز نشان می‌دهد که به‌طور متوسط دشواری این آزمون بیشتر از متوسط توانمندی ریاضی افراد است. به دلیل ماهیت تشریحی سؤالات این آزمون عامل شانس صفر بوده ( $g=0$ ) و احتمال پاسخ شانس وجود ندارد. بررسی سطح معنادار مقادیر  $\chi^2$  دو نیز نشان داد که همگی سؤالات با مدل دو پارامتری برازش کامل دادند.

در این آزمون مدت زمان لازم برای زیرمقیاس شمارش ۲ دقیقه و ۴۵ ثانیه به‌دست آمد. این نکته با یافته‌های جوردن و مونتانی (۱۹۹۷)، لندرل و همکاران (۲۰۰۴) و باترورث (۲۰۰۳) هماهنگ است. بر اساس یافته‌های این پژوهشگران اگر در آزمون‌های تشخیصی ریاضی (خرده آزمون‌های شمارش) زمان لحاظ نشود کودکان حساب نارسا نیز مانند کودکان عادی عمل خواهند کرد. در نتیجه زمانی این آزمون‌های تشخیصی دارای حساسیت کافی خواهند بود که زمانمند باشند.

بررسی منحنی آگاهی این آزمون نشان داد که بیشترین میزان آگاهی دهندگی این آزمون در فاصله سطح توانایی ۰/۲ تا ۰/۲+ بوده که نشان دهنده آن است که این آزمون برای افرادی که در این سطح توانایی قرار دارند بهترین و مناسب‌ترین کاربرد را خواهد داشت. میزان آگاهی دهندگی این آزمون در این سطوح توانایی حدود ۶ و پایین‌ترین میزان در سطوح توانایی بسیار پایین (۴-) و بسیار بالا (۴+) بوده که در این سطوح این میزان به کمتر از ۲ واحد می‌رسد و با توجه به متوسط تابع آگاهی این آزمون که توسط رایانه ۴/۴۹ به دست آمده، منحنی این آزمون یک تابع آگاهی متوسط به بالا را نشان می‌دهد. بنابراین این آزمون دارای قدرت لازم جهت سنجش متغیر زیر بنایی (خصیصه مکنون ریاضی) بوده و می‌توان بر اساس آن به ارزشیابی نسبتاً دقیق از وضعیت پیشرفت ریاضی و تفکیک افراد دارای ناتوانی ریاضی از افرادی عادی پرداخت. موارد استفاده از این آزمون با کاربرد آزمون ریاضیات ایرانی کی‌مت (محمد اسماعیل و هومن، ۱۳۸۱)، آزمون پیشرفت ریاضی (چین‌آوه و همکاران، ۱۳۸۳) هماهنگ است. می‌توان از این آزمون در تعیین نقاط ضعف و قوت در حیطه‌های مختلف ریاضی دانش آموزان پایه پنجم و نیز تفکیک دقیق افراد حساب نارسا و عادی سود جست.

- شریفی، حسن پاشا (۱۳۷۲). *اصول روان‌سنجی و روان‌سنجی و روان‌آزمایی*. تهران: انتشارات رشد.
- عریضی، حمیدرضا، کاووسیان، جواد و کدیور، پروین (۱۳۸۳). پایایی‌سنجی و اعتباریابی مقیاس شایستگی ریاضی کودکان اوتراخت در دانش‌آموزان دختر و پسر پیش‌دبستانی و دبستانی شهر اصفهان. *فصلنامه نوآوری‌های آموزشی*، سال ۳، شماره ۹، ۶۵-۴۹.
- کاپلن، هارولد و سادوک، بنیامین، جی (۱۳۸۳). خلاصه *روان‌پزشکی* (جلد سوم) (ترجمه حسن رفیعی). تهران: انتشارات ارجمند.
- محمد اسماعیل، الهه، و هومن، حیدرعلی (۱۳۸۱). انطباق و هنجاریابی آزمون ریاضیات ایران کی‌مت. *مجله پژوهش در حیطه کودکان استثنایی*، ۴، ۶، ۳۳۲-۳۲۳.
- مستامورنن، جری (۱۳۸۵). *مبانی نظری آزمون و آزمون‌سازی* (ترجمه شیده کامکار و اسدالله اسرایی). تهران: انتشارات بهینه.
- میرهاشمی، محمد (۱۳۷۱). *هنجاریابی آزمون تصاویر در هم‌اندزری در دانش‌آموزان ۷ تا ۱۵ سال و بزرگسالان شهر تهران*. پایان‌نامه کارشناسی ارشد. دانشگاه آزاد کرج.
- هومن، حیدرعلی (۱۳۷۱). *اندازه‌گیری‌های روانی و تربیتی (فن تهیه تست)*. تهران: انتشارات درسا.
- وکسلر، دیوید (۱۳۷۳). *مقیاس تجدید نظر شده هوش وکسلر برای کودکان* (انطباق و هنجاریابی: سیما شهیم). انتشارات دانشگاه شیراز.
- Ashcraft, M. H., Yamashita, T. S., & Aram, D. M. (1999). Mathematics performance in left and right brain-lesioned children. *Brain & Cognition*, 19, 208-252.
- Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology & Psychiatry*, 46, 3-18.
- Bloom, B. S. (1965). Taxonomy of educational objectives. *New York: McKay*.
- Butterworth, B. (2003). *Dyscalculia Screener*. London: Infer Nelson.
- Fuchs, L., S., & Fuchs, D. (2005). Enhancing mathematical problem solving for Student with disabilities. *The Journal of Special Education*, 39, 45-57.
- Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, 37, 4-15.
- Jordan, N. C., & Montani, T. O. (1997). Cognitive arithmetic & problem solving: a Comparison of احدی، حسن، و کاکاوند، علی‌رضا (۱۳۸۲). *اختلال‌های یادگیری (از نظریه تا عمل)*. تهران: انتشارات ارسباران.
- آلن، مری، جی، و وین، وندی، ام (۱۳۷۴). *مقدمه‌ای بر نظریه‌های اندازه‌گیری (روان‌سنجی)* (ترجمه: علی دل‌اور). انتشارات سمت.
- آناستازی، آن (۱۳۷۳). *روان‌آزمایی* (ترجمه محمد نقی براهنی). انتشارات دانشگاه تهران.
- تبریزی، مصطفی (۱۳۷۹). *درمان اختلال ریاضیات*. تهران: انتشارات فراروان.
- چین‌آوه، محبوبه، عرب‌زاده، زهرا، و منشی‌زاده، فریبا (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 خوراسگان اصفهان، ۳، ۳۲-۱۵.
- حمید، نجمه (۱۳۸۵). *بررسی اختلال یادگیری ریاضی در دانش‌آموزان دختر و پسر دوره ابتدایی ناحیه یک شهر تهران و اثر آموزش کاربردی، تقویت ژتونی و آرامش عضلانی در کاهش اختلال یادگیری ریاضی در آنان*. *مجله علوم تربیتی و روان‌شناسی دانشگاه شهید چمران اهواز*، ۱۳۸۵. دوره سوم، سال ۱۳، شماره ۲، ۱۳۶-۱۱۹.
- دادستان، پریخ، رضانی، مژگان، و بشاورد، سیمین (۱۳۸۰). *بررسی تحولی اختلال ویژه اکتساب حساب (حساب نارسایی)*. تهران: پژوهشکده کودکان استثنایی.
- رضانی، مژگان (۱۳۸۰). *بررسی میزان شیوع حساب نارسایی در دانش‌آموزان پایه‌های چهارم و پنجم مدارس ابتدایی شهر اصفهان*. تهران: پژوهشکده کودکان استثنایی.
- رضانی، مژگان (۱۳۸۲). *تحلیل خطاهای محاسباتی دانش‌آموزان حساب نارسا*. پژوهش در حیطه کودکان استثنایی، ۹ و ۱۰، ۳ و ۴، ۲۴۴-۲۲۳.
- ثرن‌دایک، رابرت، ال (۱۳۶۹). *روان‌سنجی کاربردی* (ترجمه حیدرعلی هومن). انتشارات دانشگاه تهران.
- ستاری، بهزاد (۱۳۸۲). *روان‌سنجی پیشرفته کاربردی*. مشهد: انتشارات آستان قدس رضوی.
- سرمد، زهره، بازرگان، عباس و حجازی، الهه. (۱۳۷۶). *روش‌های تحقیق در علوم رفتاری*. تهران: انتشارات آگاه.
- سیف، علی‌اکبر (۱۳۸۷). *اندازه‌گیری، سنجش و ارزشیابی آموزشی (ویرایش چهارم)*. تهران: نشر دوران.

- children with specific and general mathematics difficulties. *Journal of Learning Disabilities*, 30, 624-634.
- Jordan, N., 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.
- Landrel, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: a study of 8 to 9 Year old students. *Cognition*, 39, 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.