

Journal of research in Persian language and literature education



Explanatory Analysis of Iranian Students' Nonresponse in PIRLS 2021

M. Kabiri

Associate Professor, Institute for Research in Schooling and Education (IRSE), Tehran, Iran.

ABSTRACT

Keywords:

Reading literacy;
PIRLS;
Item nonresponse;
Constructed-response items.

1. Corresponding author
✉ maskabiri@yahoo.com

Background and Objectives: TIMSS and PIRLS results indicate that Iranian students exhibit higher rates of item nonresponse (blank responses), than the average of participating countries in these international large-scale assessments. The present study aims to identify patterns and factors associated with the occurrence of item nonresponse in the PIRLS 2021, using data from 5,962 fourth-grade students in Iran. **Methods**: Differences in item nonresponse rates were examined from three perspectives: comparative (Iran in comparison with other countries), trends (across different PIRLS cycles), and within-test (based on item characteristics). The results showed that the proportion of item nonresponse among Iranian students was higher than the international average, with particularly pronounced differences in constructed-response items. In addition, the likelihood of blank responses increased with greater text difficulty and higher levels of cognitive complexity in reading processes. **Findings**: Longitudinal analyses revealed that the highest rate of blank responses occurred in 2021, while the lowest was observed in 2011, indicating an overall increasing trend since 2011. Inferential analyses using regression modeling identified response format (constructed-response items) and item difficulty as the most important predictors of item nonresponse, together explaining a substantial proportion of the variance in item nonresponse. **Conclusion**: Overall, the findings suggest that item nonresponse among Iranian students are more strongly associated with limitations in cognitive ability and the higher mental demands of items. These results underscore the importance of strengthening higher-order reading skills and implementing targeted interventions aimed at enhancing students' competencies in order to reduce item nonresponse rates in international assessments.

Received: 1404/11/08

Reviewed: 1405/02/27

Accepted: 1405/03/09

PP: 12

Citation (APA): Kabiri, M. (2026). Explanatory Analysis of Iranian Students' Nonresponse in PIRLS 2021. *Research in Teaching Persian Language and Literature*, 7 (1), 85 – 97.

 <https://doi.org/10.48310/rpllp.2026.22217.1313>



واکاوی تبیینی پاسخ‌های سفید دانش‌آموزان ایرانی در مطالعهٔ پرلز ۲۰۲۱

مقاله پژوهشی

مسعود کبیری

دانشیار پژوهشگاه مطالعات آموزش و پرورش

چکیده

پیشینه و اهداف: شواهد نشان می‌دهد که دانش‌آموزان ایرانی نرخ بالاتری از عدم پاسخ به سؤالات یا پاسخ‌های سفید را نسبت به میانگین کشورهای شرکت‌کننده ثبت می‌کنند. پژوهش حاضر با هدف شناسایی الگوها و عوامل مرتبط با بروز پاسخ‌های سفید در مطالعهٔ پرلز ۲۰۲۱، بر داده‌های ۵۹۶۲ دانش‌آموز پایهٔ چهارم ایران با استفاده از روش تحلیل ثانویه مبتنی است. **روش‌ها:** در این مطالعه، تفاوت میزان پاسخ‌های سفید به‌صورت تطبیقی (ایران در مقایسه با سایر کشورها)، روندی (در مقایسهٔ ادوار مختلف پرلز) و درون‌آزمونی (بر اساس ویژگی‌های سؤالات) بررسی شد. نتایج نشان داد که نسبت پاسخ‌های سفید دانش‌آموزان ایرانی بیش از میانگین سایر کشورهاست و این تفاوت به‌ویژه در سؤالات تشریحی برجسته‌تر است. همچنین، با افزایش درجهٔ دشواری متن و پیچیدگی فرایندهای شناختی خواندن، احتمال بروز پاسخ سفید افزایش می‌یابد. **یافته‌ها:** تحلیل روندی حاکی از آن است که بیشترین میزان پاسخ سفید در سال ۲۰۲۱ و کمترین آن در سال ۲۰۱۱ مشاهده شده و از سال ۲۰۱۱ به بعد روندی افزایشی داشته است. در تحلیل استنباطی با استفاده از رگرسیون، نوع پاسخ‌گویی (تشریحی بودن سؤال) و دشواری سؤال به‌عنوان مهم‌ترین پیش‌بین‌های بروز پاسخ سفید شناسایی شدند و این دو متغیر بخش قابل‌توجهی از واریانس عدم پاسخ را تبیین کردند. **نتیجه‌گیری:** در مجموع، یافته‌ها نشان می‌دهد که پاسخ‌های سفید در میان دانش‌آموزان ایرانی احتمالاً به دلیل محدودیت‌های توانایی شناختی و بار ذهنی بالاتر سؤالات مرتبط است. این نتایج بر آموزش مهارت‌های خواندن سطح بالا و مداخلات هدفمند در راستای ارتقای توانمندی‌های دانش‌آموزان و افزایش تمرکز آنان برای کاهش نرخ عدم پاسخ در مطالعات بین‌المللی تأکید می‌کند.

<https://doi.org/10.48310/rpllp.2026.22217.1313>

واژه‌های کلیدی:

پاسخ سفید،
دانش‌آموزان ایران،
سواد خواندن،
سؤالات تشریحی،
مطالعهٔ پرلز.

۱. نویسنده مسئول

maskabiri@yahoo.com

تاریخ دریافت: ۱۴۰۴/۱۱/۰۸

تاریخ بازنگری: ۱۴۰۵/۰۲/۲۷

تاریخ پذیرش: ۱۴۰۵/۰۳/۰۹

شماره صفحات: ۱۳

مقدمه

در سنجش و اندازه‌گیری پیشرفت تحصیلی، عدم پاسخ به سؤال یا پاسخ سفید بخشی از بررسی‌ها و مدل‌سازی‌ها را به خود اختصاص می‌دهد. درک پدیده عدم پاسخ موضوع بسیار مهمی است زیرا بر کیفیت روان‌سنجی ابزار تأثیر گذاشته و به داده‌های غیرمعرف و بروز سوگیری^۱ منجر می‌شود و روایی و اعتبار آزمون را دچار مشکل می‌کند (مارک و بریکن^۲، ۲۰۲۴). از این‌رو، گزارش میزان پاسخ سفید در بررسی وضعیت عملکرد پاسخ‌دهندگان مهم است. به‌طور مثال، در گزارش مربوط به برنامه بین‌المللی سنجش دانش‌آموزان (پیزا)^۳ ۲۰۰۹ نشان داده شد که نرخ پاسخ سفید به‌طور میانگین برابر با ۱۰ درصد بود و در بخش خواندن مطالعه ملی پیل آموزشی^۴ نیز که در آلمان برگزار می‌شود، ۱۸/۸۳ از سؤالات پاسخ داده نشده بودند که البته ۱۳/۴۶ درصد آن مربوط به سؤالاتی بود که دانش‌آموزان زمان پاسخ‌گویی به آن را پیدا نکرده بودند (لو و وانگ^۵، ۲۰۲۰). چنین نکته‌ای در آزمون‌ها و مطالعات دیگر نیز گزارش است؛ به‌طوری که در آزمون مهارت‌های اصلی کوئیزلند^۶ میزان پاسخ سفید در سؤالات کوتاه پاسخ به ۶ درصد رسیده بود (مترز و بارنت^۷، ۲۰۰۳) و در اجرای رایانه‌ای برنامه بین‌المللی سنجشی شایستگی‌های بزرگسالان (پی‌اک)^۸ نرخ کلی پاسخ سفید شیلی ۱۶/۵ درصد بود (لو و وانگ^۹، ۲۰۲۰). این گزارش‌ها نشان می‌دهند که نرخ پاسخ سفید یکی از ویژگی‌های جدی گزارش‌ها است که ذهن دست‌اندرکاران آزمون‌سازی را به خود جلب نموده است. اینکه چرا عدم پاسخ به سؤال اتفاق می‌افتد و اینکه در چه نوع از سؤالات یا زیرگروه‌هایی از جامعه احتمال مشاهده پاسخ سفید بیشتر است و در برخورد با پاسخ‌های سفید چه مواجهه اجرایی یا روان‌سنجی باید مدنظر قرار گیرد، جزو همین دسته از موضوعات هستند که در این بخش از مقاله به آن خواهیم پرداخت.

نرخ عدم پاسخ یکی از مشکلات اصلی دانش‌آموزان ایران در مطالعات تیمز و پرلز است. بر اساس گزارش بین‌المللی نتایج تیمز ۲۰۲۳، متوسط پاسخ سفید بین دانش‌آموزان پایه چهارم ایرانی در ریاضی ۲۲ درصد و در علوم ۱۷ درصد است که این نرخ در ریاضی با اختلاف بیشترین میزان و در علوم دومین میزان عدم پاسخ در میان کشورهای شرکت‌کننده در مطالعه به شمار می‌رود (ون‌داویر و همکاران^۹، ۲۰۲۴). با وجود اینکه عملکرد ایران در مطالعه تیمز ۲۰۲۳ چندان مطلوب نبود ولی میزان عدم پاسخ دانش‌آموزان حتی در قیاس با عملکرد خود ایران نیز بسیار زیاد است. بدیهی است که رفع این مشکل یکی از روش‌های بهبود عملکرد ایران در این گونه مطالعات به شمار می‌آید. برای اجرای هر گونه اقدامی ابتدا لازم است دلایل بروز این پدیده در جامعه ایرانی شناخته شود. این مطالعه به منظور شناخت همین موضوع با تأکید بر مطالعه پرلز طرح‌ریزی شده است و پس از توصیف تفاوت میان نرخ عدم پاسخ میان انواع سؤالات همچون نوع سؤالات، انواع متون، فرایندهای پاسخ به سؤال به صورت درون کشوری، بین کشوری و تطبیقی میزان احتمال هر یک را روی میزان عدم پاسخ بررسی خواهد کرد.

چارچوب نظری و پیشینه پژوهشی

تفاوت میان پاسخ سفید در انواع سؤالات می‌تواند تا اندازه‌ای دلایل بروز عدم پاسخ به سؤالات را تبیین کند. در سنجش ملی پیشرفت‌های آموزشی آمریکا^{۱۰} در سال ۲۰۱۳، نسبت پاسخ سفید در سؤالات دویخشی ریاضی از ۰/۹۹ درصد تا ۳۸/۰۶ درصد و در سؤالات چند بخشی از ۰/۷۳ درصد تا ۱۵/۸۵ درصد در نوسان بود. در تحلیلی هم که چالدز^{۱۱} و

1. bias
2. Marcq & Braeken
3. Programme for International Student Assessment (PISA)
4. National Educational Panel Study (NEPS)
5. Lu & Wang
6. Queensland Core Skills (QCS) Test
7. Matters & Burnett
8. Programme for the International Assessment of Adult Competencies (PIAAC)
9. von Davier
10. The National Assessment of Educational Progress (NAEP)
11. Childs

دیگران (۲۰۱۸) با استفاده از روش تحلیل کلاس مکنون انجام دادند بین افرادی که هم سؤالات چندگزینه‌ای و هم سؤالات بازپاسخ را سفید می‌گذاشتند و افرادی که تنها سؤالات بازپاسخ را سفید می‌گذاشتند تمایز ایجاد شده بود. به نظر می‌رسد در سؤالات پاسخ‌گزین در مقابل سؤالات پاسخ‌ساز احتمال کمتری برای عدم پاسخ وجود دارد که احتمالاً به‌خاطر درگیری ذهنی کمتر آن باشد. با این حال، زمانی که سؤالات پاسخ‌گزین (همانند سؤالات چندگزینه‌ای) با جریمه ناشی از حدس (نمره منفی) همراه می‌شود، این قاعده متفاوت می‌شود.

پاسخ‌های سفید در حالت معمول به دلیل آنکه امتیازی دریافت نمی‌کنند معادل با پاسخ‌های نادرست تلقی می‌شوند و برخوردی شبیه با پاسخ‌های نادرست با آن‌ها می‌شود. با این وجود، در سنجش‌های کلان‌مقیاس مواجهه با پاسخ‌های سفید کمی متفاوت است. در مطالعات تیمز^۱ و پرلز^۲ سؤالات پاسخ داده‌نشده در انتهای دفترچه آزمون به عنوان سؤالات به پایان نرسیده در نظر گرفته می‌شوند. معیار تعیین، محاسبه سؤالات بعد از سه پاسخ سفید متوالی است. به عنوان مثال، اگر در بخشی از یک دفترچه ۲۶ سؤال وجود داشته باشد و آزمون‌دهنده‌ای از سؤال ۱۷ به بعد را سفید گذاشته باشد، سؤالات ۲۰ تا ۲۶ به عنوان سؤالات «به پایان نرسیده»^۳ تعیین می‌شوند (کوولیک^۴ و همکاران، ۲۰۲۰). در مرحله محاسبات نمره‌دهی، سؤالات سفید به عنوان نادرست در نظر گرفته می‌شوند ولی سؤالات «به پایان نرسیده» نادرست در نظر گرفته نمی‌شود و با آنان همانند سؤالاتی رفتار می‌شود که به آزمون‌دهندگان ارائه نشده‌اند. این کار برای برآورد بهتر و دقیق‌تر پارامترهای سؤال است. در مطالعه پیزا هم همین قاعده پی‌گرفته می‌شود ولی از نسبت سؤالات بدون پاسخ به عنوان پیش‌بین در مدل رگرسیونی استفاده می‌شود (روبیترز^۵، ۲۰۲۳).

شیوه دیگر برخورد با پدیده عدم پاسخ مربوط به مدل‌سازی‌های روان‌سنجی است. موضوع موردبحث در این منظر، تأثیر بررسی رفتار بی‌ارتباط با سازه^۶ شبیه سبک‌های پاسخ است که پاسخ‌های سفید بخشی از آن را تشکیل می‌دهد. برای درک خمیرمایه انواع مدل‌سازی‌های روان‌سنجی مختلف، دو آزمون‌شونده را در نظر بگیرید که سرعت یکسان و توانایی یکسانی دارند ولی در تمایل به پاسخ متفاوت هستند؛ به این صورت که یکی متمایل به پاسخ به همه سؤالات و دیگری متمایل به سفید گذاشتن سؤالات است. آنکه تمایل زیادی به پاسخ را نشان می‌دهد، به همه سؤالات پاسخ می‌دهد و دیگری به برخی از سؤالات آزمون پاسخ سفید می‌دهد. اگر سؤالات سفید را به عنوان نادرست در نظر بگیریم، نمره این دو نفر به این صورت خواهد بود که برای اولی توانایی و برای دومی ترکیبی از توانایی و تمایل به پاسخ موجود است. از این‌رو، چنانچه عدم پاسخ به سؤالات را نادیده بگیریم، برآوردهای توانایی سوگیر خواهند شد (پوهل^۷ و همکاران، ۲۰۲۱).

جدا از نگاه روان‌سنجی، دلیل بروز عدم پاسخ نیز موضوعی مهم است، زیرا در آزمون‌های سرنوشت‌سازی که بخش زیادی از سرنوشت تحصیلی افراد به عملکرد آزمون مرتبط است هم میزان قابل توجهی از عدم پاسخ گزارش شده است (مترز و بارت، ۲۰۰۳). به عنوان اولین حدس، به خصوص در آزمون‌های غیرسرنوشت‌ساز، کاهش انگیزه به عنوان علت عدم پاسخ معرفی می‌شود. انگیزه سازه‌ای چندمؤلفه‌ای است که نظریه انتظار - ارزش^۸ بهتر می‌تواند این ویژگی را تبیین کند. بر اساس این نظریه، رفتار فرد تابعی از انتظارات (باور به توانایی افراد برای تکمیل رفتار) و ارزش (ارزیابی افراد از تکلیف مشخص) است. انتظارات باورهای افراد در مورد این است که چگونه به طور موفقیت‌آمیزی تکلیف معین را انجام خواهند داد. ارزش به دو بخش ارزش درونی^۹ و ارزش کاربردی^{۱۰} تقسیم می‌شود. مفهوم ارزش درونی مشخص است، ولی ارزش کاربردی، سودمندی دریافت‌شده یک تکلیف در رابطه با به‌دست آوردن پاداش یا تسهیل پیشرفت

1. Trends in International Mathematics and Science Study (TIMSS)

2. Progress in International Reading Literacy Study (PIRLS)

3. not-reached

4. Kowolik

5. Robitzsch

6. construct-irrelevant behaviors

7. Pohl

8. Expectancy-value theory

9. intrinsic value

10. utility value

سایر هدف‌های فوری یا طولانی‌مدت است (دی‌چیاچیو^۱ و همکاران، ۲۰۱۶). مطابق با این نظریه، آزمون‌شوندگان به دلیل اینکه یا ارزشی برای آزمون قائل نیستند یا توانایی خود را برای پاسخ‌دادن کم می‌بینند به سؤال پاسخ نمی‌دهند. هرچند ممکن است برخی از عدم پاسخ‌ها به انگیزه پایین قابل انتساب باشد، برخی مطالعاتی نشان داده‌اند که بسیاری از دانش‌آموزان کم‌انگیزه، حتی به آزمون‌های غیرسرنوشت‌ساز هم پاسخ داده ولی این کار را بدون تلاش انجام می‌دهند. هنگامی که آزمون‌ها به صورت رایانه‌ای اجرا می‌شود برخی از این پاسخ‌های بدون تلاش تشخیص داده می‌شوند، زیرا دانش‌آموزان سؤال را در زمان کمتری نسبت به آنچه به طور معمول برای خواندن لازم است پاسخ می‌دادند. این موضوع که برخی دانش‌آموزان سؤال را نخوانده پاسخ می‌دهند، نشان می‌دهد که تصمیم درباره پاسخ دادن یا پاسخ ندادن ضرورتاً درباره یک سؤال و مربوط به زمان مشخصی گرفته نمی‌شود و به نظر می‌رسد تصمیم این گروه از دانش‌آموزان این گونه است که «آیا به همه سؤالات پاسخ دهم؟» (چالدز و دیگران، ۲۰۱۸).

علاوه بر موضوع انگیزه، ارتباط با توانایی‌های شناختی هم عامل مهم دیگری برای توضیح دلیل بروز عدم پاسخ به سؤالات است. برخی از پژوهش‌ها نشان دادند دانش‌آموزان با عملکرد پایین، سؤالات سفید بیشتری را داشتند. با این وجود، ارتباط میان عملکرد و نرخ عدم پاسخ آن‌چنان زیاد نیست که رابطه قطعی بین این دو متغیر فرض شود (ون‌شرادر و انسلی^۲، ۲۰۰۶؛ چالدز و دیگران، ۲۰۱۸). همین موضوع می‌تواند از نظرگاه اطمینان افراد به توانایی‌های خود نیز ملاحظه شود. به عبارت دیگر، ادراکی که دانش‌آموزان از توانایی‌های خود دارند و سؤالی را دشوار یا آسان ارزیابی می‌کنند در میزان عدم پاسخ آنان تأثیرگذار است. در مطالعه‌ای افرادی که به طور کلی تصمیم به عدم ادامه پاسخ به سؤالات داشتند را از بین افرادی یافتند که خودکارآمدی علوم پایینی را نشان دادند (دی‌چیاچیو و همکاران، ۲۰۱۶).

در کنار عوامل فوق به ویژگی‌های روانشناختی افرادی که پاسخ سفید بیشتری داشتند، اشاره می‌شود. تفاوت‌های شخصیتی نیز ممکن است نقش مهمی در بروز عدم پاسخ داشته باشند. در این میان تمایل به ریسک^۳ به خصوص برای سؤالات چندگزینه‌ای (ون‌شرادر و انسلی، ۲۰۰۶) قابل ذکر است. جدای از این، مطالعه‌ای نشان داده است که دانش‌آموزان سفیدپوست با احتمال بیشتری به همه سؤالات چهارگزینه‌ای جواب می‌دادند و پس از آن‌ها دانش‌آموزان آسیایی و لاتین قرار داشتند و دانش‌آموزان سیاه‌پوست بالاترین نرخ عدم پاسخ را نشان دادند (چالدز و همکاران، ۲۰۱۸). جنسیت نیز عامل مداخله‌ای دیگری است. ون‌شرادر و انسلی (۲۰۰۶) مجموعه‌ای از پژوهش‌ها را ذکر کردند که تمایل بیشتر دختران برای ارائه پاسخ سفید را در مقایسه با پسران، حتی در هنگام هم‌سطح بودن توانایی آن‌ها وجود دارد. با این وجود، در تحلیل‌های بعدی تأثیر این متغیرها چندان بالا ارزیابی نشد. از جمله آنکه مشخص شد که ۳۰ درصد از دانش‌آموزان یک یا چند سؤال را سفید گذاشته بودند ولی هیچ‌یک از متغیرهای دموگرافیکی همچون جنسیت یا نژاد حتی ۵/۰ درصد از واریانس عدم پاسخ را پیش‌بینی نکردند (براون و همکاران، ۲۰۱۴، به نقل از چالدز و همکاران، ۲۰۱۸).

روش پژوهش

نمونه

پاسخ‌های دانش‌آموزان شرکت‌کننده در مطالعه پرلز ۲۰۲۱ برای تحلیل در این مطالعه موردبررسی قرار گرفتند. پاسخ‌های ۵۹۶۲ دانش‌آموز پایه چهارمی در ایران بر اساس نمونه‌گیری خوشه‌ای طبقه‌ای دو مرحله‌ای در ۲۱۸ مدرسه در کل ایران مبنای این مطالعه بود.

ابزار

در مطالعه پرلز ۲۰۲۱، ۱۸ متن برای اندازه‌گیری درک مطلب دانش‌آموزان استفاده شدند. نیمی از متون بر اساس هدف خواندن اطلاعاتی و نیمی دیگر ادبی به شمار می‌رفتند. همچنین، دسته‌بندی متون به صورت ساده، متوسط و دشوار هم مساوی بود (۶ متن برای هر یک). برای هر یک از متون، مجموعه‌ای از سؤالات در نظر گرفته شد (حدود ۱۵ سؤال) که

1. Di Chiacchio
2. von Schrader & Ansley
3. risk-taking tendency

دانش‌آموزان پس از خواندن متن به آن پاسخ می‌دادند. این سؤالات بر اساس فرایندهای لازم خواندن در یکی از چهار فرایند تمرکز و بازیابی اطلاعاتی صریحاً بیان شده؛ استنباط مستقیم؛ تلفیق و تفسیر ایده‌ها و اطلاعات؛ و ارزشیابی و نقد محتوا و عناصر متنی را دربر می‌گرفتند. به طور کلی در مطالعه پرلز ۲۷۵ سؤال برای اندازه‌گیری خواندن به کار برده شد. پس از تحلیل سؤالات نسبت پاسخ به هر یک از اجزای سؤال برای تک تک کشورهای شرکت‌کننده گزارش شد. این اجزا شامل درصد پاسخ به هر یک گزینه‌ها (در سؤالات چندگزینه‌ای) و درصد پاسخ نادرست، ناقص و کامل (در سؤالات تشریحی) بود. علاوه بر این، درصد پاسخ‌های سفید و درصد دانش‌آموزانی که فرصت پاسخگویی پیدا نکرده‌اند نیز به تفکیک مشخص شد. هر دوی این درصدها از عدم پاسخ دانش‌آموزان حاصل می‌شوند ولی تمامی پاسخ‌های سفید پس از سه پاسخ سفید انتهایی به عنوان درصد عدم فرصت پاسخگویی یا به پایان نرسیده محاسبه می‌شوند. برای مثال، چنانچه فردی از سؤال ششم به بعد را پاسخ نداده باشد، سؤال شش تا نه او پاسخ سفید و سؤال ۱۰ تا آخر به عنوان عدم فرصت پاسخگویی در نظر گرفته می‌شود.

طرح پژوهش

این مطالعه از روش تحلیل ثانویه^۱ داده‌های مطالعه پرلز ۲۰۲۱ استفاده می‌کند. بخشی از نتایج از طریق مقایسه توصیفی (بررسی تطبیقی میزان پاسخ ایران در مقایسه با سایر کشورها، بررسی روندی وضعیت در مقایسه ادوار مختلف پرلز و مقایسه بر اساس ویژگی‌های سؤال) حاصل شد و تحلیل استنباطی از طریق کاربرد روش رگرسیون چندمتغیره روی سؤالات مطالعه پرلز ۲۰۲۱ اجرا شد. واحد تحلیل در روش رگرسیون (و سایر تحلیل‌های قبلی) سؤالات مطالعه پرلز بودند.

متغیرها

ویژگی‌های سؤال: هر سؤال ویژگی‌های متخلفی دارد. در این پژوهش، چگونگی پاسخ‌دهی با سؤال (چندگزینه‌ای یا تشریحی)، درجه دشواری متن بر اساس درجه‌بندی مرکز بین‌المللی مطالعه پرلز (در سه سطح آسان، متوسط و دشوار)، نوع متن بر اساس چارچوب مطالعه پرلز (متن به منظور کسب اطلاعات یا کسب تجربه ادبی)، فرایندهای درک مطلب یا سطوح پیچیدگی و شناختی سؤال (در چهار تمرکز و بازیابی اطلاعاتی صریحاً بیان شده، استنباط مستقیم، تلفیق و تفسیر ایده‌ها و اطلاعات و ارزشیابی و نقد محتوا و عناصر متنی) و سطح دشواری سؤال (بر اساس پارامتر دشواری در برآورهای نظریه پرسش پاسخ)^۲ به عنوان ویژگی‌های سؤال در نظر گرفته شد. پاسخ سفید: درصد عدم پاسخ به سؤال به عنوان پاسخ سفید در نظر گرفته شد. در این تعریف، پاسخ‌هایی که به عنوان «پاسخ‌های نرسیده» کدگذاری شده بودند از پاسخ‌های سفید کنار گذاشته شد.

نتایج

مقایسه بین‌المللی میزان پاسخ‌های سفید: به عنوان معیاری برای ارزیابی نسبت پاسخ‌های سفید دانش‌آموزان ایرانی از میانگین کشورهای شرکت‌کننده در مطالعه پرلز ۲۰۲۱ استفاده شد. بدین منظور ۳۲ کشور از کشورهای شرکت‌کننده در این مطالعه که در شکل کاغذی آزمون شرکت کرده بودند، انتخاب و با ایران مقایسه شدند. نتایج کلی نشان داد در حالی که به طور متوسط ۱۲ درصد از دانش‌آموزان ایرانی به پاسخ‌های مطالعه پرلز ۲۰۲۱ پاسخ نداده بودند، این نسبت برای سایر کشورها برابر با ۶/۴ درصد بود. این مقایسه ساده نشان داد که دانش‌آموزان ایرانی در مقایسه با دانش‌آموزان سایر کشورهای شرکت‌کننده نسبت بالاتری از سؤالات را بدون پاسخ باقی می‌گذارند. در جدول و نمودار ۱ بین میزان پاسخ سفید در ایران و سایر کشورها بر اساس ویژگی‌های سؤال مقایسه شده است.

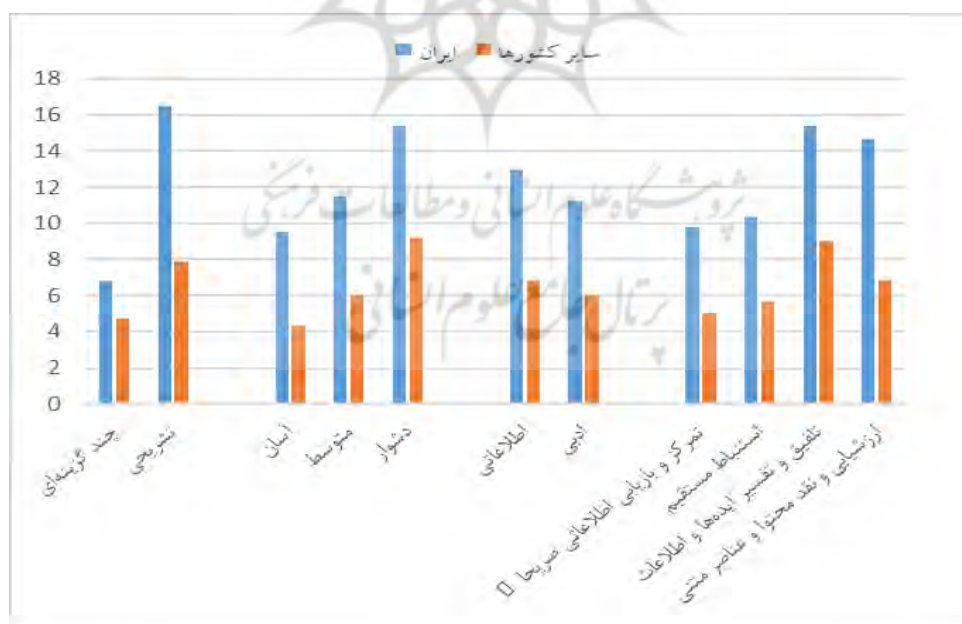
جدول ۱. مقایسه میزان پاسخ‌های سفید در ایران و سایر کشورها به تفکیک ویژگی‌های سؤال

ویژگی	ایران	سایر کشورها
نوع سؤال		

1. secondary analysis
2. item-response theory

۴/۷	۶/۸	چندگزینه‌ای
۷/۹	۱۶/۵	تشریحی
		درجه دشواری متن
۴/۳	۹/۵	آسان
۶	۱۱/۵	متوسط
۹/۲	۱۵/۴	دشوار
		نوع متن
۶/۹	۱۳	اطلاعاتی
۶	۱۱/۲	ادبی
		فرایندهای درک مطلب
۵	۹/۸	تمرکز و بازیابی اطلاعاتی صریحاً بیان شده
۵/۷	۱۰/۴	استنباط مستقیم
۹	۱۵/۴	تلفیق و تفسیر ایده‌ها و اطلاعات
۶/۹	۱۴/۷	ارزشیابی و نقد محتوا و عناصر متنی

بر اساس این جدول، در برخی از ویژگی‌ها، مثل میزان دشواری متن داستانی و نوع متن (ادبی یا اطلاعاتی)، بین میزان پاسخ سفید دانش‌آموزان ایرانی و سایر کشورها در سطوح مختلف متغیر تفاوت چندانی وجود ندارد. در فرایندهای خواندن هم تقریباً همین قاعده دیده می‌شود و تنها در فرایند ارزشیابی و نقد محتوا و عناصر متنی تفاوت کمی بیشتر از سه فرایند قبلی است. با این وجود، نوع سؤال در میزان پاسخ‌های سفید تفاوت ایجاد می‌کند؛ به این صورت که با اینکه پاسخ‌های سفید سؤالات چند گزینه‌ای در دانش‌آموزان ایرانی ۲ درصد بیشتر از میانگین سایر کشورها است ولی در پاسخ‌های تشریحی، میزان تفاوت پاسخ سفید دانش‌آموزان ایرانی و سایر کشورها به بیش از ۸ درصد می‌رسد. این تحلیل به صورت گرافیکی هم قابل نمایش است.



نمودار ۱. مقایسه میزان پاسخ‌های سفید در ایران و سایر کشورها به تفکیک ویژگی‌های سؤال

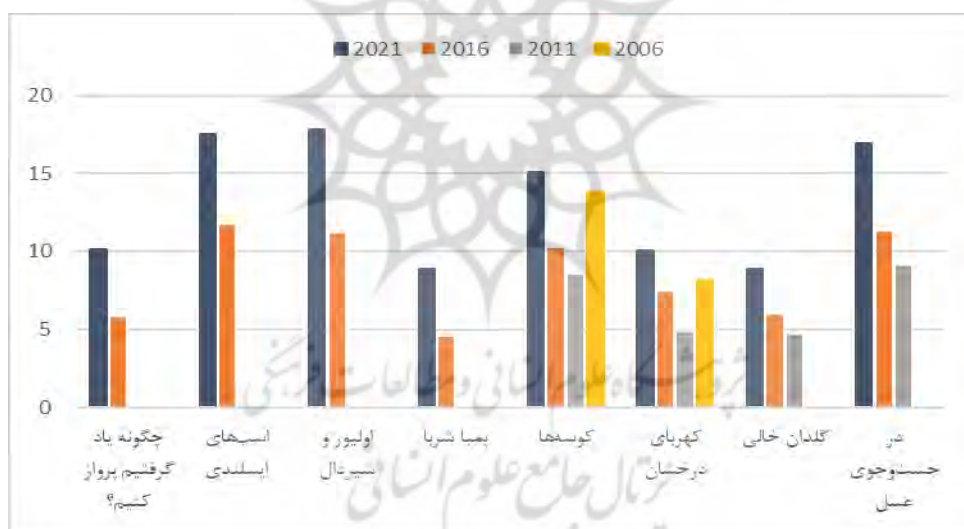
روند تغییرات پاسخ‌های سفید دانش‌آموزان: برای بررسی تغییرات پاسخ‌های سفید در دوره‌های مختلف، متون مشترک بین ادوار مطالعه پرلز انتخاب شدند. از بین ۱۸ متن مورد استفاده در مطالعه پرلز ۲۰۲۱، ۴ متن بین دوره‌های ۲۰۲۱ و ۲۰۱۶ مشترک بودند. هم‌چنین، دو متن نیز از سال ۲۰۱۱ تا ۲۰۲۱ استفاده شده و دو متن دیگر هم از سال

۲۰۰۶ تا ۲۰۲۱ استفاده شدند و بنابراین معیار مناسبی برای مقایسه روند به شمار می‌روند. جدول ۲ میزان پاسخ‌های سفید بین دوره‌های مختلف در متون مشترک را نشان می‌دهد.

جدول ۲. روند تغییرات پاسخ‌های سفید در دوره‌های مختلف مطالعه پزلز

متن	۲۰۲۱	۲۰۱۶	۲۰۱۱	۲۰۰۶
چگونه یاد گرفتیم پرواز کنیم؟	۱۰/۱	۵/۸		
اسب‌های ایسلندی	۱۷/۶	۱۱/۷		
اولیور و شیردال	۱۷/۹	۱۱/۲		
پمبا شربا	۸/۹	۴/۶		
گلدان خالی	۸/۹	۵/۹	۴/۷	
در جست‌وجوی عسل	۱۶/۹	۱۱/۲	۹/۱	
کوسه‌ها	۱۵/۱	۱۰/۲	۸/۶	۱۳/۹
کهربای درخشان	۱۰/۱	۷/۴	۴/۸	۸/۲

چنانچه در این جدول نشان داده می‌شود، در سال ۲۰۲۱ بیشترین میزان پاسخ سفید دانش‌آموزان مشاهده می‌شود. در مقایسه، در سال ۲۰۱۱ نیز کمترین میزان پاسخ سفید گزارش شده است. همچنین، میزان پاسخ‌های سفید دانش‌آموزان در سال ۲۰۰۶ به نسبت‌های مشاهده‌شده در سال ۲۰۲۱ نزدیک هستند. نمودار ۲ همین مقایسه را به صورت تصویری نشان می‌دهد.



نمودار ۲. روند تغییرات نسبت پاسخ‌های سفید در ادوار مطالعه پزلز

مقایسه میزان پاسخ سفید در انواع سؤالات: میزان پاسخ‌های سفید دانش‌آموزان در سؤالات مختلف متفاوت است. در مطالعه پزلز ۲۰۲۱ میزان پاسخ سفید دانش‌آموزان از ۰/۹ درصد تا ۵۴/۳ درصد در نوسان بود. برای اینکه مشخص شود در هر یک از انواع سؤالات به چه میزان پاسخ سفید وجود دارد، نسبت پاسخ سفید را در دسته‌های مختلف سؤالات بررسی کرده‌ایم. جدول ۳ این مقایسه را میان سؤالات از لحاظ نوع سؤال، نوع متن، درجه دشواری متن و فرایندهای درک مطلب نشان می‌دهد.

جدول ۳. مقایسه بین میزان پاسخ‌های سفید در انواع سؤالات

ویژگی‌های سؤال	درصد پاسخ سفید
نوع پاسخگویی	

۶/۸	چند گزینه‌ای
۱۶/۵	تشریحی
	درجه دشواری متن
۹/۵	آسان
۱۱/۵	متوسط
۱۵/۴	دشوار
	نوع متن
۱۳	اطلاعاتی
۱۱/۲	ادبی
	فرایندهای درک مطلب
۹/۸	تمرکز و بازایی اطلاعات صریحاً بیان شده
۱۰/۴	استنباط مستقیم
۱۵/۴	تلفیق و تفسیر ایده‌ها و اطلاعات
۱۴/۷	ارزشیابی و نقد محتوا و عناصر متنی

چنانچه این نمودار نشان می‌دهد، سؤالات تشریحی بیشتر از سؤالات چندگزینه‌ای سفید گذاشته می‌شوند. به‌طور متوسط سؤالات تشریحی توسط ۱۶/۵ درصد از دانش‌آموزان پاسخ داده نمی‌شوند ولی میزان متناظر برای سؤالات چندگزینه‌ای ۶/۸ درصد است. همچنین هرچه متن دشوارتر می‌شود، به احتمال بیشتری با پاسخ‌های سفید روبه‌رو می‌شود؛ به‌طوری که متون ساده به‌طور متوسط با ۹/۵ درصد پاسخ سفید روبه‌رو می‌شوند ولی با تغییر درجه دشواری متن به متوسط و دشوار درصد پاسخ‌های سفید به ۱۱/۵ و ۱۵/۴ درصد می‌رسد. در انواع متون از لحاظ هدف خواندن تفاوت چندانی ملاحظه نمی‌شود و با اینکه متون اطلاعاتی اندکی بیش از متون ادبی با پاسخ سفید مواجه می‌شوند (۱۳ در مقابل ۱۱/۲ درصد) ولی تفاوت ناچیز است. در نهایت، مشخص شد که با افزایش پیچیدگی سؤالات از لحاظ شناختی نیز میزان پاسخ سفید بیشتر می‌شود. این تفاوت به خصوص بین مهارت‌های ساده‌تر درک مطلب (تمرکز و بازایی اطلاعات صریحاً بیان شده و استنباط مستقیم به ترتیب با ۹/۸ و ۱۰/۴ درصد) و مهارت‌های سطح بالاتر (تفسیر و تلفیق ایده‌ها و اطلاعات و ارزیابی و نقد محتوا و عناصر متنی به ترتیب با ۱۵/۴ و ۱۴/۷ درصد) دیده می‌شود. مقایسه نسبت پاسخ‌های سفید در نمودار ۳ نشان داده شده است.



نمودار ۳. مقایسه نسبت پاسخ‌های سفید در انواع سؤالات

تحلیل تأثیر عوامل مرتبط با سؤال بر بروز پاسخ‌های سفید: برای اینکه مشخص شود سؤالات واجد چه ویژگی‌هایی هستند که بیشتر منجر به پاسخ‌های سفید دانش‌آموزان می‌شوند از تحلیل رگرسیون استفاده شد. در این تحلیل درصد

پاسخ‌های سفید به عنوان متغیر ملاک و ویژگی‌های سؤال به عنوان متغیرهای پیش‌بین استفاده شدند. این ویژگی‌ها شامل درجه دشواری سؤالات (بر مبنای پارامتر جایگاه در نظریه پرسش پاسخ)، نوع پاسخگویی به سؤالات (چندگزینه‌ای یا تشریحی)، نوع متن (اطلاعاتی یا ادبی)، درجه دشواری متن (ساده، متوسط و دشوار بر اساس ارزیابی اولیه مرکز بین‌المللی مطالعه) و فرایندهای شناختی سؤالات (تمرکز و بازیابی اطلاعات صریحاً بیان‌شده، استنباط مستقیم، تفسیر و تلفیق ایده‌ها و اطلاعات، و ارزیابی و نقد محتوا و عناصر متنی) بودند. به جز درجه دشواری سؤالات که مقیاس متغیر پیوسته بود بقیه متغیرهای پیش‌بین در سطح اسمی بودند. برای تحلیل آن‌ها از کدگذاری تصنعی استفاده شد و طبقات سؤالات چندگزینه‌ای، سطح تمرکز و بازیابی اطلاعات صریحاً بیان‌شده، متون ادبی، میزان دشواری ساده متون به عنوان طبقات مرجع در نظر گرفته شدند. نتایج تحلیل در جدول ۴ ارائه شده است.

جدول ۴. نتایج تحلیل رگرسیون عوامل مرتبط به سؤالات بر پاسخ سفید

پارامتر	ضریب استاندارد نشده	خطای استاندارد	بتا	t
عرض از مبدأ	۱۸/۱۳	۰/۶۴		
نوع سؤالات (تشریحی)	-۹/۱۲	۰/۸۷	-۰/۵۰۴	-۱۰/۴۹*
دشواری سؤالات	۳/۳۰	۰/۵۰	۰/۳۱۷	۶/۶*

تحلیل رگرسیون با استفاده از روش ورود پیش‌رونده اجرا شد و با دو متغیر پیش‌بین معنادار ۰/۳۸۵ از واریانس متغیر ملاک را تبیین کرد. مدل رگرسیون با دو پیش‌بین با مقدار F برابر با ۸۴/۳۳ معنادار بود که به معنای مناسب بودن مدل ارزیابی است. نتایج نشان داد که از میان تمامی متغیرهای موردبررسی، دو متغیر تشریحی بودن سؤالات و میزان دشواری سؤالات در بروز پاسخ سفید دانش‌آموزان ایران در مطالعه پزلز ۲۰۲۱ نقش معناداری دارند. از لحاظ مقایسه این دو متغیر، متغیر سؤالات تشریحی اهمیت بیشتری در مقایسه با دشواری سؤالات داشت که به معنای آن است که مهم‌ترین عامل در سفید باقی گذاشتن سؤالات توسط دانش‌آموزان ایرانی در مطالعه پزلز ۲۰۲۱ تشریحی بودن سؤال است.

بحث و نتیجه‌گیری

موضوع نرخ عدم پاسخ دانش‌آموزان در مطالعات بین‌المللی تیمز و پزلز یکی از موضوعات مهم تحلیل وضعیت عملکرد دانش‌آموزان ایرانی است. این اهمیت از نرخ بالای ۲۲ درصدی عدم پاسخ در مطالعه تیمز ۲۰۲۳ قابل‌درک است. بدیهی است که ریشه‌شناسی علت عملکرد ضعیف دانش‌آموزان ایرانی در مطالعات بین‌المللی بدون توجه به وضعیت عدم پاسخ‌گویی دانش‌آموزان به سؤالات راه به جایی نخواهد برد. از این‌رو، در مطالعه حاضر به بررسی علت عدم پاسخ دانش‌آموزان به سؤالات مطالعه پزلز ۲۰۲۱ پرداخته شده است. این تحلیل به مقایسه توصیفی عدم پاسخ به سؤالات در انواع سؤالات (چندگزینه‌ای و بازپاسخ)، فرایندهای درک مطلب سؤال (تمرکز و بازیابی اطلاعاتی صریحاً بیان‌شده، استنباط مستقیم، تلفیق و تفسیر ایده‌ها و اطلاعات و ارزشیابی و نقد محتوا و عناصر متنی)، اهداف متن (ادبی و اطلاعاتی) و درجه دشواری متن توجه کرده است. تحلیل توصیفی در سه شکل تطبیقی (مقایسه ایران با سایر کشورها)، روندی (مقایسه میان متون در ادوار قبلی) و بین سؤالی (مقایسه بر اساس ویژگی‌های سؤال) صورت پذیرفته است. سپس با اجرای تحلیل رگرسیونی عوامل مؤثر بر عدم پاسخ به صورت استنباطی برآورد شده است. از داده‌های مطالعه پزلز ۲۰۲۱ برای این تحلیل کمک گرفته شده است.

نتایج بررسی نشان داد که هر چند میزان پاسخ سفید دانش‌آموزان ایرانی در مقایسه با سایر کشورهای شرکت‌کننده در مطالعه پزلز ۲۰۲۱ بیشتر بوده است ولی در فرایند ارزشیابی و نقد محتوا و عناصر متنی که پیچیده‌ترین فرایند خواندن به‌شمار می‌رود بیش از بقیه بود. همچنین، تفاوت بین میزان پاسخ سفید دانش‌آموزان ایران در مقایسه با سایر کشورها در سؤالات تشریحی بیش از سؤالات چندگزینه‌ای است. در بررسی روندی نیز مشخص شد که میزان پاسخ سفید دانش‌آموزان

در سال ۲۰۲۱ بیش از همه و در سال ۲۰۱۱ کمتر از همه بوده است. به این شکل که از سال ۲۰۱۱ روند تدریجی عدم پاسخ شروع شده است. مقایسه میزان پاسخ سفید در تحلیل داده‌های مطالعه پرلز ۲۰۲۱ نیز نشان داد که سؤالات تشریحی در مقابل سؤالات چندگزینه‌ای، متون دشوار در مقابل متون ساده و فرایندهای پیچیده‌تر خواندن در مقابل فرایندهای ساده‌تر تفاوت‌هایی از لحاظ بروز پاسخ سفید میان دانش‌آموزان را پدید می‌آورد.

در تحلیل استنباطی که درصد پاسخ‌های سفید به عنوان متغیر ملاک و ویژگی‌های سؤال شامل درجه دشواری سؤالات، نوع پاسخگویی به سؤالات، نوع متن از لحاظ هدف، درجه دشواری متن و فرایندهای شناختی سؤالات به عنوان متغیرهای پیش‌بین در نظر گرفته شدند، نتایج نشان داد که نوع پاسخ‌گویی (تشریحی) و دشواری سؤالات پیش‌بین‌های عدم پاسخ دانش‌آموزان به شمار می‌روند.

در این مطالعه عدم پاسخ دانش‌آموزان به سؤالات از منظر ویژگی‌های سؤالات پرداخته شده است. به عبارت دیگر، این سؤال مطرح بوده است که دانش‌آموزان ایرانی به چه نوع سؤالاتی با احتمال بیشتری پاسخ نخواهند داد. تحلیل توصیفی درون کشوری و تطبیقی به همراه تحلیل رگرسیونی نشان داد که پاسخ سفید دانش‌آموزان عموماً در سؤالات تشریحی، دشوار یا پیچیده از لحاظ فرایندهای خواندن اتفاق می‌افتد. هر سه ویژگی بیان‌شده پیچیدگی‌های شناختی و درگیری ذهنی بیشتری را از دانش‌آموزان طلب می‌کند. به عبارت ساده‌تر، می‌توان این گونه بیان کرد که دانش‌آموزان ایرانی با وجود اینکه میزان عدم پاسخ بیشتری در مقایسه با دانش‌آموزان سایر کشورهای شرکت‌کننده در مطالعه پرلز دارند، ولی در سؤالاتی که بار ذهنی بیشتری دارند، احتمال سفید گذاشتن سؤالات در آن‌ها بیشتر می‌شود.

شواهد دیگری نیز وجود دارد که نشان می‌دهد دانش‌آموزان سؤالات را بیشتر به دلیل عدم توانایی شناختی بدون پاسخ می‌گذارند تا به دلیل عدم انگیزه. در تحلیل میزان همبستگی دوره‌های نقطه‌ای بین پاسخ سفید به سؤال و توانایی دانش‌آموزان در مطالعه پرلز ۲۰۲۱ مشخص شد که عموم همبستگی‌های منفی بود. میانگین همبستگی در همه سؤالات برابر با ۰/۱۵- بود. این موضوع نشان می‌داد افرادی که سؤال‌های مطالعه را سفید باقی می‌گذاشتند توانایی درک مطلب ضعیفی داشتند.

با اینکه در بافت آموزش ابتدایی ایران کاربرد سنجش توصیفی مورد تأکید قرار گرفته است ولی ممکن است در عمل طیف وسیع آزمون‌ها به خصوص آزمون‌های عملکردی که بر تکالیف نوشتاری متمرکز هستند کمتر به کار برده شود. در سنجش درک مطلب نیز احتمالاً همین شرایط وجود دارد. در نتیجه احتمالاً درخواست کمتری از دانش‌آموزان برای نوشتن و استفاده از سؤالات تشریحی به خصوص سؤالات تشریحی بلندپاسخ وجود دارد.

چنانچه پیچیدگی و تشریحی بودن سؤال را به عنوان عامل‌های اصلی عدم پاسخ دانش‌آموزان در نظر داشته باشیم، نیاز است که در مدارس روی تغییر شکل آموزش و سنجش دانش‌آموزان توجه ویژه‌ای داشته باشیم. وادار کردن دانش‌آموزان به نوشتن، فاصله گرفتن از آزمون‌های انتخابی، آموزش راهبردهای مواجهه با سؤالات تشریحی (مثل بازنویسی سؤال، طرح‌ریزی پاسخ، مدیریت زمان) و همچنین در نظر گرفتن تدریجی سؤالات پیچیده‌تر از لحاظ بار شناختی به خصوص در سنجش‌های خواندن می‌تواند از جمله این تمهیدات باشد.

محدودیت‌ها

اگرچه این مطالعه بینشی جدید در مورد تبیین پاسخ‌های سفید دانش‌آموزان در اختیار قرار می‌دهد که پیش از این در ادبیات پژوهشی و سنجش داخل کشور برجسته نشده بود، ولی هنوز به دلیل محدودیت‌های روشی قابل بررسی مجدد است. از جمله استفاده از روش‌های همبستگی که قابلیت تبیین‌های علی و معلولی را به میزان روش‌های آزمایشی سلب می‌کند. مضاف بر آن، هر چند در مورد انگیزه دانش‌آموزان تفسیری ارائه کرد ولی به دلیل اینکه از تحلیل ثانویه استفاده شده بود، برآوردی از انگیزه دانش‌آموزان وجود نداشت. همچنین، داده‌های مورد بررسی مربوط به مطالعه پرلز ۲۰۲۱ بودند که هم از لحاظ سنجش موضوع مطالعه و سال مورد بررسی دچار محدودیت‌هایی است که ظرفیت تعمیم‌پذیری نتایج را محدود می‌سازد. در پایان تحلیل روی سؤالات استوار شده بود. چنانچه مطالعه با مدل‌سازی

پاسخ‌های سفید روی دانش‌آموزان و ویژگی‌های آنان انجام شود، می‌تواند زوایای جدیدی را از این موضوع روشن کند. اینکه چه گروهی از دانش‌آموزان و با چه ویژگی‌هایی بیشتر مرتکب پاسخ سفید می‌شوند.

مشارکت نویسندگان

تمام مراحل انجام پژوهش شامل طرح‌ریزی، گردآوری داده‌ها و تالیف مقاله توسط نویسنده انجام شده است.

تشکر و قدردانی

نویسنده از تمامی داوران صمیمانه سپاسگزاری می‌کند.

تعارض منافع

هیچ‌گونه تعارض منافع مالی، علمی یا شخصی در ارتباط با نگارش و انتشار این مقاله وجود ندارد.



COPYRIGHTS
©2021 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

References

- Childs, R. A., Broomes, O., & Herbert, M. B. (2018). Deciding Whether to Respond: An Analysis of Nonresponse on Ontario's Grade 9 Assessment of Mathematics, *Alberta Journal of Educational Research*, 64.1, 70-87. <https://doi.org/10.55016/ojs/ajer.v64i1.56472>
- Di Chiacchio, C., De Stasio, S., & Fiorilli, C. (2016). Examining how motivation toward science contributes to omitting behaviours in the Italian PISA 2006 sample, *Learning and Individual Differences*, 50, 56-63. Doi: <http://dx.doi.org/10.1016/j.lindif.2016.06.025>
- Kowolik, K., Cockle, M., & Tavena, M. (2020). Creating the TIMSS 2019 international database. In M. O. Martin, M. von Davier, & I. V. S. Mullis (Eds.), *Methods and Procedures: TIMSS 2019 Technical Report* (pp. 8.1-8.17). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/methods/chapter-8.html>
- Marcq, K., & Braeken, J. (2024). Where nonresponse is at its loudest: Cross-country and individual differences in item nonresponse across the PISA 2018 student questionnaire, *International Journal of Testing*, Doi: <https://doi.org/10.1080/15305058.2024.2377550>
- Matters, G. & Burnett, P.C. (2003). Psychological predictors of the propensity to omit short-response items on a high-states achievement test, *Educational and Psychological Measurement*, 63(2), 239-256. Doi: 10.1177/0013164402250988
- Pohl1, S., Ulitzsch, E., & von Davier, M. (2021). Reframing rankings in educational assessments, *Science*, 372(6540), 338-342. Doi: 10.1126/science.abd3300
- Robitzsch, A. (2023). Nonignorable Consequences of (Partially) Ignoring Missing Item Responses: Students Omit (Constructed Response) Items Due to a Lack of Knowledge, *Knowledge*, 3, 215-231. Doi: <https://doi.org/10.3390/knowledge3020015>
- Lu, J. & Wang, C. (2020). Response Time Process Model for Not-Reached and Omitted. *Journal of Educational Measurement*, Doi: 10.1111/jedm.12270
- von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 International Results in*

Mathematics and Science. Boston College, TIMSS & PIRLS International Study Center.
<https://doi.org/10.6017/lse.tpisc.timss.rs6460>
von Schrader, S., & Ansley, T. (2006). Sex Differences in the Tendency to Omit Items on Multiple-Choice Tests: 1980–2000, *Applied Measurement in Education*, 19(1), 41–65.

