

## برنامه بین‌المللی سنجش دانش‌آموزان (پیزا)

تاریخ دریافت: ۱۳۹۳/۶/۱۵

تاریخ پذیرش و انتشار: ۱۳۹۷/۱۰/۱۵

مسعود کبیری<sup>۱</sup>

### تاریخچه و اهداف

شکل‌گیری «برنامه بین‌المللی سنجش دانش‌آموزان» (پیزا)<sup>۲</sup> از انتظارات عمیق‌تر نسبت به آموزش و پرورش در قرن بیست و یکم نشأت گرفته است. این رویکرد بر مبنای تحولات حاضر در اقتصادهای پیشرفته است که در آن توجه به افراد بر این اساس صورت می‌پذیرد که با آموخته‌های خود چه کاری می‌توانند انجام دهند، نه اینکه چه چیزی می‌دانند (سازمان توسعه و همکاری اقتصادی<sup>۳</sup>، ۲۰۱۴ ب). با وجود تغییر رویکرد به بازده‌های آموزشی، حساسیت سیاست‌گذاران آموزشی به این مسئله جلب شده است که آیا نظام‌های آموزشی به گونه‌ای عمل کرده‌اند که دانش‌آموزان آنچه را که یاد گرفته‌اند در محیط‌های ناآشنا، چه در مدرسه و چه در بیرون از آن، به کار گیرند. تغییر نگاه به برونداد آموزش از مدل پویای یادگیری در جریان زندگی پیروی می‌کند که در آن فرض می‌شود دانش و مهارت‌ها باید جهت انطباق مناسب با شرایط در حال تغییر شخصی، اقتصادی و اجتماعی در طول زندگی کسب گردند (پسچار<sup>۴</sup>، ۲۰۰۴).

با وجود دغدغه‌های جدید برای سیاست‌گذاران آموزشی اندیشه ایجاد سازوکار مشخصی برای سنجش دانش و مهارت‌های لازم برای مشارکت کامل در جوامع پیشرفته از سال ۱۹۹۴ تحت عنوان مطالعه پیزا به وجود آمد (ریتزن<sup>۵</sup>، ۲۰۱۳)، ولی اولین اجرای مطالعه در سال ۲۰۰۰ انجام شد. کنسرسیومی از شورای پژوهش آموزشی استرالیا<sup>۶</sup>، مؤسسه ملی اندازه‌گیری آموزشی<sup>۷</sup> در هلند، دو مؤسسه خدمات آزمون آموزشی<sup>۸</sup> و وی‌استات<sup>۹</sup> در آمریکا و مؤسسه ملی پژوهش آموزشی ژاپن<sup>۱۰</sup> طراحی و اجرای مطالعه پیزا را برعهده گرفتند (لیتون<sup>۱۱</sup>، ۲۰۱۳). پیزا برنامه مستمری است که در زمینه سیاست‌ها و اقدامات آموزش و پرورش شواهد تجربی تولید می‌کند. این یافته‌ها به سیاست‌گذاران سرتاسر دنیا مجال می‌دهد که دانش و مهارت‌های دانش‌آموزان کشور خود را در مقایسه

<sup>۱</sup> عضو هیئت علمی پژوهشگاه مطالعات آموزش و پرورش وابسته به سازمان پژوهش و برنامه ریزی آموزشی

<sup>۲</sup> The Programme for International Student Assessment (PISA)

<sup>۳</sup> Organisation for Economic Co-operation and Development

<sup>۴</sup> Peschar

<sup>۵</sup> Ritzen

<sup>۶</sup> Australian Council for Educational Research (ACER)

<sup>۷</sup> National Institute for Educational Measurement (CITO)

<sup>۸</sup> Educational Testing Service (ETS)

<sup>۹</sup> Westat

<sup>۱۰</sup> National Institute for Educational Research (NIER)

<sup>۱۱</sup> Leighton

با سایر کشورها اندازه‌گیری کرده تا بتوانند سیاست‌های آموزشی خود را از طریق مقایسه با سیاست‌های ثمربخش در سایر نظام‌های آموزشی تعیین کرده و از آنان درس بیاموزند.

هرچند برنامه پیزا نمی‌تواند روابط علت و معلولی را بین سیاست‌ها یا اقدامات و بازده دانش‌آموزان برقرار سازد، ولی می‌تواند به آموزشگران، سیاست‌گذاران و افراد علاقه‌مند در دنیا نشان دهد که نظام‌های آموزشی از چه جنبه‌هایی مشابه یا متفاوت هستند. این کارکرد باعث می‌شود که قبل از اینکه سیاست یا اقدامی به کار گرفته شود از طریق وضعیت تأثیر آن در کشورهای دیگر پیش‌بینی دقیق‌تری از میزان تأثیرگذاری اقدام خود داشته باشند. به‌طور کلی سنجش پیزا سه نوع خروجی را فراهم می‌کند:

- نشانگرهایی برای نشان دادن نیم‌رخ خط پایه در زمینه دانش و مهارت‌های دانش‌آموزان هر کشور،
- نشانگرهای نشان‌دهنده ارتباط دانش و مهارت‌ها با عوامل جمعیت‌شناختی، اجتماعی، اقتصادی و آموزشی،
- نشانگرهایی در زمینه روندهای تغییرات در متغیرهای سطح دانش‌آموز و مدرسه (سازمان توسعه و همکاری اقتصادی، ۲۰۱۴ الف).

### ابعاد سنجش و گستره فعالیت

سؤال کلیدی که طراحی سنجش پیزا حول آن شکل گرفته این است که دانش‌آموزان در پایان تحصیلات عمومی، به کدام یک از مهارت‌ها برای انجام وظایف خود به عنوان شهروند نیاز دارند؟ برای پاسخ به این سؤال نیاز بود که از مرزهای معمول برنامه‌های درسی فراتر رفت و گستره کلی‌تری از آنچه که دانش‌آموزان می‌دانند و می‌توانند انجام دهند را مورد توجه قرار داد (اوون، استفنز، موسکوویتز و جیل، ۲۰۰۴). به این مهارت‌ها عناوینی چون مهارت‌های زندگی، دستاوردهای محدود نشده به برنامه درسی یا شایستگی‌های بین برنامه‌ای نیز اطلاق می‌شود (کلیمه<sup>۲</sup>، ۲۰۰۴).

چنین مفهوم‌سازی به‌طور مستقیم در طراحی مطالعه منعکس شد؛ به‌طوری که حوزه‌های مورد سنجش بر اساس دانش و مهارت‌های لازم برای زندگی بزرگسالی، نه برای تسلط بر برنامه، سنجیده می‌شود. بدین منظور از اصطلاح شایستگی<sup>۳</sup> به جای توانایی<sup>۴</sup> استفاده می‌شود تا مهارت‌های کلی و مستقل از زمینه در موقعیت‌های زندگی واقعی مورد توجه قرار گیرند. این شایستگی‌ها به فرایندهای غیرروتین و سطح بالای شناختی متکی است که مستلزم حل مسئله است (کلیمه، ۲۰۰۴). بنابراین، حوزه‌های سنجش مطالعه پیزا حول ایده‌های وسیع‌تری شکل گرفته‌اند.

<sup>۱</sup> Owen, Stephens, Moskowitz & Gil,

<sup>۲</sup> Klieme

<sup>۳</sup> competency

<sup>۴</sup> ability

به عنوان مثال به جای اینکه حیطه محتوایی هندسه و جبر وجود داشته باشد، حوزه فضا و شکل در نظر گرفته شده است. البته این نگاه مورد نقد نیز قرار گرفته است. به طوری که برخی امکان سنجش موارد خارج از کلاس را مطرح کرده‌اند. دشواری انجام تغییرات بر اثر نتایج نقد دیگر وارد شده است زیرا محتوای سنجش ارتباط کمی با برنامه‌های درسی دارد (اوون و همکاران، ۲۰۰۴). همچنین، این شیوه طراحی سنجش به زعم عده‌ای نشانه رد آشکار تأثیر برنامه درسی قلمداد شده است (ویجمیکرا، ۲۰۱۴).

با اینکه در طراحی سنجش پیزا مفاهیم کلی در نظر گرفته شده بودند ولی تفکیک موضوعات اصالت خود را حفظ کرده است. از سال ۲۰۰۰ که اولین دور برنامه پیزا اجرا شد، سه موضوع ریاضی، خواندن و علوم به عنوان ماده‌های اصلی مدنظر بودند، با این حال، به جای «ماده‌های درسی»، «سواد»های مختلف مورد آزمون قرار گرفتند. سواد به عنوان بلکه مهارتی در نظر گرفته می‌شود که در طول زندگی کسب یا استفاده می‌شود. علاوه بر این، حوزه‌های دیگری به صورت بین‌برنامه‌ای مورد آزمون قرار گرفتند که از جمله آن می‌توان به حل مسئله، سواد فن‌آوری و اخیراً سواد مالی اشاره کرد (سازمان توسعه و همکاری اقتصادی، ۲۰۱۳ ب). پوشش همه مواد آزمون فوق در یک دور سنجش بسیار دشوار است، به همین منظور در طراحی مطالعه به گونه‌ای عمل شد که از سه موضوع مورد سنجش یک حوزه به عنوان موضوع اصلی و بقیه حوزه‌ها به عنوان موضوعات فرعی در نظر گرفته شوند. با توجه به چرخه‌های تکرار سه ساله، در هر دوره درس اصلی از بین سه موضوع کلی تغییر می‌کند. در اولین دور مطالعه در سال ۲۰۰۰، خواندن به عنوان درس اصلی در نظر گرفته شد. هر یک از حوزه‌های مورد سنجش عبارت بودند از:

*سواد خواندن*: سواد خواندن به عنوان درک، استفاده و تأمل<sup>۲</sup> بر متون نوشتاری در رابطه با رسیدن به اهداف، توسعه دانش و قابلیت‌ها و همچنین مشارکت در جامعه تعریف شده است (سازمان توسعه و همکاری اقتصادی، ۲۰۰۶، ۲۰۰۹). علاوه بر درک مطلب و استفاده از آن، تمرکز بر ظرفیت فرد در تأمل است که بخش فعالی در معنی‌سازی دارد. خواننده بین متن، خواننده و بافت تعامل دارد. برای رسیدن به اهداف خواندن، راهبردها و برای معنادگی و پاسخ به متن، دانش قبلی را به کار می‌گیرد (لافونتاین<sup>۳</sup>، ۲۰۰۴). در چارچوب سنجش خواندن پیزا، بر خلاف مطالعات قبل همچون پرلز، به جای فرایندها، جنبه‌های<sup>۵</sup> خواندن مفهوم‌پردازی شدند که شامل بازیابی اطلاعات، شکل دادن به درک کلی گسترده، توسعه تفسیر، تأمل و ارزیابی محتوای متن و همچنین تأمل و ارزیابی شکل متن می‌گردد (سازمان توسعه و همکاری اقتصادی، ۲۰۱۴ الف).

<sup>۱</sup> Wagemaker

<sup>۲</sup> reflection

<sup>۳</sup> Lafontaine

<sup>۴</sup> processes

<sup>۵</sup> aspects

سواد ریاضی: سواد ریاضی شایستگی افراد در اندازه‌گیری ظرفیت افراد برای چارچوب‌بندی کردن، به‌کارگیری و تفسیر ریاضی در گستره‌ای از بافت‌ها است که استدلال ریاضی و استفاده از مفاهیم، رویه‌ها، حقایق و ابزارهای ریاضی برای توصیف، تبیین و پیش‌بینی پدیده‌ها را دربر می‌گیرد. برای سنجش سواد ریاضی چالش یا مسئله‌ای در بافت دنیای واقعی در نظر گرفته می‌شود که شامل چهار مقوله شخصی، اجتماعی، حرفه‌ای و علمی است (سازمان توسعه و همکاری اقتصادی، ۲۰۱۳ الف). علاوه بر این، چهار طبقه محتوایی کمیّت، ابهام و داده‌ها، شانس و روابط، و فضا و شکل در نظر گرفته شده است. در طبقات فرایندی سه فرایند چارچوب‌بندی کردن موقعیت‌ها به شکل ریاضی؛ به‌کار بردن مفاهیم، حقایق، رویه‌ها و استدلال ریاضی؛ و همچنین تفسیر، به‌کارگیری و ارزیابی دستاوردهای ریاضی در ظرفیت‌های واقعی ریاضی برآورده می‌شود. علاوه بر دسته‌بندی‌های فوق هفت ظرفیت بنیادی ریاضی نیز در سنجش ریاضی پیزا مورد توجه قرار گرفته است که شامل ارتباط؛ ریاضی‌سازی؛ نمایش؛ استدلال و مباحثه؛ راهبردهای حل مسئله؛ استفاده از زبان و عملیات نمادین، رسمی و فنی؛ و استفاده از ابزارهای ریاضی است (سازمان توسعه و همکاری اقتصادی، ۲۰۱۴ الف).

سواد علوم: سواد علمی شامل سه مؤلفه تعیین موضوعات علمی، تبیین پدیده‌ها به‌طور علمی و استفاده از شواهد علمی است. دو نوع دانش لازم شمرده می‌شود که شامل دانش علوم و دانش در مورد علوم است. در دانش علوم، مباحث لازم برای درک دنیای طبیعی و معنی‌دارسازی تجربیات در موقعیت‌های شخصی، اجتماعی و جهانی مورد توجه است. دانش‌آموزان باید درک خود را از مفاهیم محتوایی در بافت مرتبط با زندگی به‌کار گیرند. در دانش در مورد علوم دو موضوع برجسته شده است؛ اول، کاوشگری علمی که به منشأ کاوشگری، هدف، آزمایش‌ها، نوع داده‌ها و ویژگی‌های نتایج می‌پردازد و دوم تبیین‌های علمی شامل انواع تبیین‌ها، شکل‌دهی، قوانین و پیامدها است (سازمان همکاری و توسعه اقتصادی، ۲۰۰۹). علاوه بر این، در نگرش نسبت به علوم به عنوان یک مؤلفه جداگانه، مقصود توسعه نگرش‌ها در جریان کسب موضوعات علمی و به‌کارگیری دانش علمی و فن‌آوری برای فواید شخصی، اجتماعی و جهانی است. نگرش دانش‌آموزان نیز در سه حوزه علاقه به علوم، حمایت از کاوشگری علمی و مسئولیت نسبت به منابع و محیط بررسی می‌شود (سازمان همکاری و توسعه اقتصادی، ۲۰۰۶).

حل مسئله: ظرفیت افراد برای استفاده از فرایندهای شناختی برای مواجهه شدن و حل موقعیت‌های بین رشته‌ای است که به‌طور فوری قابل حل نباشد (سازمان توسعه و همکاری اقتصادی، ۲۰۱۳ ج). حوزه حل مسئله برای بررسی شایستگی‌های بین برنامه‌ای مورد توجه بوده است و لذا اندازه‌گیری آن‌ها در حین بررسی زیرمؤلفه‌های سنجش رخ داده است. برای مثال، دو خوشه بازپدیدآوری و ارتباطات تکالیف غیرمسئله‌ای و خوشه تأملات تکالیف مسئله‌ای را در سنجش ریاضی اندازه‌گیری می‌کرد که خوشه آخر محملی برای اندازه‌گیری حل مسئله را

فراهم می‌کند. همین موضوع در تمایز فرایندهای درک کاوشگری علمی و تفسیر شواهد و نتیجه‌گیری علمی در سنجش علوم موضوعیت دارد (کلیمه، ۲۰۰۴).

به دلیل آنکه مهارت‌های زندگی روزمره در پیزا سنجیده می‌شود، جای شگفتی نیست که سنجش به صورت رایانه‌ای انجام شود. چنین تمهیدی نه تنها نه تنها برای اندازه‌گیری مهارت‌های سواد رایانه‌ای، بلکه برای اجازه به محدوده وسیع‌تری از تکالیف پویا و تعاملی و به‌کارگیری روش‌های مؤثرتر برای انجام آزمون‌های دانشی و مهارتی اتخاذ شده است (بلر<sup>۱</sup>، ۲۰۱۳). سوالات از طریق به‌کارگیری طرح سنجش اصلی / فرعی به دلیل کاهش تعداد سوالات در دروس فرعی ارائه می‌شوند. در دروسی که به عنوان اصلی انتخاب شده‌اند، حدود ۱۲۰ سؤال در ۷ بلوک قرار گرفته‌اند (ویکس، ون‌داویر و یاماموتو<sup>۲</sup>، ۲۰۱۴). سوالات به سه شکل کلی چند گزینه‌ای، کوتاه پاسخ و سوالات باز پاسخ تنظیم شده‌اند.

هم‌گام با سایر ویژگی‌های که با منطق اصلی سنجش مطابقت دارد، طرح نمونه‌گیری سن محور به‌کار گرفته شده است تا آزمون مهارت-محوری (و نه برنامه درسی محور) طراحی گردد (ویجمیکر، ۲۰۱۴). بدین صورت که برای مقایسه بهتر عملکرد دانش‌آموزان از لحاظ بین‌المللی، سن خاصی برای دانش‌آموزان هدف‌گذاری شده است. دانش‌آموزان شرکت‌کننده در این مطالعه باید در هنگام سنجش بین ۱۵ سال و سه ماه تا ۱۶ سال و دو ماه داشته و حداقل شش سال از تحصیلات رسمی را تکمیل کرده باشند. این دانش‌آموزان می‌توانند در هر نوع مؤسسه یا مدرسه‌ای ثبت‌نام کرده باشند. با استفاده از این سن امکان مقایسه متجانس بین دانش و مهارت‌های افراد متولد شده در یک سن (۱۵ ساله‌ها)، علی‌رغم گستردگی آموزش و پرورش در بین کشورها و در زمان‌های مختلف به‌وجود می‌آید. این انتخاب به دلیل تفاوت بین کشورها در ماهیت و میزان آموزش، فعالیت‌های پیش از دبستان در سن ورود به تحصیلات رسمی، ساختار نظام مدرسه و وضعیت تکرار پایه است.

امروزه برنامه پیزا به عنوان یک سنجش تطبیقی در بسیاری از مناطق دنیا مورد پذیرش قرار گرفته است، به طوری که در سال ۲۰۱۸، ۷۹ کشور و اقتصاد در این برنامه شرکت کرده‌اند. از کشورهای شرکت‌کننده منطقه خاورمیانه در این دوره می‌توان به گرجستان، قزاقستان، اردن، قطر، لبنان، عربستان سعودی، ترکیه، امارات متحده عربی، و آذربایجان (شهر باکو) اشاره نمود. از آنجایی که طرح و ترجمه آزمون، به علاوه نمونه‌گیری و گردآوری داده‌ها با کنترل کیفیت جدی همراه است، یافته‌های برنامه پیزا با روایی و اعتبار بالایی عجین شده است. به ویژه آنکه تلاش و منابع قابل ملاحظه‌ای برای رسیدن به توازن فرهنگی و زبانی در مواد سنجشی انجام می‌شود.

<sup>۱</sup> Beller

<sup>۲</sup> Weeks, von Davier, & Yamamoto

## دستاورد‌های آزمون

اهمیت برنامه پیزا تا اندازه‌ای به دلیل نتایج متنوع و چالش‌برانگیز آن مربوط بوده است. آخرین نتایج منتشر شده مربوط به پیزا ۲۰۱۵ نشان داد که سنگاپور بالاترین نمره را در علوم کسب نموده است. پس از آن ژاپن، استونی، چین تایپه، فنلاند، ماکائو (چین)، کانادا، ویتنام، هنگ‌کنگ، و چهار استان از کشور چین (پکن، شانگهای، جیانگسو، و گوانگ‌دونگ) ده کشور بالا از لحاظ عملکرد در علوم به شمار می‌آیند (سازمان توسعه و همکاری اقتصادی، ۲۰۱۶). همچنین نشان داده شد که بیشتر کشورها و اقتصادها با داده‌های قابل مقایسه از ۲۰۰۳ تا ۲۰۱۵ به سوی مدارس مجهزتر و با کادر آموزشی بهتر حرکت کرده‌اند.

علاوه بر این برابری امکانات و آموزش سهم زیادی در توفیق کشورهای برتر در سنجش پیزا دارد: کشورهای با عملکرد بالا تمایل دارند منابع خود را به‌طور مساوی در بین مدارس تقسیم کنند. با این وجود، نتایج پیزا نشان داد که در بسیاری از نظام‌های آموزشی منابع به‌طور مساوی اختصاص پیدا نمی‌کنند: مدارس نابرخوردار کلاس‌های کمتری دارند، از ضعف معلمان، یا نامناسب بودن مواد آموزشی یا زیرساخت‌های فیزیکی رنج می‌برند. در همین راستا، در کشورهایی که بیشترین بهبود عملکرد را داشتند (شیلی، پرو، آلبانی و اندونزی)، رابطه ضعیفی بین وضعیت اقتصادی-اجتماعی و بازده یادگیری به‌دست آمد که نشان‌دهنده عملکرد مثبت نظام آموزشی در کاهش نابرابری آموزشی است (ویجمیکر، ۲۰۱۴).

از جمله نتایج دیگر پیزا آن بود که اگر امکان انتخاب مدارس برای والدین وجود داشته باشد، والدین ملاک‌هایی شبیه «محیط مدرسه امن» و «شهرت خوب مدرسه» را از «پیشرفت تحصیلی مناسب دانش‌آموزان در مدرسه» مهم‌تر می‌دانستند. علاوه بر نتایج مؤید انتظارت قبلی، به چالش کشیدن برخی از پنداشت‌های قبلی از جمله ویژگی‌های آزمون‌های بین‌المللی است. به عنوان مثال، یافته‌های پیزا نشان داد که یادگیری مشارکتی لزوماً بر یادگیری رقابتی برتری ندارد (ویجمیکر، ۲۰۱۴).

## تأثیرات آزمون بر برنامه درسی

پیامدهای پیزا بر نظام‌های آموزشی قابل توجه بوده است که در چند مقوله قابل دسته‌بندی است. این پیامدها شامل تغییر در گفتمان آموزش به سوی پژوهش‌های تجربی و توجه جدی به پایش استانداردهای آموزشی مرتبط با کیفیت و پاسخگویی در آلمان، معرفی سنجش‌های ملی در زیست‌شناسی، شیمی و فیزیک در سوئد (بریکسپیر، ۲۰۱۲)، توجه اصلاحات آموزشی مخصوصاً در زمینه بهبود عملکرد معلم و انجام ارزشیابی در مکزیک، توجه افزایش روزهای آموزشی برای مدارس و میزان زمان آموزش در ژاپن و توقف برخی از سیاست‌های قبلی در سوئیس (ویجمیکر، ۲۰۱۴) در تغییر بود.

به دلیل اینکه پیزا بر سطح/پایه خاصی از برنامه درسی تأکید نمی‌کند، بسیار دشوار است اثر مستقیم آن را بر برنامه درسی تعیین کرد. با این وجود، بیشترین تأثیر پیزا بر برنامه درسی به ارتقای استانداردهای برنامه درسی برمی‌گردد. به عنوان مثال در اسلواکی شایستگی‌های کلیدی موجود در برنامه پیزا که در برنامه‌های درسی قبلی این کشور وجود نداشت، به استانداردهای برنامه درسی تلفیق شد (بریکسپیر<sup>۱</sup>، ۲۰۱۲) و در ایرلند تغییراتی در برنامه درسی آموزش متوسطه اول رخ داد (ویجمیکر، ۲۰۱۴). در قرقیزستان توسعه استانداردهای جدید و برنامه درسی، کاهش بار آموزشی معلم، مدرن‌سازی زیرساخت‌ها و تجهیزات مدرسه و بهبود استانداردهای تدریس و عملکرد را می‌توان از جمله پیامدهای پیزا برشمرد. در نهایت تغییرات در توازن و تسلسل برنامه‌های درسی ملی را می‌توان از جمله تأثیرات مستقیم پیزا بر برنامه درسی کشورهای شرکت کننده قلمداد نمود (ویجمیکر، ۲۰۱۴).

## منابع

- Beller, M. (۲۰۱۳). Technologies in large-scale assessments: New directions, challenges, and opportunities. In M. v. Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), *The role of international large-scale assessments: perspectives from technology, economy, and educational research* (pp. ۲۵-۴۶). New York: Springer.
- Breakspear, S. (۲۰۱۲). *The Policy Impact of PISA: An exploration of normative effects of international benchmarking in school system performance*. Paris: OECD Publishing.
- Klieme, E. (۲۰۰۴). Assessment of cross-curricular problem-solving competencies. In J. H. Moskowitz & M. Stephens (Eds.), *Comparing learning outcomes: International assessments and education policy* (pp. ۸۱-۱۰۷). London: RoutledgeFalmer.
- Lafontaine, D. (۲۰۰۴). From comprehension to literacy: Thirty years of reading assessment. In J. H. Moskowitz & M. Stephens (Eds.), *Comparing learning outcomes: International assessments and education policy* (pp. ۲۴-۴۵). London: RoutledgeFalmer.
- Leighton, J. P. (۲۰۱۳). Large-scale assessment design and development for the measurement of student cognition. In M. Simon, K. Ercikan, & M. Rousseau (Eds.), *Improving large-scale assessment in education: Theory, issues, and practice* (pp. ۱۳-۲۶). New York: Taylor & Francis.
- OECD. (۲۰۰۶). *Assessing scientific, reading and mathematical literacy: A framework for PISA ۲۰۰۶*. Paris: OECD.
- OECD. (۲۰۰۹). *PISA ۲۰۰۹ assessment framework: Key competencies in reading, mathematics and science*. Paris: OECD.
- OECD, P. (۲۰۱۳a). *PISA ۲۰۱۲ results: Ready to learn students' engagement, drive and self-beliefs* (Vol. III). Paris: OECD Publishing.
- OECD, P. (۲۰۱۳b). *PISA ۲۰۱۲ results: What makes schools successful? Resources, policies and practices* (Vol. IV). Paris: OECD Publishing.
- OECD. (۲۰۱۳c). *PISA ۲۰۱۲ assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris: OECD Publishing.
- OECD. (۲۰۱۴). *PISA ۲۰۱۲ results: What students know and can do student performance in mathematics, reading and science* (Vol. I revised edition, ۲۰۱۴). Paris: OECD Publishing.

---

<sup>۱</sup> Breakspear

- OECD, P. (٢٠١٤). *Results: Creative problem solving—Students' skills in tackling real-life problems* (Vol. V). Paris: OECD Publishing.
- OECD (٢٠١٦), PISA ٢٠١٥ Results (Volume I): *Excellence and Equity in Education*, Paris: OECD Publishing.
- Owen, E. H., Stephens, M., Moskowitz, J. H., & Gil, G. (٢٠٠٤). Comparing learning outcomes: International assessments and education policy. In J. H. Moskowitz & M. Stephens (Eds.), *Comparing learning outcomes: International assessments and education policy* (pp. ٣-٢٣). London: RoutledgeFalmer.
- Peschar, J. L. (٢٠٠٤). Cross-curricular competencies: developments in a new area of education outcome indicators. In J. H. Moskowitz & M. Stephens (Eds.), *Comparing learning outcomes: International assessments and education policy* (pp. ٥٩-٨٠). London: RoutledgeFalmer.
- Ritzen, J. (٢٠١٣). International large-scale assessments as change agents. In M. v. Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), *The role of international large-scale assessments: perspectives from technology, economy, and educational research* (pp. ١٣-٢٤). New York: Springer.
- Wagemaker, H. (٢٠١٤). International Large-Scale Assessments: From Research to Policy. In L. Rutkowski, M. v. Davier & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: background, technical issues, and methods of data analysis* (pp. ١١-٣٦). Boca Raton, FL: CRC Press.
- Weeks, J. P., Davier, M. v., & Yamamoto, K. (٢٠١٤). Design considerations for the program for international student assessment. In L. Rutkowski, M. v. Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: background, technical issues, and methods of data analysis* (pp. ١١-٣٦). (pp. ٢٥٩-٢٧٦). Boca Raton, FL :CRC Press.