

Teachers' Understanding and Use Of Authentic Assessment In the Teaching-Learning Process

Hana Ahmadi¹, Naser Shirbagi*², Sanam Shirbagi³

پندیرش مقاله: ۱۴۰۲/۱۱/۲۷

دریافت مقاله: ۱۴۰۲/۰۵/۲۰

Accepted Date: 2024/02/16

Received Date: 2023/08/11

Extended Abstract

Objectives and questions: In recent years, interest and support for a type of measurement that is less common has been aroused, and that is original measurement. The original measurement is different in its form and implementation, but it is not different from the common forms of measurement, density and formation measurement that are used today. Authentic assessment focuses and relies on the existence of connections between the real world outside the school and the school environment itself. Authentic assessment is an assessment task that has value and meaning for the student beyond the grade obtained. Authentic assessment is a concept popularized by among others that students engage in the use and application of skills and knowledge to address real-world problems or potential situations and give the assignment a sense of authenticity. With the concept of authentic measurement, its tools and types, there is a need to conduct research in this regard, considering the novelty of its concepts. The purpose of the current research is to develop and validate two questionnaires on the extent of teachers' use of authentic assessment and teachers' perception of authentic assessment. Therefore, the following research questions have been raised: 1) what is the level of perception of high school teachers about the concept of authentic assessment? 2) What is the rate of use of the principles of authentic assessment by high school teachers? 3) What difference do demographic variables such as gender, teaching experience, level of education, service area, and type of employment make in teachers' perception or use of the concept and principles of authentic assessment?

Research method: In this research, a quantitative research approach with a survey strategy was used. The statistical population consists of all teachers of the theoretical second secondary level (certain formal, formal experimental, contract and teaching

1. MA of Educational Administration. Department of Education, University of Kurdistan. Sanandaj. Iran.

2. Professor of Educational Administration, Department of Education, University of Kurdistan. Sanandaj. Iran.

*Corresponding Author: Email: nshirbagi@uok.ac.ir

3. MA of Education,, Faculty of Psychology and Educational Sciences, Shahid Behashi University, Tehran. Iran.

rights) working in the dual education and training areas of Sanandaj city in the academic year 1401-1400. The sample people in this research were selected by two-stage cluster sampling method, so that in the first stage, 20 schools were randomly selected as a cluster out of 83 secondary schools, and then they were randomly selected from this school according to the number of their teachers, in order to estimate the sample size of Cochran's formula was used, and the number of 200 subjects was determined. Descriptive statistics and inferential statistics methods were used to analyze the quantitative data collected by questionnaire. In the descriptive statistics section, the frequency and percentage as well as statistical indicators such as mean, standard deviation, minimum, median, maximum and skewness coefficient have been used, and in the inferential statistics section, the one-sample mean test using Student's t-distribution, comparison test The average of two independent groups was used using Student's t distribution and the comparison test of the average of several independent groups was used using one-way analysis of variance and Pearson's correlation coefficient test.

Data collection tools: To collect data from two questionnaires created by the researcher under the headings A- Questionnaire on the extent of teachers' use of authentic assessment and exploratory factor analysis using principal component analysis method with Varimax rotation, and the results of the above table show that 8 factors have special value There are more than one, the highest eigenvalue is equal to 14.62 and the lowest eigenvalue is 1.08. The first factor accounts for 28.66% of the variance and the eighth factor accounts for 2.12% of the variance, and eight factors together account for 68.56% of the variance of the questionnaire. B- The teachers' perception questionnaire was used in the classroom. Two methods of test-retest and Cronbach's alpha have been used for reliability. Exploratory factor analysis was performed using principal component analysis with varimax rotation. The results show that 7 factors have an eigenvalue greater than one, the highest eigenvalue is 11.22 and the lowest eigenvalue is 1.09. The first factor accounts for 28.78% of the variance and the seventh factor accounts for 1.092% of the variance, and eight factors together account for 70.75% of the variance of the questionnaire.

Findings: The level of perception of high school teachers about the concept of authentic assessment was more than average and with the findings of the researches that expressed the attitude and understanding of teachers about authentic assessment, it is significant. The amount of use of the selected sample teachers of authentic assessment principles was also more than average and it was in line with the findings of previous researches, he reached the conclusion in his study in Indonesia that the use of authentic assessment among teachers is more than average.

Discussion and conclusion: The obtained results indicated that the level of perception and use of authentic assessment of male and female teachers was equal, that is, gender had no effect on it. The results of this finding are also in line with the results of

previous findings which showed that gender has no effect on teachers' perception and use of authentic assessment. The level of perception and the amount of use of authentic assessment was also the same in the teachers of district one and two of Sanandaj city, which indicates that the district did not affect the level of perception and the amount of use of authentic assessment. The level of teachers' perception and use of authentic assessment was equal in teachers with associate, bachelor, master and doctorate degrees and teachers with different employment methods. That is, the educational qualification and type of employment of teachers did not affect the level of perception and the amount of use of authentic assessment. In the examination of the work experience of the teachers, the result is that there is no significant linear relationship between the work experience of the teachers and the level of perception and the amount of use of authentic assessment, that is, the amount of work experience of the teachers did not affect the level of perception and the amount of use of authentic assessment. The results of this finding are also in line with previous researches that stated that work history and experience have no effect on professors' attitude and use of authentic assessment.

Keywords: Authentic Assessment, Survey Method, Utilization, Teachers' Perception, Students

درک و میزان بهره‌گیری معلمان از «سنجش اصیل» در فرایند یاددهی- یادگیری

هانا احمدی^۱، ناصر شیربگی*^۲، صنم شیربگی^۳

چکیده

تدوین و اعتبارسنجی دو پرسشنامه محقق ساخته «میزان استفاده معلمان از سنجش اصیل» و «ادراک معلمان از سنجش اصیل» و ترویج مفهوم سنجش اصیل در بین دبیران دبیرستان‌های شهر سنندج هدف اصلی تحقیق حاضر بود. از این رو، محققان درصدد کشف نیمرخ وضعیت ادراک معلمان و میزان استفاده احتمالی آنان از اصول سنجش اصیل بودند. پژوهش از نوع کمی و توصیفی-پیمایشی بود. نمونه آماری متشکل از ۲۰۰ نفر از دبیران دبیرستان‌های شهر سنندج در سال تحصیلی ۱۴۰۰-۱۴۰۱ بودند که با روش نمونه‌گیری خوشه‌ای دومرحله‌ای انتخاب شدند. نتایج مطالعه نشان داد که میزان استفاده و ادراک معلمان از سنجش اصیل در دبیرستان‌های شهر سنندج بیش از حد متوسط بود. متغیرهای جنسیت، سابقه تدریس، میزان تحصیلات، ناحیه محل خدمت و نوع استخدام معلمان با میزان استفاده و ادراک آنان از سنجش اصیل همبستگی نداشت. معلمان فارغ از جنسیت، سابقه کار، ناحیه خدمت و نوع استخدام در محیطی مشابه تحصیل، کار و زندگی می‌کنند و در معرض تجربه‌های یکسان، آموزش‌ها و دوره‌های مشابه قرار گرفته بودند درک یکسانی از سنجش اصیل داشتند و میزان استفاده آنان نیز از سنجش اصیل در فرایند یاددهی-یادگیری مشابه بود. به نظر می‌رسد که نیاز به آموزش‌های بیشتر و برگزاری کارگاه‌ها و دوره‌های آموزشی و... برای آشنایی و توجیه متصدیان مربوطه، معلمان و همچنین دانش‌آموزان و والدین آنان با مفهوم سنجش اصیل وجود دارد.

کلیدواژه‌ها: سنجش اصیل، روش پیمایشی، میزان استفاده، ادراک معلمان، دانش‌آموزان

۱. کارشناسی ارشد مدیریت آموزشی، گروه علوم تربیتی، دانشکده علوم انسانی و اجتماعی، دانشگاه کردستان، سنندج، ایران.

۲. استاد مدیریت آموزشی، گروه علوم تربیتی، دانشکده علوم انسانی و اجتماعی، دانشگاه کردستان، سنندج، ایران.

Email: nshirbagi@uok.ac.ir

*نویسنده مسئول مکاتبات:

۳. کارشناسی ارشد مدیریت آموزشی، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه شهید بهشتی، تهران، ایران.

مقدمه

در دهه گذشته، بسیاری از کشورهای پیشرفته دنیا، توجه خود را به‌طور روزافزونی بر بهبود مدارس متمرکز کرده‌اند. این دولت‌ها به دلیل رقابت‌های موجود در بازار جهانی، از مدارس تقاضا دارند تا نیروی کاری را پرورش دهند که توانایی سازگاری با تغییرات سریع فناوری‌های تولید و اختراعات، کاربرد فناوری‌های پیچیده در تمام مشاغل، از امور بانکی گرفته تا تولید مواد غذایی، تعمیرات، خدمات پزشکی و ورزش را داشته باشند. آن‌ها هم‌چنین با وضع قوانین و اجرای خط‌مشی‌ها، خواهان موفقیت تحصیلی تمام یادگیرندگان هستند. به‌علاوه، برخی از این نظام‌ها معلمان و مسئولان مدارس را مسئول ناکامی تحصیلی یادگیرندگان می‌دانند تا جایی که ممکن است افرادی را که عملکردی ضعیف دارند، از کار بازدارند. (Sergiovanni & Starratt, 2010).

اقدامات اصلاح مدارس، بیشتر توجه خود را بر یادگیری از موضوع درسی معطوف کرده‌اند و اثبات این یادگیری از طریق عملکرد یادگیرندگان در آزمون‌ها بوده است. چنین تمرکزی باعث ایجاد تغییر در کار معلمان و سرپرستان شده است، به این طریق که از آن‌ها می‌خواهند تا به آنچه دانش‌آموزان در واقع یاد می‌گیرند، نحوه سنجش و نحوه بازسازی تدریس، زمانی که عملکرد یادگیرندگان ضعیف است توجه خاصی داشته باشند. دولت‌ها خواستار آن‌اند که یادگیرندگان به استانداردهای سطح بالای پیشرفت تحصیلی دست یابند.

یادگیری و سنجش «دوروی یک سکه‌اند» (Gulikers et al., 2004) و سنجش مستمر یادگیری در کلاس درس یکی از جنبه‌های اساسی تدریس اثربخش است (MacTighe & Ferraras, 2021). رویه‌های نوین جهانی در اصلاحات آموزشی، جایگاه سنجش در مدارس را موردبازنگری قرار داده است (Kamens & Benavot, 2011; Looney, 2018). سنجش یک اصطلاح گسترده است که به فرایند جمع‌آوری و ترکیب اطلاعات برای درک و توصیف بهتر یادگیری اشاره دارد. معلمان می‌توانند از انواع روش‌های سنجش برای تشخیص نقاط قوت و نیازهای دانش‌آموزان، برنامه‌ریزی، تنظیم دستورالعمل، و ارائه بازخورد به دانش‌آموزان و والدین درباره‌ی پیشرفت و دستاوردها استفاده کنند. هدف اولیه از سنجش کلاسی آگاهی از آموزش و بهبود یادگیری است نه صرفاً تصدیق نمره یا دسته‌بندی و انتخاب دانش‌آموزان. سنجش آموزشی فرایندی چندبعدی است و انتخاب روش‌های سنجش خاص با توجه به هدف سنجش، اهداف یادگیری هدفمند و سطوح سنی دانش‌آموزان متفاوت خواهد بود جالب است که بدانیم اصطلاح سنجش از ریشه‌ی لاتین *assidere* به معنای کنار نشستن گرفته شده است. اگرچه ممکن است این معنای اصلی با تصاویر آزمون‌های استاندارد، برگه‌های "حباب‌دار" چندگزینه‌ای، محدودیت‌های زمانی سفت‌وسخت و کار بی‌صدا در تضاد به نظر برسد، اما با مجموعه روش‌های سنجش که معمولاً توسط معلمان برای ارزیابی دانش‌آموزان استفاده می‌شود، مطابقت دارد. در واقع، *assidere* پیشنهاد می‌کند که علاوه بر آزمون‌ها و سنجش‌های عملکردی، سنجش‌های کلاسی شامل روش‌های غیررسمی "نشستن در کنار دانش‌آموزان" مشاهده و گفتگو با دانش‌آموزان به‌عنوان ابزاری برای درک و توصیف آنچه آن‌ها می‌دانند و می‌توانند انجام دهند، است (MacTighe & Ferrara, 2021).

(Wiggins, 1998) بیان می‌کند که سنجش پیشرفت تحصیلی نشان‌دهنده عملکرد یکنواخت و به‌طور شوک‌آوری ضعیف دانش‌آموزان حتی با نمرات بالا در انجام تکالیف بوده است برای مثال، دانش‌آموزان علی‌رغم کسب دانش پایه در چند دهه گذشته، توانایی کاربرد دانش، طراحی یک آزمایش یا قدرت استدلال «نامیدکننده‌ای» داشته‌اند. در واقع، تدریس برای آزمون‌های ساده دستاوردهای شگرف را غیرمحمول می‌سازد. در پی این انتقادات نهضت «ارزش سنجش اصیل» به وجود آمد. هدف، تدوین آزمون‌های استاندارد است که خروجی‌های پیچیده، قابل توجه و مرتبط با زندگی واقعی را سنجش کند. برای این رویکرد اصطلاحات «سنجش مستقیم»، «سنجش عملکرد» یا «سنجش جایگزین» نیز استفاده می‌شود. این معادل‌ها به روندهای آزمون‌های اشاره می‌کنند که تحت عنوان جایگزین آزمون‌های استاندارد چندگزینه‌ای سنتی‌اند، به این دلیل که این آزمون‌ها عملکرد یادگیرندگان را در فعالیت‌های زندگی واقعی و به‌طور مستقیم مورد ارزیابی قرار می‌دهند (Hoy & Hoy, 2013).

به‌عنوان یک جزء جدایی‌ناپذیر از فرایند آموزش، سنجش با فراهم کردن فرصتی برای نشان دادن مهارت‌ها و دانش به یادگیرندگان و درعین حال تعیین دستاوردهای حرفه‌ای و تحصیلی آن‌ها، از یادگیری حمایت می‌کند (Ashford_Rowe et al., 2014). (Boud, 1995) اشاره می‌کند که سنجش مهم‌ترین عامل برای یادگیری است و چیزی است که دانش‌آموزان آن را مهم می‌دانند زیرا آن‌ها را به‌عنوان دانش‌آموز تعریف می‌کند و در قلب یادگیری دانش‌آموزان است (Brown & Knight, 1994) و به همین دلیل برای طراحان آموزشی اهمیت زیادی دارد. این دیدگاه توسط (Brown et al, 1997) پشتیبانی می‌شود که ادعا می‌کنند اگر می‌خواهید یادگیری دانش‌آموزان را تغییر دهید، روش‌های سنجش را تغییر دهید.

سنجش محدود به شناسایی میزان دستاوردهای موارد درسی دانش‌آموزان نیست، بلکه هدایت و یادگیری دانش‌آموز به سمت توسعه مهارت‌های تفکر مرتبه بالاتر و به‌کارگیری دانش برای حل مشکلات مختلف زندگی است و این به‌نوبه‌ی خود باعث شده که سنجش اصیل با ارائه ابزارهای اندازه‌گیری مناسب برای موقعیت‌های یادگیری اصیل جای خود را در فرایند آموزش پیدا کند (Naser et al., 2017).

در سال‌های اخیر علاقه و حمایت از یک نوع سنجش که کمتر رایج است برانگیخته شده است و آن سنجش اصیل است. سنجش اصیل در شکل و اجرای آن هر دو متفاوت است ولی از اشکال رایج سنجش، سنجش تراکمی و تکوینی که امروزه استفاده می‌شود متفاوت نیست (Lundahl, 2009) سنجش اصیل متمرکز و متکی بر وجود ارتباط بین دنیای واقعی بیرون از مدرسه و خود محیط مدرسه است (Allen et al., 2012). سنجش اصیل یک تکلیف سنجش است که برای دانش‌آموز ارزش و معنایی فراتر از نمره کسب‌شده داشته باشد. سنجش اصیل مفهومی است که توسط (Wiggins, 1989) در میان دیگران رواج داده شد که به‌موجب آن دانش‌آموزان درگیر استفاده و به‌کارگیری مهارت‌ها و دانش برای پرداختن به مشکلات دنیای واقعی (Lund, 2009) یا موقعیت‌های محتمل هستند و به تکلیف احساس اصالت می‌دهند. هدف این است که چیزی فراتر از توانایی به یادآوردن پاسخ‌ها یا پاسخ دادن از روی قاعده به سؤالات تنظیم شده باشد، همان‌طور که ممکن است در قالب آزمون‌های سنتی انجام بگیرد (Ramsden, 2003). تکالیف سنجش اصیل زمینه معناداری را فراهم می‌کند که بین زندگی دانش‌آموزان و تجارب یادگیری

آن‌ها ارتباط برقراری کند. باین‌حال مفهوم اصالت یک مفهوم نسبی است و می‌تواند به‌عنوان رشته یا حوزه‌ی موضوعی در نظر گرفته شود (Sambell et al., 2013). یا می‌تواند به‌علاقه شخصی دانش‌آموزان یا مسائلی که برای خود شناسایی کرده‌اند مرتبط باشد (Davison, 2011).

بنابراین سنجش اصیل دانش‌آموزان را تشویق می‌کند تا فعالیت‌های سنجش را ذاتاً معنادار، جالب، مرتبط و دارای ارزش بلندمدت ببینند. تحقیقات (Sambell et al., 1997) در مورد تأثیر سنجش جایگزین نشان داد که اگر دانش‌آموزان احساس کنند تکالیف سنجش آن‌ها ذاتاً مفید و ارزشمند است به‌عنوان مثال، آن‌ها را ملزم به برقراری ارتباط، بحث، دفاع از ایده‌های خود یا تولید چیزی می‌کنند که برای آن‌ها فراتر از کسب فقط نمرات باشد، آن‌ها به‌احتمال زیاد تلاش خود را برای یادگیری عمیق، واقعی و پایدار صرف می‌کردند، به‌جای اینکه از طریق طرح‌ها حرکت کنند، یا هر آنچه را که فکر می‌کردند معلمانشان می‌خواهند بشنوند یا ببینند ارائه دهند. به‌این ترتیب، تکالیف سنجش که ما تعیین می‌کنیم می‌توانند پیام‌های قوی در مورد نوع یادگیری عمیق، پیچیده، و با مرتبه بالایی که ما در آموزش برای آن ارزش قائل هستیم که شامل معنا سازی، انتقادپذیری و ارتباط است، ارسال کند. بنابراین به‌طور بالقوه روش‌هایی را که در آن یادگیرندگان دنیا را می‌بینند و با آن ارتباط برقرار می‌کنند را تغییر می‌دهد (Ashwin, 2020).

بنابراین سنجش‌های اصیل و مبتنی بر یادگیری می‌توانند تأثیر مثبت قدرتمندی بر رویکردهای دانش‌آموزان به یادگیری و نحوه نگرش آن‌ها به مطالعات خود داشته باشند (Sambell et al., 2013). یکی از ویژگی‌های سنجش اصیل "پایبندی" است در برخی زمینه‌های مرتبط با رشته‌ی تحصیلی پایبندی به عملکرد حرفه‌ای مبتنی بر کار و خواسته‌های شغلی خاص ممکن است به‌ناچار بسیار مهم باشد (2004 Gulikers et al., و می‌تواند به‌طور کلی یادگیرندگان را برای پیچیدگی‌های عملکرد حرفه‌ای در سطح فارغ‌التحصیلی تجهیز کند که اغلب یک هدف کلان آموزشی است (Villarroel et al., 2018).

سنجش‌های اصیل از دانش‌آموزان می‌خواهد که به‌جای پاسخ دادن به یک سؤال با استفاده از رعایت دقیق "دستورالعمل"، دانش یا مهارت‌های خود را در موقعیت‌های جدید به‌کارگیرند، یا دانش را به‌طور خلاقانه ادغام، استفاده یا ترکیب کنند (Ashford_Rowe et al., 2014). بنابراین ویژگی‌های پیچیده و تفکر برانگیز سنجش اصیل به‌عنوان کاتالیزور برای رویکردهای عمیق و نه سطحی به یادگیری عمل می‌کند (Ramsden, 2003). افزون بر این، علاوه بر تقویت یادگیری عمیق "سنجش می‌تواند نیروی قدرتمندی در شکل‌دهی هویت انفرادی و احساس ارزشمندی دانش‌آموزان باشد و در باب این است که آن‌ها چه کسی هستند و قرار است در جامعه ما چه کسانی باشند" (McArthur, 2021).

بنابراین سنجش‌هایی که به‌طور مناسب چارچوب‌بندی مجدد شده‌اند، اصیل و مبتنی بر یادگیری‌اند می‌توانند نیروی قدرتمندی در تقویت رشد شخصی دانش‌آموز، پیشرفت، احساس تعلق و بودن در جامعه مربوط به رشته‌ی تحصیلی، و همچنین هویت، اعتمادبه‌نفس و عزت‌نفس در دانش‌آموز باشند (Arnold, 2019; Brown, 2019). و تکالیف اصیل می‌تواند نقش کلیدی در رشد دانش‌آموزان نه‌تنها به‌عنوان افراد، بلکه به‌عنوان شهروندانی فعال و مفید برای مثال از طریق پرورش انتقادپذیری، تفکر متفاوت، پویایی، انگیزه

برای زیر سؤال بردن مفروضات رایج، خلاقیت و احساس غرور و ارزش در درگیر شدن عمیق با دانش پیچیده، یا مشارکت‌های بالقوه به صلاح اجتماع داشته باشد (McArthur, 2021).

یکی از مشکلات قابل توجه آموزش این حقیقت است که بین تدریس و جهان واقعی و بین تکالیف سنجش و آنچه در دنیای کار در حال وقوع است فاصله وجود دارد (Boud, 1990). خواسته ما صرفاً این نیست که یادگیرندگان مطالب رشته تحصیلی را در هنگام فارغ‌التحصیلی آموخته باشند، ما در واقع می‌خواهیم آنان قادر به به‌کارگیری دانش و مهارت‌های آموخته‌شده باشند (Mueller, 2005). مطالعات در چندین کشور نشان می‌دهد که دانش‌آموختگان برای ورود به دنیای واقعی آماده نیستند (Gulikers et al., 2006). در همین راستا، در دهه آخر سده گذشته آموزش از دانش‌بنیان به آموزش مبتنی بر شایستگی تغییر جهت داده و تمرکز اهداف آموزشی بر توسعه یادگیرندگان شایسته و کارکنان آینده است (Segers et al., 2003). منتقدان بر این باورند که در سنجش یادگیری از یادگیرندگان خواسته می‌شود مسائلی را حل کنند و به سؤالاتی پاسخ دهند که هیچ‌وقت در دنیای بیرونی و واقعی با آن برخورد نمی‌کنند. آن‌ها مجبورند در محدودیت‌های زمانی با مسائل برخورد کنند درحالی‌که در زندگی واقعی این‌طور نیست و حل مسائل مهم، وقت‌گیر و مستلزم به‌کارگیری منابع، مشورت با دیگران و تلفیق مهارت‌های بنیادین از طریق خلاقیت و تفکر سطح بالا است) سلیمانی و همکاران به نقل از هوی و هوی، (۱۳۹۶).

Zu'bi (2013) مطالعه‌ای با هدف بررسی دانش و میزان استفاده از سنجش اصیل توسط معلمان ریاضی در پایه‌ی راهنمایی انجام داد و نتایج نشان داد که میزان استفاده از سنجش اصیل توسط معلمان در حد متوسط است و هیچ رابطه‌ای میان جنسیت، مدرک تحصیلی و تجربه با میزان استفاده از سنجش اصیل وجود ندارد. (Naser et al (2017) در مطالعه‌ی خود با عنوان " میزان استفاده از ابزارهای سنجش اصیل در دانشگاه قدس و ارتباط آن با نگرش آن‌ها نسبت به سنجش اصیل " پرداخت و نتایج مطالعه نشان داد که میزان استفاده از سنجش اصیل و نگرششان نسبت به آن متوسط است و ارتباطی میان جنسیت، تجربه و رتبه علمی با میزان استفاده از سنجش اصیل و نگرش آنان نسبت به آن وجود ندارد. (Mohammad 2020) Seifodin, در مطالعه‌ی خود که در دبیرستان‌های شهر مالانگ در شرق اندونزی انجام داد به این نتیجه دست یافت که استفاده از سنجش اصیل در بین معلمان بیشتر از حد متوسط است، همان‌گونه که در برنامه‌های یادگیری طراحی شده بود. همچنین مطالعات نشان‌دهنده عملکرد یکنواخت و ضعیف دانش‌آموزان حتی با نمرات بالا است. دانش‌آموزان باوجود کسب دانش، توانایی کاربرد دانش خود در دنیای واقعی را ندارند و تکالیف و سنجش‌های انجام‌شده در راستای دنیای کار، جهان واقعی و زندگی روزمره‌ی دانش‌آموزان نیست. همین مشکل نیز در مدارس ایران و در بین دانش‌آموختگان دیده می‌شود، به‌طور مثال دانش‌آموزان با نمرات عالی دروسی مانند زبان انگلیسی، زبان عربی و... را در مدارس می‌گذرانند درحالی‌که در دنیای واقعی قادر به کاربرد دانش خود نیستند و حتی نمی‌توانند مکالمه‌ای چنددقیقه‌ای را در زبان آموخته‌شده به‌کارگیرند همین مسئله نیز در دروس دیگر به‌وضوح مشاهده می‌گردد. این مشکل به دلیل وجود فاصله میان تکالیف، دانش آموخته‌شده و سنجش‌های صورت گرفته با دنیای واقعی و کاراست و

سنجش اصیل با استفاده از تکالیف، آزمون‌ها و سنجش‌های مرتبط به زمینه‌ی اجتماعی، فیزیکی، و دنیای واقعی در پاسخ به حل این مشکل به وجود آمد.

علی‌رغم فراوانی نظریه‌ها و الگوهای یادگیری در چند دهه‌ی گذشته، اقدامات ناچیزی برای اصلاح شیوه‌های سنجش در مدارس صورت گرفته است. در واقع سنجش، همواره در نیمکت آخر برنامه‌ها و دستورهای کاری اصلاحات آموزشی بوده است (Shinn, 2012; Suleiman et al., 2021). اهمیت سنجش اصیل در سال‌های اخیر مورد توجه قرار گرفته و آموزش‌دهندگان در بسیاری از کشورها از آزمون‌های استاندارد سنتی به سمت سنجش‌های اصیل مبتنی بر برنامه‌ی درسی حرکت کرده‌اند به‌طور مثال در برنامه درسی ۲۰۱۳ اندونزی، سنجش اصیل به‌عنوان یک ابزار سنجش در فرایند یادگیری گنجانده شده و اعتقاد بر این است که این سنجش می‌تواند به‌طور مناسب پیشرفت و دستاوردهای دانش‌آموزان را در تمام زمینه‌های شایستگی ارزیابی و اندازه‌گیری کند. در این رابطه همچنین، بسیاری از مدارس در سنگاپور بازنگری‌های اساسی در برنامه درسی و ساختارهای سنجش را انجام داده‌اند و معتقدند همان‌گونه که فرایندهای آموزشی تغییر کرده‌اند، پارادایم‌های ارزیابی نیز که زیربنای آن است باید تغییر کند (Ashford_ et al., 2014). نظر به ضرورت آشناسازی و افزایش آگاهی معلمان، دانش‌آموزان و متصدیان آموزشی با مفهوم سنجش اصیل، ابزارها و انواع آن نیاز به انجام تحقیقات در این رابطه، با توجه به جدید بودن مفاهیم آن نیاز است. هدف پژوهش حاضر تدوین و اعتبارسنجی دو پرسشنامه میزان استفاده معلمان از سنجش اصیل و ادراک معلمان از سنجش اصیل است. از این‌رو پرسش‌های پژوهش ذیل مطرح شده‌اند: (۱) میزان ادراک دبیران دبیرستان‌های شهر سنندج از مفهوم سنجش اصیل چگونه است؟ (۲) میزان استفاده دبیران دبیرستان‌های شهر سنندج از اصول سنجش اصیل چیست؟ (۳) متغیرهای جمعیت شناختی مانند جنسیت، سابقه تدریس، میزان تحصیلات، ناحیه خدمت و نوع استخدام چه تفاوتی در میزان ادراک یا استفاده معلمان از مفهوم و اصول سنجش اصیل ایجاد می‌کنند؟

روش پژوهش

در این پژوهش از رویکرد تحقیق کمی با راهبرد پیمایشی استفاده شده است. جامعه آماری عبارت است از تمام دبیران مقطع متوسطه دوم نظری (رسمی قطعی، رسمی آزمایشی، پیمانی و حق‌التدریسی) شاغل در نواحی دوگانه آموزش و پرورش شهر سنندج در سال تحصیلی ۱۴۰۱-۱۴۰۰ است. افراد نمونه در این پژوهش، با روش نمونه‌گیری خوشه‌ای دو مرحله‌ای انتخاب شده‌اند به‌طوری‌که در مرحله اول از ۸۳ مدرسه مقطع متوسطه دوم نظری شهر سنندج ۲۰ مدرسه به‌عنوان خوشه به‌صورت تصادفی انتخاب و سپس از این مدرسه به‌تناسب تعداد معلمانشان به‌صورت تصادفی انتخاب گردیدند، برای برآورد حجم نمونه از فرمول کوکران استفاده شده است که تعداد ۲۰۰ آزمودنی مشخص شده است. برای تجزیه و تحلیل داده‌های کمی گردآوری شده با پرسشنامه از روش‌های آمار توصیفی و آمار استنباطی استفاده شده است. در قسمت آمار توصیفی از فراوانی و درصد و همچنین از شاخص‌های آماری مانند میانگین، انحراف معیار، کمترین، میانه، بیشترین و ضریب چولگی بهره گرفته شده است و در بخش آمار استنباطی از آزمون میانگین تک نمونه‌ای با استفاده از توزیع t -استودنت، آزمون مقایسه میانگین دو گروه مستقل با استفاده از توزیع t -استودنت و

آزمون مقایسه میانگین چند گروه مستقل با استفاده از روش تحلیل واریانس یک‌راهه و آزمون ضریب همبستگی پیرسون استفاده شده است.

ابزار گردآوری داده‌ها: برای گردآوری داده‌ها از دو پرسشنامه محقق ساخته تحت عناوین الف- پرسشنامه میزان استفاده معلمان از سنجش اصیل و ب- پرسشنامه ادراک معلمان از سنجش اصیل در کلاس درس استفاده شد. برای پایایی از دو روش آزمون- بازآزمون و آلفای کرونباخ استفاده شده است. در روش آزمون- بازآزمون پرسشنامه به ۱۵ نفر داده شد و آن‌ها پرسشنامه را تکمیل کردند. سپس بعد از دو هفته پرسشنامه بر روی همان مشارکت‌کنندگان اجرا شد. سپس میزان همبستگی پاسخ‌ها با استفاده از ضریب همبستگی اسپیرمن محاسبه شد. با توجه به این که آماره آزمون برابر است با $t = 0.847$ و معنی‌داری برابر با $Sig = 0.000$ بود پس بین پاسخ‌های افراد در طی یک هفته رابطه خطی وجود دارد. یعنی اینکه پایایی پرسشنامه‌ها در حد مطلوبی است. از روش آلفای کرونباخ هم برای سازگاری درونی بهره گرفته شد که نتایج به شرح زیر است:

جدول ۱: محاسبه پایایی پرسشنامه

عنوان پرسشنامه	تعداد گویه	ضریب آلفای کرونباخ
ادراک معلمان از ارزشیابی اصیل	۳۹	۰,۸۳۱
میزان استفاده از ارزشیابی اصیل	۵۱	۰,۹۲۲

برای بررسی روایی پرسشنامه پس از تأیید پرسشنامه توسط اساتید که روایی محتوایی آن مورد تأیید قرار گرفت سپس با استفاده از تحلیل عاملی اکتشافی روایی سازه مورد بررسی قرار گرفت که نتایج در زیر آمده است:

جدول ۲: شناسایی عامل‌های پرسشنامه «میزان استفاده از ارزشیابی اصیل»

عامل	گویه‌ها	بار عاملی	درصد واریانس	مقدار ویژه
۱- معیارهای شفاف	دانش‌آموزان باید برای بخش‌های مهم کارشان خود-ارزیابی ارائه	۰,۵۶۸	۲۸,۶۶	۱۴,۶۲
	به دانش‌آموزان شیوه خود-ارزیابی و تطبیق با معیارهای عملکرد	۰,۴۹۷		
	از دانش‌آموزان می‌خواهم اجراها و کارکردهای هنری، سیاسی یا	۰,۶۹۵		
	از دانش‌آموزان می‌خواهم بر آفرینش محصول، اجرای پروژه یا فعالیت	۰,۵۸۰		
	تکالیف انجام‌شده دانش‌آموزان را به کمک خودشان نقد می‌کنیم.	۰,۵۹۴		
	تکالیف انجام‌شده به خاطر اشتباهات دستوری یا روش‌شناختی	۰,۵۲۱		
	تکالیف انجام‌شده اگر نامرتب، بی‌روح یا عاری از خلاقیت باشند	۰,۵۲۱		
	دانش‌آموزان را بر اساس پرسش‌های مطرح‌شده توسط آنان در کلاس	۰,۷۱۷		
	دانش‌آموزان را بر اساس ارائه نقد از باورهای پذیرفته‌شده ارزیابی	۰,۷۴۶		
	دانش‌آموزان را بر اساس پژوهش در مورد مسائل مغفول ارزیابی	۰,۷۱۸		
	دانش‌آموزان را بر اساس ارائه دیدگاه‌های متفاوت و جایگزین ارزیابی	۰,۶۹۴		
دانش‌آموزان را آگاه می‌سازم تا به ماهیت نسبی و غیرقطعی علم و	۰,۷۵۲			

		۰,۵۶۷	دانش‌آموزان را راهنمایی می‌کنم از مطلق‌گرایی در هر رشته علمی	
		۰,۷۳۹	دانش‌آموزان را با تاریخچه رشته‌های لمی آشنا می‌کنم.	
		۰,۵۴۱	دانش‌آموزانم را با موضوعات مورد اختلاف در رشته‌های علمی آشنا	
		۰,۶۶۶	دانش‌آموزان باید پروژه‌ها را از همان ابتدای انجام مستندسازی و به	
		۰,۷۳۰	دانش‌آموزان باید صداقت علمی خود را در انجام تکالیف و پروژه‌ها	
		۰,۶۳۳	دانش‌آموزان می‌دانند که عادت‌های ذهنی آن‌ها را ارزیابی می‌کنم.	
		۰,۶۶۴	تکالیفی که به دانش‌آموزان می‌دهم همواره در محیط شبیه‌سازی شده	
		۰,۵۱۱	تکالیفی که به دانش‌آموزان می‌دهم همواره به تولید محصول یا اثر	
۸,۷۴	۱۷,۱۳	۰,۹۹۴	نظام سنجش و ارزیابی من برای دانش‌آموزان شفاف است.	۲- خودارزیابی
		۰,۹۹۷	دانش‌آموزان می‌دانند که نظام سنجش من وزارتی و غیر اختیاری	
		۰,۹۵۶	در نظام ارزیابی من خود-ارزیابی دانش‌آموزان هسته اصلی کار را	
		۰,۹۶۹	در نظام ارزیابی من خود-ارزیابی دانش‌آموزان رویدادی متداول و	
		۰,۹۷۱	در نظام ارزیابی من دانش‌آموز باید خود-ارزیابی را در همه تکالیف	
		۰,۹۷۹	تکالیفی که به دانش‌آموزان می‌دهم نیازمند استفاده از دانش در	
		۰,۹۷۱	تکالیفی که به دانش‌آموزان می‌دهم بازخوردهای مفید و تشخیصی	
		۰,۹۶۱	تکالیفی که به دانش‌آموزان می‌دهم شبیه مسائل پیش‌رو در محیط	
۳,۳۷	۶,۶۱	۰,۷۴۵	از دانش‌آموزان می‌خواهم که صرفاً دیدگاه‌های متداولی را که	۳- یادگیری اصیل
		۰,۷۱۲	از دانش‌آموزان می‌خواهم فنونی را که آموخته‌اند بدون تفکر و درون	
		۰,۶۷۱	از دانش‌آموزان می‌خواهم از تولیدات خود (مانند: مقاله /گزارش	
		۰,۷۶۱	از دانش‌آموزان می‌خواهم به انتقادات من و همکلاسان از تکالیفشان	
		۰,۷۹۶	از نظر من دانش‌آموزان به افرادی به‌عنوان الگوهای موفق نیاز دارند.	
۲,۷	۵,۱۶	۰,۵۵۴	تکالیف دانش‌آموزان دربرگیرنده کارهای ضروری تکرارشونده هستند.	۴- تکالیف فکورانه
		۰,۸۴۲	تکالیفی دانش‌آموزان بر اساس مهارت‌ها و موقعیت‌های قابل	
		۰,۶۱۷	تکالیفی که به دانش‌آموزان می‌دهم به قضاوت و نوآوری نیاز دارند.	
		۰,۸۳۴	تکالیفی که به دانش‌آموزان می‌دهم نیازمند خلق اثر، توجیه و عملکرد	
		۰,۷۰۰	تکالیفی که به دانش‌آموزان می‌دهم به‌صورت مستقیم به سراغ نتیجه	
۱,۹۱	۳,۷۵	۰,۶۹۲	دانش‌آموزان به فرصتی برای تقلید و سرمشق‌گیری از الگوهای موفق نیاز دارند.	۵- الگودهی
		۰,۶۶۰	نتایج و هر محصول تولیدی دانش‌آموزان باید سطح مناسبی از	
		۰,۷۸۱	دانش‌آموز باید محصولی را تولید کند که از فعالیت‌های یادگیری قبلی الگوبرداری کرده است.	
		۰,۷۶۴	با دانش‌آموزان باید بر اساس دسترسی به الگوها، بازخورد یادگیری و	
		۰,۷۶۸	دانش‌آموزان معیارهای نمره دهی من را از قبل می‌دانند.	
		۰,۶۷۴	دانش‌آموزان الگوهایی برای (نوشتن مقاله و یا اجرای ضبط‌شده) را از	
۱,۳۸	۲,۷۱	۰,۷۰۵	دانش‌آموزان بر پایه معیارها و استانداردهای شناخته‌شده ارزیابی	۶- استاندارد شناخته‌شده
		۰,۵۸۸	در نظام ارزیابی من به‌درستی خود-ارزیابی دانش‌آموز نمره اختصاص	
		۰,۶۳۰	تکالیفی که به دانش‌آموزان می‌دهم از زمینه و موقعیت‌های واقعی	
۱,۲۳	۲,۴۲	۰,۵۱۳	از دانش‌آموزان می‌خواهم درک و مهارت خود را از آنچه یاد گرفته‌اند	۷- توجیه یادگیری
		۰,۵۷۶	تکالیفی که به دانش‌آموزان ارائه می‌دهم پیوندی قوی با دنیای واقعی	
		۰,۵۸۲	دانش‌آموزان اجازه دارند که از گونه‌های جایگزین برای انجام تکالیف	

۱,۰۸	۲,۱۲	۰,۶۶۱	دانش‌آموزان نمونه سؤالاتی که احتمال می‌رود در امتحان طرح شوند
------	------	-------	---

تحلیل عاملی اکتشافی به روش تحلیل مؤلفه‌های اصلی^۱ با چرخش واریماکس^۲ انجام گرفته است و نتایج جدول فوق بیانگر این است که ۸ عامل دارای مقدار ویژه بیشتر از یک هستند که بیشترین مقدار ویژه برابر ۱۴,۶۲ و کمترین مقدار ویژه ۱,۰۸ است. عامل اول ۲۸,۶۶ درصد از واریانس و عامل هشتم ۲,۱۲ درصد از واریانس و هشت عامل باهم ۶۸,۵۶ درصد واریانس پرسشنامه را به خود اختصاص داده‌اند.

جدول ۳: شناسایی عامل‌های پرسشنامه «دراک از ارزشیابی اصیل»

مقدار ویژه	درصد واریانس	بار عاملی	گویه‌ها	عامل
۱۱,۲۲	۲۸,۷۸	۰,۸۷۵	ارزیابی از عملکرد دانش‌آموزان باید راهی مؤثر برای ارزیابی مهارت‌های	۱- شفاف‌سازی
		۰,۹۰۲	ارزیابی از عملکرد دانش‌آموزان باید با شفاف‌سازی انتظارات معلم قبل از	
		۰,۵۰۱	ارزیابی از عملکرد دانش‌آموزان باید دانش و مهارت‌های حرفه‌ای را در	
		۰,۹۴۸	آزمون‌های من شامل بیان دوباره، به خاطر سپاری، و کپی کردن دانش	
		۰,۹۶۷	سؤالات آزمون‌های من باید تا موقع امتحان مخفی نگه‌داشته شوند*	
		۰,۹۷۷	منابع آزمون‌های من در طول امتحان از دسترس دانش‌آموز دورنگه داشته	
		۰,۹۷۰	آزمون‌های من بر تولید اثر و معیارهای شناخته‌شده باکیفیت بالا متمرکز	
۶,۰۴	۱۵,۵۰	۰,۶۴۵	ارزیابی از عملکرد دانش‌آموز باید بر مهارت‌های پیشرفته و سطح بالاتر تمرکز داشته باشد.	۲- معیارهای واقعی
		۰,۶۱۸	ارزیابی از عملکرد دانش‌آموز باید همواره بر روی صلاحیت‌های مرتبط با دنیای کار متمرکز باشد.	
		۰,۶۰۹	ارزیابی از عملکرد دانش‌آموز باید بر کاربرد ذهنی مهارت‌های آموخته‌شده در یک محیط جدید متمرکز باشد.	
		۰,۵۳۰	یک ویژگی کلیدی ارزیابی از عملکرد دانش‌آموز باید «قابل اجرا بودن در دنیای واقعی» باشد.	
		۰,۵۱۸	آزمون‌های من فقط به پاسخ‌های صحیح نیاز دارند.*	
		۰,۷۵۳	آزمون‌های من سؤالاتشان تا قبل از اجرا ناشناخته می‌مانند.*	
		۰,۷۶۰	آزمون‌های من پاسخ‌هایشان شامل یک حدس یا گمان هستند.*	
		۰,۷۳۶	آزمون‌های من شامل سؤالاتی است که به تشخیص پاسخ یا مهارت‌های شناخته‌شده نیاز دارد.*	
		۰,۷۱۴	آزمون‌های من فهم دانش‌آموز را از پیوند بین چیزی که ارزیابی شده و آنچه باید ارزیابی شود استنباط می‌کند.*	
		۰,۷۵۸	آزمون‌های من همواره نمره ارائه می‌دهند.*	
		۰,۶۲۴	آزمون‌های من برای این که معتبر نمره‌گذاری شوند آسان می‌شوند.*	
۰,۷۸۲	آزمون‌های من از زمینه و موقعیت‌های واقعی جدا هستند.*			

1. Principal Component Analysis

2. Varimax Rotation

۳,۹۳	۱۰,۰۸	۰,۸۸۰	ارزیابی از عملکرد دانش‌آموز باید با حرفه آتی که ممکن است انتخاب کند	۳- اصالت
		۰,۹۴۲	عنصر اصلی ارزیابی از عملکرد دانش‌آموزان باید اصالت باشد.	
		۰,۹۳۷	ارزیابی از عملکرد دانش‌آموز باید مستلزم به‌کارگیری مشورت با دیگران باشد.	
		۰,۹۴۸	ارزیابی از عملکرد دانش‌آموز باید مهارت‌های پایه را با خلاقیت و تفکر سطح بالا	
		۰,۹۳۵	ارزیابی از عملکرد دانش‌آموز باید نتیجه‌اش به شکل عملکرد یا خلق یک اثر	
۲,۶۷	۶,۸۵	۰,۷۱۳	ارزیابی از عملکرد دانش‌آموز باید آنان را برای حرفه‌های آینده آماده کند.	۴- تمرکز بر حرفه آینده
		۰,۶۵۱	ارزیابی از عملکرد دانش‌آموز باید شبیه کار افراد در دنیای واقعی باشد.	
		۰,۶۳۰	ارزیابی از عملکرد دانش‌آموز باید از کار و وظایف واقعی افراد در جامعه متفاوت	
		۰,۷۸۵	ارزیابی از عملکرد دانش‌آموز باید کار حرفه‌ای افراد در جامعه را شبیه‌سازی	
		۰,۷۷۵	ارزیابی از عملکرد دانش‌آموز باید مشابه نتایجی باشد که افراد شاغل در جامعه	
۱,۵۲	۳,۸۹	۰,۵۳۳	ارزیابی از عملکرد دانش‌آموز باید میزان درک و دانش او را در موقعیت‌های	۵- واقع‌گرایی
		۰,۷۲۵	ارزیابی از عملکرد دانش‌آموز باید فرصت دریافت بازخورد، اصلاح عملکرد و	
		۰,۸۰۰	ارزیابی از عملکرد دانش‌آموز باید واقع‌گرا باشد.	
		۰,۸۵۳	ارزیابی از عملکرد دانش‌آموز باید شامل بیان دوباره و به خاطر سپاری دانش	
		۰,۷۶۳	ارزیابی از عملکرد دانش‌آموز باید فرصت‌های مناسب برای تمرین و کنکاش	
۱,۱۲	۲,۸۶	۰,۵۶۶	ارزیابی از عملکرد دانش‌آموز باید شامل پاسخ به پرسش‌هایی باشد که هیچ‌وقت در دنیای واقعی با آن برخورد نمی‌کنند.*	۶- چالش انگیزی
		۰,۶۹۰	ارزیابی از عملکرد دانش‌آموز باید چالش‌برانگیز باشد.	
۱,۰۹	۲,۷۹	۰,۴۷۰	ارزیابی از عملکرد دانش‌آموز باید مشابه معیاری باشد که در عمل مورد نیاز است.	۷- کاربردی بودن
		۰,۶۲۴	ارزیابی از عملکرد دانش‌آموز باید دارای معیارهای کاملاً شفاف باشد.	
		۰,۵۷۴	ارزیابی از عملکرد دانش‌آموز باید بر نیازمندی حرفه‌های واقعی اجتماع تمرکز	

تحلیل عاملی اکتشافی به روش تحلیل مؤلفه‌های اصلی با چرخش واریماکس انجام گرفت. نتایج بیانگر این است که ۷ عامل دارای مقدار ویژه بیشتر از مقدار یک هستند که بیشترین مقدار ویژه برابر ۱۱,۲۲ و کمترین مقدار ویژه ۱,۰۹ است. عامل اول ۲۸,۷۸ درصد از واریانس و عامل هفتم ۱,۰۹۲ درصد از واریانس و هشت عامل باهم ۷۰,۷۵ درصد واریانس پرسشنامه را به خود اختصاص داده‌اند.

یافته‌های پژوهش

فرضیه اول: سطح ادراک دبیران دبیرستان‌های شهر سنندج از مفهوم سنجش اصیل بیش‌ازحد متوسط است.

جدول ۴: آماره‌های آزمون مقایسه میانگین متغیر ادراک معلمان با عدد ثابت

Sig	آماره آزمون t	انحراف معیار	میانگین	تعداد	متغیر
۰,۰۰۰	۱۵,۸۰۸	۸,۹۱	۱۳۶,۹۶	۲۰۰	ادراک

همان‌گونه که ملاحظه می‌شود مقدار آماره آزمون t استودنت برابر است با $t=15.808$ و با توجه به مقدار $Sig=0.000$ چون $Sig < 0.05$ فرض H_0 را در سطح معنی‌داری $\alpha=0.05$ رد می‌شود. می‌توان گفت سطح ادراک دبیران دبیرستان‌های شهر سنندج از مفهوم سنجش اصیل بیش‌ازحد متوسط است.

فرضیه دوم: میزان استفاده دبیران دبیرستان‌های شهر سنندج از اصول سنجش اصیل بیش از حد متوسط است.

جدول ۵: آماره‌های آزمون مقایسه میانگین متغیر میزان استفاده معلمان با عدد ثابت

متغیر	تعداد	میانگین	انحراف معیار	آماره آزمون t	Sig
میزان استفاده	۲۰۰	۱۷۳,۰۵	۲۸,۰۷	۱۰,۱	۰,۰۰۰

همان‌گونه که ملاحظه می‌شود مقدار آماره آزمون t استودنت برابر است با $t=10.1$ و با توجه به مقدار $Sig=0.000$ چون $Sig < 0.05$ فرض H_0 را در سطح معنی‌داری $\alpha=0.05$ رد می‌کنیم. می‌توان گفت میزان استفاده دبیران دبیرستان‌های شهر سنندج از اصول سنجش اصیل بیش از حد متوسط است. برای بررسی اثر جنسیت بر ادراک معلمان از سنجش اصیل، با توجه به اینکه هم تعداد مردان و هم تعداد زنان بیشتر از ۳۰ است قضیه حد مرکزی در اینجا صدق می‌کند و می‌توان از آزمون t-استودنت استفاده کرد.

جدول ۶: آماره‌های توصیفی متغیر ادراک به تفکیک جنسیت

گروه	تعداد	میانگین	انحراف معیار	کمترین	میانه	بیشترین	ضریب چولگی
زن	۷۹	۱۲۷,۳	۹,۶	۱۰۱	۱۲۹	۱۴۹	-۰,۴۴۸
مرد	۱۲۱	۱۲۶,۷	۸,۵	۱۰۰	۱۲۷	۱۴۷	-۰,۴۰۰

برای انجام آزمون فوق از روش مقایسه میانگین دو جامعه با استفاده از آزمون t-استودنت استفاده شد که نتایج به‌قرار زیر است:

جدول ۷: آماره‌های آزمون مقایسه میانگین متغیر ادراک در دبیران مرد و زن

آزمون مقایسه میانگین دو جامعه مستقل						
متغیر	آزمون هم‌واریانس لیون			آزمون برابری میانگین‌ها		
	Sig	آماره آزمون F	Sig	آماره آزمون t	درجه آزادی	Sig
ادراک	۰,۱۴۶	۲,۱۳۴	۰,۱۴۶	۰,۴۲۴	۱۹۸	۰,۶۷۲
	۰,۴۱۳			۱۵۱,۶۰۶	۰,۶۸۱	

ابتدا باید فرض هم‌واریانس بررسی شود که آماره آزمون لیون برابر است با $F=2.134$ و با توجه به مقدار $Sig=0.146$ چون $Sig > 0.05$ فرض H_0 را در سطح معنی‌داری $\alpha=0.05$ رد نمی‌شود یعنی اینکه توزیع متغیر ادراک در گروه مردان و زنان هم‌واریانس هستند و به همین خاطر از آماره t استودنت استفاده شد که ملاحظه می‌شود مقدار آماره آزمون t استودنت برابر است با $t=0.424$ و با توجه به مقدار $Sig=0.672$ چون $Sig > 0.05$ فرض H_0 را در سطح معنی‌داری $\alpha=0.05$ رد نمی‌شود. یعنی می‌توان گفت سطح متغیر ادراک بین معلمان زن و مرد برابر است به عبارت دیگر جنسیت روی ادراک معلمان اثرگذار نبوده است.

برای بررسی اثر ناحیه بر سطح ادراک سنجش اصیل نیز از آزمون t-استودنت استفاده شد. حال توصیفی از متغیر ادراک در دو ناحیه را نشان داده می‌شود:

جدول ۸: آماره‌های توصیفی متغیر ادراک به تفکیک ناحیه

گروه	تعداد	میانگین	انحراف معیار	کمترین	میانه	بیشترین	ضریب چولگی
ناحیه ۱	۹۳	۱۲۷,۷	۸,۸	۱۰۱	۱۲۹	۱۴۹	-۰,۳۹۲
ناحیه ۲	۱۰۷	۱۲۶,۳	۹,۰	۱۰۰	۱۲۷	۱۴۷	-۰,۴۲۶

برای انجام آزمون فوق از روش مقایسه میانگین دو جامعه با استفاده از آزمون t - استودنت استفاده شد که نتایج به‌قرار زیر است:

جدول ۹: آماره‌های آزمون مقایسه میانگین متغیر ادراک در ناحیه یک و دو

آزمون مقایسه میانگین دو جامعه مستقل						
متغیر	آزمون هم‌واریانس لیون			آزمون برابری میانگین‌ها		
	آماره آزمون F	Sig	آماره آزمون t	درجه آزادی	Sig	
ادراک	۰,۰۶۴	۰,۸۰۰	۱,۱۵۸	۱۹۸	۰,۲۴۸	واریانس‌ها برابر
			۱,۱۶۰	۱۹۵,۴۶۶	۰,۲۴۷	واریانس‌ها نابرابر

برای بررسی آزمون فرض فوق ابتدا باید فرض هم‌واریانسی بررسی شود که آماره آزمون لیون برابر است با $F=0.064$ و با توجه به مقدار $Sig=0.800$ چون $Sig > 0.05$ فرض H_0 را در سطح معنی‌داری $\alpha=0.05$ رد نمی‌شود یعنی اینکه توزیع متغیر ادراک در ناحیه یک و دو هم‌واریانس هستند و به همین خاطر از آماره t استودنت استفاده می‌شود که ملاحظه می‌شود مقدار آماره آزمون t استودنت برابر است با $t=1.158$ و با توجه به مقدار $Sig=0.248$ چون $Sig > 0.05$ فرض H_0 را در سطح معنی‌داری $\alpha=0.05$ رد نمی‌شود. یعنی اینکه می‌توان گفت سطح متغیر ادراک بین معلمان ناحیه یک و دو برابر است به عبارت دیگر ناحیه روی ادراک معلمان اثرگذار نبوده است.

اثر مدرک تحصیلی را بر سطح ادراک سنجش اصیل بررسی شد. با توجه به اینکه تعداد معلمان دارای مدرک تحصیلی کاردانی و دکتری کمتر از ۳۰ است قضیه حد مرکزی در اینجا صدق نمی‌کند پس باید فرض نرمال بودن داده‌های مربوط به متغیر ادراک را برای مدرک تحصیلی بررسی شود و اگر داده‌ها برای هر چهار گروه نرمال باشند می‌توان از تحلیل واریانس یک‌طرفه استفاده کرد در غیر این صورت باید از آزمون نا پارامتری کروسکال-والیس استفاده کرد. حال توصیفی از متغیر ادراک در چهار گروه مدرک تحصیلی را نشان داده شده است:

جدول ۱۰: آماره‌های توصیفی متغیر ادراک به تفکیک مدرک تحصیلی

گروه	تعداد	میانگین	انحراف معیار	کمترین	میانه	بیشترین	ضریب چولگی
کاردانی	۵	۱۲۵,۶	۱۴,۷	۱۰,۸	۱۲۸	۱۴۲	-۰,۲۰۰
کارشناسی	۱۳۳	۱۲۷,۴	۸,۵	۱۰,۱	۱۲۸	۱۴۹	-۰,۴۸۰
کارشناسی ارشد	۵۶	۱۲۵,۶	۹,۸	۱۰,۰	۱۲۶	۱۴۷	-۰,۲۲۴
دکتری	۶	۱۲۷,۰	۴,۲	۱۲۱	۱۲۸,۵	۱۳۱	-۰,۶۷۶

برای بررسی نرمال بودن داده‌ها در چهار گروه مدرک تحصیلی از آماره آزمون شاپیرو-ویلک استفاده شد:

جدول ۱۱: آزمون نرمال بودن داده‌های متغیر ادراک به تفکیک مدرک تحصیلی

مدرک تحصیلی	آماره شاپیرو-ویلک	درجه آزادی	سطح معنی‌داری
کاردانی	۰,۹۲۳	۵	۰,۵۵۰
کارشناسی	۰,۹۸۲	۱۳۳	۰,۰۷۵
کارشناسی ارشد	۰,۹۹۱	۵۶	۰,۹۵۷
دکتری	۰,۸۳۶	۶	۰,۱۲۲

با مشاهده مقادیر آماره آزمون شاپیرو-ویلک سطوح معنی‌داری متناظر با توجه به اینکه مقادیر سطوح معنی‌داری همگی بیشتر از ۰,۰۵ هستند می‌توان گفت که توزیع متغیر ادراک برای هر چهار مقطع تحصیلی نرمال است. بنابراین می‌توان از آزمون ANOVA استفاده کرد که نتایج در زیر آمده است:

جدول ۱۲: آماره‌های تحلیل واریانس متغیر ادراک برای مدرک تحصیلی

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون F	Sig
بین گروهی	۹۰,۱۱۸	۳	۳۰,۰۳۹	۰,۳۷۵	۰,۷۷۱
درون گروهی	۱۵۷۰۹,۵۶۲	۱۹۶	۸۰,۱۵۱		
کل	۱۵۷۹۹,۶۸۰	۱۹۹			

با توجه به مقدار آماره آزمون فیشر $F=0.375$ و مقدار سطح معنی‌داری متناظر $Sig=0.771$ چون مقدار سطح معنی‌داری از ۰,۰۵ بزرگ‌تر است پس فرض صفر رد نمی‌شود یعنی می‌توان گفت که میزان ادراک معلمان با مدرک تحصیلی کاردانی، کارشناسی، کارشناسی ارشد و دکتری باهم برابر است و تفاوت معنی‌داری بین آن‌ها مشاهده نمی‌شود.

برای بررسی اثر نوع استخدام بر سطح ادراک سنجش اصیل با توجه به اینکه تعداد معلمان حق‌التدریس کمتر از ۳۰ نفر است قضیه حد مرکزی در اینجا صدق نمی‌کند پس باید فرض نرمال بودن داده‌های مربوط به متغیر ادراک را برای نوع استخدام بررسی کرد. اگر داده‌ها برای هر چهار گروه نرمال باشند می‌توان از تحلیل واریانس یک‌طرفه استفاده کرد.

جدول ۱۳: آماره‌های توصیفی متغیر ادراک به تفکیک نوع استخدام

گروه	تعداد	میانگین	انحراف معیار	کمترین	میان	بیشترین	چولگی
پیمانی	۵۳	۱۲۶,۶	۸,۱	۱۰۷	۱۲۶	۱۴۲	-۰,۲۶۰
رسمی آزمایشی	۶۲	۱۲۶,۲	۹,۱	۱۰۱	۱۲۸	۱۴۳	-۰,۵۷۲
رسمی قطعی	۷۴	۱۲۸,۰	۹,۵	۱۰۰	۱۲۸	۱۴۹	-۰,۴۲۱
حق‌التدریس	۱۱	۱۲۶,۴	۷,۶	۱۱۳	۱۲۹	۱۳۵	-۰,۷۶۷

برای بررسی نرمال بودن داده‌ها در چهار گروه نوع استخدام از آماره آزمون شاپیرو-ویلک استفاده شد:

جدول ۱۴: آزمون نرمال بودن داده‌های متغیر ادراک به تفکیک نوع استخدام

نوع استخدام	آماره شاپیرو-ویلک	درجه آزادی	سطح معنی‌داری
پیمانی	۰,۹۸۳	۵۳	۰,۶۵۹
رسمی آزمایشی	۰,۹۷۱	۶۲	۰,۱۴۵
رسمی قطعی	۰,۹۸۳	۷۴	۰,۴۴۰
حق‌التدریس	۰,۸۹۳	۱۱	۰,۱۵۱

با مشاهده مقادیر آماره آزمون شاپیرو-ویلک سطوح معنی‌داری متناظر با توجه به اینکه مقادیر سطوح معنی‌داری همگی بیشتر از ۰,۰۵ هستند می‌توان گفت که توزیع متغیر ادراک برای هر چهار نوع استخدامی نرمال است. بنابراین می‌توان از آزمون ANOVA استفاده کرد که نتایج در زیر آمده است:

جدول ۱۵: آماره‌های تحلیل واریانس متغیر ادراک برای نوع استخدام

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون F	Sig
بین گروهی	۱۲۴,۴۹۲	۳	۴۱,۴۹۷	۰,۵۱۹	۰,۶۷۰
درون گروهی	۱۵۶۷۵,۱۸۸	۱۹۶	۷۹,۹۷۵		
کل	۱۵۷۹۹,۶۸۰	۱۹۹			

با توجه به مقدار آماره آزمون فیشر $F=0.519$ و مقدار سطح معنی‌داری متناظر $Sig=0.670$ چون مقدار سطح معنی‌داری از ۰,۰۵ بزرگ‌تر است پس فرض صفر رد نمی‌شود یعنی می‌توان گفت که میزان ادراک

معلمان پیمانی، رسمی آزمایشی، رسمی قطعی و حق التدریس باهم برابر است و تفاوت معنی داری بین آن‌ها مشاهده نمی‌شود.

حال می‌خواهیم رابطه بین سابقه کار و ادراک معلمان را بسنجیم. برای این کار آزمون فرض زیر را انجام می‌دهیم:

بین "سطح ادراک سنجش اصیل" معلمان و سابقه کار آنان رابطه خطی وجود ندارد: H_0

بین "سطح ادراک سنجش اصیل" معلمان و سابقه کار آنان رابطه خطی وجود دارد: H_1

جدول ۱۶: ضریب همبستگی بین سابقه کار و ادراک معلمان

سابقه کار	ادراک	ضریب همبستگی پیرسون	تعداد	سطح معنی داری
		۰,۰۴	۲۰۰	۰,۵۷۴

با توجه به این که آماره آزمون برابر است با $r = 0.04$ و مقدار سطح معنی داری برابر است با $Sig = 0.574$ و با توجه به اینکه $Sig > 0.05$ پس فرض H_0 را در سطح معنی داری پنج درصد رد نمی‌شود یعنی اینکه بین ادراک و سابقه کار معلمان رابطه خطی وجود ندارد.

برای بررسی اثر جنسیت بر میزان استفاده از اصول سنجش اصیل از آزمون t -استودنت استفاده شد. حال توصیفی از متغیر "استفاده" در دو گروه مردان و زنان نشان داده می‌شود:

جدول ۱۷: آماره‌های توصیفی متغیر میزان استفاده به تفکیک جنسیت

گروه	تعداد	میانگین	انحراف معیار	کمترین	میانه	بیشترین	چولگی
زن	۷۹	۱۷۴,۳	۳۰,۵	۶۵	۱۷۴	۲۴۳	-۰,۳۶۹
مرد	۱۲۱	۱۷۲,۲	۲۶,۵	۱۰۵	۱۷۱	۲۳۹	-۰,۰۶۹

برای انجام آزمون فوق از روش مقایسه میانگین دو جامعه با استفاده از آزمون t -استودنت استفاده شد که نتایج به‌قرار زیر است:

جدول ۱۸: آماره‌های آزمون مقایسه میانگین متغیر میزان استفاده در مردان و زنان

آزمون مقایسه میانگین دو جامعه مستقل						
متغیر	آزمون هم‌واریانس لیون			آزمون برابری میانگین‌ها		
	آماره آزمون F	Sig	آماره آزمون t	درجه آزادی	Sig	
استفاده	۱,۳۹۰	۰,۲۴۰	۰,۵۰۱	۱۹۸	۰,۶۱۷	
	واریانس‌ها برابر		۰,۴۸۷	۱۵۰,۱۲۰	۰,۶۲۷	

برای بررسی آزمون فرض فوق ابتدا باید فرض هم‌واریانس بررسی شود. آماره آزمون لیون برابر است با $F = 1.390$ و با توجه به مقدار $Sig = 0.240$ چون $Sig > 0.05$ فرض H_0 را در سطح معنی داری $\alpha = 0.05$ رد نمی‌شود یعنی توزیع متغیر استفاده در گروه مردان و زنان هم‌واریانس هستند و به همین خاطر از آماره

- t استودنت استفاده شد. ملاحظه می‌شود مقدار آماره آزمون t استودنت برابر است با $t=0.501$ و با توجه به مقدار $Sig=0.617$ چون $Sig > 0.05$ فرض H_0 را در سطح معنی‌داری $\alpha=0.05$ رد نمی‌کنیم. می‌توان گفت سطح متغیر استفاده بین معلمان زن و مرد برابر است به عبارت دیگر جنسیت روی میزان استفاده از اصول سنجش اصیل معلمان اثرگذار نبوده است.

برای بررسی اثر مدرک تحصیلی بر "میزان استفاده از سنجش اصیل" با توجه به اینکه تعداد معلمان دارای مدرک تحصیلی کاردانی و دکتری کمتر از ۳۰ است قضیه حد مرکزی در اینجا صدق نمی‌کند پس باید فرض نرمال بودن داده‌های مربوط به متغیر "استفاده" را در برای مدرک تحصیلی بررسی شود و اگر داده‌ها برای هر چهار گروه نرمال باشند می‌توان از تحلیل واریانس یک‌طرفه استفاده کرد. حال توصیفی از متغیر "استفاده" در چهار گروه مدرک تحصیلی را نشان داده می‌شود.

جدول ۱۹: آماره‌های توصیفی متغیر میزان استفاده به تفکیک مدرک تحصیلی

گروه	تعداد	میانگین	انحراف معیار	کمترین	میانه	بیشترین	چولگی
کاردانی	۵	۱۷۶۰٫۶	۲۲٫۱	۱۴۵	۱۷۶	۱۹۳	-۰٫۳۰۶
کارشناسی	۱۳۳	۱۷۱٫۴	۲۹٫۴	۶۵	۱۶۸	۲۴۳	-۰٫۱۵۱
کارشناسی ارشد	۵۶	۱۷۶٫۸	۲۴٫۹	۱۰۵	۱۷۴	۲۳۹	-۰٫۱۹۶
دکتری	۶	۱۷۶٫۳	۳۳٫۳	۱۲۴	۱۸۰	۲۰۸	-۰٫۶۷۴

برای بررسی نرمال بودن داده‌ها در چهار گروه مدرک تحصیلی از آماره آزمون شاپیرو-ویلک استفاده شد:

جدول ۲۰: آزمون نرمال بودن داده‌های متغیر میزان استفاده به تفکیک مدرک تحصیلی

مدرک تحصیلی	آماره شاپیرو-ویلک	درجه آزادی	سطح معنی‌داری
کاردانی	۰٫۸۷۶	۵	۰٫۲۹۰
کارشناسی	۰٫۹۹۰	۱۳۳	۰٫۴۳۶
کارشناسی ارشد	۰٫۹۷۸	۵۶	۰٫۴۰۸
دکتری	۰٫۸۹۸	۶	۰٫۳۶۴

با مشاهده مقادیر آماره آزمون شاپیرو-ویلک سطوح معنی‌داری متناظر با توجه به اینکه مقادیر سطوح معنی‌داری همگی بیشتر از ۰٫۰۵ هستند می‌توان گفت که توزیع متغیر استفاده برای هر چهار مقطع تحصیلی نرمال است. بنابراین می‌توان از آزمون ANOVA استفاده کرد که نتایج در زیر آمده است:

جدول ۲۱: آماره‌های تحلیل واریانس متغیر ادراک برای مدرک تحصیلی

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون F	Sig
بین گروهی	۱۲۳۲٫۳۷۷	۳	۴۱۰٫۷۹۲	۰٫۵۱۸	۰٫۶۷۱
درون گروهی	۱۵۵۵۲۶٫۲۱۸	۱۹۶	۷۹۳٫۵۰۱		

			۱۹۹	۱۵۶۷۵۸,۵۹۵	کل
--	--	--	-----	------------	----

با توجه به مقدار آماره فیشر $F=0.518$ و مقدار سطح معنی‌داری متناظر $Sig=0.671$ چون مقدار سطح معنی‌داری از 0.05 بزرگ‌تر است پس فرض صفر رد نمی‌شود. میزان استفاده معلمان از سنجش اصیل با مدرک تحصیلی کاردانی، کارشناسی، کارشناسی ارشد و دکتری با هم برابر است و تفاوت معنی‌داری بین آن‌ها مشاهده نمی‌شود.

برای بررسی اثر نوع استخدام بر میزان استفاده از سنجش اصیل نیز با توجه به اینکه تعداد معلمان حق‌التدریس کمتر از ۳۰ است و قضیه حد مرکزی صدق نمی‌کند باید فرض نرمال بودن داده‌های مربوط به متغیر استفاده را برای نوع استخدام بررسی کنیم و اگر داده‌ها برای هر چهار گروه نرمال باشند می‌توان از تحلیل واریانس یک‌طرفه استفاده کرد در ادامه توصیفی از متغیر "میزان استفاده" در چهار گروه نوع استخدام را نشان داده می‌شود:

جدول ۲۲: آماره‌های توصیفی متغیر میزان استفاده به تفکیک نوع استخدام

گروه	تعداد	میانگین	انحراف معیار	کمترین	میان	بیشترین	ضریب چولگی
پیمانی	۵۳	۱۷۵,۰	۲۶,۶	۱۲۰	۱۷۵	۲۳۸	-۰,۰۵۴
رسمی آزمایشی	۶۲	۱۷۰,۳	۲۹,۷	۶۵	۱۷۱	۲۴۳	-۰,۵۹۹
رسمی قطعی	۷۴	۱۷۳,۴	۲۸,۷	۱۰۵	۱۷۰	۲۳۹	۰,۰۷۵
حق‌التدریس	۱۱	۱۷۵,۰	۲۳,۲	۱۴۰	۱۷۱	۲۲۱	۰,۵۳۱

برای بررسی نرمال بودن داده‌ها در چهار گروه نوع استخدام از آماره آزمون شاپیرو-ویلک استفاده شد:

جدول ۲۳: آزمون نرمال بودن داده‌های متغیر میزان استفاده به تفکیک نوع استخدام

مدرک تحصیلی	آماره شاپیرو-ویلک	درجه آزادی	سطح معنی‌داری
پیمانی	۰,۹۹۰	۵۳	۰,۹۲۴
رسمی آزمایشی	۰,۹۶۶	۶۲	۰,۰۸۳
رسمی قطعی	۰,۹۸۹	۷۴	۰,۷۶۵
حق‌التدریس	۰,۹۶۸	۱۱	۰,۸۷۰

با مشاهده مقادیر آماره آزمون شاپیرو-ویلک سطوح معنی‌داری متناظر با توجه به اینکه مقادیر سطوح معنی‌داری همگی بیشتر از 0.05 هستند می‌توان گفت که توزیع متغیر استفاده برای هر چهار نوع استخدامی نرمال است. بنابراین، می‌توان از آزمون ANOVA استفاده کرد که نتایج در زیر آمده است:

جدول ۲۴: آماره‌های تحلیل واریانس متغیر ادراک برای نوع استخدام

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون F	Sig
بین گروهی	۷۴۷,۴۲۴	۳	۲۴۹,۱۴۱	۰,۳۱۳	۰,۸۱۶
درون گروهی	۱۵۶۰۱۱,۱۷۱	۱۹۶	۷۹۵,۹۷۵		
کل	۱۵۶۷۵۸,۵۹۵	۱۹۹			

با توجه به مقدار آماره آزمون فیشر $F=0.313$ و مقدار سطح معنی‌داری متناظر $Sig=0.816$ از 0.05 بزرگ‌تر است پس فرض صفر رد نمی‌شود. میزان استفاده معلمان پیمانی، رسمی آزمایشی، رسمی قطعی و حق‌التدریس از سنجش اصیل باهم برابر است و تفاوت معنی‌داری بین آن‌ها مشاهده نمی‌شود.

در بررسی رابطه بین سابقه کار و استفاده معلمان از سنجش اصیل فرض زیر را در نظر گرفته شد:

بین میزان استفاده از سنجش اصیل معلمان و سابقه کار آنان رابطه خطی وجود ندارد: H_0 :

بین میزان استفاده از سنجش اصیل معلمان و سابقه کار آنان رابطه خطی وجود دارد: H_1 :

جدول ۲۵: ضریب همبستگی بین سابقه کار و میزان استفاده از سنجش اصیل

استفاده	ضریب همبستگی پیرسن	تعداد	سطح معنی‌داری
سابقه کار	۰,۰۱	۲۰۰	۰,۹۹۵

با توجه به این‌که آماره آزمون برابر است با $r = 0.01$ و مقدار سطح معنی‌داری برابر است با $Sig=0.995$ و با توجه به این‌که $Sig>0.05$ پس فرض H_0 را در سطح پنج درصد رد نمی‌شود یعنی اینکه بین استفاده معلمان از سنجش اصیل و سابقه کار معلمان رابطه خطی وجود ندارد.

بحث و نتیجه‌گیری

سطح ادراک دبیران دبیرستان‌های شهر سنندج از مفهوم سنجش اصیل بیش‌ازحد متوسط بود و با یافته‌های حاصل از پژوهش‌های (Naser et al., 2017) که نگرش و درک مدرسان از سنجش اصیل را قابل توجه بیان کردند. میزان استفاده معلمان نمونه انتخابی شهر سنندج از اصول سنجش اصیل نیز بیش‌ازحد متوسط بود و با یافته‌های حاصل از پژوهش (Mohammad Seifodin (2020 همسویی داشت، او در مطالعه خود در آندونزی به این نتیجه دست‌یافت که استفاده از سنجش اصیل در بین معلمان بیشتر از حد متوسط است، این نتایج همچنین با پژوهش‌های (Naser et al و (Zu'bi 2013 نیز همسویی دارد.

نتایج به‌دست‌آمده بیانگر این بود که سطح ادراک و استفاده از سنجش اصیل معلمان زن و مرد باهم برابر بود یعنی جنسیت روی آن تأثیری نداشته است. نتایج این یافته نیز با نتایج یافته‌های (Naser 2017) et al که نشان دادند جنسیت تأثیری بر روی ادراک و میزان استفاده مدرسان از سنجش اصیل ندارد همسویی دارد. سطح ادراک و میزان استفاده از سنجش اصیل نیز در معلمان ناحیه یک و دو شهر سنندج باهم برابر بود و بیانگر این است که ناحیه روی سطح ادراک و میزان استفاده از سنجش اصیل تأثیری نداشته است. سطح ادراک و استفاده معلمان از سنجش اصیل در معلمان دارای مدرک تحصیلی کاردانی، کارشناسی، کارشناسی ارشد و دکتری و معلمان با شیوه‌های مختلف استخدام باهم برابر بود. یعنی مدرک تحصیلی و نوع استخدام معلمان روی سطح ادراک و میزان استفاده از سنجش اصیل تأثیری نداشت. در بررسی میزان سابقه کار معلمان نتیجه از آن است که بین سابقه کار معلمان و سطح ادراک و میزان استفاده از سنجش اصیل آنان رابطه خطی معنی‌داری وجود ندارد یعنی میزان سابقه کار معلمان روی سطح ادراک و میزان استفاده از سنجش اصیل تأثیری نداشته است. نتایج این یافته نیز با پژوهش (Naser et al (2017

که اظهار کردند سابقه کار و تجربه تأثیری بر روی نگرش و استفاده اساتید از سنجش اصیل ندارد همسویی دارد.

معلمان فارغ از جنسیت، سابقه کار و ناحیه خدمت و نوع استخدام در یک محیط (شهر به طور مثال) زندگی می کنند و در معرض تجربه های یکسان، آموزش ها و دوره های مشابه هستند، بنابراین نگرش و درک یکسان از سنجش اصیل دارند و میزان استفاده ی آنان نیز مشابه است و این متغیرها تأثیری بر روی دیدگاه های آنان نسبت به سنجش اصیل ندارند.

پرسشنامه نخست یعنی میزان استفاده معلمان از سنجش اصیل هشت عاملی و دارای ۵۱ گویه بود. بر اساس نتایج تحلیل عاملی اکتشافی همخوان با الگوی پنج بعدی (Gulikers et al., 2004) یک مدل هشت عاملی که همپوشانی با ابعاد الگوی ذکر شده را دارد و عامل های معیارهای شفاف، خودارزیابی، یادگیری اصیل، تکالیف فکورانه، الگودهی، استانداردهای شناخته شده، توجیه یادگیری، و تکالیف جایگزین را شامل می شود استخراج گردید. عامل ها و الگوی استخراج شده با مفاهیم و الگوی پنج بعدی ارائه شده سنجش اصیل توسط گولیکرز و همکاران مشابه بود:

۱. عامل اول، تحت عنوان معیارهای شفاف دارای (۲۰ گویه) است. آنچه Wiggins (1998) آن را اصیل می داند در وهله نخست شامل چهار معیار است. نخست، کارهایی که واقعاً معرف عملکرد در زمینه و بافت یاددهی - یادگیری معیارهای مورد استفاده در سنجش باشد. گویه: تکالیفی که به دانش آموزان ارائه می دهد همواره در محیط شبیه سازی شده واقعی انجام می شوند مصداقی از این عملکرد در زمینه است. دوم، خودارزیابی بیشتری نسبت به آزمون های مرسوم صورت گیرد، به طور مثال گویه: دانش آموزان برای بخش های مهم کارشان خودارزیابی ارائه می دهند و گویه: به دانش آموزان شیوه خودارزیابی و تطبیق با معیارها و عملکرد آموزش داده می شود نمونه از این مورد هستند. سوم، دانش آموز کار خود را شفاهی یا کتبی ارائه دهد. در زندگی واقعی افراد معمولاً می دانند در چه معیارهای عملکردشان قضاوت می شود. بعضی معیارها باید مربوط به یک نتیجه واقعی، توضیح ویژگی ها یا احتیاجات محصول، عملکرد یا راه حل هایی که دانش آموزان خلق می کنند، و درگیر توسعه مهارت های حرفه ای مرتبط در موقعیت های واقعی زندگی باشند (Gulikers et al., 2004) گویه های که می توانند این موارد را بیان کنند عبارت اند از: من دانش آموزانم را بر اساس ارائه پذیرفته شده ارزیابی می کنم و گویه: تأکید بر آفرینش یک محصول، اجرای یک پروژه یا فعالیت توسط خود دانش آموز گویه: مستندسازی و به نام خود ثبت کردن پروژه ها توسط خود دانش آموز از همان ابتدا. در نهایت دانش آموز باید از کار خود چه شفاهی یا کتبی دفاع کند، گویه: که تکالیف انجام شده دانش آموزان را به کمک خودشان می خوانیم و نقد می کنم شامل دفاع دانش آموز از کار خود می شود. در واقع می توان نتیجه گرفت که عامل های معیارهای شفاف می تواند ابعاد الگوی گولیکرز و همکارانش را پوشش دهد، یعنی می توان اظهار داشت که بین معیارهای شفاف و ابعاد پنج گانه گولیکرز و همکارانش ارتباط وجود دارد. گولیکرز و همکارانش نیز رابطه متقابل بین معیارهای سنجش و چهار بُعد دیگر الگوی

خود را تأیید می‌کنند که به‌طور مثال، اگر بافت فیزیکی زمان پنج ساعت را برای اتمام یک تکلیف در نظر بگیرد یک معیار باید این باشد که تکلیف در مدت پنج ساعت به اتمام برسد. این عامل همچنین دارای دو خرده مقیاس صداقت علمی و سبک و بیان است. گویه: دانش‌آموزان من باید صداقت علمی خود را در انجام تکالیف و پروژه‌ها نشان دهند. گویه: دانش‌آموزان من می‌دانند که عادت‌های ذهنی آن‌ها را ارزیابی می‌کنم اشاره‌ای به خود سنجش و تفاوت قائل شدن بین آن چیزی است که یادگیرندگان می‌دانند و آن چیزی است که جعل می‌کنند. وادار ساختن دانش‌آموزان به ثبت و مستند کردن پروژه‌ها از آغاز نیز به همین دلیل است. همچنین آموزش باید سبک و بیان دانش‌آموز را رشد دهد و نتایج کار یادگیرندگان اگر اشتباهات دستوری یا روش‌شناختی داشته باشد یا اگر کسل‌کننده، نامرتب، بی‌روح و عاری از خلاقیت دانش‌آموز باشد برگشت می‌خورد. نقد همکلاسی‌ها و سنجش همتایان در اینجا کمک‌کننده است.

۲. عامل دوم، تحت عنوان خودارزیابی دارای (هشت گویه) است. بر اساس مبانی نظری سنجش اصیل برای خودارزیابی دانش‌آموز از عملکرد خود، نظام سنجش باید شفاف، واضح و غیر اختیاری باشد (گویه: نظام ارزیابی من برای دانش‌آموزانم شفاف است و گویه: دانش‌آموزان می‌دانند که نظام سنجش من وزارتی و غیر اختیاری است). برای شفاف بودن ارزیابی می‌توان از سنجش‌ها، مقیاس‌های رتبه‌بندی و چک‌لیست‌ها استفاده کرد. خودارزیابی باید در مرکزیت قرار بگیرد و رویدادی متداول و عادی در نظام سنجش اصیل باشد. تکالیفی نیز که به دانش‌آموزان ارائه می‌شود باید شبیه مسائل و مشکلاتی باشد که افراد واقعی در زندگی واقعی با آن‌ها روبرو می‌شوند. می‌توان گفت که عامل خودارزیابی با معیارهای سنجش و زمینه فیزیکی و اجتماعی الگوی پنج بعدی گولیکرز مرتبط است.

۳. عامل سوم، تحت عنوان یادگیری اصیل دارای (۵ گویه) است، که از دانش‌آموزان می‌خواهد از بازگویی صرف دیدگاه‌های متداول و فزونی که آموخته‌اند بدون تفکر و خارج از بافت خودداری کنند که همسو با نظر (Resnick, 1987) است که تأکید می‌کند یادگیری و عملکرد خارج از مدرسه اغلب در یک نظام اجتماعی اتفاق می‌افتد. بنابراین، یک الگو برای سنجش اصیل باید فرایندهای اجتماعی را در نظر بگیرد که در زمینه‌های واقعی وجود دارند این گویه‌ها با بافت اجتماعی الگوی پنج‌بعدی گولیکرز همخوان است. عامل یادگیری اصیل همچنین از دانش‌آموزان می‌خواهد از تولیدات خود (چه یک مقاله یا گزارش آزمایشگاه باشد) دفاع کنند و به انتقادات معلم و همکلاسی‌هایشان از تکالیفشان پاسخ دهند، این گویه‌ها نیز با بعد شکل و نتیجه سنجش گولیکرز و همکارانش همسو است که به نقل از (Wiggins, 1989) اظهار می‌دارد که مهم است دانش‌آموز برای اطمینان از این که تسلطش واقعی است از کارش دفاع کند. عامل یادگیری اصیل همچنین به نیاز به افرادی به‌عنوان الگوهای موفق برای دانش‌آموزان اشاره کرده است.

۴. عامل چهارم، تحت عنوان تکالیف فکورانه که شامل (۵ گویه) است. (Gulikers et al, 2004) بیان کردند که تکالیف باید دانش‌آموزان را در انجام فعالیت‌هایی که در عمل حرفه‌ای انجام می‌شوند درگیر کنند و آن‌ها را ملزم به ادغام دانش، مهارت‌ها و نگرش‌ها همان‌گونه که افراد حرفه‌ای انجام می‌دهند، کنند. گویه‌های تکالیف فکورانه همسو با مفهوم تکالیف در الگوی پنج‌بعدی گولیکرز و همکارانش هستند. این گویه‌ها بیان می‌کنند که تکالیف نیازمند خلق اثر، توجیه و عملکرد باکیفیت هستند، بر اساس مهارت‌ها و

موقعیت‌های قابل پیش‌بینی تنظیم می‌شوند. نیاز به قضاوت و نوآوری دارند. همچنین تکالیف دربرگیرنده کارهای ضروری تکرارشونده هستند و باید به صورت مستقیم به سراغ نتیجه دلخواه بروند. این تکالیف چالش‌های یکپارچه‌ای هستند که نیازمند استفاده خلاقانه از دانش هستند. تکالیف می‌توانند نقش کلیدی در رشد دانش‌آموزان به عنوان شهروندانی فعال و مفید از طریق پرورش انتقادپذیری، تفکر متفاوت، پویایی و انگیزه برای پرسش از مفروضات رایج، خلاقیت و احساس غرور و ارزش در درگیر شدن عمیق با دانش پیچیده یا مشارکت‌های بالقوه به صلاح اجتماع داشته باشد (MacArthur, 2021).

۵. عامل پنجم، الگودهی شامل ۶ گویه است. گویه‌های این عامل بیان می‌کنند که دانش‌آموزان به فرصتی برای تقلید و سرمشق‌گیری از الگوهای موفق نیاز دارند و محصولاتی را باید تولید کنند که از فعالیت‌های یادگیری قبلی آنان الگوبرداری شده است. همچنین باید با دانش‌آموزان بر اساس دسترسی به الگوها و بازخوردها در یادگیری با آنان برخورد شود. دانش‌آموزان الگوها را به طور مثال برای نوشتن مقاله و یا اجرای ضبط‌شده را از قبل دیده باشند. این عامل‌ها همخوان با بُعد زمینه اجتماعی الگوی گولیکرز و همکاران است که اظهار کرده‌اند که تکالیف سنجش اصیل باید شامل فرایندهای اجتماعی که معادل شرایط و موقعیت‌های موجود در زندگی واقعی‌اند باشند (با چه کسی باید آن‌ها را انجام داد)، زمینه اجتماعی اصالت سنجش را تحت تأثیر قرار می‌دهد و در زندگی واقعی کار کردن باهم به جای استثناء یک قاعده است.

۶. عامل ششم، استانداردهای شناخته‌شده شامل ۲ گویه است که تأکید می‌کنند ارزیابی دانش‌آموزان باید بر پایه معیارها و استانداردهای شناخته‌شده باشد و این که از معیارهای اصیل طبق گفته (1993) Wiggins می‌تواند این باشد که به درستی خودارزیابی دانش‌آموز نمره اختصاص داده شود. این عامل همچنین با بُعد معیارهای سنجش الگوی گولیکرز و همکاران مرتبط است.

۷. عامل هفتم، توجیه یادگیری شامل ۲ گویه است که از دانش‌آموزان می‌خواهد درک و مهارت خود را از آنچه یاد گرفته‌اند توجیه کنند و همچنین تکالیفی که به دانش‌آموزان ارائه می‌شود باید پیوندی قوی با دنیای واقعی داشته باشند.

۸. عامل هشتم، تکالیف جایگزین شامل ۲ گویه است که در واقع می‌توان آن‌ها را جزء عامل تکالیف فکورانه دانست. گویه‌های تکالیف جایگزین اشاره به امکان ارائه گونه‌های جایگزین برای انجام یک تکالیف اشاره دارند. همچنین عنوان می‌کنند که می‌توان فهرستی از سؤالاتی که احتمال می‌رود برای امتحان انتخاب شوند در اختیار دانش‌آموزان قرار داد. البته باید به این نکته دقت کرد که منظور از ارائه سؤالات قبل از امتحان این نیست که همانند آزمون‌های سنتی تعدادی سؤال چندگزینه‌ای و یا انتخاب پاسخ در اختیار دانش‌آموزان قرار داد، چرا که سنجش اصیل در نقد به این گونه آزمون‌ها به وجود آمد.

نتایج حاکی از مشابهت مفهومی میان این پرسشنامه و الگوی پنج بعدی Gulikers et al است. گویه‌های پرسش‌نامه همخوانی لازم با این الگو را دارند. بر اساس نتایج مطالعه می‌توان نتیجه گرفت که در مجموع پرسشنامه ۵۱ گوی‌های میزان استفاده معلمان از سنجش اصیل از اعتبار و روایی مناسب برای ارزیابی میزان استفاده از سنجش اصیل در جامعه ایرانی برخوردار است.

پرسشنامه دوم ادراک معلمان از سنجش اصیل دارای ۷ عامل شامل شفاف‌سازی، معیارهای واقعی، اصالت، عمل‌گرایی، واقع‌گرایی، چالش‌انگیزی و کاربردی بودن است. این پرسشنامه درک معلمان از مفهوم سنجش اصیل را می‌سنجید. در تحلیل داده‌ها میزان ادراک معلمان از ارزش‌یابی اصیل بالاتر از حد متوسط برآورد شد. اما با دقت در پاسخ‌های گویه‌های پرسشنامه می‌توان به این استنباط رسید که آشنایی معلمان با مفهوم سنجش اصیل در مواردی از مبانی نظری که بدیهی به نظر می‌رسند کافی و مناسب نیست، برای مثال در عامل نخست (شفاف‌سازی) در پاسخ به گویه: آزمون‌های من شامل بیان دوباره، به خاطر سپاری، و کپی کردن دانش است، ۵۷،۵ درصد از معلمان گزینه موافق و کاملاً موافق را انتخاب کرده‌اند و فقط ۱۶،۵ درصد از معلمان مخالف بوده‌اند. همچنین در پاسخ به گویه: آزمون‌های من باید سؤالاتش تا موقع امتحان مخفی نگه‌داشته شوند نیز همانند مورد قبل ۵۷،۵ درصد از معلمان با آن موافق و کاملاً موافق بوده‌اند درحالی‌که این موارد در آزمون‌های سنتی به کار گرفته می‌شوند و سنجش اصیل با آن‌ها موافق نیست.

عامل دوم (معیارهای واقعی)، در پاسخ به گویه: ارزیابی از عملکرد دانش‌آموز باید تمرکز بر مهارت‌های پیشرفته و سطح بالاتر باشد فقط ۲۶ درصد از معلمان موافق و کاملاً موافق بوده‌اند و ۵۳،۵ درصد از آنان مخالف و کاملاً مخالف بوده‌اند و در پاسخ به گویه: ارزیابی از عملکرد دانش‌آموز باید تمرکز بر همواره بر روی صلاحیت‌های مرتبط با دنیای کار باشد ۳۵ درصد از معلمان با این گویه موافق و کاملاً موافق بوده‌اند و ۳۰ درصد از آنان نیز مخالف و کاملاً مخالف بوده‌اند. با توجه به پاسخ‌ها، مشاهده می‌کنیم که معلمان با این دو گویه که جزء اصول و مبانی سنجش اصیل هستند بیشتر از این‌که موافق باشند مخالف هستند. درحالی‌که با گویه: آزمون‌هایی که از دانش‌آموزان می‌گیرم فقط به پاسخ‌های صحیح نیاز دارند و گویه: آزمون‌هایی که از دانش‌آموزان می‌گیرم پاسخ‌هایشان شامل یک حدس یا گمان هستند و گویه: آزمون‌هایی من برای این‌که معتبر نمره‌گذاری شوند آسان می‌شوند که در تقابل با مفاهیم سنجش اصیل هستند به ترتیب بیشتر از ۶۸ و ۴۷ و ۵۸،۵ درصد اعلام موافقت کرده‌اند و کمتر از ۲۰ درصد مخالفت کرده‌اند. می‌توان استنباط کرد که معلمان شرکت‌کننده با اصول و مفاهیم اولیه سنجش اصیل آشنایی کافی و کاملی ندارند. در عامل چهارم (ارزیابی مبتنی بر حرفه آینده) نیز در پاسخ به گویه معکوس: ارزیابی عملکرد دانش‌آموز باید از کار و وظایف واقعی افراد در جامعه متفاوت باشد، ۷۱،۵ درصد از پاسخ‌دهندگان با این گویه موافق و کاملاً موافق بوده‌اند و فقط چهار درصد از آنان مخالف و کاملاً مخالف بوده‌اند. در عامل پنجم (واقع‌گرایی) در پاسخ به گویه: ارزیابی از عملکرد دانش‌آموزان باید میزان درک و ارائه دانش آن‌ها را در موقعیت‌های واقعی زندگی بسنجد فقط ۱۴،۵ درصد با این گویه موافق و کاملاً موافق بوده‌اند و ۷۸ درصد از آنان مخالف و کاملاً مخالف بوده‌اند. و در عامل ششم (چالش‌انگیزی) نیز در پاسخ به گویه: ارزیابی از عملکرد دانش‌آموزان باید شامل پاسخ به پرسش‌هایی باشد که هیچ‌وقت در دنیای واقعی با آن برخورد نمی‌کنند ۵۵،۵ درصد اعلام کرده‌اند که با این گویه معکوس موافق و کاملاً موافق هستند. درحالی‌که هدف سنجش اصیل تدوین آزمون‌های استاندارد است که خروجی‌های پیچیده، قابل توجه و مرتبط با زندگی واقعی را سنجش کند. این آزمون‌ها عملکرد یادگیرندگان را در فعالیت‌های زندگی واقعی و به‌طور مستقیم

مورد ارزیابی قرار می‌دهند. (Wiggins 1998) بیان کرده که سنجش زمانی اصیل و واقعی است که بجای یادآوری و پاسخ دادن صرف به سؤالات آسان به آنچه افراد واقعی به‌طور معمول در زندگی انجام می‌دهند تغییر یابد. می‌توان نتیجه گرفت که نیاز به آموزش‌های بیشتر با برگزاری کارگاه‌های آموزشی و آشناسازی معلمان و همچنین دانش‌آموزان و والدین آن‌ها با مفاهیم، اصول و استانداردهای سنجش اصیل برای به‌کارگیری پرسشنامه‌ها، اجرای سنجش اصیل و حصول نتایج بهتر ضروری به نظر می‌رسد.

در اجرای سنجش اصیل باید در نظر داشت که به همکاری گسترده و فعال دانش‌آموزان در این امر نیاز است. گویه‌های پرسشنامه میزان استفاده معلمان از سنجش اصیل از نُه اصل (Wiggins 1993) استخراج گردیدند. با توجه به مفاهیم اصول نه‌گانه به ترتیب می‌توان گفت که این اصول راجع به دفاع از محصول یادگیری، در نظر گرفتن فراگیر به‌مثابه هنرمند آزاداندیش کارآموز، معیارهای سنجش شفاف، خودارزیابی دانش‌آموز، عاملیت دانش‌آموز، سبک و بیان دانش‌آموز، پرسش‌گری دانش‌آموز، ماهیت نسبی دانش و صداقت علمی دانش‌آموز هستند، که می‌توان این ۹ اصل را تجمیع و به ۳ عامل اصلی: دانش‌آموز، معلم، و ماهیت دانش دسته‌بندی کرد. درواقع، می‌توان نتیجه گرفت که بخش اصلی سنجش اصیل به دانش‌آموز و خودارزیابی دانش‌آموزان و یا ارزیابی از هم‌کلاسی‌ها (سنجش همتایان) برمی‌گردد. معلم به‌عنوان عامل دیگر در سنجش اصیل به راهنمایی دانش‌آموزان (به‌طور مثال با استفاده از معیارهای سنجش) برای پیشرفت در فرایند یادگیری می‌پردازد و هدفش تشویق دانش‌آموز برای مشارکت فعالانه است. در این فرایند دانش‌آموزان با عملکرد گذشته خود و نه دیگر دانش‌آموزان مقایسه می‌شوند و بر روی نقاط قوت و آنچه دانش‌آموز می‌تواند انجام بدهد تأکید می‌شود. عامل بعدی ماهیت دانش است که اشاره به نسبی بودن و عدم قطعیت علم و نظام‌های دانش دارد و از معلم می‌خواهد که دانش‌آموزان را راهنمایی کند تا از مطلق‌گرایی هر رشته علمی خاص اجتناب کنند و آن‌ها را با موضوعات مورد اختلاف در رشته‌های علمی آشنا سازد. یکی از اشارات کاربردی این پژوهش آن است که اجرای سنجش اصیل به دلیل ناآشنایی معلمان، دانش‌آموزان و والدین باید با احتیاط انجام شود. دانش‌آموزان، معلمان و والدین از لحاظ ذهنی می‌بایست آمادگی و توجیه مقدماتی برای پذیرش آن را کسب کنند.

References

- Allen, J. P., Frey, B. B., & Schmitt, V. L. (2012). Defining Authentic Classroom Assessment. *Practical Assessment, research & Evaluation*, 17(2), 18.
- Arnold, L. (2019). Advancing Authentic Assessment Practice. Arise lecture, 27/11/2019, Edinburgh Napier University <https://staff.napier.ac.uk/services/dlte/Pages/DLTE-Past-Events-201920.aspx>.
- Ashford_ Rowe, k., Herrington, J., & Brown, c. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, 39(2), 205-222.
- Ashwin, P. (2020). *Transforming university education: a manifesto*. Bloomsbury Publishing.

Biggs, J. B., & Tang, C.S. (2007). *In: Teaching for Quality Learning at University: What the Students Does*, third ed. McGraw-Hill/Society for Research in to Higher Education & Open University Press, Maidenhead.

Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education* 15(1), 101_111

Boud, D., & Associates. (2010). Assessment 2020: Seven propositions for the assessment reform in higher education.

Boud, D., 1995. Assessment and learning: contradictory or complementary? In: knight, P. (Ed.), *Assessment for Learning in Higher Education*. Kogan Page, London, p.189.

Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing Student Learning in Higher Education*. London, New York: Routledge.

Brown, S. (2019). Using assessment and feedback to empower students and enhance their learning, in Bryan, C. and Clegg, K. (eds), *Innovative Assessment in Higher Education*. 2nd edn. Oxford: Routledge, pp. 50–63.

Center for Innovative Teaching and Learning. (2021). Authentic assessment. Available online. Indian University Bloomington <http://citl.indiana.edu/>

Davison, G. (2011) Investigating the relationships between authentic assessment and the development of learner autonomy. University of Northumbria at Newcastle (United Kingdom).

Gulikers, J. T. M. (2006). Authenticity is in the eye of the beholder: Beliefs and perceptions of authentic assessment and the impact on student learning. Interuniversity Centre for Educational Research, pp. 1-160.

Gulikers, J. T. M., Bastiaens, Th. J., & Kirschner, P.A. (2006). Authentic assessment, student and teacher perception: the practical value of the five-dimensional framework. *Journal of vocational Education and Training*, 58, 337_357

Gulikers, J.T.M., Bastiaens, T.J., & Kirschner, P.A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Developmant*, 52(3).67_86.

Hoy, A, W., & Hoy, K. W. (2013). *Instructional Leadership: A Research Based Guide to Learning in Schools*. Translated by: Nader Soleimami, Mahmmoud Safari, Morteza Nazari (2017). Tehran: Arasbaran. [In Persian]

Kamens, D., & Benavot, A. (2011). National, reginal and international learning assessment: Trends among developing countries, 1960_2009. *Globalization, societies and Education*, 9(2), 285_300

Kinash, S., Mc Gillivray, L., & Crane, L. (2018). Do university students, alumni, educators and employers link assessment and graduate employability? *Higher Education Research & Development*, 37(2),301_315

Looney, A., Cumming, J., Van Der Kleij, F., & Harris, K. (2018). Reconceptualizing the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy & Practice*, 25(5), 442_467

Lund, J. (1997). Authentic assessment: Its development & applications. *Journal of Physical Education, Recreation & Dance*, 68(7), pp.25-28.

Lundahl, B. (2009). Engelsk språkdiraktik – Texter, kommunikation, språkutveckling. Lund: Studentlitteratur.

- Mac Tighe, J., & Ferrara, S. (2021). *Assessing student learning by design: principles and practices for teachers and school leaders*. New York: Teachers College Press
- McArthur, J. (2021). For assessment to count as authentic it must mean something to students. [WONKHEhttps://wonkhe.com/blogs/forassessment-to-count-as-authentic-it-must-mean-something-to-students/](https://wonkhe.com/blogs/forassessment-to-count-as-authentic-it-must-mean-something-to-students/)
- Mueller, J., (2005). The authentic assessment toolbox: enhancing student learning through online faculty development, *Journal of Online Learning and Teaching* 1(1),1_7
- Naser, I.A.S., Arman, I.M.A., & Nafi, J.S.I, (2017).The Degree Of Faculty's Use Of Authentic Assessment Tools At AL_Quds Towards Them, *International Journal Of Education, Learning and Development*. Vol.5. No.7, PP.44_63.
- Ramsden, P. (2003). *Learning to Teach in Higher Education*, 2nd edition. Abingdon: Routledge.
- Sambell, K., McDowell, L., & Montgomery, C. (2013). *Assessment for learning in higher education*. Abingdon: Routledge.
- Sambell, K., McDowell, L. and Brown, S. (1997) "But is it fair?": an exploratory study of student perceptions of the consequential validity of assessment. *Studies in educational evaluation*, 23(4), pp.349-371.
- Segers, M., Dochy, F., & Cascallar, E. (2003). The era of assessment engineering: changing perspectives on teaching and learning and the role of new model of assessment, in: M. Sergers, F. Dochy & Cascallar E. (Eds.) *Optimizing new modes of assessment: In search of qualities and standards*
- Shinn, C. (2012). Teacher education reform in Palestine: policy challenges amid donor expectations. *Comparative Education Review*, 56(4), 608_633
- The assessment National Team. (2004). The strategies of Evaluation and tools, Administration Exams and tests, Jordan, *The Ministry of Education*.
- Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 43(5), pp.840-854
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 70,703_713
- Wiggins, G. (1993). *Assessing student performance*. San Francisco; Jessey _ Bass, pp.47_67
- Wiggins, G. (1998). *Educative assessment: Designing assessment to inform and improve student performance*. San Francisco: Jossey Bass
- Zubi, A. A. (2013). Knowledge and Utilization Degree of Authentic Assessment Strategies and Tools by Mathematical Teachers. *Journal of the Islamic University for Education and Psychology*, 12(3): 165-197. Retrieved from <http://www.iugaza.edu.ps/ar/period>