



بررسی استفاده از روش‌های مختلف سنجش و سنجش‌های مبتنی بر یادگیری پایدار

پوراندهخت پورنادری¹، احمدرضا نصر²، محمدرضا نیلی³ و زهره باقری⁴

¹ کارشناس ارشد، برنامه‌ریزی آموزشی، دانشگاه اصفهان (نویسنده مسئول)، پست الکترونیکی: arnasr@edu.ui.ac.ir

² استناد، دانشکده علوم تربیتی و روانشناسی دانشگاه اصفهان

³ استادیار، دانشکده علوم تربیتی، دانشگاه اصفهان

⁴ کارشناس، مطالعات و برنامه‌ریزی آموزشی، دانشگاه اصفهان

چکیده: آموخته‌ها برای اعطای مجوز صلاحیت به دانشجویان و افزایش یادگیری‌های آنان از جمله اهداف سنجش در آموزش عالی است. اخیراً صاحب نظران بر این باورند که این اهداف باید در جهت توجه و تأکید بیشتر به یادگیری‌های پایدار در طول زندگی توسعه یابد. انتخاب روش‌هایی برای سنجش دانشجویان که بتواند اهداف مختلف سنجش را همزمان و به‌طور اثر بخش مورد توجه قرار دهد و محقق کند، از جمله چالش‌های سنجش است. این مقاله ضمن بررسی انواع شیوه‌های متداول سنجش دانشجویان، میزان تحقق اهداف یادگیری پایدار دانشجویان به کمک این شیوه‌ها را مورد ارزیابی قرار داده است. پژوهش از نوع تحلیلی و توصیفی - پیمایشی است و جامعه‌ی آماری آن عبارت از دانشجویان شش رشته علوم انسانی دانشگاه اصفهان است که افراد نمونه به شیوه نمونه‌گیری تصادفی طبقه‌ای انتخاب شدند. ابزار پژوهش پرسش‌نامه محقق ساخته بوده که با استفاده از مرور متون تخصصی و دیدگاه‌های صاحب نظران در مقیاس لیکرت تدوین شده است. با اطمینان از روایی و پایایی پرسش‌نامه، داده‌های حاصل از آن با استفاده از آمار توصیفی و استنباطی تحلیل گردیده است. بر اساس یافته‌ها، روش‌های سنجش موجود غیر کاربردی، تراکمی، محدود و تا حد زیادی مبتنی بر امتحانات پایان ترم است و خود سنجی و هم‌تاسنجی هنوز در آن جایی ندارد. همچنین یافته‌ها نشان داد که اقدامات سنجش، در جهت افزایش یادگیری‌های پایدار دانشجویان طرح‌ریزی نمی‌شوند و نیز برای هدف اندازه‌گیری آموخته‌ها از دقت و صحت کافی برخوردار نیست. انتخاب روش‌های سنجش بر اساس پیامدهای یادگیری، خصوصاً یادگیری‌های پایدار و فعال ساختن دانشجویان در فرآیند سنجش از جمله پیشنهادهاست.

واژگان کلیدی: سنجش سنتی، سنجش جایگزین، سنجش پایدار، خودسنجی، هم‌تاسنجی

Evaluation of the Implementing Different Methods of Measurement and Sustainable Learning Based Measurements

P. Pornaderi¹, A.R. Nasr², M.R. Nili³ and Z. Bagheri⁴

¹M.A., Faculty of Human Sciences, Isfahan University, Isfahan, Iran, arnasr@edu.ui.ac.ir

² Professors, Faculty of Human Sciences, Isfahan University, Isfahan, Iran

³ Assistant Professor, Faculty of Human Sciences, Isfahan University, Isfahan, Iran

⁴M.A., Faculty of Human Sciences, Isfahan University, Isfahan, Iran

Abstract: The aim of higher education is granted jurisdiction to the students and increase their learning. In the recent years, experts believe that these aims must pay more attention to sustainable learning in the life. Selecting the suitable evaluation methods for the students, which can access different evaluation aims simultaneously, is the biggest challenge for evaluation. This research considers the different common methods of evaluation and achieving the goals of sustainable learning of the students with the help of these methods. The research is an analytical and descriptive- survey. The statistical society of this research is the students of six majors of humanity in Isfahan University. The students were selected randomly. The research tool was researcher made questionnaire which has been prepared by reviewing the special text and experts view points on a Likert scale. Ensuring the reliability and validity of a questionnaire the data were analyzed by the descriptive and inferential statistic. According to the findings the evaluating method which has been used was inapplicable, compressed, limited and they were highly according to the student's final exams. There is no place for self and peer assessment. Also finding showed that actions related to the evaluation are not programming on the line with the student sustainable learning and they have not suitable accuracy to measure the learning aims. Measurement method selection according to the learning outcome, especially sustainable learning and making student to have an active role in measurement process are among the suggestions.

Keywords: Traditional measurement, Alternative measurement, Sustainable Measurement, Self Assessment, Peer Assessment.

1 - مقدمه

کرد روش‌های سنتی سنجش در دانشگاه‌ها از تنوع برخوردار نبوده است و تأکید بیش از حدی بر امتحانات و تکرار سؤال‌ها در هر سال وجود داشته است [8]. نقدهایی نیز بر اقدامات سنجش تکوینی وجود دارد که بیشتر به چگونگی اجرای سنجش تکوینی و نه اصول آن، مربوط می‌گردد. از جمله این که تأکید بیش از حدی در این نوع سنجش بر یادگیری‌های آنی و زودگذر مربوط به یک محدوده‌ی زمانی خاص یا یک تکلیف خاص وجود دارد که ظرفیت‌های بالقوه‌ی سنجش تکوینی را کاهش داده است [9].

در سنجش تکوینی علیرغم این که تأکید زیادی بر یادگیری وجود دارد، توانمندی دانشجو در ارزیابی کردن آنچه یاد گرفته است و برنامه‌ریزی بر طبق آن مورد غفلت قرار می‌گیرد. اگر چه دستیابی به یادگیری‌های طول دوره لازم است اما آینده‌گرایی کافی در آن وجود ندارد. در صورت خوب انجام گرفتن، سنجش تکوینی فقط در کوتاه مدت مؤثر است چرا که بخش بسیار مهم و وسیعی از برنامه سنجش یعنی کمک به دانشجویان در این که چگونه بازیگران فعالی در مدیریت یادگیری خود و در سنجش خود بعد از فارغ‌التحصیلی باشند را نادیده می‌گیرد [1].

در آموزش عالی تأکید زیادی بر دیدگاه‌های یادگیری و استقلال، تفسیر و سطوح بالای شناختی وجود دارد، ولی طریقی که دوره‌ها به دانشجویان ارائه می‌گردد و ماهیت امتحانات، چیزی جز این را نشان می‌دهد [10]. نیروی کار قرن بیست و یکم به‌طور روزافزون به مهارت‌های عام مانند ارتباطات، حل مسأله، همکاری و تفکر انتقادی نیاز دارد، اما روش‌های سنجش فعلی همچون امتحانات، کمتر به توسعه این مهارت‌ها می‌پردازد. دانشجویان ممکن است در این روش‌ها به نمره‌ی خوب دست پیدا کنند ولی این به معنای کوشش زیاد و عمیق آنان نیست [11]. یافته‌های پژوهش‌های فازی نشان داد دانشجویان سال اول دانشگاه استقلال بیشتری نسبت به سال دومی‌ها و سال سومی‌ها احساس می‌کنند. فازی چنین نتیجه گرفت که آموزش عالی به جای تربیت دانشجویان با اعتماد به نفس، مستقل و خودگردان، موجب تضعیف این کیفیت‌ها در یادگیرندگان در طول تحصیلات دانشگاهی می‌گردد [12]. یادگیری دانشجویان بعد از زمان فارغ‌التحصیلی دیگر تحت هدایت معلمان و

بر اساس گفته‌ی باد چالش اساسی سنجش مربوط به وظایف و کارکردهای دوگانه‌ی آن است [1]. وظیفه‌ی نمره دهی و اعطای مجوز صلاحیت به سطحی از اکتسابات یا کارکرد تراکمی سنجش از یک سو، وظیفه یادگیری یا کارکرد تکوینی از سوی دیگر، اندازه‌گیری آموخته‌های دانشجویان در برابر ملاک‌ها و استانداردها از یک سو و ایجاد رقابت بین افراد و هنجار مرجعی از سوی دیگر، سنجیدن آنچه اندازه‌گیری آن آسان است و یا می‌تواند به طور پایا اندازه‌گیری شود، در برابر سنجیدن آنچه مهم و با ارزش است، از جمله انتظارات از سنجش است که اصلاح آن را مشکل می‌سازد. کارکرد تراکمی سنجش منطبق با الگوی «سنجش به‌عنوان اندازه‌گیری» سرافینی است [2]. که غالباً با آزمون‌های استاندارد شده مدادکاغذی و هنجار مرجع همراه است. در این الگو، سنجش با پایایی¹ و روایی²، سوگیری و عادلانه بودن، تقلب، کیفیت و استانداردها ارتباط می‌یابد. از جمله ویژگی‌های آن، کمی‌گرایی، استفاده محدود از روش‌ها و تکنیک‌های سنجش، دانش غیر وابسته به زمینه و بی ارتباط با تجارب دانشجو، مخفی بودن ملاک‌ها و عدم اقتدار یادگیرنده است [3].

هدف دیگر سنجش، تسهیل یادگیری از طریق تهیه و ارائه بازخورد در انواع آزمون‌ها یا تکالیف است که با اقدامات سنجش تکوینی همراه می‌شود. نقدهایی درباره‌ی محدودیت‌های هر دو نوع سنجش وجود دارد.

نایت می‌گوید که سنجش تراکمی در مسائلی همچون اعتبار و سنجیدن یادگیری‌های پیچیده که در آموزش عالی مورد تأکید است، دچار مشکل است [4]. بر اساس مطالعه براون، رایس و اسمیت سنجش معمول و سنتی در دانشگاه‌ها هشتاد درصد به شکل مقاله، گزارش و امتحانات تشریحی است [5]. در مطالعه‌ی سمبل، مکداول و براون برداشت دانشجویان از روش‌های سنتی سنجش با روش‌های جایگزین همچون کارنما³، خودسنجی، همتاسنجی و شبیه‌سازی، مورد مقایسه قرار گرفت [6].

از نظر این دانشجویان، سنجش سنتی در اندازه‌گیری آموخته‌ها دقیق و عادلانه نبوده است. همچنین در یک مطالعه‌ی تحلیلی که از سال 1993 تا 2001، توسط آژانس تضمین کیفیت⁴ (QAA) انجام گرفت [7]. آژانس گزارش

و سنجش دیگرانی که در تشخیص چالش‌ها تخصص دارند، نیز اساس سنجش پایدار است. دانشجویان در موارد بسیاری نیاز به دیگران دارند و معلمین همیشه قابل دسترس و لزوماً مفیدترین اشخاص در این گونه موارد نیستند. «سنجش پایدار» نوع جدیدی از سنجش نیست بلکه راهی است که با استفاده از سنجش تراکمی و تکوینی در جهت

پرورش اهداف دراز مدت به کار برده می‌شود. مهارت‌هایی که در یادگیری طول زندگی سهم دارند باید از راه‌های مختلف، که فعالیت‌های سنجش فقط یکی از آنهاست توسعه یابند. برای پر کردن شکاف بین یادگیری درون دانشگاه و بیرون دانشگاه، نیازی به ابداع یا ابداعاتی در زمینه یادگیری نیست، به جای آن، یک بازاندیشی در همه فعالیت‌های دوره‌ها و به خصوص در فعالیت‌های سنجش که به‌طور روزمره بدیهی پنداشته می‌شود لازم است [13]. همچنین باید دانست که «سنجش پایدار» نظریه‌ای نیست که بتوان آن را در فعالیت خاصی ارائه داد.

فعالیت‌های این سنجش مستقل از محیط یادگیری نیست و باید به‌وسیله‌ی معلمین و یادگیرندگان در طول زمان به‌طور مداوم، از نو خلق و مفهوم‌سازی مجدد گردد. قضاوت تراکمی معلم برای نمره دادن و اندازه‌گیری عملکردها و پاسخ‌گویی و به تصویر درآوردن آموخته‌ها با اهمیت است و از میان برداشتن آن نه ممکن است و نه مطلوب، اما در فرآیند توجه به مدرک، نباید هدف یادگیری و اقدامات سنجشی

که برای تحقق آن لازم است، فراموش گردد [1]. درایت قابل توجهی برای تعادل بین اهداف سنجش در جهت پرورش دانشجویان به عنوان ارزیابان طول زندگی مورد نیاز است. طرح‌ریزی فرصت‌های یادگیری و انتخاب روش‌های سنجش متناسب با آنها از جمله راهکارهاست. در جدول 1 برخی از فرصت‌ها و پیامدهای یادگیری، همراه با روش‌های سنجش مناسب آنها، برگرفته از نایتینگل و دیگران و مؤسسه تضمین کیفیت برای آموزش عالی خلاصه شده است. نویسندگان مقاله مذکور ضمن آنکه به نیاز هدف مدرک دهی توسط سنجش توجه دارد، درگیر شدن فراگیران در تکالیفی که به پیامدها و مهارت‌های مورد نیاز طول زندگی آنان منجر شود را ضروری می‌داند [17]. با توجه به نیاز به بررسی‌های بیشتر در خصوص مباحث

برنامه درسی نیست، از این رو بهتر است یادگیرندگان، ابتکار عمل و مسئولیت یادگیری خود را برای روبرو شدن با دنیای ناشناخته و فوق‌پیچیده و نیاز کار کردن برای یادگرفتن و یادگرفتن برای کار کردن، به دست بگیرند [1]. در محل کار، دانش‌آموختگان معمولاً امتحان نخواهند داد یا مقاله‌ی آکادمیک نمی‌نویسند، محل کار عناصر بسیار اضافه‌تری دارد و شامل مخاطرات بیشتری است. آنها باید بتوانند در محیط کار در برابر آنچه کار خوب به حساب می‌آید به ارزیابی کار خود بپردازند [13]. انتظارات کارفرمایان از دانش‌آموختگان بر اساس تحقیقات مختلف انجام شده، توسط نایت به ترتیب شامل: ابداع و خلاقیت، سازگاری و انعطاف، میل به یادگیری، کار مستقل، خود سامانی و گرفتن مسئولیت، توانایی کار گروهی، مدیریت خود و دیگران، مدیریت زمان و کار فشرده، ارتباط خوب شفاهی، توانایی ارتباط نوشتاری برای اهداف یا مخاطبین مختلف، توجه به جزئیات، برنامه‌ریزی، هماهنگ‌سازی و سازماندهی است [14].

از این رو، رانتری هدف دیگری برای سنجش تعریف کرد. او گفت سنجش باید دانشجویان را برای زندگی آماده کند [4]. یاد این هدف را با مفهوم سازی آن به عنوان «سنجش پایدار»⁵، توسعه بخشید [1].

ویژگی مهم این هدف این است که دانشجو هر چه بیشتر نقش فعال‌تری در فرآیند سنجش، نسبت به آنچه در سنجش تراکمی و تکوینی دارد، بازی کند. دانشجویان همواره در فرآیند سنجش دریافت‌کننده‌ی اعمال دیگرانند نه عاملان فعال آن. آنها باید آزمون‌ها را بپذیرند و بازخوردی که به آنها داده می‌شود در مواردی است که معلم اهمیت آن را تشخیص می‌دهد. چنین مفهومی از سنجش نه تنها مناسب یادگیری‌های پایدار نیست، یادگیری‌های جاری را نیز محدود می‌کند [15].

در بررسی‌های وسیع بلک و ویلیام دربارهی سنجی تکوینی، آنها به نکات مهمی در این باره اشاره کرده‌اند که نقطه آغاز مهمی در ساختن چهار چوبی برای سنجش مداوم (پایدار) گردید که توسط یاد مفهوم سازی شد [16]. از جمله این موارد تأکید بر خودسنجی، بعد از جستجوی ملاک‌ها و درگیری فعال در تکلیف است. یادگیرندگان باید پیشرفت خود به سمت اهداف را ارزیابی کنند. همتاسنجی

مذکور، پرسش‌هایی که پژوهش قصد بررسی آنها را دارد به شرح زیر است: 1- انواع سنجش متداول به لحاظ زمان انجام سنجش، ارزیابان، نوع سؤال‌ها و نوع تکلیف‌ها چیست؟، 2- از نظر دانشجویان، سنجش در هدف اندازه‌گیری خود چقدر از دقت و صحت برخوردار است؟ 3- از نظر دانشجویان میزان حمایت سنجش از یادگیری‌های پایدار طول زندگی آنان تا چه حد است؟

2- روش تحقیق

این پژوهش از نوع تحلیلی و توصیفی - پیمایشی است. جامعه آماری این پژوهش شامل دانشجویان سال آخر مقطع کارشناسی و کارشناسی ارشد رشته‌های علوم تربیتی، روان‌شناسی، تاریخ، علوم اجتماعی، مدیریت و علوم سیاسی دانشگاه اصفهان در سال تحصیلی 88-1387 هستند. آمار دانشجویان سال آخر کارشناسی رشته‌های مذکور، 1061 نفر و در مقطع کارشناسی ارشد، 204 نفر

بوده است. علت انتخاب رشته‌های علوم انسانی برای موضوع این پژوهش، آشنا بودن پژوهشگر با روش‌های سنجش این رشته‌ها و ساختار دانش آن است. علت انتخاب شش رشته خاص از سه دانشکده علوم انسانی، ماهیت نسبتاً همگون این رشته‌ها از حیث ساختار دانش و مخاطبان است. انتخاب دانشجویان سال آخر کارشناسی و دانشجویان کارشناسی ارشد نیز به این دلیل بوده است که این دانشجویان درس‌های تخصصی بیشتری را گذرانده‌اند و بهتر می‌توانند درباره چگونگی سنجش این درس‌ها قضاوت کنند. فرض بر این بود که دانشجویان کارشناسی ارشد در پاسخ به سؤال‌های پرسش‌نامه، تجربه‌های خود از مراحل تهیه پایان‌نامه را به‌عنوان یک تکلیف سنجش، در نظر بگیرند. حجم نمونه در این مطالعه پس از اجرای آزمایشی ابزار پژوهش، و فرمول برآورد حجم نمونه 224 نفر تعیین گردید که افراد نمونه با روش نمونه‌گیری تصادفی طبقه‌ای، متناسب با حجم دانشجویان در رشته‌ها، معدل‌ها و مقاطع تحصیلی مختلف، انتخاب شدند.

جدول 1 انتخاب روش سنجش بر اساس پیامدها و فرصت‌های یادگیری

پیامدها و فرصت یادگیری برای دانشجویان	نمونه‌ای از تکالیف و سنجش‌ها
تفکر انتقادی و تصمیم‌گیری (توسعه توان قضاوت، سنجش، ارزشیابی و داوری)	مقایسه، گزارش، ثبت وقایع، نوشتن یک مقاله روزنامه‌ای، اظهار نظر روی ابعاد نظری یک مقاله، نوشتن یک نامه و نظر درباره یک سیاست، مسائل بهداشت همگانی و...
حل مسأله و تدوین برنامه، کار روی چند پرسش و مشکل و کسب اطلاعات درباره آن‌ها، با دادن وقت بیشتر برای تفکر و رجوع به منابع و نوشته‌ها و تکیه کمتر بر حافظه، نسبت به امتحانات سنتی	امتحانات کتاب باز (open book)، امتحانات پاسخ در خانه (take home)، وامتحاناتی که سؤالات آن قبلاً به دانشجویان داده شده است. تدوین سناریوی مشکل، کارگروهی، کار مسأله محور، تحلیل موردی، پیش‌نویس یک پژوهش و کنفرانس دادن.
ایجاد تبحر در مستندسازی نظامند مشاهدات، تجارب، افکار در هنگام درگیری در کار یا تکلیف جاری و بررسی آنچه از آن می‌توان یاد گرفت، اجرای روش کارها و نشان دادن توانایی انجام فنون (محاسبه، کاربرد تجهیزات، انجام دستورالعمل‌ها، دنبال کردن روش کارهای آزمایشگاه)	پروژه، کارپوشه (کارنما)، شرح حال نویسی، ثبت متفکرانه وقایع، یادداشت روزانه استودیو و آزمایشگاه، ارائه، ایفای نقش، ساخت ویدئو (نوشتن اسکرپ و ساخت آن)، تهیه پوستر، دنبال کردن و بررسی یک تمرین حرفه‌ای واقعی یا شبیه‌سازی شده و آماده کردن یک دستورالعمل توضیحی برای کاربرد تجهیزات برای مخاطبان خاص.
مرور این که چطور به‌طور اثر بخش یاد گرفته‌اند و توان تشخیص و اصلاح قوت‌ها، شکاف‌ها و بدفهمی‌های خود از طریق نظامند را به‌دست آورده‌اند. مدیریت و توسعه خود (کار مشترک، کارکردن مستقل، یادگیری مستقل، خودراهبری، مدیریت زمان، مدیریت تکالیف، سازماندهی)	خودسنجی بر اساس رایانه، کاربرد سؤالات چهارگزینه‌ای و سایر اشکال آزمون‌های روی خط. کارنما، شرح وقایع نویسی و کار گروهی.
مدیریت اطلاعات تحقیق، بررسی، سازماندهی، جمع‌آوری داده‌ها، جستجو و مشاهده و تفسیر	پروژه، پایان‌نامه، تکالیف عملی و مسائل عملی
عرضه دانش و درک (یادآوری، شرح دادن، گزارش، بازشماری، تشخیص، تعریف، بازگویی)	امتحانات نوشتاری و شفاهی، مقاله، گزارش، سؤالات کوتاه جواب، چندگزینه‌ای، صحیح - غلط، نوشتن پاسخ برای سؤال یک مشتری یا متقاضی، درست کردن یک لغت نامه
طراحی، آفرینش، اجرا (اندیشیدن، مجسم کردن، طراحی، تولید، خلق کردن، ابداع، اجرا)	کارنما، پروژه، سنجش عملکرد، ارائه
برای اینکه یاد بگیرند اطلاعات و نظرات خود را به صورت غیر نوشتاری و در حضور شنوندگان حاضر بیان کنند. ارتباط (ارتباط یک‌سویه یا دو سویه ارتباط با یک گروه ارتباط کلامی، نوشتاری، و غیر کلامی. بحث، توصیف، طرفداری، مصاحبه، مذاکره، کاربرد اشکال نوشتاری)	سمینار، پوستر، ویدئو کنفرانس یا سایر ارائه‌های چند رسانه‌ای. ارائه‌های کتبی (مقاله، گزارش)، ارائه‌های شفاهی، کارگروهی، ایفای نقش، بحث، دنبال کردن و بررسی یک تمرین حرفه‌ای واقعی یا شبیه‌سازی شده، شرکت در یک محکمه پرسش و سؤال.
برای رسیدن به درک بهتر از قوت‌ها و چالش‌های کار، به صورت گروهی و مشارکتی، همچنین کسب بینش در این که چطور دیگران به سؤالات می‌پردازند، مشکلات را تعریف و و ظایف را سازماندهی می‌کنند و نظرات را منتقل می‌نمایند	حل مسأله و ارائه‌های گروهی، ایفای نقش، پروژه‌های مشارکتی، بررسی مشارکتی کتاب
برای گرفتن نقش فعال‌تر در بخشی از سنجش و در نتیجه افزایش توانایی آن‌ها در بررسی و کاربرد ملاک‌های سنجش، برای قضاوت کیفیت کار خود و دیگران، گرفتن بازخورد سازنده برای بهبود کار در آینده	بازخورد هم‌تایم روی تکالیف خودسنجی یک کار، نمره دادن همتا به گزارش آزمایشگاه و ایجاد ملاک توسط دانشجویان برای یک پروژه اجتماعی.

دانشجویان، اظهار داشته اند تکالیف سنجش در حد زیاد و خیلی زیاد به صورت ارائه‌های شفاهی (کنفرانس) بوده است. سنجش از طریق نوشتن مقاله، طبق بیان ۵۴/۲٪ دانشجویان و تکالیف‌های کتاب باز و کار در خارج از کلاس، بر اساس پاسخ ۷۱/۴٪ آنان در حد کم و خیلی کم انجام گرفته است. بخش زیادی از دانشجویان (۸۳٪)، تکالیف‌های کاربردی و عملکردی را در حد کم و خیلی کم برآورد کرده‌اند.

بیشترین استفاده از سؤالات تشریحی بوده است که توسط ۷۱/۸ درصد دانشجویان در حد زیاد و خیلی زیاد برآورد شده در حالی که استفاده از سؤالات چهارگزینه‌ای از نظر ۴۲/۴٪ دانشجویان در حد زیاد و خیلی زیاد بوده است. استفاده از سؤال‌های کوتاه پاسخ و تکمیل کردنی بر طبق پاسخ ۶۴/۴٪ دانشجویان و استفاده از سؤالات شفاهی از نظر ۵۱/۹٪ دانشجویان در حد کم و خیلی کم بوده است. ۹۴/۵٪ دانشجویان گفته‌اند که آزمون‌های پایان ترم در حد زیاد و خیلی زیاد انجام شده است. امتحان گرفتن در طول ترم با میانگین (۲/۷۳)، کمتر از حد متوسط مورد توجه است.

انجام سنجش آغازین از نظر ۵۱/۷٪ پاسخ‌گویان در حد کم و خیلی کم بوده است. هفتاد و هشت درصد دانشجویان خودسنجی و ۷۴٪ هم‌تاسنجی را در حد کم و خیلی کم برآورد کرده‌اند. مقایسه‌ی انواع روش‌های سنجش در دو مقطع کارشناسی و کارشناسی ارشد نشان داد میانگین گویه‌های مربوط به استفاده از تکالیف‌های «کتاب باز» و «کار در خارج از کلاس»، سؤال‌های شفاهی، انجام سنجش در آغاز ترم، خودسنجی و هم‌تاسنجی در مقطع کارشناسی ارشد به ترتیب ۲/۵، ۲/۸۲، ۲/۷۷، ۲/۴۲ و ۲/۵۵ بوده است که از میانگین آنها در مقطع کارشناسی (۱/۸۸، ۲/۴۱، ۲/۳۴، ۱/۷۲ و ۱/۷۸) به‌طور معنادار بیشتر است.

میانگین پاسخ‌ها در استفاده از تکالیف‌های به شکل مقاله، ارائه‌های شفاهی، و استفاده از سؤال‌های تشریحی نیز به‌طور معنادار در مقطع کارشناسی ارشد (به ترتیب با میانگین‌های ۳/۸۰، ۴/۲۰ و ۴/۴۷) بیشتر از مقطع کارشناسی (به ترتیب با میانگین‌های ۲/۱۶، ۳/۳۸ و ۳/۸۰) است. استفاده از تکالیف‌های کاربردی در مقطع ارشد با میانگین ۱/۶۰ تفاوت معناداری با مقطع کارشناسی با

با همه کوششی که برای دریافت کامل پرسش‌نامه‌ها به عمل آمد، به دلیل مخدوش بودن و یا عدم پاسخ‌گویی کامل برخی از دانشجویان به سؤالات پرسش‌نامه، یازده پرسش‌نامه مورد استفاده واقع نشد. لذا در مجموع ۲۱۳ پرسش‌نامه تجزیه و تحلیل شد. تعداد پاسخ‌گویان مقطع کارشناسی ۱۷۳ نفر و تعداد پاسخ‌گویان در مقطع کارشناسی ارشد ۴۰ نفر بوده است.

ابزار پژوهش: پس از تدوین پرسش‌نامه اولیه‌ی پژوهش، با مراجعه به داوران متخصص و نیز مصاحبه‌های اکتشافی که با دانشجویان انجام گردید، اصلاحات لازم در خصوص پرسش‌نامه صورت گرفت و روایی محتوایی آن محرز گردید. پرسش‌نامه بر مبنای طیف پنج درجه‌ای لیکرت و شامل سه بخش است.

بخش اول شامل ۱۳ گویه و مربوط به سؤال اول پژوهش است و انواع سنجش به لحاظ انواع تکالیف (گویه‌های ۱ تا ۴)، انواع سؤال‌ها (گویه‌های ۵ تا ۸)، زمان سنجش (گویه‌های ۹ تا ۱۱) و ارزیابان (گویه‌های ۱۲ و ۱۳) را مورد بررسی قرار می‌دهد.

بخش دوم مربوط به مؤلفه‌ی دقت و صحت سنجش و شامل گویه‌های ۱۴ تا ۲۰ است، که برای پاسخ سؤال دوم پژوهش تدوین شده است. بخش سوم شامل گویه‌های ۲۱ تا ۳۲ و مربوط به مؤلفه‌ی یادگیری‌های طول زندگی است که به سؤال سوم پژوهش می‌پردازد. پرسش‌نامه همچنین دارای ۳ سؤال جمعیت شناختی است. ضریب پایایی پرسش‌نامه به کمک روش آلفای کرونباخ در مؤلفه‌ی دقت و صحت سنجش ۰/۷۲ و در مؤلفه‌ی یادگیری‌ها ۰/۸۸ محاسبه گردید. برای تجزیه و تحلیل داده‌ها در سطح توصیفی از میانگین، درصد فراوانی و انحراف معیار و در سطح استنباطی از آزمون‌های t تک‌متغیره، آزمون t مستقل، آزمون تحلیل‌واریانس (ANOVA) و آزمون‌های-تعقیبی توکی^۱ استفاده شده است.

۳- نتایج و بحث

سؤال اول پژوهش: انواع سنجش به لحاظ زمان انجام،

ارزیابان، نوع سؤال‌ها و نوع تکالیف‌ها چیست؟

در جدول ۲، یافته‌های توصیفی گویه‌های مربوط به انواع سنجش ارائه شده است. بر اساس یافته‌ها ۵۷/۸ درصد

(۲/۶۰) از میانگین رشته علوم تربیتی (۱/۷۳) بیشتر است. انجام آزمون‌های طول ترم، با میانگین ۳/۶۰ مربوط به رشته علوم اجتماعی است که از میانگین رشته روان‌شناسی و علوم تربیتی (۲/۲۵ و ۲/۲۶) بیشتر است. میانگین استفاده از خودسنجی و هم‌تاسنجی در رشته علوم سیاسی (۲/۳۶) و (۲/۴۶) از میانگین رشته‌های علوم تربیتی، روان‌شناسی و مدیریت بیشتر است. در استفاده از سؤالات چندگزینه‌ای میانگین رشته روان‌شناسی (۴/۰۰) از میانگین رشته علوم اجتماعی (۱/۶۳) بیشتر است. انجام خودسنجی و هم‌تاسنجی در رشته علوم سیاسی دارای میانگین ۲/۳۶ و ۲/۴۶ بوده که تفاوت معناداری با میانگین رشته‌های علوم تربیتی، روان‌شناسی و مدیریت در این گویه‌ها داشته است. در همه رشته‌ها آزمون‌های پایان ترم بسیار زیاد مورد توجه است و از این نظر اختلاف زیادی بین آنها مشاهده نمی‌شود. بیشترین میانگین در استفاده از تکلیف‌های کاربردی (۲/۰۲) در رشته علوم تربیتی بود که تفاوت معناداری با سایر رشته‌ها دارد.

میانگین ۱/۵۲ نداشته است. اما میانگین استفاده از آزمون طول ترم، سؤال‌های چهارگزینه‌ای و کوتاه پاسخ در مقطع کارشناسی (به ترتیب با میانگین‌های ۲/۹۲، ۳/۲۵ و ۲/۳۴) به طور معنادار از کارشناسی ارشد بیشتر است. در میزان استفاده از آزمون پایان ترم بین دو مقطع تحصیلی تفاوتی مشاهده نمی‌شود (میانگین در هر دو مقطع ۴/۶۲). بین میانگین پاسخ دانشجویان با معدل‌های متفاوت، در این بخش تفاوت معنادار وجود ندارد. مطابق آزمون تحلیل واریانس یک‌راهه و آزمون تعقیبی توکی تفاوت میانگین میان رشته‌های مختلف در برخی از انواع روش‌های سنجش بدین شرح معنادار بوده است: در استفاده از سؤال‌های شفاهی میانگین رشته روان‌شناسی (۳/۰۸) از میانگین رشته‌های علوم تربیتی (۲/۲۶)، علوم اجتماعی و مدیریت (۲/۲۸) بیشتر است. میانگین رشته علوم اجتماعی (۳/۸۷) در تکالیف به شکل کنفرانس و گزارش از میانگین رشته مدیریت (۳/۰۴) بیشتر است. در تکالیف به صورت کتاب باز و انجام در خانه، میانگین رشته علوم اجتماعی

جدول ۲ درصد و میانگین پاسخ دانشجویان به گویه‌های انواع سنجش

میانگین	زیاد و خیلی زیاد	تاحدودی	کم و خیلی کم	وجود انواع روش‌ها در تجربه‌ی سنجش
۲/۴۷	۲۱/۷	۲۳/۱	۵۴/۲	۱- سنجش به صورت نوشتن مقاله
۳/۵۳	۵۷/۸	۲۲/۱	۲۰/۲	۲- سنجش به صورت ارائه‌های شفاهی (کنفرانس)
۲/۰۰	۱۳/۲	۱۶/۴	۷۱/۴	۳- سنجش با تکلیف‌های «کتاب باز» و «کار در خارج از کلاس» (Open book و Take home)
۱/۵۴	۴/۲	۱۲/۷	۸۳/۰	۴- سنجش با تکلیف‌های کاربردی و عملکردی مانند کارهای میدانی، کارگاهی و آزمایشگاهی
۲/۴۹	۱۹/۴	۲۸/۸	۵۱/۹	۵- سنجش با سؤال‌های شفاهی
۳/۹۲	۷۱/۸	۱۷/۸	۱۰/۴	۶- سنجش با سؤال‌های تشریحی
۳/۰۲	۴۲/۴	۲۲/۴	۳۵/۲	۷- سنجش با سؤال‌های چندگزینه‌ای
۲/۲۲	۱۳/۸	۲۱/۸	۶۴/۴	۸- سنجش با سؤال‌های کوتاه پاسخ و یا تکمیل کردنی
۲/۴۲	۱۲/۶	۳۵/۷	۵۱/۷	۹- انجام سنجش در آغاز ترم
۲/۷۳	۲۸/۲	۳۰/۰	۴۱/۸	۱۰- برگزاری امتحان در طول ترم
۴/۶۱	۹۴/۵	۴/۲	۴/۳	۱۱- برگزاری امتحان در پایان ترم
۱/۸۵	۶/۱	۱۶/۰	۷۸/۰	۱۲- استفاده از خودسنجی در سنجش دانشجویان
۱/۹۲	۹/۵	۱۶/۵	۷۴/۰	۱۳- استفاده از هم‌تاسنجی (سنجش دانشجویان توسط یکدیگر)

جدول ۳ درصد و میانگین پاسخ دانشجویان به مؤلفه‌ی دقت و صحت سنجش

میانگین	زیاد و خیلی زیاد	تاحدودی	کم و خیلی کم	وجود موارد در تجربه‌ی سنجش
۳/۳۴	۴۶/۳	۳۵/۳	۱۸/۴	۱۴- نمره کل به چند ملاک تقسیم می‌شد
۲/۹۶	۲۹/۱	۴۱/۸	۲۹/۱	۱۵- تکالیف و امتحان‌ها با اهداف مهم درس، مطابقت داشت
۲/۷۲	۲۲/۱	۳۵/۷	۴۰/۴	۱۶- اصول علمی در ساختن سؤال‌های امتحان (مانند مبهم و دو پهلو نبودن سؤال‌ها)، رعایت می‌شد
۳/۱۳	۴۰/۳	۳۳/۲	۲۶/۶	۱۷- بخش زیادی از سرفصل‌های مهم درس مورد سنجش قرار می‌گرفت
۲/۶۷	۲۷/۰	۲۸/۴	۴۴/۶	۱۸- تدابیری برای کاهش امکان تقلب وجود داشت (مانند تکراری نبودن سؤال‌ها در هر سال)
۲/۷۰	۲۱/۲	۴۱/۰	۳۷/۷	۱۹- استادان در تصحیح اوراق و بررسی تکالیف دقیق بودند
۲/۴۹	۱۴/۵	۴۰/۴	۴۵/۱	۲۰- در کل، نمره‌های پایان ترم دانشجویان عادلانه بودند

جدول ۴ درصد و میانگین پاسخ دانشجویان به مؤلفه‌ی یادگیری‌های پایدار

میانگین	زیاد و خیلی زیاد	تاحدودی	کم و خیلی کم	وجود موارد در تجربه‌ی سنجش
۲/۶۷	۲۱/۲	۳۵/۲	۴۳/۶	۲۱- سبب شد از ماهیت رشته تحصیلی خود، آگاه‌تر شوم
۲/۲۶	۹/۸	۳۳/۳	۵۶/۸	۲۲- مرا در کاربرد آموخته‌ها، برای حل مسائل حرفه‌ای و شغلی مربوط به رشته‌ام، توانمند ساخت
۲/۵۱	۲۰/۱	۲۹/۶	۵۰/۲	۲۳- توان مطالعه و یادگیری مستقل پس از فارغ‌التحصیلی را، در من افزایش داد
۲/۴۵	۱۵/۴	۳۴/۳	۵۰/۳	۲۴- توانایی همکاری با دیگران به عنوان عضوی از گروه را، در من تقویت کرد
۲/۶۷	۲۶/۳	۳۰/۵	۴۳/۱	۲۵- توان تحقیق (دنبال کردن ایده، جمع‌آوری اطلاعات و حل مسأله) را در من افزایش داد
۲/۶۸	۲۳/۰	۳۱/۰	۴۶/۰	۲۶- احساس اعتماد بنفس مرا در ارتباط با حل مسائل ناشناخته، افزایش داد
۲/۶۵	۲۲/۰	۳۳/۳	۴۴/۶	۲۷- در من شور و انگیزه تداوم یادگیری، کنجکاوی و کاوشگری ایجاد کرد
۲/۳۸	۱۴/۶	۲۷/۷	۵۷/۸	۲۸- مرا به سوی تفکر خلاق و ارائه راه‌حل‌های جدید تشویق کرد
۲/۶۳	۱۹/۷	۳۷/۶	۴۲/۷	۲۹- مرا به سمت استفاده از تکنولوژی و فناوری روز هدایت کرد
۲/۶۹	۲۴/۹	۳۱/۹	۴۳/۲	۳۰- در کسب مهارت‌های خودتنظیمی مانند مدیریت زمان، به من کمک کرد
۲/۸۸	۳۰/۵	۳۴/۷	۳۴/۷	۳۱- به تقویت ارتباطات شفاهی و نوشتاری من کمک کرد
۲/۷۷	۲۶/۲	۳۲/۴	۴۱/۳	۳۲- سبب شد بدون نیاز به معلم، بتوانم درباره یادگیری‌ام قضاوت کنم

سؤال دوم پژوهش: از نظر دانشجویان، سنجش در هدف

اندازه‌گیری خود چقدر از دقت و صحت برخوردار است؟ داشته است. به عبارت دیگر، روایی سنجش تنها تا حدودی رعایت شده

گویه‌های ۱۴ تا ۲۰ در جدول ۳ مربوط به مؤلفه‌ی دقت است. در گویه ۲۰ با میانگین ۲/۴۹ عادلانه بودن نمره‌ها، تنها توسط صحت سنجش است. بر اساس نتایج این جدول بیشترین درصد دانشجویان بیشتر از حد متوسط برآورد شده است

میانگین پاسخ‌ها (۳/۳۴، ۳/۱۳) مربوط به گویه ۱۴ و رعایت اصول علمی در ساختن سؤال‌های امتحان را، ۴۰/۴

است که بیانگر این است که تقسیم نمره به ملاک‌های دانشجویان در حد کم و خیلی کم و ۳۵/۷ دانشجویان، آن را در حد متعدد و سنجش همه‌ی سرفصل‌های مهم مورد توجه است برآورد کرده‌اند. اتخاذ تدابیر برای کاهش تقلب، از استادان بوده است. طبق پاسخ ۴۲ درصد دانشجویان، ۴۴/۶ دانشجویان در حد کم و خیلی کم بوده است. چهل و یک

تعقیبی توکی نشان می‌دهد که دقت و صحت سنجش در رشته مدیریت با رشته تاریخ دارای اختلاف میانگین معنی‌دار بوده و میانگین رشته مدیریت از رشته تاریخ بیشتر است.

سؤال سوم پژوهش: از نظر دانشجویان میزان حمایت سنجش از یادگیری‌های پایدار طول زندگی آنان تا چه حد است؟

گویه‌های 21 تا 32 مربوط به مؤلفه‌ی یادگیری‌های پایدار است. بر طبق یافته‌های جدول 4، همه گویه‌های این مؤلفه، میانگینی کمتر از میانگین فرضی 3 داشته است.

بیشترین میانگین در این مؤلفه (2/88)، مربوط به گویه 31 «تقویت ارتباطات شفاهی و نوشتاری» و کمترین میانگین (2/26) مربوط به گویه 22 «تقویت کاربرد آموخته‌ها، برای حل مسائل حرفه‌ای و شغلی» و پس از آن گویه 28 «تشویق به تفکر خلاق و ارائه راه‌حل‌های جدید» با میانگین 2/38 است که به ترتیب 56/8 و 57/8 درصد دانشجویان آن را در حد کم و خیلی کم برآورد کرده‌اند. براساس یافته‌های جدول 5، در سطح خطای 5 درصد، میزان حمایت سنجش از یادگیری‌های پایدار، به طور معنادار کمتر از سطح متوسط مقیاس (3) است. به عبارت دیگر، سنجش در مهارت‌ها و یادگیری پایدار دانشجویان و یا تقویت آن به حد کافی موفق نبوده است. بر اساس یافته‌های جدول 6 در مؤلفه‌ی یادگیری پایدار تفاوت میانگین پاسخ دانشجویان دو مقطع تحصیلی دارای تفاوت معنادار است. بدین ترتیب که میانگین در مقطع کارشناسی ارشد به طور معنادار بیشتر از مقطع کارشناسی بوده است. درباره میزان حمایت سنجش از یادگیری پایدار دانشجویان، بر حسب رشته تحصیلی جدول 7 و معدل تفاوت معنادار وجود نداشته است.

درصد دانشجویان، تصحیح اوراق امتحانی و تکالیف را تا حدودی دقیق دانسته‌اند و 37/7% دانشجویان این میزان را در حد کم و خیلی کم دانسته‌اند. براساس یافته‌های جدول 5، در سطح خطای 5 درصد، میزان دقت و صحت سنجش (2/86) به طور معنادار کمتر از سطح متوسط (3) است. بنابراین سنجش در هدف اندازه‌گیری خود از دقت و صحت کافی برخوردار نبوده است.

جدول 5 مقایسه میانگین پاسخ‌ها به مؤلفه‌ها با میانگین فرضی 3

مؤلفه	\bar{X}	S	Se	t	Sig
دقت و صحت سنجش	2/86	0/648	0/044	3/08	0/00
یادگیری‌های درازمدت	2/61	0/739	0/050	7/64	0/00

بر اساس یافته‌های جدول 6 بین پاسخ دانشجویان در مؤلفه‌ی دقت و صحت سنجش، بر حسب مقطع تحصیلی تفاوت معناداری ملاحظه نمی‌شود. همچنین میانگین پاسخ‌ها در این مؤلفه بر حسب معدل نیز معنادار نبوده است.

جدول 6 مقایسه میانگین پاسخ دانشجویان به مؤلفه‌ها بر حسب مقطع تحصیلی

گویه‌ها سنجش	دقت و صحت سنجش	مؤلفه یادگیری‌های درازمدت	کارشناسی		کارشناسی ارشد		t	p
			S	\bar{X}	S	\bar{X}		
			0/638	3/04	0/670	3/04	1/92	0/056
			0/723	3/01	0/678	3/01	3/94	0/00

براساس یافته‌های جدول 7 تفاوت میانگین پاسخ‌ها بر حسب رشته تحصیلی، در مؤلفه صحت و دقت سنجش در سطح خطای 5 درصد، معنادار بوده است. یافته‌های آزمون

جدول 7 مقایسه میانگین پاسخ دانشجویان به مؤلفه‌ها بر حسب رشته تحصیلی

P	F	مدیریت n=42		سیاسی n=33		علوم اجتماعی n=33		تاریخ n=31		روانشناسی n=36		علوم تربیتی n=38		مؤلفه‌ها
		S	\bar{X}	S	\bar{X}	S	\bar{X}	S	\bar{X}	S	\bar{X}	S	\bar{X}	
0/01	3/03	0/58	3/06	0/66	3/06	0/60	2/92	0/65	2/61	0/61	2/77	0/68	2/71	دقت و صحت سنجش
0/098	1/94	0/626	2/47	0/634	2/75	0/705	2/87	0/813	2/46	0/799	2/65	0/800	2/46	یادگیری‌های درازمدت

دهد که از نظر دانشجویان، سنجش به حد کافی از دقت و صحت برخوردار نیست و همه‌ی گویه‌هایی که با پایایی سنجش ارتباط می‌یابد مانند رعایت اصول علمی در ساخت آزمون‌ها، کنترل فرآیند سنجش برای جلوگیری از تقلب، دقت استادان در تصحیح اوراق و بررسی تکالیف، میانگینی کمتر از حد متوسط داشته است. براساس یافته‌های این مطالعه، اقدامات و روش‌های سنجش متداول از یادگیری-های پایدار دانشجویان، به حد کافی حمایت نمی‌کند. همه گویه‌های این مؤلفه، میانگینی کمتر از میانگین فرضی 3 داشته است. استفاده نکردن از تکالیف کاربردی و سنجش اصیل⁷ که بر اساس نیازهای واقعی طرح ریزی شده باشند این نتیجه را در پی دارد که دانشجویان اظهار داشته‌اند سنجش به توان کاربرد آموخته‌های دانشجویان کمک زیادی نکرده است. سنجش همچنین به لحاظ تشویق دانشجویان به سوی تفکر خلاق، افزایش توان مطالعه و یادگیری مستقل پس از فارغ‌التحصیلی و تقویت توانایی همکاری با دیگران و کارهای گروهی، هم موفق نبوده است. اغلب معلمان اینگونه گمان می‌کنند که مشارکت سبب تقلب می‌شود لذا دانشجویان را به کار گروهی تشویق نمی‌کنند. سنجش برای یادگیری پایدار و مادام‌العمر یک روش یا فن نیست، بلکه راهی بر مبنای تفکر دوباره درباره‌ی همه‌ی ابعاد اقدامات سنجش است. این تفکر باید در راستای توجه به تأثیرات قصد نشده‌ی سنجش تراکمی و یافتن راهی برای آن صورت پذیرد.

پی‌نوشت

¹ Reliability

² Validity

³ Portfolio

⁴ Quality Assurance Agency

⁵ Sustainable assessment

⁶ Tukey

⁷ Authentic assessment

مراجع

- [1] Boud D., *Sustainable assessment: Rethinking assessment for the learning society*, Studies in Continuing Education, Vol.22, No.2, 2000, pp.151-167.
- [2] Serafini F., *Three paradigms of assessment: Measurement, procedure, and enquiry*, The Reading Teacher, 2000.

4 - نتیجه‌گیری

از یافته‌های مطالعه‌ی حاضر، متنوع نبودن روش‌های سنجش، استفاده زیاد از امتحانات پایان ترم که فقط در یک روز برگزار می‌شوند و نیز استفاده‌ی زیاد از سؤالات تشریحی است. این نوع سنجش ممکن است به دانشجویانی که در انتهای ترم با عجله خود را برای امتحان آماده می‌کنند، نسبت به آنان که در طول ترم تلاش بیشتری کرده‌اند پاداش دهد. همچنین در استفاده زیاد از سؤالات تشریحی امکان صرف نظر کردن از حجم وسیعی از محتوا و باز هم گرفتن نمره‌ی خوب وجود دارد که از دقت و عدالت سنجش می‌کاهد. متنوع نبودن روش‌های سنجش نیز از عمده‌ترین عوامل کاهش پایایی و روایی سنجش است، زیرا ابعاد مختلف یادگیری را با روش‌های محدود سنجش نمی‌توان اندازه گرفت. استفاده از انواع بیشتر روش‌های سنجش سبب می‌شود که برداشت دانشجویان از سنجش به تجربه امتحانات چهار گزینه‌ای و تشریحی محدود نشود. عدم استفاده از خود سنجی و همتاسنجی که از یافته‌های دیگر مطالعه‌ی حاضر است نشان می‌دهد معلمان تنها ارزیابان سنجش هستند. این خود دلیلی بر عدم وضوح ملاک‌های سنجش و احتمالاً دلیل دیگری بر ناعادلانه بودن سنجش است. به کمک خودسنجی و همتا سنجی ملاک‌های نهفته در ذهن معلمان، برای دانشجویان شفاف می‌گردد. در تکالیف تشریحی نیز فقط توانایی به صف کشیدن فهرستی از حقایق و جزئیات سنجیده می‌شود. در حالی که سنجش با روش‌های جایگزین کیفیت‌ها و مهارت‌ها و صلاحیت‌های مختلف را در زمینه‌های گوناگون اندازه می‌گیرد. بر اساس یافته‌های مطالعه‌ی حاضر، عمده‌ترین روش سنجش، امتحانات پایان ترم است، و اکثراً تکالیف سنجش بیشتر به کنفرانس و ارائه گزارش محدود بوده است. سایر تکالیف‌های سنجش، همچون تکالیف‌های کاربردی و عملکردی، نوشتن مقاله، تکالیف‌های کتاب باز و کار در خارج از کلاس، کمتر از حد متوسط مورد استفاده قرار گرفته است. انجام نشدن سنجش‌های آغازین، کافی نبودن سنجش‌های طول ترم و تأکید بر امتحانات پایان ترم که از یافته‌های مطالعه‌ی حاضر است، بیانگر تکوینی نبودن سنجش، تراکمی بودن آن و خالی بودن از بازخورد سازنده برای دانشجو و استاد است. یافته‌های این مطالعه نشان می‌-

- [11] Boud D., Foreword in D. and Carless G., *How assessment supporting learning?*, *Learning-oriented assessment in action*, 2006, pp.ix-x.
- [12] Taras M., *Using assessment for learning and learning for assessment*, *Assessment and Evaluation in Higher Education*, Vol.27, No.6, 2002, pp.502-508.
- [13] Boud D., and Falchikov N., *Aligning assessment with long-term learning*, *Assessment and Evaluation in Higher Education*, Vol.31, No.4, 2006, pp.399-413.
- [14] Carless D., *Learning-oriented assessment, outcomes and trust*, *Invited seminar*, HKUST February, Available at: celt.ust.hk/obe/download/.../hkust_obe_Assessment_handout.pdf. 2008.
- [15] Rowntree D., *Assessing Students: How shall we know them?* (2nd edition), London, Kogan Page.1987.
- [16] Black P. and Wiliam D., *Inside the Blck Box*, Lodon: Kings College, 1998.
- [17] Nightingale P., Te Wiata I., Toohey S., Ryan G., Hughes C., and Magin D., *Assessing learning in universities*, Sydney: University of New South Wales Press, 1996.
- [3] Birenbaum M., and Dochy F.J.R.C., *Alternatives in assessment of achievements, learning processes and prior knowledge*, 1996.
- [4] Gibbs G., assessment supports students' learning, *Learning and Teaching in Higher Education*, No.1, 2004, pp.3-31.
- [5] Brown S., Race P. and Smith B., *500 Tips on assessment*, London, Kogan Page, 1996.
- [6] Sambell K., McDowell L. and Brown S., *But is it fair?: An exploratory study of student perception of the consequential validity of assessment*, *Studies in Educational Evaluation*, Vol.23, No.4,1997, pp.349- 371.
- [7] Quality Assurance Agency for Higher Education (QAA), *Blending assignments and assessments for high-quality learning*. Scottish Enhancement Themes: Guides to Integrative Assessment, Vol.3, 2007.
- [8] Baty P., *Flaws in marking revealed by QAA*. Available at: <http://www.time-higher-education.co.uk/story.asp?storyCode=186852§ioncode>, 2004.
- [9] Boud D. and Falchikov N., *Redesigning assessment for learning beyond higher education*, *Proceedings HERDSA Conference*, 2005, pp.34-41.
- [10] Entwistle A. and Entwistle N., *Experiences of understanding in revising for degree examinations*, *Learning and Instruction*, Vol.2, 1992, pp.1-22.