Studies in Educational Evaluation 40 (2014) 63–68

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/stueduc

Assessing university students' achievements by means of standard score (*Z* score) and its effect on the learning climate

Joseph Klein *

Bar-Ilan University, School of Education, Ramat-Gan 52900, Israel

ARTICLE INFO

Article history:
Received 7 June 2013
Received in revised form 7 November 2013
Accepted 15 December 2013
Available online 4 January 2014

Keywords:
Student assessment
Standard scores
Cooperative learning

ABSTRACT

Scholastic achievements are reported in sequential scores. The hypothesis examined was that incorporating a standard score in the achievement record would resolve these problems, increase interpersonal competition and decrease cooperative learning among students.

182 students in Economics, Education and Social Work, who their achievements were evaluated on a scale of 0–100, completed a questionnaire of learning climate and personal learning style. Afterwards they were told of plans to include a standard score in their grade record with an explanation of its implications regarding student ranking. The participants then completed the same questionnaire a second time.

When the standard score was included in the evaluation, the climate became significantly more competitive in all three majors.

Introduction

Institutions of higher education report academic achievements on a scale of grades ranging from A–F, 0–100 or 1–20. These grades reflect factual results without any statistical transformation or conversion, which is why they are called "raw scores." In the absence of comparative information about the achievements of other students these grades can be misinterpreted. In order to facilitate a more comprehensive judgment of academic results, it has been proposed that a standard score be added to the achievement record in order to designate the relative standing of learners within their cohort. A growing number of institutions of higher education inform students of their comparative achievements but few of them publish this information on official grade transcripts. Such a measure might be of use to employers and other academic institutions in their evaluation and ranking of candidates, especially in an era of grade inflation.

However, before deciding to include a standard score in official grade records, its potential effects on academic and social interrelations among students should be examined. Such an evaluation is necessary because of the unique nature of the standard score, which limits the number of students that can be categorized as outstanding.

Review of the literature

The background for recommending the incorporation of a standard score in grade reports in higher education:

> Monitoring achievements, an integral part of the teaching–learning process, has three main functions: to provide information about the nature of the learning that transpired (Burton, & Ramsit, 2001, 1998); to encourage future learning (Cole, 1993); and to examine how well the teacher has carried out his or her task (Beller, 2010; Mao & Zakrajesk, 1993). Such monitoring examines proven academic achievements and includes impressionistic judgments when factual results do not sufficiently reflect learners' performance and abilities (Angoff, 1971, chap. 15).
>
> In the literature, analysis of factual learning achievement is called "measurement", while judgmental impressions are termed "assessment." The decision of whether to determine a student's grade on the basis of measurement alone or to incorporate some element of assessment as well depends on the approach of the lecturers who plan the learning tasks and on the policies of the learning institutions. Assessment is prevalent in the early stages of the educational process in school as a means of encouraging and fostering motivation to learn. In higher education the accepted practice is to assign the main weight of grade designation to factual results (Koretz & Hamilton, 2006; Wikström, 2005). This age distinction is not dichotomous and assessment can be found in post-high school studies as well.

* Tel.: +972 3 5317734; fax: +972 3 7384029.
E-mail address: Joseph.Klein@biu.ac.il

0191-491X/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.stueduc.2013.12.002

This is especially evident in bachelor degree studies (Mao & Zakrajesk, 1993) and has become more widespread in master's degree studies, mainly in programs that do not include research theses (Kfir, Fresco, & Benjamin-Paul, 2003).

The lack of uniform policy regarding measurement and assessment in higher education can engender inequality in achievement opportunities for students. This is evident, for instance, in parallel courses given by different lecturers who test their students independently, as well as in differences in overall grade levels in different departments (Baird, 1988; Cameron & Ettington, 1988). Measurement validity has also come under criticism because of grade inflation. This is seen when students score high on internal examinations but do not exhibit parallel achievements in external indicators (Cluskey & Griffin, 1997; Koretz, 2005; Marshall, 1997). Several explanations can be given for this phenomenon. One is based on the perception of students as consumers and customers, and the surveys that schools conduct to evaluate "customer" satisfaction. Lecturers striving to boost their popularity ratings in surveys sometimes hope to improve their standing by repressing criticism of students and their learning. Some lecturers continue to believe that awarding high grades will raise their own ratings, despite research showing no correlation between lenient academic requirements and improved learner ratings of their teachers. In fact, results may be contrary to the lecturers' hopes (Hativa, 2000). High grades help students when they graduate from one learning framework or level to another. High grade inflation raises the specter of low level teaching and an intentional bias in grades, making it more difficult to distinguish between students' learning levels so that appropriate differential assistance can be provided. Grade inflation motivates students to prefer courses known for their high average grades and it creates obstacles for employers who expect grade records to reflect clearly the abilities of applicants for work (Cole, 1993; Sonner, 2000). In the United States, average grades in colleges rose by 15–20% in the 1990s. At Princeton University, for example, the frequency of A grades rose from 33% in 1981 to 40% in the 1990s. A similar trend was seen in Harvard and Stanford. The change was also accompanied by a narrowing of the grade distribution range (Bilimoria, 1995). Nor are high schools immune to grade inflation. As a result, candidates registering for universities have higher high school average grades, and only during the first year of academic studies do the low performance levels of some of these students become apparent (Koretz, 2005; Marshall, 1997).

Means for improving grade validity

Awareness of the shortcomings entailed in measuring achievements has spurred academic institutions to encourage lecturers to participate in workshops for developing skills in valid and reliable measurement (Germain & Scandura, 2005). Proposals include making B− and C+ the median grade and including on official records a measure indicating the relative achievement of a given learner in comparison to other students in the same learning framework (Dorsey & Colliver, 1995; Salvia & Ysseldke, 1995). Two ways of listing relative achievements were discussed: one entailed publication of the relative standing of students on a sequential scale such as percentiles. The second proposal advocated the integration of a standard score in the official achievement report. Such a measure describes the extent to which a given student's raw score deviates from the average of the entire learning cohort in terms of standard deviations. One advantage of the standard score is its simplicity and intuitiveness (Walhout, 1997). It provides a criterion that can be used for comparing departments that employ different measurement and assessment criteria (Baird, 1988; Cameron & Ettington, 1988; Entwise & Tait, 1990). However, before a decision is made whether to update achievement records by incorporating new monitoring measures, it is essential to examine whether such a step is merely a technical alteration or whether it also entails implications for learning and for social ties. When a standard score is employed, only a small number of students can be ranked as outstanding in the normal distribution of achievements, thus creating the potential for greater competition and a decrease in academic cooperation.

Academic cooperation and competition among students

Academic studies are characterized by cooperation and competition. The balance between them is affected by students' personality traits and learning habits, local culture, and the type of task involved. Academic collaboration increases intellectual productivity and cooperative thinking (Attle & Baker, 2007; Marwell & Schmitt, 1975). Team thinking helps when large quantities of information must be processed; it promotes brainstorming and thus provides a rich perspective of the subjects under discussion (Blinder & Morgan, 2005; Kerr & Tindale, 2004; Tjosvold, Sun, & Wan, 2005). Cooperative thinking stimulates a methodical approach because all those involved are obliged to explain their proposals convincingly to the others (Davis & Toseland, 1987; Kameda, Tindale, & Davis, 2003; Komiya, Kusumi, & Watabe, 2007). It helps to foster good interpersonal relations and a positive learning climate, while also raising motivation to succeed (Johnson, Johnson & Holubec, 1994). Group learning creates a classroom climate that is characterized by mutual assistance and cooperation (Dyson & Grineski, 2001). Few studies have examined the link between classroom climate and academic achievements in higher education, but its potential can be inferred from studies conducted in lower levels of schooling. It has been found that classroom climate and cooperative learning are significant predictors of learning products (Attle & Baker, 2007; Dorman, 2009; Fraser, 2007; Goh & Khine, 2002). A study that included 82 classes found that perception of classroom climate explained a high percentage of variance for cognitive and affective achievements (Haertel, Walberg & Haertel, 1981). A meta-analysis conducted by Johnson, Maruyama, Johnson, Nelson, and Skon (1981) surveyed 122 studies that compared individualistic and cooperative learning effectiveness in competitive and non-competitive conditions. They found that cooperative learning contributed to higher learning achievements than did individualistic and interpersonally competitive learning.

According to one economic social norm, competition itself is perceived as a creator of motivation and an impetus to excellence. Competition in learning has three forms: direct, indirect and shared, each of which has positive and negative aspects (Graham, 1976; Johnson & Johnson, 1999). Lam, Yim, Law, and Cheung (2004) and Michaels (1977) reported the contribution of competition to fostering motivation and learning achievements in class. In contrast, other studies found that excessive competition caused negative side effects, such as fear of failure among learners and hostility toward competitors (Johnson & Johnson, 1974; Nichols & Berliner, 2007). Umbach and Potter (2002) examined differences in learning culture in several departments of institutions of higher education and found that more selective departments were marked by a more competitive atmosphere and less cooperation. The choice of a communal or individualistic learning style is also affected by intrinsic factors such as self-regulated learning (SRL), which results when learners attribute significance to the learning materials and set goals and strategies for learning them. SRL is further intensified when learners self-monitor the quality of their learning and compare it to the desired results by means of detailed