

Measuring student development using points

Munir Mandviwalla, David Schuff, Laurel Miller, Manoj Chacko

Department of Management Information Systems, Fox School of Business, Temple University, Philadelphia, USA.

Abstract

While higher education is effective at measuring the acquisition of knowledge, it is less successful in quantifying other types of learning such as learning to know, live together, and to be. This is a problem because it makes it difficult for institutions to implement and sustain student development programs. In this paper, we describe how to measure student development using points. Using data from 623 students over 7.5 years, we show how our points system was used to improve student development over time by focusing on the quantity, variety, quality, and distribution of development activities. Based on our findings, we recommend avenues for additional research.

Keywords: Higher education; student development; measurement.

1. Introduction

Higher education is typically associated with physical, cognitive, and personality development of young adults (Kail and Cavanaugh, 2018). Yet, a major United Nations study concluded that “formal education systems tend to emphasize the acquisition of knowledge to the detriment of other types of learning; but it is vital now to conceive education in a more encompassing fashion” (Delors et al. 1996, p. 37). However, higher education still primarily focuses on one aspect of human development – academics – as measured by GPA.

Delors et al. (1996) place human development at the core of higher education, the principle of which has been widely adopted by universities (Kilpatrick 2019). We interpret this to mean going beyond pedagogy (e.g., teaching tips, GPA), scale (e.g., MOOCs), and reach (e.g., distance learning) to focus on value generation activities that develop students. Student development has always been a major goal of universities (Kilpatrick 2019), and availability of development activities is typically included in institutional assessment (Zilvinskis et al. 2017). Yet, to date there are no systematic and scalable measures of student development. Therefore, our research question is: *How can we measure student development?* This question is important because for student development to become a core activity and move beyond a nice-to-have but difficult-to-attain goal, it must be measurable.

2. Student development

According to Delors et al. (1996), learning can be broadly conceptualized as: *learning to do* - the acquisition of skills and competencies, *learning to know* - the ability to think and integrate new information, *learning to live together* - understanding others, managing conflicts, and *learning to be* - developing one’s personality and judgment. We adopt this broad view of student development. Student development is also related to forming an identity. Chickering and Reisser (1993) conceptualize establishing identity by proposing seven vectors that include competence, emotions, autonomy and interdependence, interpersonal relationships, identity, purpose, and integrity. Student development and identity relate to employability, which Yorke and Knight (2006, p. 8) define as “A set of achievements – skills, understandings, and personal attributes – that make individuals more likely to gain employment and be successful in their chosen occupations, which benefits themselves, the workforce, the community, and the economy.”

Another important stream of research focuses on student involvement (also termed engagement), which is “the amount of physical and psychological energy that the student devotes to the academic experience” (Astin 1984, p. 518). According to Astin (1984), student learning and personal development are a function of the quality and quantity of student involvement. Today, involvement is typically seen as an indicator of institutional excellence as well as the effectiveness of education policy and practice (Axelson and Flick 2011).

Involvement includes traditional academic activities such as studying - learning to do - as well as experiences such as discovery, engagement, and feedback of ideas, cultures, places, and others that relate to Delors et al.'s other forms of learning. Zilvinskis et al. (2017) show that these different forms of student engagement increase the Delors et al. expanded view of learning. While the above research is important for understanding the meaning and importance of student development, it provides little guidance on measuring, implementing, and weighing different forms of development.

Integrating the above literature, we define **student development** as the *achievement of Delors et al. (1996) outcomes of learning to do, know, live together, and be which establish identity and increase employability*. To establish a boundary and to acknowledge its formative and emergent nature, we conceptualize *development as a longitudinal process that involves internal and external activities of varying depth and value, which in turn differentially influence aspects of the above outcomes*.

3. Measuring student development

Engagement, involvement, and development activities are currently measured with survey instruments upon graduation (Astin 1984; Kuh 2001), while others have explored proxy measures such as learning management system (LMS) log files. These are cross-sectional perceptual measures. We instead follow a process view and operationalize student development as a series of activities accomplished at different times measured by points. The point value of each activity is associated with its development value.

Points are common in gamified systems (Liu et al. 2017). They are tangible and provide feedback, which in turn influences motivation by fostering competence, relatedness, and autonomy (Ryan and Deci 2000). Point-based rewards can also produce recurrent behavior (Liu et al. 2017). Points communicate recognition and a sense of accomplishment for activity and task completion, and the quantification enables social comparison and competition. Points are also more objective than the perceptual surveys used in the involvement literature (Astin 1984; Kuh 2001), and related survey measures of acceptance, usage, and satisfaction.

Overall, points work well as an individual measure that can quantify the number and weight of development activities. For example, if a student completes an internship, they might be awarded 300 points, while attending a talk by a speaker might yield only 25 points - reflecting the differential developmental value of each respective activity. The total points that a student has earned per term (points per term) and over the course of their degree program (total points per student) provide a summary of an individual student's development. The average of total points per student provides an aggregate measure of performance, at the academic department (our focus) or institutional university level.

The above measures are important because they focus on the *quantity* of student development activities, but they are not sufficient. Point totals do not address *variety* - the number of different activities (e.g., completing a project and an internship vs. attending two lectures), *quality* – the intrinsic development value of different activities (e.g., leadership vs. attendance activities), and *distribution* (e.g., all the activities in one term vs. spread out across multiple terms). For example, a student earning 500 points may seem very good, but perhaps they earned those points by completing only a few, and potentially similar, high-value activities in their last term before graduation. Addressing variety, quality, and distribution is important to ensure that development is holistic in addressing different aspects of learning (e.g., learning to live together vs. learning to know). Further, our goal is to measure development as a *process*, in which students develop over time by engaging in a variety of high-quality activities.

4. A field experiment in a living lab

We implemented points as a student development measure through a novel web-based self-service technology platform. The platform manages and records the points earned by each student for activities such as experiences (e.g., internships), career awareness (e.g., career fairs), leadership (e.g., officer position in club), enrichment (e.g., study abroad, competitions), communication (e.g., conferences, social activities), team work (e.g., community service, team projects), and workplace readiness (e.g., mentoring, site visits).

The authors' home department served as a living lab to implement and study the use of points as a measure for development. Living labs are a real world test and experimentation environment which enable co-creation of innovation among stakeholders and creators. Living labs offer incremental and visible improvements that reduce fear of failure and co-opt sources of resistance into co-designers (Hyysalo and Hakkarainen 2014; Mandviwalla et al. 2008) by exposing stakeholders to successive prototypes (Mandviwalla 2015). The stakeholders included students, friends and family, faculty, staff, college and university administration, and employers. The stakeholders participated in the development of each iteration of the platform through feedback and use.

The project evolved considerably over a decade of incremental improvements involving more than 7500+ students. To the stakeholders, we positioned student development as a program of learning that *complements but is separate from academics* in which students were expected to gain 1000 points prior to graduation. Over time, we integrated variety, quality, and distribution into the system as follows. *Variety* – we placed restrictions on how many times an activity can be repeated for credit, promoted activities offered across the university and the local community, and adjusted point values so that a student can only meet point expectations by participating in more than a few activities. We also allowed students to

propose new activities. *Quality* - We adjusted the point value of activities to reflect their development potential (e.g., internship receives more points than attending a lecture). *Distribution* - We restricted how many times a student can get credit for an activity (e.g., receive points for attending a lecture only once a term).

5. Results

We collected data from 623 students over a 7.5-year period of using the platform. For each student, we recorded the points they earned for the eight consecutive terms leading up to their graduation as well as the total number of points they earned upon graduation. Table 1 shows the percentage of students attaining point levels in the years following the platform's implementation, in which each year includes the group scheduled to graduate that year (only). As the program was designed for students to achieve 1,000 points before graduation, we chose point cutoffs to designate low, average, above average, advanced, and very advanced achievers. The results show that over time a greater percentage of students achieve higher point levels with 100% achieving the 1000-point expectation by the fifth year. Close to 100% of graduating seniors achieved at least 1,000 points by year 4, compared to a little over 9% in year 1. Over the course of the five years, the percentage above 1,400 points grew from 1.3% to 12.6%.

Table 1. Percentages of students and points levels

Point levels/Year	1	2	3	4	5
<i>Poor (0-300)</i>	76.0%	6.4%	0.0%	2.0%	0.0%
<i>Low (300-999)</i>	14.7%	7.3%	1.1%	0.0%	0.0%
<i>Average (1000-1200)</i>	6.7%	49.1%	59.8%	65.3%	80.0%
<i>Advanced (1201-1400)</i>	1.3%	10.0%	7.6%	8.2%	7.4%
<i>Very Advanced (1400+)</i>	1.3%	27.3%	31.5%	24.5%	12.6%

The results suggest that (a) students will, over time, adopt and embrace new measures of development, and (b) that once you start measuring an activity, perceptions change, and development becomes more frequent. Further, given the structures we put in place, students engaged in a variety of high quality development activities. We measured variety by reviewing the average number of activities. In addition, we adjusted the point values so that to meet expectations students had to participate in high value point activities (quality).

However, it was unclear if the activities were distributed over time or bunched together. Figure 1 shows activity across eight terms (T1 – T8) which is about 2.5 years. The data was standardized to enable comparison so that for a student graduating in any year, T1 is when

they started participating (eight terms in the past). The later years (3, 4, and 5) show more activity with less variance compared to the earlier years. The figure suggests that compared to year 1, year 5 graduates are more developed because they completed more activities each term at a consistent level, reflected by the lower variance.

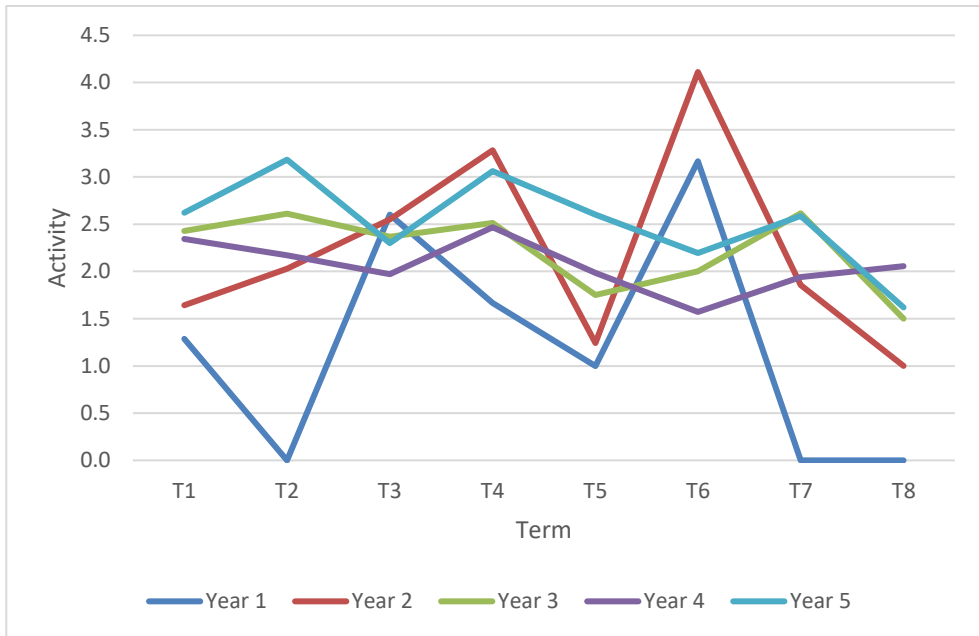


Figure 1: Development activity

While illuminating, the above analysis requires qualitative interpretation of graphs. We drew inspiration from Shannon’s diversity index (Shannon 1948), which has been used to assess species diversity in biology and ecology (Spellerberg and Fedor 2003). We developed a novel application of the Shannon Diversity Index to measure Development Distribution (DD). DD is expressed as: $-\sum_{i=1}^N p_i \log_2 p_i$, where N is the number of terms and p_i is the proportion of activities completed in a particular term. DD is useful because summing the number of activities only measures what is termed richness, as opposed to diversity, which is the function of the relative frequency of different species (Keylock 2005, p. 203). In our case, DD calculates the activity distribution across terms, i.e., the development process. So that higher DD implies activities are more evenly spread out across more terms. For example, consider six activities completed over five terms. Completing two activities every alternative term across 5 terms (2,0,2,0,2) generates a DD score of 1.585. In contrast completing two activities in the first term and one in each term thereafter (2,1,1,1,1) generates a DD score of 2.251. Another scenario where activities are concentrated in terms 2 through 4 (0,1,4,1,0) generates an DD score of only 1.252. DD is not affected, however, by the total number of activities if the distribution remains the same. Consider a scenario where the distribution of

activities is (3,0,3,0,3). Even though there are now nine activities instead of six, the DD score remains the same (1.585). This means that DD is a more useful measure when combined with total points earned. In sum, DD rewards better distribution, meaning more time to absorb, reflect on, apply, and experience the benefits from each activity as well as apply what was learned to the next activity (e.g., an officer position one term, followed by taking the lead role in a competition in the following term).

Table 2. DD scores over five year period

Year	DD
1	1.57
2	1.84
3	1.97
4	2.09
5	2.18

Table 2 summarizes DD scores of 623 graduates participating in 6,474 activities totaling 617,558 points across five years, in which the minimum possible score is 0 and 4.31 is the theoretical maximum. The gradual increase in DD scores suggests that students were completing activities that were more evenly distributed across terms. This result matches what is observable in figure 1, providing an intuitive validation of the efficacy of the measure.

DD needs to be further improved since it treats all activities as equal, even though it is likely that at different times, for different students, the importance of learning to do, know, live, and be will vary. In addition, we don't know which factors motivate and influence the trajectory of student development. Finally, we also need additional research on how to benchmark points and DD scores at the individual and institutional level.

6. Conclusion

In this study, we show how we developed a measure of student development in a living lab using a technology platform. For student development to move from the nice-to-have to the essential activity conceptualized by Delors et al. (1996), it must be measurable. The results show that points can serve as a measure of student development, including the quantity, variety, quality, and distribution of development activities. Overall, as far as we know, we are the first to measure student development across time using direct rather than perceptual or proxy measures.

References

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of college student personnel*, 25(4), 297-308.

- Axelson, R. D., & Flick, A. (2010). Defining student engagement. *Change: The magazine of higher learning*, 43(1), 38-43.
- Chickering, A. W., & Reisser, L. (1993). *Education and Identity. The Jossey-Bass Higher and Adult Education Series*. Jossey-Bass Inc., Publishers, 350 Sansome St., San Francisco, CA 94104.
- Delors, J., Amagi, I., Carneiro, R., Chung, F., Geremek, B., Gorham, W., ... & Suhr, M. W. (1996). *Learning. The treasure Within*. Paris.
- Hyysalo, S., & Hakkarainen, L. (2014). What difference does a living lab make? Comparing two health technology innovation projects. *CoDesign*, 10(3-4), 191-208.
- Kail, R. V., & Cavanaugh, J. C. (2018). *Human development: A life-span view*. Cengage Learning.
- Keylock, C. J. (2005). Simpson diversity and the Shannon–Wiener index as special cases of a generalized entropy. *Oikos*, 109(1), 203-207.
- Kilpatrick, S. W. (2019). Visualizing the Future: Work-Integrated Learning and the Psychosocial and Professional Development of Undergraduates.
- Kuh, G. D. (2001). The National Survey of Student Engagement: Conceptual framework and overview of psychometric properties.
- Liu, D., Santhanam, R., & Webster, J. (2017). Toward Meaningful Engagement: A Framework for Design and Research of Gamified Information Systems. *MIS Quarterly*, 41(4).
- Mandviwalla, M. (2015). Generating and Justifying Design Theory. *Journal of the Association for Information Systems*, 16(5), 314.
- Mandviwalla, M., Patnayakuni, R., & Schuff, D. (2008). Improving the peer review process with information technology. *Decision Support Systems*, 46(1), 29-40.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1), 68.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.
- Yorke, M., & Knight, P. T. (2006). *Embedding employability into the curriculum* (Vol. 3). York: Higher Education Academy.
- Zilvinskis, I. F., & Fedor, P. J. (2003). A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon–Wiener’ Index. *Global ecology and biogeography*, 12(3), 177-179.
- Zilvinskis, J., Masseria, A. A., & Pike, G. R. (2017). Student engagement and student learning: Examining the convergent and discriminant validity of the revised national survey of student engagement. *Research in Higher Education*, 58(8), 880-903.