Research in Educational Administration &Leadership

Volume: 2, Issue: 2/ December 2017

How can Student Learning Data at Institutional Level Support Decision-Making for Educational Improvement for Academic Programme? A Case Study in A Hong Kong University

Joseph Chow

The Hong Kong Polytechnic University, Hong Kong, China

Abstract

Student surveys have been commonly used by university administrators for assessment of student learning outcomes in higher education. The information collected from such stakeholders' surveys has only been to use as supporting documents for the effectiveness of university education. More recently, there have been calls for utilizing student feedback information to improve quality of education at the programme level and guide educational leaders and teachers about implementing data-based decision-making. This study aimed to explore the learning gains in the undergraduate education perceived by its alumni in one of Hong Kong universities. The study reported the procedures in examining the assessment scores and discussed the interpretation of assessment results at the programme level. The findings showed that interpretation of the student ratings of perceived learning gains could not be appropriately aggregated at the programme level. Implications for theory and practice of educational data use and for educational administration are discussed.

Article Info

Article History:

Received
March 17,2017

Revision received
June 4, 2017

Accepted
June 5, 2017

Keywords:
Student learning data
decision-making
case study
Hong Kong

Cite as:

Chow, J. (2017). How can student learning data at institutional level support decision-making for educational improvement for academic programme? A case study in a Hong Kong University. *Research in Educational Administration & Leadership*, *2* (2), 144-169.

Assessment of Student Learning for Multiple Purposes

In accountability era, student learning outcomes are of important focus in demonstrating effectiveness of higher education. Universities are encouraged to demonstrate their graduates acquired certain level of proficiency of learning outcomes through measurement, assessment and evaluation efforts (Coates, 2016; Melguizo & Coates, 2017; Secolsky & Denison, 2012). In particular, information obtained from student learning assessment serves three major purposes: accountability use and improvement use, and instruction use (Ebbeler, Poortman, Schildkamp, & Pieters, 2016), which are briefly described as follows.

The first purpose of using student assessment data is for providing indicators for documenting the effectiveness in various learning outcomes and experience. In particular, assessment data of the student perceived learning gains in terms of graduate attributes is an emphasis in higher education in Hong Kong. Universities are held accountable to make data-informed review of the student learning outcomes and effectiveness of programme of study. Higher education administrators are faced with accountability pressure on demonstrating learning gains by the students. The second purpose of using student assessment data is for educational improvement use. Such improvement that involves use of data includes reviewing and revising the existing curriculum, setting and adjusting the key performance indicators (KPIs) for units, and formulating improvement goals in specific areas. The third purpose of using student assessment data is for instructional use. Examples of using data to support instructional use include making data-driven adjustment to classroom instruction, identifying student learning needs, monitoring student mastery upon completion of a course.

Completing the Assessment Cycle

In higher education settings, student learning data from multiple sources including student surveys are often used for institutional research and evaluation of accountability (Chatman, 2007; Douglass, Thomson & Zhao, 2012) and expected to support planning for improvements at various levels. In terms of guiding educational improvement efforts in higher education institutions, the resulting data-based activities are often framed as going through iterative steps in completing assessment cycles. For example, Liu (2017) mentioned the recent trend on using assessment data for internal improvement beyond traditional accountability purposes in higher education institutions. Musekamp and Pearce (2015) have also discussed a case study using low-stake assessment results to inform and make evidence-based decisions to realize educational improvement in engineering higher education.

Rationale of This Study

Previous studies have been reported from a perspective of using the assessment results for quality assurance purposes at the institutional level. This study has a particular focus on examining whether using assessment results can support educational improvement at the programme level, and in particular how educational administrators should promote internal improvements via appropriate data-based decision-making and data use. In school settings, there has been growing studies of appropriate data use in school guiding for school improvement (e.g., Marsh, 2012). For example, Thompson, Adie and Klenowski (2017) questioned the legitimacy of using the NAPLAN results for making valid comparisons between schools, and bring the discussion to educational policy and administration. However, relatively less studies were reported on using data for improvement in higher education institutions. Some of the existing studies are on comparison of institutional effectiveness against the international

benchmark or comparison with other higher education institutions that are of similar contexts (e.g., Kuh, Jankowski, Ikenberry, & Kinzie, 2014).

McKenow and Ercikan (2017) investigated the multi-level validity for documenting student learning at the programme level based on the National Survey of Student Engagement, also known as NSSE, a well-known instrument used internationally for quality assurance of university education. For local instruments that collect student data for local use at higher education institutions, there is relatively less coverage in the literature. This study will investigate how such instruments as alumni survey is developed and used locally by the university.

This study will explore results of the alumni survey from views of recent graduates of the university. This adds values to the conventional studies on university student learning which focused on perception of the current students, for example, either freshmen or seniors, who are still studying in the university. The alumni, in contrast, feature themselves by perceiving relatively lower gains in the graduate attributes then current students who have not graduated. Common reasons reported showed that while graduating students feel quite prepared for entry level employment, employers have concerns about graduates' preparedness for work (Hart Research Associates, 2015) and that students downplay the value of their degree in preparing them for their future careers when asked about this six or more months post-graduation (Rayner & Pakakonstantinou, 2015).

While it is recognized the three purposes of collecting assessment data of student learning (i.e., accountability, improvement, and instruction) are inter-related, this study will mainly focus on the relationship between using data for accountability and using data for improvement purposes. It is common to see due to policy and accountability pressure that stakeholders, especially university leaders, who expect student learning data such as alumni survey data

is useful for multiple purposes, including educational accountability purposes as well as for educational improvement purposes. For accountability purposes, the focus is often on using the data to provide evidence to external stakeholders for documenting the institutional effectiveness. For educational improvement purposes, the focus is often on using the data to support decision making for policy and implementation at the local programme level.

Despite the common assumption of wide applicability of student survey data, the validity of interpretation and use of assessment results at the programme level, however, is seldom examined prior making conclusions at the programme level. This study will illustrate the importance of investigating the legitimacy of making inferences for planning and implementing educational improvement initiatives based on assessment results. It aims to contribute to the understanding of theory and practice of data use for quality assurance and improvement in higher education.

Alumni Perception of Educational Experience

Student surveys are administered to students or graduates for two major purposes. First, it serves as a stakeholder survey to enable universities to collect views from the students or graduates as one of the university stakeholders about their study. Second, it also allows the university to collect information to document the achievement of the student learning outcomes and experience.

The alumni survey in this study is designed and administered with the assessment goals of collecting alumni's views about their overall learning experience of the university education over their study period. In particular, such information collected will provide the basis for assessment data interpretation and use to support making evidence-based decisions and practice in quality assurance and improvement in both the institutional (i.e. the university level) and unit levels (e.g., programme level). The ratings from the alumni are expected to be used to continuously improve university's

programmes as the faculties and departments review the group-level data to revise their approaches. This study will focus on the usefulness of the assessment data for informing decision-making and planning improvement efforts at the programme level. The compatibility of the assessment goals valued by stakeholders at different levels of the institution will also be examined and discussed.

Variation Across Programme of Study

Using assessment results to guide educational improvement decision and efforts might be more complex to be done than it is intended. There are multiple issues to consider, such as representativeness of the sample, the criterion, and the measures for driving improvement (Judd & Keith, 2012). In particular, the issues of appropriate level of aggregation and analysis are important for making score-based decision regarding programme review and educational improvement (Griffith, 2000). Previous research showed that appropriateness of aggregation could be variable dependent. For example, McKeown and Ercikan (2017) reported perceived general learning outcomes are reliable at the programme level for some variables but not the others.

Previous research have been focusing on the perspective the students who are currently enrolled in the university. For example, McKeown and Ercikan (2017) focus on the first-year student and the fourth-year students' response on NSSE. Similar in the more well-known international survey such as NSSE, this requirement of validity of assessment data is also applicable to local student learning surveys such as the alumni survey under investigation in this study.

Data-Use Theory of Action

Data-use theory of action framework was introduced and applied using data to collect and use evidence to support improvement in educational organizations (Coburn & Turner, 2011; Marsh, 2012). Such a framework (Figure 1) has several features: 1) it is embedded in

the organizational context in which data use takes place; 2) it involves interaction of characteristics and perception of multiple layers of stakeholders or components (e.g., educational leader, middle-managers, and the data itself); 3) it is an iterative process of using data to support improving educational outcomes (Ebbeler et al., 2016).

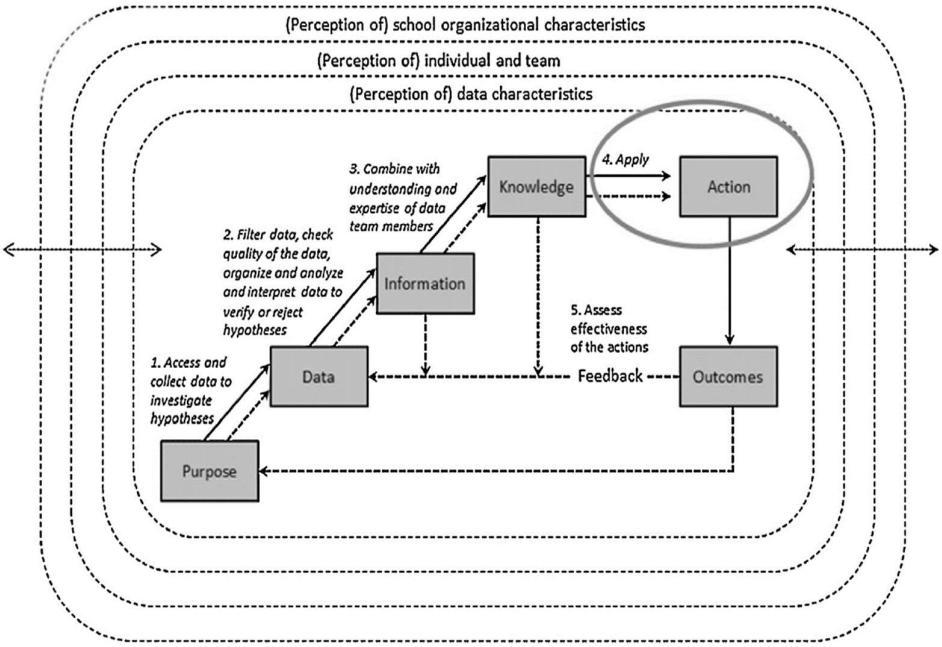
The data use process involves several components. It starts with asking the purpose of collecting data, then the data collected should go through quality check and be analyzed to produce meaningful information and knowledge for guiding action for educational improvement to achieve desired outcomes (Schildkamp & Poortman, 2015).

This study focuses on how the quality of data may be a limit to realize the educational leaders' aspiration of using same data for multiple purposes. In particular, the variation of student learning across programmes of study as shown in the alumni survey in this study will be used to illustrate the importance of checking the appropriate level of aggregation when there are multiple assessment goals motivated by different plausible purposes of educational data use.

It is argued in this study that unclear purposes of data use may lead to conflicting theories of action on data use at different levels within the educational institution that results in ineffective decision and planning for improvement. Using a case study with the alumni survey, this study will illustrate the importance of careful and informed application of data use theory of action for internal educational improvement.

Figure 1

Data Use Theory of Action, and Factors Influencing Data Use (Ebbeler et al., 2016; based on Coburn & Turner, 2011a, 2011b; Ikemoto &Marsh, 2007; Lai & Schildkamp, 2013; Mandinach, Honey, Light, & Brunner, 2008; Marsh, 2012, p. 4; Schildkamp & Kuiper, 2010; Schildkamp & Lai, 2013; Schildkamp & Poortman, 2015; Supovitz, 2010).



Method

Instrument

The alumni survey has a total of 67 items to collect student views on their learning gains, learning experience, and how study at the university has prepared their post-university life. It was administered to recent graduates within 18 to 24 months after graduation. In particular, alumni's perceived gains in learning outcomes were assessed in terms of the six areas of generic skills.

Measures

The university undergraduate programmes normally take four academic years to complete. It is the university's strategic initiatives to nurture the development of all-round students with competent levels of generic skills. In particular, it is expected that a typical graduate (or alumni) from undergraduate programme should possess and be able to demonstrate competence of generic skills in the six broad areas: competent professional, critical thinker, effective communicator, innovative problem solver, lifelong learner, and ethical leader (See Table 1). These attributes are usually called graduate attributes on which students were assessed with the assessment results helping the university to monitor student learning progress and the effectiveness of the university education.

In terms of student learning outcomes, the graduate attributes are considered as an important learning outcome of undergraduate students. Alumni's perception of their learning gains in these areas as completion of university education was collected in the alumni survey. Scales were constructed based on the items making up of the six graduate attributes separately.

The present study was based on the student perception of the learning gains as captured in the alumni survey. These measures are selected because these are deemed to be important learning outcomes of completing undergraduate studies in the university. After 18 to 24 months since graduation, most of the students are able to report

retrospectively their learning outcomes with reference to their lives beyond completion of university education.

Table 1

Six Areas of Generic Skills

Generic skills	**Description**
Competent Professional	Graduates should be able to integrate and apply in practice the fundamental knowledge and skills required for functioning effectively as an entry-level professional
Critical Thinker	Graduates should be able to examine and critique the validity of information, arguments, and different viewpoints, and reach a sound judgment on the basis of credible evidence and logical reasoning
Effective Communicator	Graduates should be able to comprehend and communicate effectively in English and Chinese, orally and in writing, in professional and daily contexts
Innovative Problem Solver	Graduates should be able to identify and define problems in professional and daily contexts, and produce creative and workable solutions to the problems
Lifelong Learner	Graduates should recognize the need for continual learning and self-development, and be able to plan, manage and improve their own learning in pursuit of self-determined development goals
Ethical Leader	Graduates should have an understanding of leadership and be prepared to lead a team, and should acknowledge their responsibilities as professionals and citizens to the society and their own nation, and be able to demonstrate ethical reasoning in professional and daily contexts

Samples

All the 2015 cohort of graduates of undergraduate programmes, including both the bachelor's degree and higher diploma degree, of the university were mailed a hard copy of the alumni survey in February 2017, a time point falling on between 18 and 24 months

from their graduation time. A souvenir was provided for respondents. In this study, only the responses of the bachelor's degree graduates were analyzed.

Programmes of Study

Alumni were grouped into their programme of study before invitation was sent for their participation in filling out the survey. Table 2 showed the distribution of the respondents by their programme of study and compositions of the sample which showed a considerable variation in sample sizes across programmes. Only programmes of study with at least five respondents were included.

Table 2

Sample of Alumni Survey Respondents and Their Programme of Study.

Programme Name	Respondent	%
Applied Psychology	20	1.8
Art and Design In Education	7	0.6
Chinese and Bilingual Studies	7	0.6
English for Business and Professional Communication	19	1.7
English Studies	14	1.3
Interactive Media	10	0.9
Accounting and Finance	14	1.3
Accountancy	60	5.4
Applied Biology with Biotechnology	13	1.2
Applied Ageing Studies	12	1.1
Bilingual Studies	20	1.8
Biomedical Engineering	6	0.5
Building Engineering and Management	13	1.2
Building Services Engineering	52	4.7
Chemical Technology	11	1.0
Civil Engineering	34	3.1
Computing	23	2.1
Convention and Event Management	11	1.0
Design	25	2.3
Electrical Engineering	17	1.5
Electronic and Information Engineering	11	1.0
Enterprise Engineering with Management	8	0.7

Enterprise Information Systems	15	1.4
Environment and Sustainable Development	25	2.3
Food Safety and Technology	7	0.6
Fashion and Textile Studies	60	5.4
Geomatics (Land Surveying)	13	1.2
Hotel Management	41	3.7
Industrial and Systems Engineering	9	0.8
Industrial Quality Management	6	0.5
Investment Science	7	0.6
Logistics Engineering and Management	10	0.9
Engineering Physics	20	1.8
Financial Services	7	0.6
Global Supply Chain Management	19	1.7
International Shipping and Transport Logistics	20	1.8
Management	8	0.7
Marketing	14	1.3
Mechanical Engineering	38	3.4
Medical Laboratory Science	15	1.4
Mental Health Nursing	8	0.7
Nursing	72	6.5
Occupational Therapy	24	2.2
Optometry	9	0.8
Physiotherapy	21	1.9
Product Analysis and Engineering Design	17	1.5
Product Engineering with Marketing	24	2.2
Property Management	11	1.0
Radiography	23	2.1
Social Policy and Administration	22	2.0
Social Work	40	3.6
Statistics and Computing	17	1.5
Surveying	43	3.9
Tourism Management	25	2.3
Transportation Systems Engineering	7	0.6
Total	1104	100.0

Analysis

Three Steps ANOVA

The variability of the student scores across programmes of study was investigated using ANOVA. A three-step ANOVA approach was used to examine the variability of the student ratings across programmes of study of the respondents. According to Griffith (2002), there are three indicators of the variability of the scores in the survey, namely, non-independence, reliability of group means, and within-group agreement.

Non-independence refers to the degree to which the variable of concern is dependable on the other variable. In this example, if the rating of perceived learning gains are influenced by the programme of study of the respondents, the mean ratings at the programme level should differ. The level of such dependability can be empirically examined by calculating the intra-class correlation (ICC) (1) values. The ICC (1) values indicated how much the variance of the perceived learning gains can be explained by the variable of programme of study. As mentioned, the sample size varied considerably across the students' programme of study in this study. In reducing the effect of extreme sample sizes across groups, the harmonic means were used (McKeown & Ercikan, 2017).

Reliability of programme-level means refers to the degree to which the rating of a single respondent within a programme would be precise estimate of the programme-level mean. The ICC (2) values indicated the level of which how many students are required to provide a reliable estimate of the programme-level mean.

Within-programme agreement refers the extent to which the respondents within a programme of study showed agreement in their ratings. For example, an individual respondent gave ratings of 1, 2, and 3 on three items on a 5-point scale, and another individual respondent gave ratings of 3, 4, and 5. As Griffith (2002) commented, a high reliability can be observed in this situation because the ratings

between respondents will be balanced. However, the within-programme agreement would be low because, for example, the rating of 1 by the first respondent corresponds to the rating of 3 by the second respondent. The r_{wg} indicates the extent to which respondents within a group gave similar ratings. This is why within-programme provides another perspective on how valid it is to interpret the survey findings at the programme level.

Results

Scale Statistics

Table 3 shows the descriptive statistics of the scales and respondents.

Table 3
Descriptive Statistics of The Scale and Respondents

Scale	Item	n	Mean values	Standard error	Cronbach's alpha
Competent professional	2	1091	3.43	0.79	0.81
Critical thinker	4	1086	3.47	0.61	0.81
Effective communicator	2	1090	3.39	0.76	0.62
Innovative problem-solver	3	1088	3.49	0.65	0.79
Lifelong learning	4	1086	3.45	0.68	0.79
Ethical leader	4	1089	3.34	0.71	0.77

ANOVA Results

Table 4 showed the results of ANOVA of the variability of the responses in the programme of study. All F values were statistically significant at the $p < .05$ level for all graduate attributes. The ICC (1) values ranged 0.03-0.08, across programmes of study, indicating a small variance between programmes of study. The graduate attribute that showed the largest variance in the response was professional competence, with an ICC (1) value of 0.08. For other graduate attributes, the ICC (1) values were around 0.03-0.06, indicating a low

dependence of rating on which programme of the respondent belongs.

The reliability measures, ICC (2) ranged from 0.11 to 0.62, which are considered very low (Griffith, 2002). The results indicated that the programme means were not reliable measures of the programme-level performance.

The within-programme agreement was reported by the statistics r_{wg}, which ranged from 0.29 to 0.49 for the six graduate attributes. The low values indicated a low level of agreement among the respondents under the programme of study.

Table 4

ANOVA Statistics of the Six Scales

Scale	*F* value	*p* value	ICC(1)	ICC(2)	r_{wg}
Competent professional	1.70	.002	.08	.62	.49
Critical thinker	1.54	.005	.04	.32	.32
Effective communicator	1.62	.004	.05	.28	.38
Innovative problem-solver	1.60	.004	.06	.19	.45
Lifelong learning	1.49	.006	.05	.25	.37
Ethical leader	1.44	.012	.03	.11	.29

Discussion

Using data to review education quality based on stakeholder surveys, including student survey, was widely practiced in universities. Assessment data of student learning outcomes in the form of self-reported learn gains in the alumni survey, was investigated in this study. More recently the focus of student survey as tool for assuring educational quality is shifting towards supporting planning and implementing educational improvement efforts.

The validity evidence for using assessment results regarding educational improvement is not commonly examined. The analyses above provided some validity evidence about the extent to which the

inference about programme-level performance is justified. The ANOVA results suggested that the assessment results were not useful for summarizing the student performance and thus making justified conclusion about the effectiveness at the programme level.

This raised the awareness of the possibility that data collected from student survey designed for informing the effectiveness of learning outcomes at the institutional level were not applicable for productive use for educational improvement purposes, such as at the programme level in university education.

While it is recognized that using data for accountability actions does not translate directly into educational improvement effects (Ebbeler et al., 2016), it does not rule out the possibility of using data for accountability purposes may also be used for educational improvement. However, when the assessment data is collected for accountability purposes, it should not be presumed that the data would be equally applicable for improvement purposes. In regards to completing the assessment cycle (Liu, 2017), this study examined and reported whether the data for educational accountability has adequate technical quality that allows interpretation and use of student assessment data for educational improvement at the programme level.

These results have shown that justification for interpretation of assessment data is important to examine because, as the results above showed, assessment results are not necessarily directly useful for interpretation and use beyond its primary purpose. Consequently, lack of validity evidence, or ignorance of collecting validity evidence, might lead to conclusions that are not well supported by evidence (O'Leary, Hattie & Griffin, 2017).

Implications for Educational Leadership and Administration

As mentioned earlier, Ebbeler and researchers (2016) mentioned that the data-based action that educational leaders can take can be categorized into three areas: accountability, instruction, and school

[educational] development, and added that 'although the data use actions are presented as three distinct categories, in reality they are intertwined' (p. 21).

Despite the desire of using the data to its full potential (e.g., Schildkamp et al., 2017), educational leaders should also be aware of the limitations regarding the nature of the educational data collected and its use. The results of the study suggested to university administrators that validity of interpretation of the assessment results has to be examined systematically prior any meaningful use of the assessment scores. University leaders should understand that accountability pressure for using data support improvement itself at the institutional level does not legitimize action for directly using assessment results at the programme level.

As mentioned, educational data use are promoted and practiced in universities for accountability reasons and improvement reasons. Educational leaders should also understand the importance of technical quality of the assessment data despite the policy and accountability pressure for data use for multiple purposes, especially when there are inadequate sources of evidence for which interpretation and use of assessment scores are made.

This study has highlighted the importance of validating the data when using it for educational improvement purposes. When assessment results are aggregated in attempt to inform evaluation of programme effectiveness, it should be cautioned that there should be adequate level of variability and across programmes before making claims about programme-level effectiveness. Implications for educational administrators include explicit examination of the technical quality of data available in drawing conclusions at appropriate levels, and leadership in building the organizational capacity for data use.

Appropriate data use depends on the organizational capacity to use data to enhance student learning. From the assessment data of the alumni survey, this study provided an example examining validity

evidence and its plausibility that it may prevent the administrators from making erroneous assumptions and unnecessary inferences. (Thompson et al., 2017). The implication for educational administration is that it suggested the need to maintain an adequate organizational capacity in understanding the score reports and its meaning to promote use of data for educational improvement. As Moss (2016, p. 248) mentioned, 'Of course, no collection of evidence at the organizational level can ensure the validity of any particular interpretation, decision or action, but evidence of capacity to use data well increases confidence in the likelihood of well-warranted decisions, including the likelihood that problematic decisions will be illuminated for further inquiry and revision.'

The results of the study highlighted the importance to build in higher education institutions capacity for data use, and to provide professional development for teachers and administrators on data literacy skills to make justified decisions in relation to educational improvements. University administrators should encourage middle-managers to access valid and reliable data and make inferences at appropriate levels to back the decisions. Leaders should also encourage more collaborative involvement within school about data use (Schildkamp et al., 2017), and recognize the collective responsibility in realizing data-based educational improvement initiatives.

Leading and planning for improvement would involve processes for engaging faculty and providing sufficient professional training and resources to faculty (Hutchings, 2010) to build assessment and data capacity at both individual and organizational levels. Marsh and Farrell (2015) suggested that leaders check the current status of the institutional assessment and data capacity, and target in data-use processes where more support is needed.

At the institutional level, the assessment strategies could be better addressed by explicitly recognizing the degree of 'incompatibility of assessment goals valued by stakeholders at different levels'

(Chatteriji, 2013, p. 303). A more informed data-based decision making process can be promoted by defining, recognizing and communicating well the purposes of the use assessment results and limitations of assessment data.

Other examples of possible efforts by university administrators include collaboration with academic development professionals to offer advice in using assessment results. The support offered by the expertise of academic developer should be included in the conversation between policy and practice of data-based decision making in higher education institutions. University leaders should build internal channels and systems that will facilitate the advice of assessment experts to be taken in actual planning and implementation of educational improvement initiatives. These include work of devising plans and channels for communicating assessments data internally for appropriate use at different levels, such as programme level, faculty level, or department level. Across the university, university leaders can also cultivate an assessment culture (Liu, 2017). In particular, Liu (2017) suggested faculty members work actively with assessment specialists, who can offer advice on the theory and practice on validity and validation (Brennan, 2006; Crooks, Kane & Cohen, 1996).

Implications for Application of Data-Use Theory of Action

The results of this study have implications for applying the data use theories of action under the framework of data-based decision making: in addition to the reliability and validity of the assessment instrument itself, it is important to check the data quality for interpretation and use at the aggregated level.

It should be noted that data-use theory of action should be examined carefully before taking actions for particular purposes. As mentioned, data use process involves multiple components and steps, including purpose, data, and knowledge. In particular, the linear relation of steps between different components, especially steps

between the components of *purpose*, *data*, and *action*, should be carefully examined. For example, using data with reference to educational improvement in student learning could be different from using data with reference to assurance of quality in education. While aspirations for using assessment data for multiple purposes—accountability, improvement, and instruction (Schildkamp et al., 2017)—should be encouraged; this study has shown that the one particular dataset might not be effective for multiple purposes, and these purposes may require some different quality of data. Failure in fulfilling the requirement of data quality may affect the validity of conclusion made based on the assessment data.

It is important to keep accountability and improvement efforts coherent, as Earl and Katz (2006) stated, 'Accountability without improvement is empty rhetoric, and improvement without accountability is whimsical action without direction' (p. 12). However, it is also important for educational institutions to recognize the challenges in realizing both accountability efforts and improvement efforts in the lack of quality data, or when the data that are available are more useful for one purpose than the others.

Conflicting theories of actions have been reported (Chatterji, 2013, p. 303) and such conflicts, particularly regarding educational data use, could be resulted from uses of student assessment results in multiple areas (some might not be the primary intended purposes) by different stakeholders. In the case shown in this study, the conflict of theories of action in data use lies on action of data use for quality assurance at the institutional level and action of data use for quality improvement at the programme level. Educational institutions should be aware of possible conflicting expectation of data use actions, and try to minimize the conflict, for example, when there are concurrent demands of data use for accountability purposes and improvement purposes.

Operationally, to complete the multiple phases of assessment cycles, including assessment development, validation work, and use

of assessment results, efforts should be aligned to function as an integrated whole, or 'a connected series of process' (Chatterji, 2013, p. 299). The link between assessment development, and use of assessment results by stakeholders via incorporating the use of validation of assessment results should be clearly recognized and strengthened.

The current study suggested *data characteristics*, one of crucial properties in data use theory of action (e.g., Schildkamp & Lai, 2013), should be understood and checked in advance for data use purposes beyond it is original design and intent; otherwise it could become a factor hindering the process of data-based decision making (Schildkamp et al., 2017).

This study has questioned that the perception that purposes of data use can be defined independently of particular sets of educational data and associated improvement actions. In applying the data use theory of action within educational institutions, it should be cautioned that one assessment dataset might not fit all purposes. As Moss (2013, 2016) mentioned, assessment information should be used together with other information to guide data use and data-based decision making in educational institutions. This is especially useful in local decision-making at universities. For example, the assessment data collected from student-reported surveys, such as alumni survey, should not be the only source of evidence of student learning. In this connection, the data of student learning outcome assessment is not the only source of evidence, there are possibly a wide range of sources of evidence including interview and discussion with students regarding their learning experience. This study has shown that that rely on one source of information could be risky, especially when the technical quality of evidence is not evaluated.

Conclusion

This study demonstrated the validation of using student learning survey results for educational improvement based on assessment data obtained from alumni perception of learning gains in graduate attributes based on the instrument designed locally for informing effectiveness of university education. The results suggested that valid interpretation of assessment results is not supported at the programme level. Data use for educational accountability purposes at one level is a necessary condition but not sufficient for data use for educational improvement purposes at another level.

This study also revealed the need for carefully examining the scope of the data use theory of action when taking associated action for educational improvement. This paper has argued that educational leaders should consider the legitimacy of making inferences for planning and implementing educational improvement initiatives based on assessment results that are intended to be reported at the institutional level. Despite the policy pressure from external stakeholders requesting to make most out of the data in decision making, institutions should be aware of the limitation of datasets available for specific purposes; otherwise, it may lead to erroneous decisions and thus ineffective improvement efforts. Recommendations to educational administrators and leaders include careful tackling possible conflicting theories of actions in data use, collaboration with academic development professionals, exploring multiple sources of evidences, as well as cultivating an assessment culture among administrators and teachers with support of assessment specialists.

References

Chatman, S. (2007). *Institutional versus academic discipline measures of student experience: A matter of relative validity (CSHE.8.07).* Berkeley: Center for Studies in Higher Education, University of California.

Chatterji, M. (2013). Insights, emerging taxonomies, and theories of action toward improving Validity. In Chatterji, M. (Ed.), *Validity and Test Use: An International Dialogue on Educational Assessment, Accountability and Equity* (pp. 273-307). Bingley: Emerald Group Publishing.

Coates, H., & Richardson, S. (2012). An international assessment of bachelor degree graduates' learning outcomes. *Higher Education Management and Policy*, 23 (3), 1-19.

Douglass, J. A., Thomson, G., & Zhao, C-M. (2012). The learning outcomes race: the value of self-reported gains in large research universities. *Higher Education Research*, 63(4), 317-335.

Erisman, W. (2009). *Measuring student learning as an indicator of institutional effectiveness: Practices, challenges, and possibilities.* Austin, TX: Texas Higher Education Policy Institute.

O'Leary, T. M., Hattie, J. A., & Griffin, P. (2017). Actual Interpretations and Use of Scores as Aspects of Validity. *Educational Measurement: Issues and Practice.*

Brennan, R. L. (2006). Perspectives on the evolution and future of educational measurement. In R. L. Brennan (Ed.), *Educational Measurement* (pp. 1-16). Westport, CT: Praeger Publishers.

Coates, H. (2016). Assessing student learning outcomes internationally: Insights and frontiers. *Assessment & Evaluation in Higher Education*, 41(5), 662-676.

Coburn, C. E., & Turner, E. O. (2011a). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research & Perspective*, 9(4), 173-206.

Coburn, C. E., & Turner, E. O. (2011b). The practice of data use: An introduction. *American Journal of Education*, 118 (2), 99-111.

Crooks, T. J., Kane, M. T., & Cohen, A. S. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*, 3(3), 265-286.

Earl, L. M., & Katz, S. (2006). *Leading schools in a data-rich world. Harnessing data for school improvement.* Thousand Oaks, CA: Corwin Press.

Ebbeler, J., Poortman, C. L., Schildkamp, K., & Pieters, J. M. (2016). Effects of a data use intervention on educators' use of knowledge and skills. *Studies in educational evaluation,* 48, 19-31.

Griffith, J. (2002). Is quality/effectiveness an empirically demonstrable school attribute? Statistical aids for determining appropriate levels of analysis. *School effectiveness and school improvement,* 13(1), 91-122.

Hart Research Associates (2015). *Falling Short? College Learning and Career Success.* Downloaded 18/11/2017 from https://www.aacu.org/sites/default/files/files/LEAP/2015employerstudentsurvey.pdf

Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the ''data driven'' mantra: Different conceptions of data-driven decision making. In P. A. Moss (Ed.), *Evidence and decision making (National Society for the Study of Education Yearbook,* Vol. 106, Issue 1). (pp.105–131). Chicago, IL: National Society for the Study of Education.

Judd, T. & Keith, B. (2012). Student learning outcomes assessment at the program and institutional levels. In C. Secolsky & B. Denison. (Eds.), *Handbook on measurement, assessment, and evaluation in higher education.* (pp.31-46). New York, NY: Taylor and Francis.

Kuh, G. D., Jankowski, N., Ikenberry, S. O., & Kinzie, J. (2014). *Knowing what students know and can do: The current state of student learning outcomes assessment in U.S. colleges and universities.* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Lai, M. K., & Schildkamp, K. (2013). Data-based decision making: An overview. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 9–21). Dordrecht, The Netherlands: Springer.

Liu, O. L. (2017). Ten Years After the Spellings Commission: From Accountability to Internal Improvement. *Educational Measurement: Issues and Practice.* Advance online publication. DOI: 10.1111/emip.12139

Mandinach, E., Honey, M., Light, D., & Brunner, C. (2008). A conceptual framework for data-driven decision-making. In E. B. Mandinach & M. Honey (Eds.), *Data-driven school improvement: Linking data and learning* (pp. 13–31). New York, NY: Teachers College Press.

Marsh, J. A. (2012). Interventions promoting educators' use of data: Research insights and gaps. *Teachers College Record*, 114(11), 1–48.

Melguizo, T., & Coates, H. (2017). The Value of Assessing Higher Education Student Learning Outcomes. *AERA Open.* Advance online publication. DOI: 10.1177/2332858417715417.

Musekamp, F., & Pearce, J. (2015). Assessing engineering competencies: the conditions for educational improvement. *Studies in Higher Education,* 40(3), 505-524.

Moss, P. A. (2013). Validity in action: Lessons from studies of data use. *Journal of Educational Measurement,* 50(1), 91-98.

Moss, P. A. (2016). Shifting the focus of validity for test use. *Assessment in Education: Principles, Policy & Practice,* 23(2), 236-251.

Rayner, G. & Pakakonstantinou, T. (2015). Student perceptions of their workplace preparedness: Making work-integrated learning more effective. *Asia-Pacific Journal of Cooperative Education,* 16(1), 13-24.

Schildkamp, K., & Kuiper, W. (2010). Data informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education,* 26, 482–496.

Schildkamp, K., & Lai, M. K. (2013). Conclusions and a data use framework. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 177-191). Dordrecht, The Netherlands: Springer.

Schildkamp, K., & Poortman, C. (2015). Factors influencing the functioning of data teams. *Teachers college record,* 117(4), 1-42.

Schildkamp, K., Poortman, C., Luyten, H., & Ebbeler, J. (2017). Factors promoting and hindering data-based decision making in schools. *School effectiveness and school improvement,* 28(2), 242-258.

Supovitz, J. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly*, 46(1), 31–56.

Secolsky, D., & Denison, B. (Eds.). (2012). *Handbook on measurement, assessment, and evaluation in higher education*. New York, NY: Routledge.

Thompson, G., Adie, L., & Klenowski, V. (2017). Validity and participation: implications for school comparison of Australia's National Assessment Program. *Journal of Education Policy*. Advance online publication. DOI: 10.1080/02680939.2017.1373407

Dr. Joseph Chow works in the Educational Development Centre of the Hong Kong Polytechnic University. His main research interests are in educational assessment, teacher assessment capacity, and evidence-based decision making with a goal to support students, teachers, and researchers with actionable information they need, value and can trust. E-mail: kfchow@polyu.edu.hk