Assessing the Reliability of Merging Chickering & Gamson’s Seven Principles for Good Practice with Merrill’s Different Levels of Instructional Strategy (DLISt7)

Dr. Syaril Izwann Jabar  
*University Selangor, Malaysia*

Prof. Peter R. Albion  
*University of Southern Queensland, Australia*

### Abstract

Based on Chickering and Gamson’s (1987) Seven Principles for Good Practice, this research project attempted to revitalize the principles by merging them with Merrill’s (2006) Different Levels of Instructional Strategy. The aim was to develop, validate, and standardize a measurement instrument (DLISt7) using a pretest-posttest Internet quasi-experiment. It was proposed that the instrument could then be used as a rubric either for facilitating the implementation of DLISt7, or as a set of unobtrusive diagnostic indicators for assessing the quality of learning experienced by students in blended and online courses. The study was conducted across five faculties at a regional Australian multi-campus university. The intent was to contribute to knowledge building by leveraging the data that had been collected, analyzed, and reported to generate awareness about the likelihood of scaffolding and scaling, varying levels of instructional strategies for communicating expectations, and relaying information. The idea was to produce a tool that would create more opportunities for more of the principles to be put to good use as an effectiveness multiplier. The findings from the analysis conducted using exploratory and confirmatory factor analysis verified the validity of DLISt7 and demonstrated excellent internal reliability values.

### Introduction

In the same way discoveries and inventions fuelled the Industrial Revolution mechanizing human society, technology has often been credited as the catalyst that modernized the way learning can be organized and teaching resources utilized (Lever-Duffy, McDonald & Mizell, 2003; Shneiderman, 1998). However, the debate continues about how successful academics have been at leveraging what computers, mobile devices and the Internet have to offer in terms of modernizing the delivery of education to learners (Lever-Duffy & McDonald, 2008). Despite the enthusiastic adoption of blended and online learning in
higher education, research into web-based methodologies have revealed contradictory results that have not yielded clear pedagogical ground rules (Garrison, Cleveland-Innes & Tak Shing Fung, 2010).

Thus, there is little value in knowing what contemporary technology has to offer when educators are hesitant about when, where and how to best use technology to support the process of learning and teaching (Syaril Izwann Jabar, 2012a). Chickering and Ehrmann (1996) noted that simply having access does not guarantee being able to effectively leverage and innovatively utilize what contemporary technology has to offer, henceforth referred to as instructional technology (IT) so as to include the use of instructional media for the process of instruction (Morrison, Ross & Kemp, 2001; Reiser, 2012).

The challenge of utilizing IT to support a wide range of pedagogical approaches borrowed from the conventional classroom is already a difficult undertaking (Albion & Redmond, 2006), and becomes even more challenging when having to make it all work in an effective, efficient and engaging manner (Merrill, 2006). Even though technology often takes centre stage, it is actually the connections made with humans that drive online learning systems. Thus, the fundamental issue that arises for online education is “how to marry the power of networked connectivity with established pedagogical principles to produce better learning outcomes” (Kehrwald, Reushle, Redmond, Cleary, Albion & Maroulis, 2005, p. 1).

Although the technologies used to facilitate learning have evolved from classroom blackboards to www.blackboard.com, the characteristics of teaching and learning have not really changed; good teaching is still good teaching (Albion & Redmond, 2006) but “what has changed is how education providers and teachers facilitate learning” (Kehrwald et al., 2005, p. 1). For example, one widely recognised formulation of such principles is the Seven Principles for Good Practice in Undergraduate Education (Seven Principles) that was first proposed for face-to-face education (Chickering & Gamson, 1987) and subsequently adapted for a variety of contexts and purposes (Chickering & Gamson, 1999). The principles affirm that good practice;

1. Encourages contact between students and teaching staff
2. Develops reciprocity and cooperation among students
3. Encourages active learning
4. Gives prompt feedback
5. Emphasizes time on task
6. Communicates high expectations
7. Respects diverse talents and ways of learning

Teachers working in online environments can transfer what they know about teaching from other environments but need to consider how to adapt their knowledge while managing technology to balance flexibility and structure, put the learner at the centre of a supportive process, and create opportunities for engagement and interaction that stimulate learning (Kehrwald et al., 2005).

The complex nature of online instructional design suggests the need for practical ground rules designed to enable the effective integration of instructional technology with good pedagogy regardless of form, delivery system, or instructional architecture (Merrill, 2008; Merrill, 2009). Teaching staff with traditional, minimal or no online teaching experience should not have to rely on the process of trial and error to learn about how to teach effectively while at the same time scrimmage with the idiosyncrasies of the modern online environment (Haughton & Romero, 2009).

Clearly, there is a need for some form of guidance that educators can adapt to their personal style, course content, student population, and available technology (Shneiderman, 1998). One source of such guidance is the Different Levels of Instructional Strategy (DLIS) proposed by Merrill (2006). He describes five levels of application based on his First Principles of Instruction, namely, Demonstration,
Application, Task-centred, Activation, and Integration. Progress along this continuum can be charted using four levels (0-3), Level 0 is Information only, Level 1 is Information plus Demonstration, Level 2 is Information plus Demonstration plus Application, and Level 3 is Task-centred integration of all three preceding levels (Merrill, 2006, 2009).

Of particular concern is the possibility of a gap in the synergy of events between cognitive presences, social presences, teaching presences, and strategies or tactics for online learning and teaching (Kehrwald et al., 2005). For example, a recommendation was made on the back of findings from a number of research projects about the widely accepted Community of Inquiry (CoI) framework that learning presence should be assimilated as a new conceptual element (Shea, Hayes, Smith, Vickers, Bidjerano, Pickett, Gozza-Cohen, Wilde & Jian, 2012).

Within the context of this research a similar recommendation was made with regards to issues associated with innovation complexity, obscurity of results, and lingering doubt about the effectiveness, efficiency, and engagement of online education (Syaril Izwann Jabar, 2007). Consequently, in order to maximise the benefits of online learning, learners need to be encouraged to accommodate strategies such as forethought, planning, mentoring, learning performance and reflection that good students are supposed to utilize while self-regulating learning presence (Shea, Hayes, Smith, Vickers, Bidjerano, Gozza-Cohen, Jian, Pickett, Wilde & Tseng, 2013).

Thus, it is proposed that in order for the science of learning and the art of teaching (Skinner, 1968) to be more effective in blended or online environments, the eclectic selection of appropriate pedagogy should consider the systematic and conscientious use of contextual engagement. This is because, as the world’s population continues to grow “a far greater part of them want an education” for which “the demand cannot be met simply by building more schools and training more teachers” (Skinner, 1968, p. 29).

Instead, “education must become more efficient. To this end curricula must be revised and simplified, and textbooks and classroom techniques improved” upon using the latest available technology (Skinner, 1968, p. 29). Teachers for instance, can help students “in every way short of teaching them” by communicating new approaches about the acquisition of knowledge, not to mention the relaying of information about the “methods and techniques of thinking, taken from logic, statistics, scientific methods, psychology, and mathematics. That’s all the ‘college education’ they need. They [can] get the rest by themselves in our libraries and laboratories” (Skinner, 1962, p. 121).

**Literature Review**

**Integrating Instructional Technology with Good Pedagogy**

As it stands, pedagogy is the action of teaching, or what teachers do when implementing their craft to assist students’ learning (Lever-Duffy, McDonald & Mizell, 2005). In an age when information is expanding and becoming more readily accessible, the most critical outcome of education is learning how to learn. The application of sound educational principles and methods that informs teachers’ practice can perhaps lead to improvements in the sequencing of educational experiences for designing, developing, implementing, and evaluating a pedagogically sound learning environment that is conducive for advancing knowledge building (Lever-Duffy & McDonald, 2008).

An understanding of learning and teaching has developed alongside technology and generations of educators have recognised that “as a mere reinforcing mechanism” the teacher might one day become antiquated, but “if the teacher is to take advantage of recent advances in the study of learning, she must
have the help of mechanical devices” (Skinner, 1968, p. 22). Such an idea was probably thought of at that point in time as science fiction because nobody yet understood how mechanical and electrical devices would evolve into modern one-to-one, one-to-many or many-to-many robust supporting communication technologies that could be skilfully used to progress through a series of successive shaping approximations that shortened the window of opportunity between reinforcement contingencies (Anderson & Dron, 2012; Case & Bereiter, 1984; Skinner, 1954). That was true for the blackboard in its time and continues to be true for the evolution from Skinnerian behaviourism to cognitivism and for successive generations of technology enhanced distance education because of the awareness that “the human organism is, if anything, more sensitive to precise contingencies” (Skinner, 1954, p. 94).

The Digitization of Education

Scientific advancement is not a smooth evolutionary process, but is more like a sequence of peaceful intervals interrupted by intellectually violent upheavals in which an individual’s attributional interpretation of the world is improved upon or replaced by another (Kuhn, 1970). Educational change exhibits similar characteristics and perhaps the time has come to re-evaluate the current paradigm of online pedagogy and set new directions for continued growth, modernization, and eventual maturity.

In light of the fact that “transitioning from teaching in the traditional classroom to the online environment is not a simple task for most faculty,” there exists the need for a set of ground rules that can be used to improve the whole experience and enable the continuation of good teaching practices integrated with instructional technology (Grant & Thornton, 2007, p. 352). Anderson and Dron (2011) articulated the idea best when they said that quality online learning experiences make the most of cognitive-behaviourist, social constructivist and connectivist pedagogies across three generations to encapsulate what distance education has evolved and matured into.

Theoretical Rationale

Learner engagement can be defined as the quality of effort, in terms of time and energy invested in purposeful interactive involvement in educational activities and conditions that are likely to contribute directly to the construction of understanding (Coates, 2006; Kuh & Hu, 2001). When achieved alongside effectiveness and efficiency, engagement can lead to an increase in learner interaction, interest, and subsequently satisfaction (Merrill, 2008; Merrill 2009). According to Bangert (2008a), “a recommended instructional practice for higher education faculty is to engage in valued and meaningful inquiry-based, learning activities” that are “designed to create knowledge structures that can be retrieved when necessary to solve problems encountered in real-world contexts” (p. 36).

An aim of online education has been to actualize the potential afforded by communication and Internet technologies via design options that would enable participants to maintain engagement in a community of learners using asynchronous interaction (Garrison & Cleveland-Innes, 2005). Another goal has been to “structure the educational experience to achieve defined learning outcomes” using interaction that is structured, systematic, critical, and reflective (Garrison & Cleveland-Innes, 2005, p. 134).

Analyzed closely, these are also the achievable learning and teaching goals that are the focus of the Seven Principles (Chickering & Gamson, 1987) and DLIS (Merrill, 2006, 2009). The outcome of an education programme is often simply based on the premise of encouraging the active construction and demonstration of understanding, or as stated by Scardamalia and Bereiter (2006) “All understandings are inventions; inventions are emergents” (p. 15). Hence, as described in this paper the Different Levels of Instructional Strategies (DLIS7) for Online Learning was designed to function intelligibly as, “a set of workable principles that could guide pedagogy in a variety of contexts” (Scardamalia & Bereiter, 2006, p. 24).
Research Design

Statement of the Problem

This study was designed to investigate whether the Seven Principles could be revitalised by merging them with DLIS to form DLISt7. The approach was to develop a measurement instrument that could be used either as a rubric for facilitating the extrinsic implementation of DLISt7, or as a set of unobtrusive diagnostic “process indicators” (Kuh, Pace & Vesper, 1997, p. 436) for intrinsically assessing the quality of learning experienced by students in blended and online courses.

The intent was to contribute by leveraging the data that had been collected, analyzed, and reported to generate awareness about the likelihood of scaffolding and scaling varying levels of instructional strategies to make informed improvements in the instructional design of future online courses. The idea was to produce a tool that would create more opportunities for more of the principles to be put to good use as an effectiveness multiplier in support of efficient and engaging online learning. The critical insight for educational administrators, teaching staff, and instructional designers is the importance of utilizing learning analytics to make informed decisions about the appropriate balance between the eclectic utilization of asynchronous or synchronous communication technology and other available online resources.

In other words, when DLIS is used as a rubric either for teaching or treatment purposes to prompt and stimulate conditional response from students - which explains the $t$ in DLISt7 - favourable online learning experiences that are consistent with the Seven Principles would be realistically achievable in ways that are familiar and unobscure (Syaril Izwann Jabar, 2012c).

Focus of the Research

An instructional principle, as defined in the context of this study, is “a relationship that is always true under appropriate conditions regardless of the methods or models which implement this principle” and whose underlying function is “to promote more effective, efficient, or engaging learning” (Merrill, 2009, p. 43). In their original form, the Seven Principles were designed to be robust so as to always be true under appropriate conditions with each principle having the capacity to “stand alone on its own, but when all are present their effects multiply” (Chickering & Gamson, 1987, p. 2).

Upon being updated, the term “instructional strategy” was integrated so as to accentuate the utility of the Seven Principles in promoting effective, efficient, and engaging learning in conjunction with “new communication and information technologies that had become major resources for teaching and learning in higher education” (Chickering & Ehrmann, 1996, p. 1). Despite their apparent simplicity and practicality, there has been a tendency for the Seven Principles to not be fully utilized (Bangert, 2004; Bangert 2008b; Batts, 2008; Chickering & Gamson, 1999; Cobbett, 2007, & Wuensch, Shahnaz, Ozan, Kishore & Tabrizi, 2009).

A review of the above literature suggests a penchant for the Seven Principles to be implemented and subsequently assessed in their standalone form instead of as a whole. Perhaps the Seven Principles could be resuscitated by being analysed from a different perspective. To echo the words of Merrill, “we need to back up and find out if there's a set of principles we can agree to and then build on these principles. Let's build on what's there instead of starting over and reinventing the wheel every single time” (as cited in Spector, Ohradza, Van Schaack & Wiley, 2005, p. 318). Thus, this study was devised to approach the Seven Principles from a different perspective, that is, by linking them with DLIS (Chickering & Gamson, 1987; Merrill, 2006).
Research Objective

The primary purpose of this research project was to obtain data that would facilitate the development, validation, and standardization of a measure for DLISit7. As a rule, a measure is said to be standardized when; (a) its rules of measurement are clear, (b) it is practical to apply, (c) it is not demanding of the administrator or respondent, and (d) its results do not depend upon the administrator (Netemeyer, Bearden & Sharma, 2003; Nunnally & Bernstein, 1994). Consequently, a measure that successfully fulfils these criteria would yield “similar results across applications (i.e., the measure is reliable), and offer scores that can be easily interpreted as low, medium [or] high” (Netemeyer et al., 2003, p. 2).

This research project also attempted to ascertain the validity of DLISit7 as a conceptual framework and the reliability of the scale. This was achieved by systematically determining the relationships for the following research questions using accumulated and integrated evidence (Cronbach, 1990). Firstly, how many principles from DLISit7 would actually load significantly? Secondly, would the factor loadings indicate construct validity? Lastly, would an assessment of the summated gain scores reveal the perceived effectiveness of DLISit7 (Dunn-Rankin, Knezek, Wallace & Zhang, 2004; Tuckman, 1999)?

Research Methodology

The study was designed as a non-equivalent pretest-posttest control group Internet quasi-experiment that would “provide substantially better control of the threats to validity than do pre-experimental designs” (Tuckman, 1999, p. 167). Such a design may be used “where better designs are not feasible” (Campbell & Stanley, 1963, p. 204) because “conditions complicate or prevent complete experimental control” (Tuckman, 1999, p. 168).

DLISit7 as a “treatment is included by selection rather than manipulation” (Tuckman, 1999, p. 181) and because of its inherent qualities can also be used as an unobtrusive measure that does not “require acceptance or awareness by the experimental subjects” (Tuckman & Harper, 2012, p. 126). Further replication of the research project and the utilization of the research instrument by others in the future to cross-validate DLISit7 would be valuable in unlocking its Rubik’s cube-like potential (Syaril Izwann Jabar, 2012b).

In light of this research being conducted over the Internet, it also qualifies as a field experiment because the research was in a real-life setting (Christensen, 1997, p. 93). The significance of in-the-field Internet experimentation cannot be overlooked because it is useful in terms of determining if a manipulation would work in the real-world (Johnson & Christensen, 2012, p. 285). Moreover, there are also the value-added advantages of speed, low cost, external validity, experimenting around the clock, a high degree of automation of the experiment (i.e., low maintenance, limited experimenter effects), and a wider sample (Reips, 2002, p. 244). However, higher than usual dropout rates [also known as differential attrition (Johnson & Christensen, 2008), differential loss or experimental mortality (Campbell & Stanley, 1963)] are a disadvantage of voluntary participation in Web experiments (Reips, 2000).

Research Sampling

Sample members were drawn using a three-stage purposive cluster sampling technique (Ary, Jacobs & Sorensen, 2010; Cochran, 1977; Johnson & Christensen, 2008). The first sampling element used was of nationality. This was followed by the second element of how far the participants had progressed in their degrees. The third sampling element was of academic affiliation across five faculties at
a regional Australian multi-campus university. Participants were recruited based on enrolment in intact courses subject to approval from Faculty. Full ethics clearance was granted by the university’s fast track Human Research Ethics Committee (H10REA016). The whole process of research sampling took sixteen months to complete before any data could be collected.

Research Instrument

Reliability Analysis

In an effort to answer the not-so-simple question to which “there are legitimate disagreements about the correct answer” regarding the issue of how are the measurement of constructs developed and validated (Nunnally & Bernstein, 1994, p. 86) the “psychometrical properties of the questionnaire, such as construct validity and reliability” are discussed in the following section (Vandewaetere & Desmet, 2009, p. 349). In extending on a previous investigation which replicated Guidera’s (2003) doctoral research project at the masters’ level (2004-2007) a variant of Ehrmann, Gamson, and Barsi’s (1989) Faculty Inventory was translated, rephrased, and adapted for use as a Student Inventory.

The objectivity and content validity of the adapted version of the research instrument was informally evaluated by a panel of experts consisting of one Associate Dean of Academic Affairs and two subject matter experts (SME), one from the Faculty of Education and Human Development, and the other from the Faculty of Languages and Communication at a teacher training university in Malaysia. The instrument included multiple items for each of the Seven Principles. Each item was presented as a statement eliciting students’ perception about the degree of success to which instructional strategies were being effectively or ineffectively used to conduct online learning using a five (5)-point Likert scale.

For the pilot study an excellent value for Cronbach’s alpha ($\alpha = 0.97, n = 74$) was obtained with individual items having alphas ranging from a lower limit (LL) of 0.972 to an upper limit (UL) of 0.974. For the main study a slightly lower but still excellent value for alpha ($\alpha = 0.94, N = 397$) was obtained with individual items having alphas ranging from 0.938 (LL) to 0.941 (UL) (George & Mallery, 2011; Syaril Izwann Jabar, 2007). No problematical items were identified requiring exclusion from the instrument (Coakes & Ong, 2011).

This was followed by an exploratory factor analysis (EFA) to determine the construct validity of the intangible constructs that constitute the Seven Principles. A principal component analysis (PCA) was conducted on the 34 items to verify construct validity. The rotated component matrix revealed that of the 34 items used, 23 were pure variables, while 11 were complex variables (Coakes, Steed & Price, 2008). However, these complex variables did not have loadings that made their structure ambiguous and interpretation difficult (Syaril Izwann Jabar, 2007).

More recent revisions to the measurement instrument being developed at the postgraduate level (2009-2013) involved attaching DLIS to the Seven Principles framework to form DLIS7. A set of four items addressing elements of DLIS were added to the beginning of the instrument addressing each of the four levels (0-3) from presentation of Information only (Level 0) to Task-centred integration (Level 3). The sets of items for each principle were then overlayed with attributes associated with students’ perception. The Likert scales were also switched to a Sentence Completion Rating scale “with descriptive statements on either end” (Tuckman & Harper, 2012, p. 229).

This adjustment was made to sidestep “the multidimensionality innate in Likert-type scales” and eliminate “the extra cognitive load associated with the use of item reversals” (Hodge, 2007, p. 289). Furthermore, the use of such a scale would be an improvement in terms of fulfilling parametric
assumptions and coping with issues such as “coarse response categories” and “equating the neutral option with a not applicable response” (Hodge & Gillespie, 2003, p. 53).

Accordingly, the perception of the participants towards DLISt7 was successfully measured using scores that can be easily interpreted as low, medium or high (Netemeyer et al., 2003). Cronbach’s alpha for the pilot study revealed an excellent coefficient ($\alpha = 0.92$, $n = 39$) with results indicating that removal of individual items would result in alphas ranging from 0.913 (LL) to 0.918 (UL). In the main study, a slightly higher alpha coefficient ($\alpha = 0.95$, $n = 283$) was obtained using a larger sample with removal of individual items resulting in alphas ranging from 0.950 (LL) to 0.952 (UL). Again, there were no problematical items identified (Coakes & Ong, 2011) and by assessing alpha coefficients from both the pilot and main study it was determined that the temporal stability of the measure was excellent (George & Mallery, 2011).

Thus, the internal consistency of the measurement instrument has held up well because when estimating the “correlation (reliability coefficient) to be expected if two independent, more or less equivalent forms of a test are applied on the same occasion,” it is expected that “the stronger the intercorrelations among a test’s items, the greater its homogeneity” (Cronbach, 1990, p. 704). Although validation can be obtained from a single study, “the ideal is a process that accumulates and integrates evidence on appropriateness of content, correlations with external variables, and hypotheses about constructs” (Cronbach, 1990, p. 707).

**Results**

**Students’ Awareness of DLISt7**

The questionnaire (Appendix) was administered online using LimeSurvey® (limesurvey.com). Faculty members responsible for selected courses at a regional Australian university cooperated by including the universal resource locator (URL) for the survey in messages to students. A total of 319 completed responses were collected for the pre-test administered at the start of the semester.

Regardless of gender, nationality, academic progress, or faculty affiliation, 194 (60.8%) indicated “yes” they were aware of DLISt7 and 125 (40.0%) indicated “no,” they were not aware. In view of DLISt7 being an unpublished conceptual framework, it is doubtful that undergraduate students from this university could have had a priori knowledge about it. Any claim made contrary to the fact could have occurred purely by chance, but is more likely to have been a combination of circumstances that cannot be isolated without further study. The following reasons are suggested to explain why there was a sixty/forty split in responses.

Firstly, it is possible that it was a case of the Hawthorne effect, which “refers to performance increments prompted by mere inclusion in an experiment” (Tuckman & Harper, 2012, p. 132). This could have come to pass because of a fumble when allocating participants to the No Treatment-Treatment conditions that might have tipped-off or aroused the suspicion of a few participants (McMillan & Schumacher, 2009).

Secondly, it is also possible that performance on the posttest was affected by experience from the pre-test (Tuckman & Harper, 2012). Problems related to testing occur because “experience of taking such a pre-test may increase the likelihood that the subjects will improve their performance on the subsequent posttest, particularly when it is identical to the pre-test” (Tuckman & Harper, 2012, p. 126).
Thirdly, is the predisposition to “provide the answer they want others to hear about themselves rather than the truth... that shows oneself in the best possible light,” known as the social desirability response bias (Tuckman & Harper, 2012, p. 265).

Lastly, there is also the “tendency to mark a single choice for all questions out of boredom, disinterest, or hostility” (Tuckman & Harper, 2012, p. 264) known as “acquiescent” response bias (Cronbach, 1990, p. 470).

**The Utilization of Communication Technology and Online Resources by Teaching Staff**

The most frequently utilized communication technology or online resource used for conveying instructional strategies was StudyDesk with 284 (89.0%) “yes” responses followed by email with 256 (80.3%) “yes” responses. This was followed in descending order by Wimba Online Classrooms (f = 82, 25.7%), Moodle Forums (f = 81, 25.4%), blogs (f = 71, 22.3%), telephone: voice (f = 32, 10.0%), Moodle Chat (f = 29, 9.1%), and instant messaging (f = 28, 8.8%). Hence, it would probably be reasonable to assume that teaching staff at this university had the tendency to rely heavily on two of the more important communication technologies (StudyDesk and email) made available to them while preferring to be parsimonious when choosing what other online resources to incorporate into their instructional repertoire. Possible reasons include lack of familiarity with less common resources or the desire to not overwhelm the students.

**The Interaction between Awareness of DLISt7, No Treatment-Treatment Group, and Gender**

Initial findings revealed that students’ Awareness of DLISt7 at the pretest stage was independent of or not related to being in the No Treatment-Treatment group. However, students’ Awareness of DLISt7 at the posttest stage was related to being in the No Treatment-Treatment group. Thus, the need arose to further investigate whether it would be probable to assume that the intervening variable was actually transmitting or mediating the effect of the treatment variable onto the dependent variable (Creswell, 2012), or was it a case of uncontrolled extraneous variables confounding the results, and “casting doubt about the validity of inferences made” (Pedhazur & Schmelkin, 1991, p. 212).

A higher order between-subjects three-way ANOVA revealed that there was no statistically significant interaction between the posttest scores for Awareness, No Treatment-Treatment group, and Gender. However, there was a statistically significant main effect for the No Treatment-Treatment group. It was also established that students’ Awareness of DLISt7 at the posttest stage was related to being in the No Treatment-Treatment group and that pre- and posttest scores were related.

Once again the researcher had to attempt to comprehend the source and nature of why the mean scores for the posttest were not significantly greater than the mean scores for the pretest (Cohen, Cohen, West & Aiken, 2003). From where did the uncertainty overshadowing “the accuracy of the inferences, interpretations, or actions made on the basis of test scores” provided by the participants from the No Treatment group originate (Johnson & Christensen, 2012, p. 597)? The following explanation was proposed to clarify how the confounding variable of group mean scores and the extraneous factor of group sample size for the No Treatment group have come together to limit the reliability and validity of the inferences derived from the findings.

In view of DLISt7 being an unpublished conceptual framework, it is doubtful that undergraduate students from the No Treatment group, could have had a priori knowledge about DLISt7. Although the probability does exist that such responses could have occurred purely by chance, logic favours the assumption that the responses were confounded, either by the Hawthorn or testing effect, together with the social desirability and acquiescent response bias. As stated by McMillan and Schumacher (2009),
“scores cannot be valid unless they are reliable…. Reliability is needed for validity; scores can be reliable but not valid” (p. 185). Hence, the only way to know for sure is to conduct further research using the Solomon four-group design (McMillan & Schumacher, 2009).

Justification for this alternative explanation was realized while conducting a detailed analysis of the mean scores for the No Treatment-Treatment group. For the pretest, the mean score for the No Treatment group was 79.85 (N = 63) but, for the posttest the mean score was 84.23 (N = 34). It was also determined that the pretest mean score for the Treatment group was 76.37 (N = 220) and the posttest mean score was 76.26 (N = 82), an insignificant difference that did not register on any of the statistical tests but could be further investigated.

The primary point of contention that warranted careful consideration was whether the posttest mean scores for the No Treatment-Treatment group (M=84.23) were representative of the population mean since they were from a sample of 34 participants made up of 29 Females (M = 85.25) and 5 Males (M = 78.26). Consequently, the posttest mean scores from the No Treatment group would appear inflated when compared to the posttest mean scores from the Treatment group (M = 76.26) which was from a larger sample of 82 participants made up of 53 Females (M = 79.18) and 29 Males (M = 70.93). As a result, the mean scores that came from the latter Treatment group and not the former No Treatment group would appear to best represent the population mean without giving the impression of being overstated.

The next point of contention is the fact that a sample of approximately 40 would have been better for invoking the central limit theorem (Field, 2009). With a sample of less than 30, the resulting sampling distribution would have a different shape compared to the parent population causing doubt about whether “the sampling distribution has a normal distribution with a mean equal to the population mean” (Field, 2009, p. 42).

According to Glass and Hopkins (1996), “the validity of the central limit theorem allows [for] statistical inferences to [be made across] a much broader range of applications than would otherwise be possible” (p. 235). This theorem applies “even when the parent population is not normal, [because] the formula \( \sigma_s = \sigma / \sqrt{n} \) accurately depicts the degree of variability in the sampling distribution” (Glass & Hopkins, 1996, p. 239).

For example, when “sample sizes are small (1, 2, 5 and 10); some degree of non-normality in the parent population continues to be evident in the sampling distributions, but progressively less so as n increases” (Glass & Hopkins, 1996, p. 239). When \( n \) was increased to 25, the theoretical standard error of the mean agrees almost perfectly with the standard deviation from the sample means despite a skewed parent population from which the sample was drawn (Glass & Hopkins, 1996).

The Validity of DLISt7 and the Reliability of Its Items

Findings from the exploratory and confirmatory factor analysis verified the validity of the intangible constructs that constitute the conceptual framework of DLISt7. Even though only seven of the possible eight factors submitted successfully loaded, there is a simple and logical explanation. According to Brown (2006), the number of observed measures (\( p \)) that are submitted for analysis limits the number of factors that can be extracted (\( m \)). Unequivocally, \( p – 1 \) is the maximum number of factors that can be extracted or “the number of parameters that are estimated in the factor solution (\( a \)) must be equal to or less than the number of elements (\( b \)) in the input correlation or covariance matrix (i.e., \( a \leq b \))” (Brown, 2006, p. 23).

The factor correlation matrix revealed an uncorrelated model with the oblique rotation producing a solution that was “virtually the same as one produced by orthogonal rotation” (Brown, 2006, p. 32).
fact, the interpretation of the oblique solution, although more complicated than the orthogonal solution, did provide results that were better (Hatcher, 2007). Together with the fact that the test-retest (temporal) coefficient for Cronbach’s alpha reliability analysis was excellent each time the research instrument was administered, it would probably be safe to assume that the items are actually measuring “the underlying construct comparably across groups” (Brown, 2006, p. 4).

Discussion

Validation as a process is unending and requires measures to be “constantly evaluated and re-evaluated to see if they are behaving as they should” (Nunnally & Bernstein, 1994, p. 84). As assured by Cronbach (1990), “the more reliable a measuring procedure is, the greater the agreement between scores obtained when the procedure is applied twice” (p. 705). Thus, there would be value in conducting further research using the revised version of the measurement instrument. Not only would replication make available a fresh reliability index based on test-retest reliability, but factor analysis could then be used to refine the precise measurement of constructs and contingencies, or in other words, the convergent or discriminate validity of DLISt7 (Netemeyer et al., 2003; Rust & Golombok, 1989; Skinner, 1954).

Hence, the real world utility of factor analysis is the ability to “summarize the interrelationships among the variables in a concise but accurate manner as an aid in conceptualization” (Gorsuch, 1983, p. 2). A conceptual framework is only as good as it can “reduce the amount of trial-and-error effort, and people who explore theories stand at the vanguard of each field of science” (Nunnally & Bernstein, 1994, p. 317). Only through such efforts can the hierarchical levels of a construct, also known as depth psychometry, be studied (Cattell & Schuerger, 1978, p. 223).

Conclusion

Based on Chickering and Gamson’s Seven Principles, this study attempted to revitalize the principles by merging them with Merrill’s DLIS. The goal was to develop, validate, and standardize a measure for assessing the effectiveness of DLISt7. As a measurement instrument, DLISt7 has been successfully standardized because; (a) its rules of measurement are clear, (b) it is practical to apply, (c) it is not demanding of the administrator or respondent, and (d) its results do not depend upon the administrator (Netemeyer et al., 2003; Nunnally & Bernstein, 1994). Consequently, DLISt7 successfully fulfils all the relevant criteria and has yielded similar scores across applications that can be easily interpreted as low, medium or high, indicating that as a measurement model it is reliable (Netemeyer et al., 2003).

Inductively, the research questions were also successfully answered using a systematic approach. Firstly, of the eight principles specified, seven loaded successfully. Secondly, from the factor loadings it was ascertained that the items utilized are measuring the appropriate constructs seeing as accumulated and integrated evidence indicate that such a conclusion would be appropriate (Cronbach, 1990). However, an assessment of the summed gain scores about the perceived effectiveness of DLISt7 was inconclusive.

Furthermore, DLISt7 also meets the terms of Nunnally and Bernstein’s (1994) three major aspects of construct validation, namely: (1) the domain of observables related to the construct have been specified, (2) the extent to which the observables tend to measure the same things has been determined, and (3) individual differences studies or experiments that attempt to determine the extent to which the supposed measures of construct are consistent with best guesses have also been performed. The resultant standardized measure can now be used as a rubric to assist the design of instruction and as a measure to evaluate the effectiveness of such instruction.
Thus, it is proposed that DLISt7 be utilized to enable the learning experienced by students to be systematically scalable to different levels of complexity. Teaching staff would conceivably have the flexibility of being eclectic in their choice of pedagogy for providing students with directed facilitation (Shea, Li & Pickett, 2006) to work their way through the pathways of knowledge to find their own answers. Successively less facilitated guidance, also known as guided instruction (Kirschner, Sweller & Clark, 2006) should be faded with each scaffolded task until students are completing complex tasks on their own.

Will any of this make a difference in bridging and connecting with what is there and not having to reinvent the wheel? Anderson and Dron (2011) summed up the idea well when they said that in order to identify the best mix of pedagogy and instructional technology, the learning and teaching experience has to be seen as a progression. Over the past three decades, many technologies have come and gone, and so has the popularity of different approaches to pedagogy. But each has built upon the shortcomings of the instructional technology left behind by its predecessor instead of replacing the first of its kind.

To recall and demonstrate an understanding of how to apply and integrate what use to be referred to as the Socratic method, a teacher has always, and will be expected to continue the tradition of setting good examples by being “clear about his objectives; he knows why he is doing what he is doing; and he chooses a technique to suit his objective” (Hyman, 1974, p. 101).

References


Syaril Izwann Jabar. (2012b). *Assessing the effectiveness of the different levels of instructional strategies [DLIS] for online learning at the University of Southern Queensland (USQ), Australia*. Paper presented at the International Conference on Quality of Teaching and Learning (ICQTL 2012) organized by the Centre of Quality and Academic Development, University Malaysia Terengganu.


ASSESSING THE EFFECTIVENESS OF THE DIFFERENT LEVELS OF INSTRUCTIONAL STRATEGIES (DLSI(7)) FOR ONLINE LEARNING BY UNDERGRADUATE STUDENTS AT THE UNIVERSITY OF SOUTHERN QUEENSLAND (USQ), AUSTRALIA

This survey is designed to measure and assess various attributes associated with students’ perception towards the effectiveness of the different levels of instructional strategies that can be used to conduct online learning.

All information provided will remain confidential and your identity will not be disclosed. By completing this survey you are consenting to being a participant in this research.

Please check the appropriate boxes or fill in the blanks to indicate your response.

a) Are you aware that there are Different Levels of Instructional Strategies that can be used to conduct online learning?
   □ Yes
   □ No

b) What is your Gender?
   □ Female
   □ Male

c) What is your Nationality?
   ____________________________

d) How far have you progressed in your degree at USQ?
   Degree : ______________________________
   Year (1st, 2nd, etc.,) : _________________
   Semester : __________________________

e) Which Faculty are you from?
   □ Arts
   □ Business
   □ Sciences
   □ Education
   □ Engineering & Surveying
f) Please check the boxes that indicate the communication technology or online resource utilized by teaching staff to convey instructional strategies for online learning. Check any that apply.

- Blogs
- Email
- StudyDesk
- Moodle Chat
- Moodle Forum
- Teleconferencing
- Videoconferencing
- Instant Messaging
- Wimba Online Classroom
- Telephone: Text Messaging
- Telephone: Voice
- Skype Video
- Skype Voice
- Skype Text
- None are utilized
- Other ____________________

The following statements use a sentence completion format to measure various attributes associated with students’ perception towards the effectiveness of the different levels of instructional strategies for online learning.

A partially completed sentence is provided, followed by a scale ranging from 1 to 10. The 1 to 10 range provides you with a continuum on which to reply, with 1 corresponding to a minimum amount of the attribute, while 10 corresponds to the maximum amount of the attribute. A 5 corresponds to an average amount of the attribute.

Please select a number along the continuum that best reflects your initial feeling.

1. DIFFERENT LEVELS OF INSTRUCTIONAL STRATEGY

1.1 I ________ noticed instances of Teaching staff trying to present information with accompanying recall questions.

<table>
<thead>
<tr>
<th>Rarely</th>
<th>Frequently</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

Factor Loading: 0.570 Very Good

1.2 I recall attempts by Teaching staff to promote higher levels of performance on complex tasks by way of presenting information, and demonstrating its application as being

<table>
<thead>
<tr>
<th>Meaningless</th>
<th>Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

Factor Rating: 0.669 Excellent

1.3 I can ________ understand why Teaching Staff would be willing to provide corrective feedback in order to promote improvement in my performance on complex tasks.

<table>
<thead>
<tr>
<th>Completely</th>
<th>Vaguely</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 9 8 7 6 5 4 3 2 1</td>
<td></td>
</tr>
</tbody>
</table>

Factor Rating: 0.481 Good

1.4 I_________ value attempts by Teaching staff to use a task-

Factor Rating: 0.577 Very good
centered approach to promote efficiency, effectiveness and engagement.

**Scarcely** | **Very much**
---|---
1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10

### 2. ENCOURAGING INTERACTION BETWEEN STUDENTS AND TEACHING STAFF

**FACTOR LOADING**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.812</td>
<td>Overlap</td>
</tr>
<tr>
<td>0.658</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.666</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.528</td>
<td>Good</td>
</tr>
<tr>
<td>0.648</td>
<td>Very good</td>
</tr>
</tbody>
</table>

#### 2.1 I __________ noticed instances of Teaching staff trying to communicate with me.

**Rarely** | **Frequently**
---|---
1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10

#### 2.2 I recall attempts by Teaching Staff to facilitate informal interaction with me as being __________.

**Meaningless** | **Significant**
---|---
1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10

#### 2.3 I can __________ understand why Teaching staff would demonstrate a willingness to serve as a mentor to me.

**Completely** | **Vaguely**
---|---
10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1

#### 2.4 I __________ value attempts by Teaching staff to contact me when I have fallen behind to discuss my study habits, schedules, and other commitments.

**Scarcely** | **Very much**
---|---
1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10

#### 2.5 I am __________ of teaching staff attempting to provide extra material or exercises if I lack the essential background knowledge or skills.

**Appreciative** | **Unappreciative**
---|---
10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1

### 3. DEVELOPING RECIPROCITY AND COOPERATION AMONG STUDENTS

**FACTOR LOADING**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.567</td>
<td>Very good</td>
</tr>
</tbody>
</table>

#### 3.1 I __________ noticed instances of Teaching staff trying to encourage me to participate in online activities.

**Rarely** | **Frequently**
---|---
1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10
3.2 I recall attempts by Teaching staff to get me to explain difficult ideas or concepts to others within an online learning group as being __________.

Meaningless Significant

1 2 3 4 5 6 7 8 9 10

0.742 Excellent

3.3 I can __________ understand why Teaching staff would demonstrate an eagerness to get me to discuss openly with colleagues through a forum about interests and backgrounds.

Completely Vaguely

10 9 8 7 6 5 4 3 2 1

0.591 Very good

3.4 I __________ value attempts by Teaching staff to utilize a Learning Management System such as USQStudyDesk to encourage learning communities in my course.

Scarcely Very much

1 2 3 4 5 6 7 8 9 10

0.667 Excellent

3.5 I am __________ of Teaching staff attempting to get me and my colleagues to work on projects together.

Appreciative Unappreciative

10 9 8 7 6 5 4 3 2 1

0.480 Good

4. ENCOURAGING ACTIVE, CONTEXTUAL AND MEANINGFUL LEARNING

FACTOR LOADING FACTOR RATING

4.1 I __________ noticed instances of Teaching staff trying to get me to apply meaningful learning by relating events that happened in real life to what was being learnt.

Rarely Frequently

1 2 3 4 5 6 7 8 9 10

0.640 Very good

4.2 I recall attempts by Teaching staff to get me to apply contextual learning by analyzing real-life contexts as being __________.

Meaningless Significant

1 2 3 4 5 6 7 8 9 10

0.598 Very good

4.3 I can __________ understand why Teaching staff would demonstrate a willingness to link me with professionals who are experts in the field of study so that opinions and ideas can be exchanged.

Completely Vaguely

10 9 8 7 6 5 4 3 2 1

0.696 Excellent

4.4 I __________ value attempts by Teaching staff to encourage me to express myself when I do not understand a particular subject matter.

Scarcely Very much

1 2 3 4 5 6 7 8 9 10

0.705 Excellent
4.5 I am _________ of attempts to include independent study assignments where I seek out information from the Internet and later discuss with Teaching staff the validity of the information and the reliability of its source.

Appreciative       Unappreciative
10  9  8  7  6  5  4  3  2  1

0.578 Very good

5. GIVING PROMPT FEEDBACK

5.1 I _________ noticed instances of Teaching staff trying to adjust their instructional strategy to include problem solving and task-centered activities that provided me with immediate feedback.

Rarely       Frequently
1  2  3  4  5  6  7  8  9  10

0.799 Excellent

5.2 I recall attempts by Teaching staff to provide corrective feedback regarding my performance on problem solving and task-centered activities as being _________.

Meaningless       Significant
1  2  3  4  5  6  7  8  9  10

0.740 Excellent

5.3 I can _________ understand why Teaching staff would demonstrate a willingness to politely inquire about my strengths and weaknesses in tutorials, quizzes and tests.

Completely       Vaguely
10  9  8  7  6  5  4  3  2  1

0.734 Excellent

5.4 I _________ value attempts by Teaching staff to get me to go online and contact them to discuss my academic progress.

Scarcely       Very much
1  2  3  4  5  6  7  8  9  10

0.614 Very good

5.5 I am _________ of attempts by Teaching staff to provide me with an evaluation of my proficiency.

Appreciative       Unappreciative
10  9  8  7  6  5  4  3  2  1

0.732 Excellent

6. EMPHASIZING TIME ON TASK

6.1 I _________ noticed instances of Teaching staff trying to communicate to me that I am expected to complete my assignments promptly.

Rarely       Frequently
1  2  3  4  5  6  7  8  9  10

0.589 Very good
6.2 I recall attempts by Teaching staff to deliver course materials, quizzes and assignments online as being __________.

Meaningless  Significant
1 2 3 4 5 6 7 8 9 10

6.3 I can __________ understand why Teaching staff would demonstrate an eagerness to emphasize to me the importance of diligence, sound self-pacing and scheduling.

Completely  Vaguely
10 9 8 7 6 5 4 3 2 1

6.4 I __________ value attempts by Teaching staff to make it clear to me the amount of time that is required to understand complex material.

Sscarcey  Very much
1 2 3 4 5 6 7 8 9 10

7. COMMUNICATING HIGH EXPECTATIONS

7.1 I __________ noticed instances of Teaching staff trying to communicate to me that I am expected to work hard.

Rarely  Frequently
1 2 3 4 5 6 7 8 9 10

7.2 I recall attempts by Teaching staff to emphasize the importance of holding on to high standards for academic achievement as being __________.

Meaningless  Significant
1 2 3 4 5 6 7 8 9 10

7.3 I can __________ understand why Teaching staff would demonstrate a willingness to share with me past experiences, attitudes and values.

Completely  Vaguely
10 9 8 7 6 5 4 3 2 1

7.4 I __________ value attempts by Teaching staff to provide me with a pre-test at the beginning of the course.

Unclear  Very much
1 2 3 4 5 6 7 8 9 10

7.5 I am __________ of attempts by Teaching staff to discuss my academic progress especially near the end of the course.

Appreciative  Unappreciative
10 9 8 7 6 5 4 3 2 1
8. RESPECTING DIVERSE TAELNTS AND WAYS OF LEARNING

<table>
<thead>
<tr>
<th>Factor</th>
<th>Factor Loading</th>
<th>Factor Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 I __________ noticed instances of Teaching staff trying to determine my learning style, interests or background at the beginning of the course.</td>
<td>0.878</td>
<td>Overlap</td>
</tr>
<tr>
<td>Rarely</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Frequently</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>8.2 I recall attempts by Teaching staff to relate learning activities to my learning style, interests or background as being</td>
<td>0.796</td>
<td>Excellent</td>
</tr>
<tr>
<td>Meaningless</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Significant</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>8.3 I can __________ understand why Teaching staff would demonstrate a willingness to use multiple methods to communicate their own expectations at the beginning of the course.</td>
<td>0.652</td>
<td>Excellent</td>
</tr>
<tr>
<td>Completely</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Vaguely</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8.4 I __________ value attempts by Teaching staff to encourage mastery learning or learning contracts as instructional strategies.</td>
<td>0.576</td>
<td>Very good</td>
</tr>
<tr>
<td>Scarcely</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Very much</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>8.5 I am __________ of attempts by Teaching staff to work with me to set challenging objectives for learning outcomes.</td>
<td>0.637</td>
<td>Very good</td>
</tr>
<tr>
<td>Appreciative</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Unappreciative</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

With regards to the effectiveness of the proposed Different Levels of Instructional Strategies for Online Learning, I would like to suggest that;

This is because I believe that such strategies would;

Thank you for completing this survey