INTERNATIONAL JOURNAL OF INSTRUCTIONAL TECHNOLOGY AND DISTANCE LEARNING

May 2013
Volume 10 Number 5

ISSN 1550-6908

PUBLISHER'S DECLARATION

Research and innovation in teaching and learning are prime topics for the *Journal of Instructional Technology and Distance Learning* (ISSN 1550-6908). The Journal was initiated in January 2004 to facilitate communication and collaboration among researchers, innovators, practitioners, and administrators of education and training involving innovative technologies and/or distance learning.

The Journal is monthly, refereed, and global. Intellectual property rights are retained by the author(s) and a Creative Commons Copyright permits replication of articles and eBooks for education related purposes. Publication is managed by DonEl Learning Inc. supported by a host of volunteer editors, referees and production staff that cross national boundaries.

IJITDL is committed to publish significant writings of high academic stature for worldwide distribution to stakeholders in distance learning and technology.

In its first eight years, the Journal logged over eight million page views and more than one and one half million downloads of Acrobat files of monthly journals and eBooks.

International Journal of
Instructional Technology & Distance Learning

Vol. 10. No. 5.

ISSN 1550-6908

Table of Contents – May 2013

Editorial

Instructional Design 3: Learning styles

Donald G. Perrin

Myers-Briggs Personality Types are the basis for career choices, leadership styles, and marriage counselling. Harvey Silver used Myers-Briggs concepts to develop his Learning Styles Inventory. Readers requested more information after the reference in last month's editorial. The following is a 1997 adaptation of his work prepared by the editor for the California Department of Education.

THE LEARNING STYLES INVENTORY

The ***Learning Styles Inventory*** identifies four learning styles--directive, inquiry, creative, and interactive. People learn in all modes, yet one style is usually dominant or preferred. The lecture is a *directive* style of learning--a step-by-step development toward a goal set by the instructor. This style is not for everyone. Students who are curious move ahead on their own. Students who are creative are frustrated by slow and linear presentation. Students who prefer team learning are stifled by lack of interaction. The result is that curious, creative, and interactive learners - non-traditional learners who act independently - may be considered to be inattentive, disruptive, disobedient, or even poor students.

Sensing

S-T		S-F
	DIRECTIVE Learners are led, guided or managed by a teacher, aide, or programmed instruction, in sequential steps, in an organized way, to achieve individual or group goals determined by the instructor.	**INTERACTIVE** Learners are involved in face-to-face communication in interdependent and collaborative ways to achieve a common goal or outcome determined by themselves and/or the instructor.
Thinking		**Feeling**
	FORMULATIVE Learners actively process or mediate learning variables, frame hypotheses, experiment, seek solutions, or critique products to achieve individual or group goals, established by themselves and/or the instructor.	**CREATIVE** Learners actively engage in divergent thinking, generate new ideas or products, synthesize an original pattern from static parts, or achieve a self-imposed individual or common group goal, or one established mutually with the instructor.
N-T		N-F

iNtuiting

Figure 2. The Learning Styles Inventory.

A Sensing-Thinking person (S-T) fits the traditional ***DIRECTIVE*** model of teaching and learning via lecture-demonstration, presentation, and tutorial. This learning style fits persons who are practical, matter-of-fact, and work oriented. This is the dominant instructional mode for adult learners and does not well serve the needs and preferences of those who are inquiry oriented, interactive, and creative.

Obviously brilliant persons like Albert Einstein, Bill Gates, and Whoopee Goldberg are dropouts from traditional education. They are considered by some to be *learning disabled* when required to learn in a traditional situations.

Silver notes that mismatch between teaching style and learning style is a source of difficulty that may be as frustrating as trying to write with your other hand!

An iNtuitive-Thinking person (N-T) has a mind that is ***FORMULATIVE*** or inquiry oriented. This learning style fits upper cognitive endowed persons (highly intelligent) who are logical, ingenious, and curious. This person learns best through exploration and experiment. This is the most natural way of learning since curiosity leads to experiment and the results are learned. This person tends to be self-directed and proceeds ahead of the instructor in examining and developing his or her own learning. What may seem to be an oppositional, very divergent, or baffling student behavior masks the powerful learning that *could* take place. In the hands of an insensitive teacher this learner is set up for discouragement and failure. The same person may excel in a science fair or a self-directed project.

A Sensing-Feeling person (S-F) is ***INTERACTIVE*** and fun to know. This learning style fits persons who are sympathetic, friendly, and cultivate group harmony. This person is gregarious, likes to work collaboratively and is a productive team member. He or she functions better as part of a team than working alone. In school this person may be ideally identified as the helper to students new to the class or experiencing some lesson difficulties; and an eager helper to the teacher in problem solving whatever the situation may be.

An iNtuiting-Feeling person (N-F) is invariably insightful, imaginative, and ***CREATIVE***. This person is recognized as an innovator, inventor and artist in his or her chosen disciplines. Creative persons often seem disorganized because their minds move rapidly and simultaneously in divergent paths. Some do not fit well in traditional learning environments and may fail or do poorly in required courses because they do not meet scheduled deadlines; have trouble organizing information; or produce products divergent to the instructor's intentions. Fortunately they are capable of taking care of their own learning even if they do not respond well to traditional methods of teaching. However, as suggested above, they often suffer consequences and may dislike or drop out of school despite high academic potential.

In traditional or directive learning, the instructor controls lesson goals and presentation. In the other three quadrants, students participate in goal-setting and assume greater responsibility for learning. Harvey Silver advocates teaching-around-the-wheel which means combining a variety of teaching-learning styles within a lesson to involve a wider range of students.

The test instrument for the Silver-Strong Learning Styles Inventory places students according to their level in each of the four areas. A student dominant in a single mode would be close to an outside corner. A student equal in all areas would be in the center. Most students will have one or two areas of dominance.

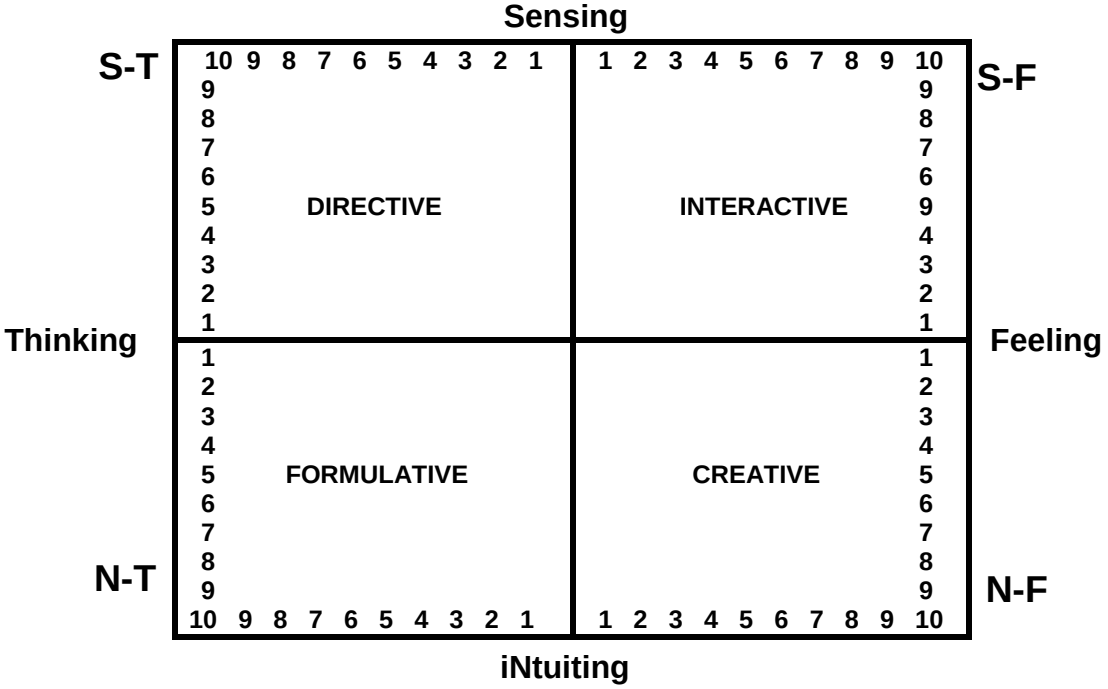

Figure 3. Quantifying levels in the Learning Styles Inventory.

BRAIN HEMISPHERES

It is not coincidence that the left and right sides of the Learning Styles Inventory correspond to the left and right hemispheres of the brain.

Left-brain dominance has a preponderance of processes that are: sequential, linear, rule-governed, and rely on previously accumulated organized information.

Right-brain dominance has a preponderance of processes that are: simultaneous, imagistic, transformative, and qualitative patterns

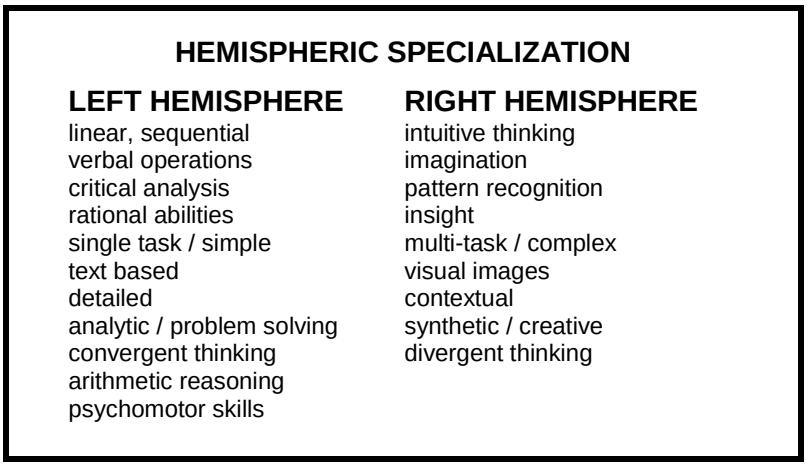

Figure 3. Hemispheres of the Brain.

Educators such as *Bruce Joyce* identified and taught instructional methods to develop learning through curiosity (Formulative), creativity (Creative) and social participation (Interactive) for persons with these learning style preferences.

Most adult learners are non-traditional learners - mature, independent, and self-reliant. They resent paternal controls and punishments that are used in traditional schools and colleges. Given a supportive learning environment, they are eager to participate in goal setting, program planning, and assume responsibility for their own learning.

In contrast to these able learners, we should recognize special needs of learners with limited English language skills, restricted opportunities for education, and learning disabilities. Initially, they may benefit from traditional teaching with sequential presentation of ideas. The ultimate goal is develop the capacity for independent learning, so assistance and incentives should be used to broaden the spectrum of learning activities - and learning styles - in a way that will promote successful learning.

References:

Joyce, Bruce and Weil, Marsha. (2008) Models of Teaching (8th Edition) Allyn & Bacon, Inc.

Myers, Steve. (2012) Myers-Briggs Personality Types. Team Technologies http://www.teamtechnology.co.uk/tt/t-articl/mb-simpl.htm

Perrin, Donald G. (1997) Child Nutrition Curriculum. California State Department of Education in collaboration with the Nutrition Department, San Jose State University.

Silver, Harvey F, Strong, Richard W. and Perini, Matthew J. (2000) So Each May Learn. Chapter 6; Teaching learning styles and multiple intelligences to students, ASCD Publications. http://www.ascd.org/publications/books/100058/chapters/Teaching-Learning-Styles-and-Multiple-Intelligences-to-Students.aspx

Return to Table of Contents

Return to Table of Contents

Editor's Note: We are reliant on test scores to measure progress, quality, and achievement. Design of testing instruments is a highly qualified task that requires input from a wide range of stakeholders: Futurists and social engineers who set goals and policies. Instructional designers, curriculum specialists, sociologists, psychologists, media producers, teachers and administrators who can translate these goals into creative and relevant educational experiences for a diversity of learners, and test gurus who design instruments to accurately measure progress and criterion achievement. The validity of a test is frequently challenged by teachers, students, administrators, and the public at large. This paper looks at ways to assess test fairness.

Test fairness in traditional versus dynamic assessment

Afsaneh Baharloo
Iran

Abstract

The focus of the present study is to unfold the conceptualizations of test fairness from two different perspectives: traditional and dynamic assessment. First, it goes over a variety of definitions presented for fairness. The paper then discusses three views regarding the relationship between test fairness and test validity in order to get better insights into the nature of the intended concept. It further investigates Kunnan's test fairness framework (TFF) as one of the most comprehensive models presented for test fairness. It tries to review and criticize this model. It is worth noting that the three views, discussed in this paper, represent fairness from a traditional perspective. Furthermore, the study elaborates on dynamic assessment and its main tenets since it intends to compare the conceptualizations of fairness within traditional and dynamic assessment. In fact, fairness is viewed from a completely different perspective in dynamic assessment in which instruction and assessment are integrated and dialectically related to form an approach which prioritizes development over measurement.

Keywords: Traditional assessment; dynamic assessment; test fairness; Kunnan's test fairness framework; language development

Introduction

Testing is a multi-faceted and intricate field in which right decision-making is very complicated. In order for any evaluation to be reliable, a number of considerations should be taken into account. In fact, evaluation usually leads into making decisions about individuals and situations; therefore, several consequences will follow as a result of the decisions. Some of these consequences are social or psychological, affecting individuals' motivation, goal, and even social status. As Bachman (1990) states, "since testing takes place in an educational or social context, we must also consider the educational and social consequences of the uses we make of tests (237)". Thus, one can easily notice that testing involves many intricacies because it eventually requires raters to judge test takers on their potentials and causes certain changes in their life path. Thus, thorough attention should be paid to consider as many relevant aspects as possible in order to make fair judgments.

Fair judgment requires measurement professionals to be aware of the concept of test fairness and its characteristics as well as other pertinent testing concepts so that they know how to observe this feature in different testing contexts as much as possible. Though many test developers and raters think that they know what 'fairness' is, they simply consider it as a test quality which pertains to a test itself and guarantees its content validity. However, one should notice that "test fairness" is a multi-faceted issue which is not confined to the content of a test and covers other aspects of testing as well. In fact, most test developers and raters attend to superficial levels and certain dimensions of test fairness which are easier to reach and do not make any attempt to achieve fairness in its full sense and at a higher level.

Fairness should not be restricted to either test development or administration. According to Willingham and Cole (1997), fairness should apply to all stages of assessment, from assessment conceptualization to the use of assessment results. One should not simply view the concept of test fairness as being confined to the test itself. As a matter of fact, having so simplistic a point of view about such an important and complex issue results in unfair testing contexts, violation of test takers' rights and finally lack of sufficient research in this area.

The concept of test fairness is so complicated and controversial that yet no agreed-upon definition is provided. In addition, some fairness models have been proposed but none lends itself easily to practical investigation of fairness. Furthermore, to the best of the researcher's knowledge, no previous study has investigated the trend that fairness follows to see how its concept differs in traditional and dynamic assessment. Therefore, the current study is an endeavor to provide a comprehensive portrait of 'test fairness' and discusses Kunnan's (2004) framework as the most comprehensive model available for test fairness. It also presents different views about fairness and elaborates on each view. Finally, the study compares fairness in traditional and dynamic assessments and hence fills the aforementioned gap.

Test fairness

Test fairness has not been paid due attention for most of the twentieth century. People believed that group differences were reflections of reality and they had few concerns about fairness. Gradually, measurement professionals began to study score differences between groups and issues of fair testing. It was almost at the beginning of 1970s that they began to pay increasing attention to test and item fairness (Cole and Zieky, 2001). Kunnan (2010) thinks that test fairness as one of the most fundamental concepts in evaluation entered the forefront of discussions in the field of language assessment from the late 1990s.

In fact, fairness is such a complicated concept that a variety of definitions has been proposed to clarify its broad and controversial nature. According to Webster's Ninth New Collegiate Dictionary (1988), 'fair' means 'free from favor toward either or any side'. Xi (2010) believes that such a definition indicates that a central focus of test fairness is the comparison of testing practices and test outcomes across different groups. Therefore, test fairness mainly arises from the way group differences are perceived and treated. Similarly, Willingham and Cole (1997) and Xi (2010) define fairness as comparable validity for all the identifiable and relevant groups.

The language testing literature has mainly treated fairness under the heading of bias. Test bias studies "are directed to identifying and where possible reducing the effect of any confounding variables on test scores, by making changes to the test" (Elder, 1997, p. 261). McNamara and Roever (2006) state that the term "bias" in assessment research conveys an unfair and skewed inclination toward one group or population to the detriment of another. Therefore, the notion of bias is highly tied to fairness in assessment: A biased judgment unduly takes into account factors other than those that should be informing it. In traditional validity terms, bias can be seen as construct-irrelevant variance that distorts the test results and hence makes conclusions based on scores less valid. Specifically, one can consider a test as biased if test takers of equal ability but from different groups score differently on the test items depending on their group membership (Angoff, 1993).

According to the Standards for Educational and Psychological Testing (AERA, APA & NCME, 1999), fairness is defined as absence of bias, equity in opportunity to learn the material in an achievement test and equitable treatment of all test takers in the testing process. However, test fairness is so broad an area that many measurement professionals consider it to encompass quality management in test design, administration and scoring, adequate coverage of relevant content, sufficient construct validation work, equal learning opportunities and access to testing, and items

measuring only the ability under investigation without being unduly influenced by construct-irrelevant variance introduced through test-taker background factors (Kunnan, 2000; McNamara and Roever, 2006; Saville, 2003, 2005; Shohamy, 2000).

Davies (2010) believes that many of the accounts of fairness go back to the philosopher John Rawls who argues that 'justice is fairness' (Rawls, 2001). He proposes two principles underlying his argument. The first is to ensure that everyone has the same claim to the basic liberties. The second is that where inequalities exist they must satisfy two conditions, that everyone should be provided with equality of opportunity, and that the least-advantaged groups should benefit most from these inequalities. Having a similar idea, Velasquez *et al*. (2008) link fairness to justice and define justice as giving every member what he or she deserves, or, in other words, giving each person his or her due.

As it was mentioned, measurement professionals have defined test fairness in different ways. Such definitions may not clarify the concept of test fairness and its relevant aspects to the extent that practitioners can observe fairness in actual testing contexts. They may need a concrete model through which fairness can really be observed and applied in testing situations. Among all the available models proposed for test fairness, Kunnan's (2004) framework can be considered as the most comprehensive fairness model. In what follows this model is explained and criticized.

The test fairness framework

A number of approaches have been proposed to investigate fairness. However, Kunnan's (2004) test fairness framework is the main concern of this study since this model has been at the forefront of attention regarding test fairness for several years. Kunnan (2010) puts forward an ethics-inspired rationale for his framework and claims that this model considers the whole system of a testing practice, not only the test itself, hence it seems to be more comprehensive than the other existing models. Kunnan's (2004) framework was the first attempt made to propose an overarching framework for fairness research in language testing (Xi, 2010). In his earlier work on test fairness (Kunnan, 2000), he considers fairness as a three-faceted concept which deals with validity, access and justice. He agrees with Jensen (1980) who thinks that "the concepts of fairness, social justice, and equal protection of the laws are moral, legal, and philosophical ideas and therefore must be evaluated in these terms" (Jensen, p. 376). Xi (2010) also thinks that such an idea mainly arises from social justice theories. However, Kunnan tried to expand his ideas and develop a more comprehensive model that was later proposed in 2004. In this framework, he views fairness as an overarching concept which includes five qualities: Validity, absence of bias, access, administration, and social consequences. According to this approach, validity of a test score interpretation, which is considered as part of the test fairness framework, can be supported through four types of evidences: *Content representativeness or coverage evidence* which refers to the adequacy with which test items represent the test domain, *construct or theory-based validity evidence* which refers to the adequacy with which test items represent the construct or the underlying trait being measured in a test, *criterion-related validity evidence* which refers to whether the test scores meet some criterion variables, and *reliability* evidence which refers to the consistency of test scores. The second quality refers to absence of any source of bias such as offensive content or language, unfair penalization based on test takers background, and disparate impact and standard setting. In fact, offensive content can cause bias for test takers from different backgrounds because it may conflict with their beliefs or it may be needlessly controversial (McNamara and Roever, 2006). A test is also biased if it causes unfair penalization due to a test taker's group membership. In addition, disparate impact and standard setting can bring about different performances by test takers from various group membership. The third quality of the fairness framework refers to test takers' access to the test in terms of educational, financial, geographical, personal, and equipment access. In other words, test takers should have opportunity

to learn the content and get familiar with the kind of tasks and cognitive demands required by the test. Furthermore, the test should be affordable for test takers and the site should also be accessible in terms of distance as well as test takers' physical and learning conditions. In addition, test takers should be familiar with the test taking equipment, procedures, and conditions. The fourth feature of test fairness framework is related to administration conditions. This quality refers to the physical conditions of test administration such as optimum light and temperature as well as uniformity and consistency across test sites and in equivalent forms. Finally, social consequences of a test should be studied as contributing to test fairness. These consequences refer either to the effect of a test on instructional practices or the remedies offered to test takers to compensate for the detrimental consequences of a test.

Although Kunnan's test fairness framework considers many relevant factors and thus seems to be a comprehensive model, it has several shortcomings. First, it mainly arises from theories and is not practical enough to provide a principled guideline to ensure all the aspects of test fairness. Having a similar idea, Xi (2010) also criticizes Kunnan's (2004) framework for not providing practical guidance on how to develop the relevant evidence to support test fairness. He thinks that although this approach may be useful in pointing to general areas of research and practice, it does not provide a mechanism to integrate all the aspects of fairness investigations into a fairness argument, nor does it offer a means to plan fairness research. Xi believes that a framework should identify and prioritize research needs so that one can gauge the progress of fairness investigations.

Second, although Kunnan (2004) claims that this fairness framework can apply to the whole system of a testing practice, it does not actually consider all the stages regarding assessment. The Code of Fair Testing Practices in Education (1988), modified in 2004, highlights the role of fairness as a test quality that pertains to the whole assessment process. According to the 1988 Code, fairness is not an isolated concept, but must be considered in all the aspects of the testing process. Therefore, fairness issues are not only associated with developing appropriate tests, administering and scoring them but also extend to the accurate reporting of individual and group test results since individuals have rights to receive feedback on their performance so that they get aware of their strength and weakness. However, there is no concern for the latter issues in Kunnan's test fairness framework; thus, it is not comprehensive enough to consider all the aspects and consequences of a testing practice. In addition, this test fairness framework does not specifically define any qualities devoted to the responsibilities of test developers and users regarding the importance of their roles. On the one hand, test developers should provide test users with sufficient information and supporting evidence to help them select appropriate tests. They should also explain how to administer and score tests correctly and fairly. On the other hand, test users should inform test takers about the nature of the test, test takers' rights and responsibilities, the appropriate use of scores, and procedures for resolving challenges that they encounter in the evaluation process (McNamara and Roever, 2006).

Finally, Kunnan's test fairness framework mainly focuses on *group differences* and the kind of bias that may stem from test takers' membership in different groups, but it does ignore the important issue of *individual differences*. This lack of attention to individual differences may result in having test items and tasks, which are more suitable and convenient for some individuals with certain learning styles; but are not appropriate for all the members of the same group. For instance, the same test given to two groups of men and women can yield different results that may be attributed to gender differences. However, one should note that there are some intra-group differences regarding the ability being tested, for example: Test taking strategies or learning styles that can bring about different performances. Highlighting the importance of individual differences, Cole and Zieky (2001) state that, "all testing data show far more individual variation of scores within groups than variation between groups. Individual variation, not group variation,

is the dominant influence on scores and should therefore be the dominant fairness concern" (p. 11). Therefore, considering the qualities and aspects included in Kunnan's test fairness framework, one eventually finds out that it does not provide a practical means to investigate test fairness in its full sense.

Xi (2010) believes that establishing a fairness framework that would be useful for practical purposes requires primary attention to the conceptualization of fairness. Fairness related theories, models, and definitions suggest that most measurement professionals study test fairness in relation to validity issues since test fairness is sometimes influenced by the interpretations of test scores. Therefore, studying about how fairness and validity are related can provide better insights into the conceptualization of fairness and its practical investigation.

Fairness and validity

Fairness has been conceptualized in various ways which result in different approaches of viewing fairness. Though these conceptual approaches may vary with regard to the degree of their emphasis on issues such as the social and political aspects of fairness, a central point on which they differ is how fairness is related to validity (Xi, 2010). Considering the relationship that may exist between fairness and validity, Xi proposes three views: whether fairness is independent of validity, subsumes it, or is a facet of it.

View 1: Fairness as an independent test quality

This view considers fairness as a relatively independent facet of test quality or general testing practices and does not have consistent and clear connections to validity. According to this view, fairness is conceptualized as a test quality that is separate from validity. The *Standards* for Fairness and Quality by Educational Testing Service (ETS, 2002) and the *Code* of Fair Testing Practices in Education (1998, 2004) are representative of this approach. The *1999 Standards* claims that "fairness requires that construct-irrelevant personal characteristics of test takers have no appreciable effect on test results or their interpretation" (p. 17). The Code primarily focuses on the partition of responsibilities between test developers and users in ensuring fair testing practices. This is also a major contribution of the Code compared to the Standards, as the division of responsibilities between test developers and users has not always been clear-cut (Shohamy, 2001), it requires both test developers and users to share joint responsibilities to ensure fairness. In addition, the ETS Standards for Fairness and Quality presents a broad list of fairness standards; but it does not provide a mechanism for weighing one piece of fairness evidence against another or for prioritizing them either. Furthermore, one of the standards, proposed in the ETS Standards for Fairness and Quality, explains that if the use of assessment results brings about unintended consequences for a studied group, the validity evidence should be investigated to see if the differential impact for the group is a result of construct-irrelevant factors or construct under-representation. Such elaboration implies the potential existence of a more consistent and coherent linkage between test fairness and validity (Xi, 2010).

View 2 – Fairness as an all-encompassing test quality

According to this view, fairness subsumes validity; in other words, fairness is viewed as an overarching test quality with different facets including validity. This view gives primacy to test fairness and defines it as a test quality, which goes beyond validity. Therefore, a test has to be valid to be fair. Kunnan's test fairness framework is a manifestation of this view since validity is considered as one of the five qualities that contribute to fairness. Kunnan sees fairness as a test quality that encompasses validity, absence of bias, access to the test, administration conditions, and test consequences (Kunnan, 2004, 2010). However, this view considers fairness as a broad concept, which consists of several facets that are related to one another. That is why Bachman (2005) criticizes Kunnan's work in which various fairness components are not necessarily con-

nected to one another. Thus, he emphasizes on the need for a mechanism to integrate them properly to support an overall fairness argument.

In addition, McNamara and Roever (2006), who are proponents of the second view, focus on the social dimensions of language testing that are evident in the investigations of item bias. Their discussion of test fairness is motivated by the desire to ensure social justice. They argue that factors, which cause advantages and disadvantages for some groups of test takers and bias their educational opportunities, lie in the larger social context. McNamara and Roever put great emphasis on the social and political aspects of fairness. They believe that the social embeddedness of testing can be dealt with in the form of procedures of fairness review and the promotion of codes of ethics. Fairness review or sensitivity review refers to the formal process of identifying and eliminating possibly biased items during the test construction process. The codes of ethics are "useful for guiding ethical decisions and protecting testers from stakeholder pressures to take actions that contravene professional conduct" (McNamara and Roever, 2006, p. 7). Therefore, adopting this view requires thorough attention to all the aspects of fairness and their relations as well as the social and political context of a testing practice.

View 3: Fairness linked directly to validity

This view arises from the *1999 Standards* which endorses three important characterizations of test fairness in the field of educational and psychological testing: fairness as lack of bias, fairness as equitable treatment of all examinees in the testing process, and fairness as equity in opportunity to learn the materials covered in an achievement test. Xi (2010) states that the *1999 Standards* explicitly rejects the popular view that fairness requires the equality of testing outcomes for different test taker groups. The *1999 Standards* argues that a more widely accepted view would hold that test takers from different groups with equal standing regarding the construct of interest should on average receive the same test score. In addition, the *1999 Standards* advocates the gathering of various pieces of evidence to ensure test fairness. The *Standards* requires the investigation of each type of validity evidence for relevant sub-groups of examinees to determine if the interpretation and meaning of test scores and the consequences of the use of assessment results may differ as a result of construct under-representation or construct irrelevant factors. Xi, who is in favor of this view, points out that the connection between discussions of fairness and validity reinforces the possibility for linking fairness back to validity in a principled way which could not be observed in Kunnan's work. Therefore, this kind of linkage would allow fairness research and practice to benefit from a well-defined framework for validity.

In fact, Xi "proposes an approach for studying fairness that links it directly to validity. Fairness is characterized as comparable validity for relevant groups that can be identified" (p. 167). She treats fairness as an aspect of validity. Therefore, anything that weakens fairness compromises the validity of a test as well. However, Davies (2010) criticizes Xi's proposed conceptualization. He believes that validity does itself pertain to all comparable groups; why then do we need to appeal to fairness? Davies argues that validity guarantees that an ability is being appropriately tested for a relevant population which will be made up of various groups but there is sufficient homogeneity across groups for them to be treated as comparable. He believes that a test that is valid for adults may not be valid for children because they belong to different populations. It is not whether such a test is fair or unfair for children: the test is just invalid for the latter group. Davis thinks that the search for fairness in language testing is chimaerical: First because it is unattainable, and second because it is unnecessary. So Davies's idea is not in line with any of the three conceptualizations presented above as he rejects the concept of fairness overall.

Fairness in traditional and dynamic assessments

All the three views, discussed regarding the relationship between fairness and validity, have been proposed within the framework of traditional assessment. Traditional testing contexts make a

clear distinction between development and measurement. Assessment usually follows instruction and is not intended to improve test takers' learning. Since such assessment is usually statistically based and grounded in psychometric principles, it considers any change in the person's performance during the administration of the assessment as a threat to those principles, and to test reliability as well (Pohner and Lantolf, 2005). In such a traditional perspective, test fairness requires providing learners with equal opportunities to learn and subsequently take part in exams. Instruction is based on a learning hierarchy composed of a sequence of increasingly difficult tasks. In fact, teachers provide all the learners with the same material without considering their needs and teach them equally since all learners are supposed to receive the same amount of input and support from the teacher to move through the predetermined hierarchy. In other words, teaching undergoes several distinct stages; therefore, learning can be investigated through traditional assessment instruments, designed equally for all the learners, at a particular point in the teaching sequence. All the learners receive the same test on which they should perform independently. They should not cooperate with their peers or teacher during the exam since the only purpose of assessment is measurement. Therefore, the central focus of fairness within a traditional framework is to provide learners with equal learning opportunities and access to a test which usually takes the form of a summative assessment and evaluates performance at the end of a program and is often used for the purposes of accountability, admission decisions, promotion and selection (Poehner and Lantolf, 2005).

However, dynamic assessment views teaching and testing from a different perspective in which the pursuit for fairness undergoes a different path. Dynamic assessment (DA) is an approach to assessment and instruction derived from Vygotsky's socio-cultural theory of mind and his focus on the Zone of Proximal Development. In this approach, assessment and instruction are integrated as a single activity that seeks to simultaneously understand and promote learners' abilities through mediated interaction in the Zone of Proximal Development (Poehner, 2008). In other words, dynamic assessment blends instruction and assessment and benefits from tutor mediation to identify and respond to the areas that students need the most support in (Shrestha, et al., 2012). Therefore, DA has two major concerns: first, teaching and testing are dialectically integrated to the extent that one cannot tell the two activities from each other at a particular point. Second, learners receive support from the teacher within their ZPD even when they are performing on a test, because this approach advocates any tools that lead to development; therefore, assessment is not mainly intended to measure learners' knowledge but to develop it. Thus, development has priority over measurement.

ZPD refers to the 'difference' between what learners can do independently and what they can do with assistance on a test. The idea of ZPD is highly associated with Vygotsky's socio-cultural theory of mind. He believes that engagement in activities that are mediated by others and by cultural objects allows individuals to develop higher forms of consciousness that are unique to humans (Vygotsky, 1978, 1986). "In Vygotsky's view, abilities do not simply mature on their own but instead result from individuals' histories of engaging in activities with others and with cultural artifacts" (Poehner, 2008, p.24). Socio-cultural theory implies that other individuals and cultural artifacts are not merely a factor of development, but they are the source of development. Dynamic assessment is not concerned with how much development can be attributed to the individual and how much to the environment. According to this approach, the individual and the environment form an inseparable dialectical unity that cannot be understood if the unity is broken. The interaction between learners and their environment helps them develop awareness of and control over their psychological functions, including attention, perception, and memory (Poehner, 2008). Newman *et al*. (1989) also believe that cognitive changes arise from the productive intrusion of other people and cultural tools in the developmental process. Kozulin (1998, 2003) considers physical, symbolic, and psychological tools as a way of conceptualizing Vygotsky's central argument that an individual's social and cultural environment is the source of

the development of higher psychological functions. In a Vygotskian view, humans relate to their world *psychologically* in much the same way as they do *physically*. Unlike physical tools, symbolic tools, which Vygotskian researchers refer to as cultural artifacts, may not only be directed outwardly to mediate individuals' relationship with the world, but also inwardly, to mediate their relationship with themselves (Vygtosky, 1994). Therefore, as Poehner and Lantolf (2005) put it, "the unit of analysis for the study of development is not the individual acting alone, but the interpersonal functional system formed by people and cultural artifacts acting jointly to bring about development" (p. 238). Wertsch (2007) explains that for Vygotsky mediation is the "hallmark of human consciousness because it is through their appropriation of the forms of mediation provided by particular cultural, historical, and institutional forces that their mental functioning is sociohistorically situated" (178). Highlighting the importance of mediation and intervention, Shrestha *et al*. (2012) believe that like the ZPD, mediation is integral to dynamic assessment. While the ZPD is about the individual's potential development, mediation provides an opportunity for such development. Mediation is defined as a process that humans employ in order to regulate the material world, others, or their own social and mental activity by using "culturally constructed artifacts, concepts and activities" (Lantolf and Thorne, 2006, p. 79).

In order to get better insights into the ways traditional and dynamic assessments differ, one can think of them as assessment *while* teaching and assessment *by* teaching respectively. According to Newman *et al.* (1989), assessment *while* teaching requires children to move through a learning hierarchy composed of a sequence of increasingly difficult tasks. Therefore, determining how successful the children are at moving through the sequence is often derived from their independent performance on traditional assessment instruments at a particular stage in the teaching sequence. But, assessment *by* teaching, which is in line with dynamic assessment, suggests that instruction is not organized according to "a neat sequence of levels to be mastered in an invariant sequence with a single correct route to mastery. Tasks and knowledge may be organized according to a teacher's assumptions about their relative complexity" (Newman *et al.*, 1989, p. 78). Therefore, once teachers and students engage in instructional activities, things can move in unanticipated directions and at unanticipated rates (Poehner and Lantolf, 2005).

In dynamic assessment, there is a shift of attention from focusing on learners' independent performance on traditional measuring instruments to focusing on the process of development through mediated interaction. Since development has priority over measurement in dynamic assessment, fairness suggests that learners should not be deprived from any tool that promotes their learning. Therefore, even a test, which has traditionally been used only for measurement purposes, should now be in service of development.

In fact, dynamic assessment views test fairness from a perspective different from the one underlying the three views proposed by Xi (2010). Those views discuss fairness within the framework of traditional assessment in which test fairness requires institutions and teachers to provide equal opportunities for all the individuals to learn the same material and consequently provide them with the same testing conditions for measurement purposes. In other words, traditional assessment is more product-oriented and seeks to measure the ultimate level that learners have reached. The views, already discussed regarding the relationship between fairness and validity, are included within this product-oriented approach. However, the third view seems to manifest some traces of dynamic assessment since in this view Xi points to some sort of equitable treatment of individuals but it still follows the main tenets of traditional assessment with regard to instruction and testing procedures. Unlike the traditional approach, dynamic assessment focuses on the developmental process and hence is considered as a process-oriented approach in which a test is a helpful tool that can both measure and promote individuals' knowledge so that they can transfer what they learn to other tasks beyond the test. Having a similar idea, Shrestha *et al*. (2012) state that, "DA is grounded in the notion of assessment as a process rather than a

product. In other words, DA is a development-oriented process which reveals a learner's current abilities in order to help them overcome any performance problems and realize their potential" (p. 5).

It is worth noting that the major difference between the ways that traditional and dynamic assessment view fairness lies in the different teaching and testing relationships within the two approaches. In the traditional sense, instruction and assessment are separate activities carried out at particular stages. All the learners are taught the same material selected based on a predetermined hierarchy and they later receive the same test on which they should perform independently since the only purpose of such a test is measurement. Therefore, in order to observe fairness, educational systems and practitioners are required to treat all the individuals equally regardless of their needs and backgrounds. But dynamic assessment takes on a different perspective in which instruction and assessment are integrated in all the stages so that one can not distinguish the two activities from one another at a single point. All the individuals do not receive the same instruction. In fact, each learner receives as much assistance as he or she deserves. Development is achieved intentionally rather than incidentally. As Poehner and Lantolf (2005) mention, "dynamic assessment is a pedagogical approach grounded in a specific theory of mind and mental development… [Therefore,] mediation cannot be offered in a haphazard, hit-or-miss fashion, but it must be tuned to those abilities that are maturing" (260). Although dynamic assessment can be carried out either formally or informally, it must be systematic. It insists upon the inseparability of assessment and instruction because they form a unity necessary for learner development. In such an approach, fairness requires providing each individual with what they deserve regarding their needs. Learners do not move through a hierarchy of tasks sequenced based on their difficulties rather each individual receives what is required for his or her development based on a theory of mind since mediated interaction and intervention should be systematic in order to be fair and beneficial. Even, their performance on a test is assisted by receiving support from others such as their peers, teachers and whatever exists in the environment. Thus, each individual receives as much assistance as he or she needs. In fact, dynamic assessment attends to individual differences in a practical sense. This is contrary to the views of traditional assessment.

Models following traditional perspectives define fairness in terms of equal treatment of individuals regarding learning opportunities and testing conditions. In such an approach, the focus is on inter-group differences rather than intra-group differences. Within this framework, tests are only used for measurement and any kind of intervention may threaten their reliability since they are often used for purposes of admission decisions, promotion, and selection. Therefore, test takers do not usually receive any kind of feedback on their performance to know which areas require more attention and practice. However, dynamic assessment employs a different view toward fairness. In this approach, teaching and testing are dialectically integrated and considered as a single activity since both aim at promoting learners' knowledge; and development has priority over measurement. In fact, this approach requires teachers to assist learners to overcome the difficulty of test tasks and master the intended knowledge being tapped by the test so that they can transfer such knowledge to other tasks beyond the test. Therefore, fairness in dynamic assessment does not only apply to the test itself or the testing process but to the whole teaching and testing activities integrated as a single unity which must ultimately lead to development. In other words, fairness implies that individuals should not be deprived of any opportunity that can promote their learning. As Poehner (2011) states, fairness in education, from a dynamic perspective, does not involve treating all learners as if they were the same, because doing so ignores that they are not. Fairness requires doing everything possible to promote learner development, with the understanding that some individuals will need more time and resources than others.

Reuven Feuerstein, a leading DA researcher, has proposed a "structural cognitive modifiability theory" to suggest that "traditional conceptualizations of the examiner/examinee roles should change in favor of a teacher-student unity that works towards the ultimate success of the student" (Feuerstein et al., 1979, p. 271). Putting this idea another way, Poehner (2011) states that the object of assessment is fully understood by actively seeking to promote a learner's knowledge. This orientation requires a shift on the part of the assessor, also referred to as a mediator, whose responsibility is no longer limited to neutrally observing learner's performance, but now involves engaging as a co-participant with learners. Feuerstein *et al.* (2002) believe that cognitive abilities are not fixed traits determined by our genetic endowments rather they can develop in a variety of ways, depending on the presence, and the quality of appropriate forms of interaction and instruction. Feuerstein *et al* (1979) state that "it is through this shift in roles that we find both the examiner and the examinee bowed over the same task, engaged in a common quest for mastery of the material" (p. 102). Following the same line of thought, Poehner (2008) thinks that teachers' intervention is necessary to help learners stretch beyond current capabilities. In other words, interpretations of learners' knowledge and abilities are broadened beyond observations of independent performance to include their contributions to, and responsiveness during engagement in joint activity with a mediator. In addition, the instructional quality of the interaction begins the process of helping learners move toward overcoming current difficulties (Poehner, 2011). Sternberg and Grigorenko (2002) also state that DA is a procedure whose outcome takes into account the results of an intervention in which the examiner teaches the examinee how to perform better on individual items or on the test as a whole. Therefore, the final score may be a learning score representing the difference between pretest (before learning) and posttest (after learning) scores, or it may be the score on the posttest considered alone. It seems that the fundamental difference between the traditional and dynamic approaches has to do with whether or not the administration of the assessment should have the expressed goal of modifying the learner performance during the assessment itself (Poehner and lantolf, 2005).

Therefore, fairness is conceptualized differently in traditional and dynamic assessment. In traditional approaches, teaching and testing are considered as distinct activities with different objectives. Hence, fairness, in the former approach, requires institutions to provide equal opportunities and conditions for all learners to learn the same material and to perform on the same traditional measuring instrument independently. However, the latter approach which blends instruction and assessment views fairness in terms of providing each individual with what he or she deserves based on need analysis and ongoing assessments used for diagnostic purposes.

Conclusion

This paper investigated test fairness conceptualizations within the perspectives of traditional and dynamic assessment. It started with presenting available definitions for fairness to provide a portrait of the concept at hand. In order to get better insights into the very nature of this test quality in a more concrete sense, it investigated Kunnan's (2004) test fairness framework as the most comprehensive available fairness model. However, a closer look at the framework and its components resulted in a number of criticisms. First, it is impractical due to the lack of guidelines on how to ensure validity. Contrary to what Kunnan claims about the comprehensiveness of his framework, it does not apply to the whole system of a testing practice since it does not indicate any concern for accurate reporting of test results and informing test takers as well as providing them with feedback on their performance with regard to their strengths and weaknesses. In addition, this model understates the important roles of test developers and test users by not clarifying their responsibilities in the testing process. Furthermore, it ignores the intra-group differences and only attends to inter-group differences. Therefore, the test fairness framework does not seem to be practical and comprehensive enough to be applied to the whole system of a testing practice appropriately. Xi (2010) suggests that establishing a useful framework for

practical research requires measurement professionals to have concerns for the conceptualizations of fairness. She proposes three views regarding how fairness is related to validity. The first view considers fairness as an independent test quality, the second view, to which Kunnan's test fairness framework belongs, sees fairness as an overarching test quality that consists of different facets including validity. The third view sees fairness as being directly linked to validity. It is worth noting that, all the views proposed by Xi are discussed from a traditional perspective toward assessment. However, fairness is viewed quite differently within the framework of dynamic assessment in which instruction and assessment are integrated in order to simultaneously measure and promote learner development. In DA, development has priority over measurement. As the paper discusses, the traditional and dynamic assessments conceptualize fairness from different perspectives. While, in the former approach, fairness requires instructors and measurement professionals to teach all the learners based on a predetermined schedule and provide them with equal opportunities to learn and take tests on which they are not assisted, the latter perspective defines fairness in terms of providing each individual with what he or she deserves. Therefore, individual differences as well as their needs and interests are taken into consideration. Although, dynamic assessment seems to view fairness from a more humanistic perspective, it requires careful attention and programming on the part of educational institutions and practitioners so that all the learners get what they really need and deserve. Employing needs analysis before beginning a course and having small classes help teachers implement the tenets of dynamic assessment and reach fairness as much as possible.

References

American Educational Research Association, American Psychological Association and National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In P. W. Holland and H. Wainer (Eds.), *Differential item functioning* (pp. 3–24). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly, 2*, 1–34.

Cole, N.S. and Zieky, M. J. (2001). The new faces of fairness. *Journal of Educational Measurement, 38*(4), 369-382.

Davies, A. (2010). Test fairness: A response. *Language Testing, 27*(2), 171–176.

Educational Testing Service (2002). *ETS standards for quality and fairness*. Princeton, NJ: Author.

Elder, C. (1997). What does test bias have to do with fairness? *Language Testing, 14*(3), 261–277.

Feuerstein, R., Rand, Y. and Hoffman, M. B. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. Baltimore: University Park Press.

Feuerstein, R., Falik, L., Rand, Y. and Feuerstein R.S. (2003). *Dynamic assessment of cognitive modifiability*. Jerusalem: ICELP Press.

Jensen, H. R. (1980). *Bias in mental testing*. New York: Free Press.

Joint Committee on Testing Practices (1988). *Code of fair testing practices in education*. Washington, DC: Author.

Joint Committee on Testing Practices (2004). *Code of fair testing practices in education*. Washington, DC: Author.

Kozulin, A. (1998). *Psychological tools: A sociocultural approach to education*. Cambridge, MA: Harvard University Press.

Kozulin, A. (2003). Psychological tools and mediated learning. In A. Kozulin, B. Gindis, V.S. Ageyev, and S.M. Miller (Eds.), *Vygotsky's Educational Theory in Cultural Context* (pp. 15-38). Cambridge: Cambridge University Press.

Kunnan, A. J. (2000). Fairness and justice for all. In A. J. Kunnan (Ed.), *Fairness and validation in language assessment* (pp. 1–14). Cambridge: Cambridge University Press.

Kunnan, A. J. (2004). Test fairness. In M. Milanovic and C. Weir (Eds.), *European language testing in a global context* (pp. 27–48). Cambridge: Cambridge University Press.

Kunnan, A. J. (2010). Test fairness and Toulmin's argument structure. *Language Testing, 27*(2), 183–189.

Lantolf, J. P., and Thorne, S. L. (2007). Sociocultural theory and second language acquisition. In B. V. Patten and J. Williams (Eds.), *Explaining second language acquisition* (pp. 201-224). Cambridge: Cambridge University Press.

McNamara, T. F. and Roever, C. (2006). *Language testing: The social dimension*. Oxford: Blackwell.

Newman, D., Griffin, P., and Cole, M. (1989). *The construction zone: working for cognitive change in school*. Cambridge: Cambridge University Press.

Poehner, M. E., and Lantolf, J. (2005). Dynamic assessment in the language classroom. *Language Teaching Research 9*(3). 233–265.

Poehner, M. E. (2008) *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Berlin: Springer.

Poehner, M. E. (2011): Dynamic assessment: Fairness through the prism of mediation. *Assessment in Education, Principles, Policy and Practice, 18*(2), 99-112

Rawls, J. (2001). *Justice as fairness: A restatement*. Cambridge, MA: Harvard University Press.

Saville, N. (2003). The process of test development and revision within UCLES EFL. In C. Weir and M. Milanovic (Eds.), *Continuity and innovation: Revising the Cambridge Proficiency in English examination 1913–2002* (pp. 57–120). Cambridge: Cambridge University Press.

Saville, N. (2005). Setting and monitoring professional standards: A QMS approach. *Cambridge ESOL Research Notes*, *22*, 2–5.

Shohamy, E. (2000). Fairness in language testing. In A. J. Kunnan (Ed.), *Fairness and validation in language assessment* (pp. 15–19). Cambridge: Cambridge University Press.

Shohamy, E. (2001). *The power of tests: A critical perspective of the uses of language tests.* London: Longman.

Shrestha, P., and Coffin, C. (2012). Dynamic assessment, tutor mediation and academic writing development. *Assessing Writing, 17*(1), 55–70.

Sternberg, R.J. and Grigorenko, E.L. (2002). *Dynamic testing. The nature and measurement of learning potential*. Cambridge: Cambridge University Press.

Velasquez, M., Andre, C., Shanks, T., and Meyer, M. J. (2008). *Justice and fairness*. California: Santa Clara University, Markkula Center for Applied Ethics.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, Massachusetts: Harvard University Press.

Vygotsky, L.S. (1986). *Thought and language*. Cambridge, MA: MIT.

Vygotsky, L.S. (1994). The problem of the cultural development of the child. In R. Van der Veer and J. Valsiner (Eds.), *The Vygotsky reader* (pp. 338-354). Cambridge, MA: Blackwell.

Webster's Ninth New Collegiate Dictionary (1988). Springfield, MA: Merriam-Webster.

Wertsch, J.V. (2007). Mediation. In H. Daniels, M. Cole, and J.V. Wertsch (Eds), *The Cambridge companion to Vygotsky* (pp. 178–92). Cambridge: Cambridge University Press.

Willingham, W. W. and Cole, N. (1997). *Gender and fair assessment*. Mahwah, NJ: Lawrence Erlbaum.

Xi, X. (2010). How do we go about investigating test fairness? *Language Testing, 27*(2) 147–170.

About the author:

Afsaneh Baharloo is a PhD candidate in TEFL at Shiraz University, Iran. She obtained her B.A. and M.A. degrees in English literature and ELT respectively from the same university.

Email Address: baharloo.afsaneh1@yahoo.com

Return to Table of Contents

Return to Table of Contents

Editor's Note: Techniques developed by advertisers and the motion picture industry to measure audience response are now widely used for data collection in a variety of disciplines and in education. Clickers and related devices enable a sampling of knowledge, comprehension and opinion at successive steps of the teaching-learning process. It requires minor redesign of the lesson for students to interact. It is also a feature of learning management systems for online learning. The data gathered can be used to guide the teacher and optimize the pace and depth of learning and to analyze the progress of each individual student.

Audience Response Systems in higher education courses: A critical review of the literature

Karly C. Good
USA

Abstract

This review explores audience response systems, otherwise known as clicker technology, usage (and impact) in higher education. Much of the research infers that there are positive effects on student assessments and student perceptions due to increased engagement through active and/or peer learning strategies. The findings are encouraging, but show limited depth in terms of pedagogy while using clicker technology in different disciplines. The quantity and quality of different question structures influenced student perceptions of technology implementation cost effectiveness. Studies showed that professional development and continued support helped alleviate regular occurrences of challenges when using clickers in the classroom. In all, clicker technology is positively viewed and could be transitioning with the digital age and mobile devices.

Keywords: audience response system, clicker technology, higher education, technology, educational technology support, literature review, student perceptions, student engagement, peer learning, active learning, pedagogy, professional development

Introduction

Educators are looking for ways to integrate interactive learning techniques into their large-size classes (Caldwell, 2007; Fitch, 2004; Hatch, Jensen, & Moore, 2005; Hoffman & Goodwin, 2006; Martyn, 2007; Mayer et al., 2009; Office of Information Technology, 2006; Plant, 2007). Many have turned to clicker technology as an educational tool that allows for scalable active learning techniques. Student learning is the focus of their instructional approaches; many instructors have attempted practices that allow the students to have voice and yet break away from public displays of knowledge.

As with any integrated technology, it takes additional work on the part of the instructor and students for the system to work. Instructors have learned to integrate more planning at the forefront of their lectures (Auras & Bix, 2007; Caldwell, 2007; Fitch, 2004; Griff & Matter, 2008; Hatch, Jensen, & Moore, 2005; Hoffman & Goodwin, 2006; MacGeorge et al., 2008; Martyn, 2007; Menon, et al., 2004; Office of Information Technology, 2006; Plant, 2007) and textbook publishers have begun to construct question banks to be used with clicker technology (Auras & Bix, 2007). Students are expected to register their clickers so they are integrated into a system and instructors are expected to prepare immediate feedback during the lecture. Students learn to discuss feedback and grades they received due to clicker technology use with their instructors (Auras & Bix, 2007; Caldwell, 2007; Edens, 2006; Kaleta & Joosten, 2007; Martyn, 2007; Mayer et al., 2009; Menon, et al., 2004; Office of Information Technology, 2006), and participate in data gathering events so results of the integration can be written about their experiences. Dissemination of experiences concludes the cycle of learning by contributing to the community and increases the knowledge base for technology integration, a worthwhile endeavor.

This critical literature review explored previous studies found throughout the pre-existing research. Each study gave a brief history of how they went about integrating the clicker technology into their course(s), curriculum, department, or university. There was a recurring need for some sort of support system if the educational technology was expected to be sustainable. Included in the integration was the process of deciding which of the technologies available would work best for the particular environment in which it was to be used. Discussions were held to decide desired factors and comparisons were made among the available technologies to help the reader understand the context, environment, desires of the instructional practice, and information that may help the reader to integrate the same system or repeat the research.

The goal of this literature review was to identify current practices, limitations, possible support techniques, and trends for successful implementation of clicker technology. For Example, in what situations could clickers help support student learning? When implementation was complete, things learned from the process was included labeled as 'best practices'. In order to fully discuss best practices of technology integration, the context in which the technology was used becomes the targeted factor (Caldwell, 2007; Edens, 2006; Mayer et al., 2009).

About clicker technology

There are a multitude of aspects that have been the focus of clicker research. The history of clicker usage seems to be widely varied. Some researchers go back as far as the 1960's claiming that response systems were electronic, but the availability of immediate feedback was lacking (Edens, 2006). During the expansion of this point in the research, it was found that clicker technologies are not emerging technologies, but ones that were adapting to the digital age, needs of instructors, and the expectations of the student bodies.

Practical considerations reported in the research

Clicker technologies included infrared, radio communication systems, wireless, and cellular/internet connections to transmit student selections/responses to the data collection location. Regularly, software in the collection computer compiled the responses and allowed communication back to the students. The technology discussed in many of the articles revealed that the radio frequency (the latest release of clicker technology before mobile usage) was the most likely to fit instructors' needs. However, a recent innovation was the use of mobile phone technology for student response devices, which may be changing the field. Earlier versions of the infrared clicker technology limited the number of clickers associated with each of the receivers and was not easy to integrate into classes larger than the receiver could manage (between 30-50 students) (Duncan, 2007; Menon et al., 2004). Radio frequency based technology was used in classrooms up to 1000 students (Edens, 2006) and seemed the stronger of the choices since 2006 (Caldwell, 2007; Duncan, 2007). In using personal response systems, like any other technology, it had noted glitches occurring. Errors from technology not working were a major frustration reported throughout the literature. To lower frustration, it was wise to prepare well ahead of time. Knowing possible issues and planning ahead seemed to relieve many of the frustrating factors included in using digital/wireless technology. Because problems were well documented, basic frustrations should no longer be issues for future users.

Theoretical perspectives that influenced clicker technology

Multiple theoretical perspectives, including, but not limited to, active and generative learning, adult learning, and social/peer learning, have influenced research on clicker technology. Generative learning is built on the idea that through active learning techniques (i.e., engagement and motivation), students can generate their own understanding of the content. Adult learning also implied that active learning was important for non-traditional learners' motivation and

retention. Furthermore, well-practiced clicker technology users found ways to include social/peer learning in conjunction with the technology. That kept anonymity in place for larger populations, and allowed students to converse and experience content while generating deeper understanding.

The variety of theories represented in the review of literature showed there were multiple perspectives and theories that could inform instructors when implementing educational technology. Previous studies referenced herein stated simply adopting technology would not result in positive reactions (Office of Information Technology, 2006) was confirmed in most of the studies. When the use of the technology helped guarantee active and constructively involved students; improvements were likely (Werner, 2008). However, clicker technology could be misused (Duncan, 2007). There was no research that specifically determined the impact on students in clicker-ready classrooms. Results in large lecture classes have shown that questioning, supported by clickers, positively impacted academic performances (Mayer et al., 2009). Kaleta & Joosten (2007) stated that one of their research questions focused on student retention and what it took to accomplish such. In this individual instance operational definitions differed from the other studies that used 'retention' as a cognition-based term rather than an attrition-based term. The terminology differed, as did the diversity of pedagogical approaches. The Office for Informational Technology at the University of Tennessee (2006) implied that professional development and continued support in implementing and continuing to use clicker technology was most beneficial.

The pedagogical and theoretical approaches found in the literature only scratched the surface of the key factors that made clicker technology a promising one. The research studies reported in this literature review failed to adequately conceptualize the complexity of the learning environment and learner and thus failed to predict how best to utilize clicker technology or to analyze research studies to determine or describe the underlying pattern of relationships. Each of these theoretical perspectives compels further discussion below.

Active learning and generative learning techniques

Several of the articles focused on the idea of active or interactive learning as the preferred method (Auras & Bix, 2007; Beatty, 2004; Caldwell, 2007; Fitch, 2004; Hoffman & Goodwin, 2006; Martyn, 2007; Mayer et al., 2009; Menon, et al., 2004; Office of Information Technology, 2006; Plant, 2007) for teaching large lecture courses (Auras & Bix, 2007; Caldwell, 2007; MacGeorge et al., 2008; Mayer et al., 2009; Menon, et al., 2004). Two theories that represented this line of thought were known as the active learning theory (Caldwell, 2007; Fitch, 2004; Martyn, 2007; Office of Information Technology, 2006) and the generative theory of learning (Mayer et al., 2009). Active learning was a generalized term where the responsibility for learning was given to the learners by being engaged with the learning process. Active learning was a part of clicker technology implementation in order for students to have the opportunity to become physically and/or cognitively engaged with the instruction as it was happening during the course (Mayer et al., 2009). Learning via the generative theory allows students to develop their own understanding of the content and how it applied to the larger discipline. Due to the nature of the generative theory of learning, it was nearly impossible to accomplish without active learning techniques. In more learner-centered environments, if students were actively engaged in the learning process they would gain a deeper understanding of the content from the instruction (Hoffman & Goodwin, 2006; Menon, et al., 2004). It was typical that larger lecture courses did not integrate pedagogy that engendered such reactions as active learning or generating connections between personal experiences and the content from the participants. In order for these theories to be applicable, the classroom teaching methods must have embraced interactive learning. Personal response systems such as clickers could enable the methods that would promote such interactions among students and between students and instructors.

Adult learning

Additional theories were needed to help make clicker technology meet the needs of an array of student groups. Adult learning theory suggested non-traditional students returning to the classroom would respond differently to classroom settings and technology than traditional full time undergraduate students. Adult learners, however, were also in need of active learning techniques during their coursework and may have been reluctant to participate without autonomy (Plant, 2007). Plant (2007) goes on to suggest that adult learners are in classroom settings and continuing to adapt to situations different from a work place environment. At this point in the literature review it is unclear whether they were applying the adult learning theory to full-time traditional students or only non-traditional students. Typically, non-traditional students are referred to as adult learners.

Social/peer learning

Research described students as preferring clicker technology and the anonymity it can provide because they do not wish to be judged by their peers (Caldwell, 2007; Hoffman & Goodwin, 2006; Martyn, 2007; Stowell & Nelson, 2007). However, when students were allowed to interact with a larger course anonymously, they regularly turned to their peers to discuss the content (Auras & Bix, 2007; Edens, 2006). This type of social interaction was not like the observational learning with which Bandura is associated (1977). It was a better representation of peer learning, based on the Vygotskian 'more knowledgeable other' and 'zone of proximal development' principles (1978), and better explained as a collaborative learning process among peers while discussing the topic at hand.

These actions showed a type of social or peer learning that encouraged peer-mentoring techniques. Some of the research included instructors that encouraged peer interactions as a part of the learning process (Duncan, 2007; Plant, 2007). In addition, these same active learning techniques, through peer interaction, were encouraged to support adult learners. Social/peer learning was usually brainstorming, mentoring, and learning from peers' expertise.

Developing knowledge from peers helped students further their academic careers. Those who were intrinsic learners would have a tendency to learn from other intrinsically motivated students. A similar statement could be expressed for those who were extrinsically motivated. This motivation type could imply major differences in student approaches to learning (and thus impact the teaching style as well). Edens (2006), in a comparison between behaviorally based approaches and metacognitive-oriented approaches, agrees that the intrinsically motivated students respond in different ways from the extrinsically motivated students. Depending on the goals of the course and the characteristics of the students, different teaching methods should be implemented to support the different types of learners. The research community could benefit from studies that show such interactions.

Literature review guiding issues

Guiding issues delineated the different research perspectives on clicker technology usage and the role clicker technology played when teaching different classes and topics. These guiding issues and examples of research questions identified in the literature review are listed below:

Pedagogical approaches, educational practices, and question structures influenced by vastness of the learning curve of implementation and differentiation of the disciplines. Example of research questions are:

> How does the additional benefit of using educational technology impact the higher learning outcomes (Martyn, 2007, pp.72)?

> Will clicker technology increase "…student-teacher interaction, …deeper cognitive processing during learning, …" and "…improvements on exam score in the course" (Mayer et al., 2009, pp. 53)?
>
> Does the use of clicker technology affect retention of concepts (Plant, 2007; Stowell & Nelson, 2007; Duncan, 2007)?

Influence of behavioral/cognitive perspectives. Example of research question is:

> "Does a pedagogy based on the behavioral perspective of operant conditioning or cognitive information processing view with a concentration on metacognitive influence achievement" (Edens, 2006, pp. 166)?

Individual learner differences and the effects on student perceptions. Example of research question is:

> "Do the outcomes of these approaches differ according to characteristics of particular individuals, such as gender, self-regulation levels, and goal orientation" (Edens, 2006, pp. 166)?

Clarity of expectations to help influence non-cognitive aspects. Example of research questions are:

> "How are other outcomes, such as level of anxiety, class preparation, and attendance influenced by the way SRS [Student Response System] is implemented" (Edens, 2006, pp. 166)?
>
> Does the use of clicker technology affect student satisfaction (Plant, 2007; Stowell & Nelson, 2007; Duncan, 2007)?

Practical Issues of implementation and limitations of technology. Example of research question is:

> Does the rapid adoption of clickers and the preparation to deal with practical issues of using new technology influence instructor's continued use of the technology (Office of Information Technology, 2006)?

Process used in the literature review

Finding published articles on the use of clicker technology was problematic. Many of the studies used different terminology in their research and used a multitude of different key terms when explaining the technology. With more than 25 synonymous terms (see Table 1) found, gathering enough research to come to generalizable conclusions was tedious.

Table 1
List of synonyms for 'clicker technology'.

Alternate Names	Reference
Audience Feedback Technology	MacGeorge et al., 2008
Audience Paced Feedback System	Auras & Bix, 2007; Caldwell, 2007
Audience Response System	Beatty, 2004; Caldwell, 2007; Edens, 2006; Hoffman & Goodwin, 2006; Menon et al., 2004; Plant, 2007; Stowell & Nelson, 2007

Audience Response Technology	MacGeorge et al., 2008
Classroom Communication Systems	Beatty, 2004; Caldwell, 2007
Classroom network	Caldwell, 2007
Classroom Performance System	Auras & Bix, 2007; Edens, 2006; Fitch, 2004; MacGeorge et al., 2008
Classroom Polling	Menon et al., 2004
Classroom Response System	Caldwell, 2007; Edens, 2006; Hoffman & Goodwin, 2006
Clickers	Auras & Bix, 2007; Caldwell, 2007; Edens, 2006; Hatch, Jensen, & Moore, 2005; Hoffman & Goodwin, 2006; Kaleta & Joosten, 2007; MacGeorge et al., 2008; Martyn, 2007; Mayer et al., 2009; Office of Information Technology, 2006; Smith et al., 2009; Stowell & Nelson, 2007;
Electronic Response System	Auras & Bix, 2007
Electronic Voting System	Caldwell, 2007
Group Response System	Auras & Bix, 2007; Caldwell, 2007
Handsets	Caldwell, 2007
Hyper-active Teaching Technology	Auras & Bix, 2007
Interactive Audience Response System	Auras & Bix, 2007
Interactive Engagement	Auras & Bix, 2007
Interactive Learning System	Auras & Bix, 2007
Interactive Student Response System	Auras & Bix, 2007
Key-pads	Caldwell, 2007; Fitch, 2004
Learnstar	Fitch, 2004
Peer Response System	Auras & Bix, 2007
Personal Response System	Auras & Bix, 2007; Caldwell, 2007; Edens, 2006; Fitch, 2004; Griff & Matter, 2008; Mayer et al., 2009; Office of Information Technology, 2006
Personal response Unit	Kaleta & Joosten, 2007
Student Response System	Auras & Bix, 2007; Edens, 2006; Martyn, 2007
Wireless Response System	Auras & Bix, 2007; Hatch, Jensen, & Moore, 2005
Zappers	Caldwell, 2007

The plethora of terminology proved to be a difficult obstacle; once identified, the process of searching for literature on the most widely used terms was accomplished. Next, the literature was narrowed to only those using clicker technologies for the classroom, to usability studies working with automated voting systems, and to practitioner's articles all published after 2003. Because there was a wide array of literature types found, practitioner's articles were included with the empirical research because it was a typical place of submission for educational technology users and professionals. Among the resources found, there were a multitude of businesses, educational institutions, training facilities, etc. using the technology. To narrow the scope of clicker technology users, this review contains only those studies from higher education, with the exception of one piece of research focused on usability testing of voting systems. The usability piece remained in the pool because no article from the higher education field found discussed systems usability and it was one of the goals of this literature review.

Major findings from the literature review

Current practices, trends of successful implementation, supporting student learning, and instructional support techniques are discussed throughout the major findings from the critical literature review. In addition, the limitations and technology advances that may influence current clicker technology implementation are discussed in the conclusion section.

Pedagogical approaches

Student activities determined by the course designer as interactive and/or generative learning ranged from regular participation (being attentive) to group discussions during class. The discussion approach was pedagogically equivalent to the social/peer learning theory. Peer-mentoring, in conjunction with clicker technology, placed the onus of understanding on the students and the quality and/or content of their discussions rather than on the teaching methods and had reports of positive effects on student learning (Auras & Bix, 2007; Edens, 2006).

As a part of active learning techniques, clicker technology could be included because students were paying attention or fully engaged, anticipating questions to answer. In return, students received feedback and reinforcements of learning by viewing logical or correct answers (Edens, 2006; Mayer et al., 2009). If questions were followed by logical discussion about the content via social/peer learning, the students were generating and understanding content-based schemata. Metacognitive skills were unconsciously being manifested during the discussion process. Students inherently reflected on what they were learning and if their methods were working (Edens, 2006; Caldwell, 2007; Hatch, Jensen, & Moore, 2005; Mayer et al., 2009; Office of Information Technology, 2006; Stowell & Nelson, 2007).

Unfortunately, some instructors used the technology as a tool that helps keep students from nodding off. In these instances, instructors were using the technology for quizzing rather than formative feedback and application (Beatty, 2004). This had the tendency to give students anxiety when using the technology (Caldwell, 2007; Edens, 2006). To reduce the anxiety, it was suggested that regular feedback follow the questions and the students perceived this feedback as positive reinforcement (Edens, 2006). Without best questioning techniques, students may not have received the highest benefit the technology had to offer. In addition, students did not use one standard process for learning, thus the universal design for learning might better have applied to instructors' methods than clicker technology. Students' learning methods were unique and needed varied techniques in order to learn (Menon, et al., 2004). The technology had to be flexible in order to respond to the diversity of learners. Clicker technology could be flexible in this manner, but that did not imply the technology was being used as such by offering different types of questions, comparisons, and pre-/post- understanding of concepts. In a review done by Fies and Marshall (2006), the authors discovered a lack of research on what made up the right

conditions for using clicker technology. The literature review conducted here was coming to similar conclusions.

Instructors took a multitude of pedagogical implementation approaches. Those who were only looking for basic classroom management needs were likely to use the clickers for attendance only. In response, the students found the clicker to be a waste of resources (Duncan, 2007). They wanted to know their investment was going to lead toward better academic standing as well as keeping them engaged during lectures (Caldwell, 2007; Duncan, 2007; Office of Information Technology, 2006). In order to accomplish this, it was likely the questions used during class time were stressing mental organization of content (Kaleta & Joosten, 2007; Mayer et al., 2009) by using higher order thinking and ill-defined question types (Hoffman & Goodwin, 2006; Mayer et al., 2009; Werner, 2008;). The questions posed were also integrating new knowledge with prior knowledge, stimulating student-instructor interactions (Mayer et al., 2009), and offering participation credit rather than points for the correct answers (Duncan, 2007). The question and response system helped identify times when re-teaching of the concepts were needed for understanding (Kaleta & Joosten, 2007) and experiencing questions from a specific instructor helped identify questioning characteristics for application during an exam (Mayer et al., 2009).

Educational practice

As far as educational practice was concerned, clicker technologies could be integrated into smaller classrooms, large lecture classrooms, or implemented campus wide. This literature review had articles including each of these integration types. They seem to be a flexible technology that could be used in a multitude of formative assessments depending on the goals of the class and the needs of the students and instructors (Caldwell, 2007; Office of Informational Technology, 2006). Before implementation was pursued, research was needed. The Office of Information Technology (2006) at the University of Tennessee reported their clicker integration program would not have been as successful without the importance instructors placed on the training and supportive aspects of technology integration and follow-up. The authors will continue providing such support based on feedback from workshop attendees taught in departmental groups mainly made up of instructors. Providing workshops in this manner supported different disciplines, as they were likely to approach the use of clicker technology from different pedagogical perspectives.

Learning how to streamline technology integration before an attempt was made to integrate was beneficial in both efficiency and satisfaction. If the technology could be integrated with software with which the faculty was already familiar, efficiency of integration increased and frustration decreased. Streamlining technology integration has similar affects with students. Levels of satisfaction have a tendency to decrease as the effort to use the technology increases (Conrad et al., 2009) and as the cost to students increases. The resulting satisfaction was also directly related to support systems for technology usage. The better the support system and training, the fewer reported negative impacts. In large lecture courses where pedagogical and classroom management challenges occur regularly, instructors should assume and prepare for student responsibilities to fail once in a while (Martyn, 2007; Menon et al., 2004). Solutions could include a system where instructors could check out a clicker technology classroom package. The package might include a laptop with the correct software, receiver, and back up clickers (Hoffman & Goodwin, 2006). Instructors with students who lose their clickers or their batteries die would be prepared for and would waste no class time trying to remedy the situation (Duncan, 2007). Since the software must be installed on the computer with the receiver, it would prove beneficial to any instructor who rarely teaches a large lecture class throughout their career. The system would help promote active learning as well as educational technology use. It would encourage instructors (Caldwell, 2007) to continue using more student-centered approaches in challenging teaching situations.

Questioning structures

Effective use of clicker technology was presented in several contexts. Research articles and practitioner's articles included the best practice techniques of peer interactions (Martyn, 2007; Mayer et al., 2009; Office of Informational Technology, 2006; Smith et al., 2009; Stowell & Nelson, 2007; Werner, 2008). Questions could be posed more than once, including a student interaction between the first and second posting (Kaleta & Joosten, 2007). Students would then include peer discussions in their thought processes before answering the second time (Caldwell, 2007). Instructors using this method saw more logical answers and attributed the change to a better understanding, and therefore a better application of the concepts, when they included peer interactions (Smith et al., 2009). Smith et al. (2009) recorded the number of responses after each question type and found more students answered the question types they didn't understand previously after a peer discussion. This type of higher order thinking and peer interaction planned with the use of the educational technology made for a flexible, collaborative approach and was based on constructivist models (Fitch, 2004). The generative learning theory directly supported higher order thinking by having the student create the connections between past experiences and knowledge with the content (Mayer et al., 2009) and offering credit for participation rather than correct answers (Duncan, 2007). The questions helped identify times when re-teaching was beneficial (Kaleta & Joosten, 2007) and when question types were similar to those on future exams (Mayer et al., 2009).

Question structure supported student understanding as an immediate formative assessment allowing the instructor to respond in a meaningful way. Instructors could then assess student mastery and adjust their lectures and assessments accordingly (Kaleta & Joosten, 2007; Menon et al., 2004). In addition, instructors found more positive reactions from students if they were informed of the reasons for using the educational technology (Duncan, 2007).

Navigating the learning curve of implementation

The study by the Office of Informational Technology at the University of Tennessee (2006) stated, "Standardization was important in helping departmental faculty become proficient (p.5)." Faculty members were able to mentor each other improving the speed of knowledge dissemination about clicker technology and pedagogical uses. The implementation process used here expressed a high level of satisfaction and repeated usage by instructors as well as positive reactions from students. Department or group implementation with support systems in place could and should be included on a 'best practices' list.

Looking from a curriculum and instruction perspective, many of the instructors have had limited training on active teaching methods and/or technology to support active learning techniques. The research including active learning implementation had a high percentage of single semester trials with no repetition in their study for validation. In other words, there was only one trial of using clickers throughout a semester. It was difficult to understand the level of impact on student learning when an instructor was still spending a good portion of time with the technology learning curve (Office of Informational Technology, 2006) and becoming fluent at integration without instructional interruption. The results of these pilot studies had a tendency to show clicker technology was performing miracles in regards to students' learning and their perceptions of technology use. Few articles moved past the introduction through the growing pains and concluded at a confident level of pedagogical integration and classroom implications. Research at this level was difficult to find (four of the 24 articles included more than two classes implementing the technology; Hatch et al., 2005; Kaleta & Joosten, 2007; MacGeorge et al., 2008; Office of Information Technology, 2006). The most valuable information for application was the need to communicate the plethora of 'best practices' accumulated to the academic world.

Differentiated use among disciplines

There were few studies that approached research from a content-driven, pedagogical perspective to impact student learning. This may have had to do with the fact that most of the research took place in courses of Science, Technology, Engineering and Mathematics (STEM) and human behavior rather than core academic courses (see Table 2).

Empirical evidence on the influence of particular pedagogical uses of the technology system on achievement thus is limited. Moreover, the interaction effects of other variables, such as student self-regulation and goal orientation with different types of pedagogical approach, are not known (Edens, 2006, pp.163).

Curriculum and instruction educators may be promoting the use of clickers through practice and publications. However, this promotion is likely to take place in practitioners' journals and not in research and development especially among the disciplines that are not well represented in this literature review.

On the other hand, in the variety of courses reviewed (see Table 2), the perspective and methodologies used and the conclusions from the research imply clicker technologies can be flexible, widely used in the classroom, and cost effective enough to implement in single classrooms as well as university wide.

Table 2
Courses reviewed for each piece of research.

Reference	Content Area
Auras & Bix, 2007	Packaging
Edens, 2006	Psychology
Fitch, 2004	Communication Disorders
Griff & Matter, 2008	Biology, Anatomy and Physiology
Hatch et al., 2005	Biology, Anatomy and Physiology, Environmental Science
Hoffman & Goodwin, 2006	Library orientations
Kaleta & Joosten, 2007	19 un-named disciplines
MacGeorge et al., 2008	Communications, Forestry and Natural Resources, Organizational Leadership and Supervision
Martyn, 2007	Intro Computer Information Systems
Mayer et al., 2009	Psychology
Menon et al., 2004	Medicine
Office of Information Technology (2006)	Biology, Biochemistry and Molecular Biology, Chemistry and Engineering, and Veterinary Medicine
Plant, 2007	Veterinary Dermatology
Smith et al, 2009	Introductory Genetics
Stowell & Nelson, 2007	Psychology

Pedagogical methods used in each of the different disciplines reach beyond the basic understanding of active and generative learning, social/peer learning techniques and adult learning theories. Because these differences were apparent to college instructors, instructional designers, and the like, focus on practices within the discipline in conjunction with technology usage needed to be better disseminated.

Student Perspectives

Students reported positive perceptions that clickers were enjoyable (Caldwell, 2007; Edens, 2006; Fitch, 2004; Hatch et al., 2005; Hoffman & Goodwin, 2006; MacGeorge et al., 2008; Martyn, 2007; Office of Information Technology, 2006; Stowell & Nelson, 2007;) and engaging (Auras & Bix, 2007; Caldwell, 2007; Edens, 2006; Fitch, 2004; Hoffman & Goodwin, 2006; Kaleta & Joosten, 2007; Martyn, 2007; Office of Information Technology, 2006; Stowell & Nelson, 2007). Students were regularly quoted with the notion clickers were 'fun' to use. There are hesitations in concluding the 'fun' factor had anything to do with student retention of content, but it did imply that student perceptions were positive. There were a few instances where the students mentioned the motivation to attend class due to the educational technology reduced their desire to skip class (Kaleta & Joosten, 2007). Instructors were not sure if they should attribute this phenomenon to the use of clicker technology itself or if it was just the novelty of a new technology in the classroom (Duncan, 2007).

Either way, the students generally liked using the technology. Their preferences rested with technology that allowed them to be anonymous to the rest of the class (Caldwell, 2007; Hoffman & Goodwin, 2006; Martyn, 2007; Stowell & Nelson, 2007). Regularly recognized as a feature of large lecture courses, student desired to remain anonymous rather than interact willingly with classmates (Plant, 2007). Anonymous or not, students felt like most of the instructors used the technology to check on their level of trivial content retention and was inappropriately appreciated by students. Reports of student perceptions showed evidence of this in their feedback given during these studies. Students liked the interaction during class, but did not see how it would positively influence their learning due to the low-level questions being asked (Mayer et al., 2009). However, those instructors who used the more conceptual building question types found that students liked understanding the material and were likely to have higher grades and feel more successful overall in those courses.

Supporting learner differences by identifying at-risk students

As one of the expectations for using clicker technology, students were responsible for registering their clickers prior to class use. E. R. Griff and S. F. Matter (2008) found a relationship between grade distribution and the order in which the students registered their clickers. Self-regulation and student responsibility reached far beyond how students engaged in course content; they were also reflected in daily activities which met course requirements. Griff and Matter (2008) stated that other learning techniques regularly implemented with clicker technologies also helped success rates of at-risk students. It follows that if a technology was already in place to help identify at-risk students as early as the second week of class, educational institutions would be employing the data collected by clicker technology to assist in helping them succeed where possible.

Clarity of instructional expectations

Duncan (2007) stated "... it is essential that you explain why you are using the clickers and what you expect from the students when they use the clickers (p. 9)." Students responded well to expectations (Auras & Bix, 2007) and therefore instructor implementation expectations should be addressed and assessed. Changes (such as expressing expectations) in courses were negligible if

the clickers were not used regularly and the instructors did not put emphasis on their importance. Clickers did, however, make a much larger impact with regular use and increased significance (Duncan, 2007). Duncan (2007) found the classes who used clickers as a substantial part of their class reported an attendance rate between 80-90%. Those who did not use clickers or did not use them regularly had attendance rates between 60-70%. There must be expectations and policies allowing each party to take responsibility for their participation (Auras & Bix, 2007).

Instructors who included expectations in syllabi and class meetings were not certain students would meet defined expectations. For instance, the students in more than half the cases were expected to purchase the technology. Once they had the clicker in hand, they needed to register the clicker with the receiving software. Students were not always responsive on their end of the expectations (Griff & Matter, 2008). This regularly reflected a negative impact on their grade (Duncan, 2007). Feedback from the automatic system helped students recognize grading inconsistencies and allowed them to react. This self-regulation, responsibility, and reaction were particularly positive for those intrinsically motivated (Edens, 2006). Using expectations to promote the technology was a perfect example of instructors understanding what their student's needs were and using the technology in problem-based situations.

In conclusion

Active learning techniques in conjunction with clicker technology seemed to positively impact student learning. When using social/peer interactions, there was a positive increase (Auras & Bix, 2007; Edens, 2006) in the generative learning opportunities due to the discussions happening with peers. Adult learning supported active learning and regular feedback to the students. The theories discussed in the literature reviewed do positively impact the use of clicker technology and students' perception of its use, but is not guaranteed to improve student learning. The results on student learning were more directly impacted by instructor's clarity of expectations, educational practices, questioning strategies, and repeated use of clicker technology to reduce technical difficulties and instructor/student frustration levels than with the simple use of clickers. However, it was not clear that studies continued research on positive influences in student improvement if the instructor gained more practice in technology use over time.

When department or group clicker technology implementation occurred, instructors had additional support with colleagues as well as with local technology support systems improving integration techniques. These support systems helped the instructors feel more confident about the technology implementation and benefited from multiple attempts at using the technology. In return, the instructors with support systems were able to reduce the amount of frustrations associated with technology implementation and were likely to continue to use the technology over multiple terms. Peer-mentoring techniques worked for instructors and students alike. Engaging in content-based discussions with peers did increase the generation of order and organization, and resulted in deeper understandings. With deeper understanding came positive increases in student learning represented by increased grades. Student grades were one type of feedback as well as feedback after questions were presented during clicker technology use. In particular, Kaleta and Joosten (2007) stated that grades in the course did increase and the qualitative comments followed suit. The conclusion stated that the "…clickers may be a contributing factor to the increase of grades (p.7)." However, it is believed the technology encouraged pedagogical change to increase active learning techniques using clickers. These changes gave more immediate feedback to students learning the content, deepened their understanding, and, in turn, had a positive influence on student grades. Implementing the technology without the pedagogical change would not have resulted in the same positive impacts.

Differentiation in questioning helped support student learning. Importance was given to question structures using differentiated learning techniques. Lists of quality question types existed from

many resources including but not limited to questions encouraging higher order thinking, integrating old and new knowledge, formative assessment, and exam-like questions. On the other hand, consistent quizzing during classroom interactions had the tendency to return negative results in student perceptions.

One way to inform faculty of student needs for clarity of expectations was through training and support. Sharing knowledge about the clicker technology helped prepare instructors for known issues, troubleshooting techniques, and anticipated frustration. Depending on the type of professional development, training can enhance the pedagogical uses across multiple disciples. It is seldom "…a question of teaching teachers how to use the technology; it's about teaching them when and why to use it" (Levin-Epstein, 2000, p.1 cited in Edens, 2006). Once those issues were worked through, most students appreciated more interactive large class setting than typical large lecture environments.

Classroom environments and teaching methods differed depending on the discipline. Each approached the implementation of clicker technology in dissimilar manners because they had different approaches to pedagogy. Content driven technology choices also differed because of the discipline in which it was employed. A contributing factor may be that the majority of the studies included behavioral and hard sciences. At this time, based on the research, findings are unexpressive of whether or not clicker technology is better suited for behavioral and hard sciences rather than other disciplines. It was likely that other disciplines may have been represented in practitioner's journals. Further exploration of differences in pedagogical approaches/disciplines and their implementation of technology would be beneficial to the body of information on educational technology.

Findings from research in application of educational technology in the different disciplines could help inform students of course expectations. The literature focused on student learning from the use of clicker technology and not the differences in implementation strategies. However, the most negative feedback presented was from the technology difficulties due to unpracticed implementation or inappropriate implementation. Students have explained in qualitative data their desire to know reasons for implementations and what benefits they may receive in return (Duncan, 2007). Understanding the purpose behind the technology can drive their positive or negative reactions to the technology.

Clicker technology might be able to reach beyond impacting student learning for intrinsically motivated students by being able to assist instructors at recognizing at-risk students earlier. At-risk students usually displayed extrinsically motivated characteristics; peers and extra facilitation by instructors usually influenced extrinsically motivated students in a positive manner. If clicker technology could help identify at-risk students, instructors could help by mentoring those students and bringing them more opportunities to strengthen their learning.

When instructors did not use the clickers for deeper understanding or use them regularly, the students had a negative perception due to lack of cost effectiveness and, in turn, had negative perceptions associated with clicker usage. For positive student perceptions, it was important to understand that multiple benefits to students were necessary. Though there was little research on the impact of classroom environments and settings for clicker usage, motivation and engagement were good reasons to use technology, however, content and student learning should be the main driving factor.

All of these areas of improvement together should encourage longevity of clicker technology use. A discussion of longevity is not well covered in the literature and could be a great study for understanding why there is not much literature available, or why clicker usage may have declined.

Limitations of clicker technology

Variables of the implementation of clickers, such as teaching methods, were not taken into account in conjunction with student learning data. This comparison needed to be tested and reported (Kaleta & Joosten, 2007). In addition, just under half of the reviewed publications mentioned there were disappointments due to regular technology problems. Expected technology issues could have been avoided had instructors done research or attended appropriate clicker technology workshops ahead of implementation. With plenty of resources noting this as an issue, it could easily be avoided by short-term preparation and piloted with large-scale integration. One solution might be to practice with the system during a colleague's class time. Another resource to help avoid known complications was to use the institution's technology support or instructional design specialists. Support and technology integration specialists often are prepared for teaching best practices and error reduction techniques to assist new users of educational technology.

The latest in clicker technology

The clicker technology allowed students more interactions with the content-based questions (Menon et al., 2004) and peers' perspectives in larger classes. Newer versions of clicker technologies now have LCD screens where answers can be reviewed and have a keyboard allowing questions to be typed and projected. These technologies have not been a prominent part of the review due to their recent induction and increased cost. Students desire technologies that are easy on their wallets and increase their understanding (Kaleta & Joosten, 2007). Though the other technology may be better for the students academically, some are too expensive to promote or implement widely. In addition to being cost effective, students also want technology that is easy to register and use (Caldwell, 2007; Conrad et al., 2009; Fitch, 2004; Hatch et al., 2005; Hoffman & Goodwin, 2006; Kaleta & Joosten, 2007; MacGeorge et al., 2008). When fewer technology glitches occurred and costs were lowered, instructors were more likely to continue using clickers after studies of the technology concluded (Caldwell, 2007; Martyn, 2007; Mayer et al., 2009; Office of Information Technology, 2006; Stowell& Nelson, 2007). Instructors continued use also depended on the amount of support that was supplied with the technology. There were several cases where support included professional development and training and thus, were included with the implementation (Auras & Bix, 2007; Edens, 2006; Hoffman & Goodwin, 2007; Kaleta & Joosten, 2007; Office of Information Technology, 2006). The instructors all had positive reflections from this training the relative ease of integration of clicker technology before mobile technology was a viable option.

New mobile technologies can now be integrated into clicker technology. However, many require a higher level of support for bring-your-own-device (BYOD) solutions. These systems might be cost effective for the students since they already are likely to own the device, and, because of the variety of devices, they may cause more frustration for the support staff as well as instructors. A good quantity of literature on BYOD devices is being published now and is a hot topic in educational technology.

Future research opportunities

Future research opportunities are readily encouraged among educational technologists. As the technology changes over time, so do the expectations by users. There are several areas where more research is encouraged to broaden the understanding of clicker technology usage. For example, identifying at-risk students with clicker registration order could potentially be a huge breakthrough for students who struggle in large lecture courses (Griff & Mayer, 2008). Other classroom environments could also benefit from a deeper understanding of how the characteristics of a classroom or study body might influence clicker usage. In addition, longitudinal studies could address concerns that might appear after a year or two including the

durability of the equipment and how often the systems need updating to keep up with evolving technology.

Separating teaching techniques between disciplines would be great advice to new adopters. Almost all (90%) of the studies included courses that were either hard sciences or behavioral sciences (refer to Table 2 above). This leaves a wide area of research open to all other disciplines. Only one research article concerned itself with the idea that previous literature mainly covered the STEM courses (MacGeorge et al., 2008). On the other hand, only one of the pieces of literature reviewed included the actual content objectives covered during research (Duncan, 2007). These two aspects of the coursework and technology integration left a lot of areas of interest in need of being examined. It should be the need for understanding complicated content that drives the use of the technology to help motivate students' technology usage for deeper learning.

Last, but not least, is the change that new technologies could promote. New portable electronic devices are making it ever easier to connect to the Internet and respond to questions presented in the classroom. What does it take to keep these systems working when all students may be using different personal mobile devices - due to BYOD initiatives at institutions? Research in this area may prove the next generation of technology already exists and is taking over what was once known as clicker technology.

References

Auras, R. & Bix, L. (2007). WAKE UP! The effectiveness of a student response system in large packaging classes. *Packaging Technology and Science,* 20, 183-195.

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice Hall.

Beatty, I. (2004). Transforming Student Learning with Classroom Communication Systems. *EDUCAUSE Center for Applied Research*, 2004(3), 1-13.

Caldwell, J. E. (2007). Clickers in the Large Classroom: Current research and best-practice tips. *CBE Life Sciences in Education*, 6(1), 9-20.

Conrad, F. G., Bederson, B. B., Lewis, B., Peytcheva, E., Traugott, M. W., Hanmer, M. J., Herrnson, P. S., & Niemi, R. G. (2009). Electronic voting eliminates hanging chads but introduces new usability challenges. *International Journal of Human-Computer Studies,* 67(2009), 111-124.

Duncan, D. (2007). Clickers: A new teaching aid with exceptional promise. *Astronomy Education Review*, 1(5), 70-88.

Edens, K. M. (2006). The Interaction of Pedagogical Approach, Gender, Self-regulation, and Goal Orientation Using Student Response System Technology. *Journal of Research on Technology in Education*, 41(2), 161-177.

Fies, C. & Marshall, J. (2006). Classroom response systems: A review of the literature. *Journal of Science Education and Technology*, 15(1), 101-110.

Fitch, J.L. (2004). Student Feedback in College Classroom: A technology solution. *Educational Technology Research and Development*, 1, 71-81.

Griff, E. R. & Matter, S. F. (2008). Early identification of at-risk students using a personal response system. *British Journal of Educational Technology*, 39(6), 1124-1130.

Hatch, J., Jensen, M., & Moore, R. (2005). Manna from Heaven or "Clickers" from Hell. *Journal of College Science Teaching*, 34(7), 36-39.

Hoffman, C. & Goodwin, S. (2006). A Clicker for Your Thoughts: Technology for active learning. *New Library World*, 107(9-10), 422-433.

Kaleta, R. & Joosten, T. (2007). Student Response Systems: A University of Wisconsin system study of clickers. *EDUCAUSE Center for Applied Research,* 2007(10), 1-12.

MacGeorge, E. L., Homan, S. R., Dunning, J. B. Jr., Elmore, D. Bodie, G. D., Evans, E., Khichadia, S., Lichti, S.M., Feng, B., & Geddes, B. (2008). Student evaluation of audience response technology in large lecture classes. *Education Technology Research Development*, 56, 125-145.

Martyn, M. (2007). Clickers in the Classroom: An active learning approach. *EDUCAUSE Quarterly,* (2), 71-74.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., Bugler, M., Campbell, J., Knight, A., & Zhang, H. (2009). Clickers in College Classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, 34, 51-57.

Menon, A.S., Moffett, S., Enriquez, M. Martinez, M. M., Dev, P., & Grappone, T. (2004). Audience Response made Easy: Using personal digital assistants as a classroom polling tool. *Journal of the American Medical Informatics Association,* 11(3), 217-220.

Office of Information Technology (2006). *Shifting Gears from Research to Adoption: A pilot study on the clicker personal response device*. Retrieved from: http://itc.utk.edu/classrooms/clickers/ClickerFinalReport.pdf March 14, 2009.

Plant, J.D. (2007). Incorporating and Audience Response system in to Veterinary Dermatology Lecture: Effect on student knowledge retention and satisfaction. *Journal of Veterinary Medicine Education,* 34(5), 674-676.

Smith, M.K., Wood, W.B., Adams, W.K., Wieman, C., Knight, J.K., Guild, N., & Su, T.T. (2009). Why Peer Discussion Improves Student Performance on In-Class Concept Questions. *Science*, 323, 122-124.

Stowell, J.R. & Nelson, J.M. (2007). Benefits of Electronic Audience Response Systems on Student Participation, Learning, and Emotion. *Teaching of Psychology*, 34(4), 253-258.

Werner, M. (2008). Audience Response Systems in Higher Education: Applications and cases. *EDUCAUSE Quarterly*, (2), 83-84.

Vygotsky, L.S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

About the Author

Karly Good (kgood@grandview.edu) has been working in instructional technology since 2003. She is currently the instructional technology specialist at Grand View University's Center for Excellence in Teaching and Learning. She has given presentations at international conferences on the mentoring of faculty in instructional technology, instructional design, and pedagogy for blended and online courses. Karly has a B. A. from the University of Northern Iowa in Biotechnology and an M.A.Ed. from Wake Forest University in Secondary Science Education. She is currently working on her Ph.D. in Curriculum and Instructional Technology from Iowa State University.

Return to Table of Contents

Editor's Note: Educational innovations have led to new theories of learning of which connectivism is the most practical and most comprehensive. This study positions the application of connectivism and learning theories in response to critics of new technologies for learning and teaching. It also shows how web 2.0 and social media overcome objections made by critics.

Online learning environments in higher education: Connectivism vs. Dissociation

Sasha A. Reese
USA

Abstract

Over the last decade, online education has emerged as a way for students and faculty to collaborate more freely, attain greater flexibility, and utilize new media to learn. The burning debate lies in whether online educational options are harmful to traditional education or offer endless benefits necessary to accommodate a 21st century learner. Supporters of virtual learning environments suggest that 21st century learners require the construction and creation capabilities offered through Web 2.0 to succeed, while critics suggest that asynchronous interactions are not engaging and rigorous enough for higher education. A balanced online environment should provide a blend of both asynchronous and synchronous opportunities, which promote communication and collaboration among classmates and instructors.

Keywords: distance learning, connectivism, hybrid, synchronous, asynchronous, blended learning, Web 2.0, new media, virtual learning, 21st century skills

Introduction

In an effort to modernize education, many institutes of higher education have adopted online courses in fully virtual and blended formats. Fully virtual courses allow students to interact with peers and instructors solely through the use of technology, whereas blended courses use online learning as a supplement to face-to-face interactions in live classrooms. The current educational trend toward the development of 21st century skills has made online learning attractive to both basic and tertiary education. Twenty-first century skills include, but are not limited to cultural and global awareness, self-direction, risk-taking, creativity, communication, reflection, and real world applications of knowledge (Green, 2010).

Controversy rests in whether this educational option is viable for both instructors and students. Research that supports the growth of online learning suggests that today's learners need collaboration, freedom to create knowledge, and an authentic audience in order to increase engagement, participation, and activity (Rheingold, 2010). This suggests that instructors in online environments should provide students with an experience which promotes both autonomy and community.

Conversely, researchers who are critics of online education mention the dissociative process that can accompany virtual learning environments, and a concern there is a disconnect in the instructor and student relationship and in the ability to build a learning community. Supporters of this perspective exhort that online learning must evolve from a delivery system of knowledge into a constructivist activity where learners engage in building knowledge (Hamilton, 2004, p. 843). This shift could help critics to view online schooling as a viable option.

Preparing for the future through Connectivism

Online education has emerged at universities across the country because of its ability to connect students to instructors, peers, and course content through flexible and asynchronous environments. Online learning creates a recipe of circumstances that can accommodate learners of the 21st century and their need for collaboration, creation, and construction (Green, 2010). According to Barbara Hoskin (2011), professor at Clemson University, adult and continuing educators are leading the way from traditional education to Web 2.0 enabled education, which promotes engagement and collaboration (p. 57). A need for a pedagogical shift is evident in the claim that 21st century learners also need to develop skills and competencies that will allow them to decipher and utilize information quickly and efficiently (Siemens, 2004). Current and future university students must learn how social media literacy applies to the real world, and instructors need to begin teaching students how to use this knowledge advantageously.

Although it has been hinted at throughout different eras, under the guise of Activity Theory and Social Learning Theory, Connectivism was coined as "the learning theory for the digital age" by George Siemens and Stephen Downes (2004). Siemens and Downes have developed distinct principles, which imply connectivisms' relevance to modern students. Connectivism is founded on individual ideas and opinions, valuing diversity in the perspectives of others, lifelong learning, building relationships, interdisciplinary connections, current information, and risk taking (Siemens, 2004). These same principles can be found in many current technologies that students use daily such as Facebook, Diigo, Wikis, YouTube, etc. In order for online learning to be successful and meaningful it must provide students with more than the transmission of data. New online environments must offer students the opportunity to become a participant and creator much like they are used to doing in everyday life. Research suggests that online learning should not just be a means of delivering content and course materials, but a catalyst to get students involved in course work with the support of the instructor and the involvement of peers.

Many online learning environments in higher education reflect stagnant, closed systems, and do not take advantage of the read/write web offered by Web 2.0 (Bell, 2011, p.99). Francis Bell (2011), suggests that previous learning theories based solely on traditional classrooms are not adequate models with which to reflect on digital learning, and that connectivism is a lens through which true reflection can occur. Connectivist theorists suggest that cognitive tasks between people and technology occur in networks, which are connections between "individuals, groups, systems, fields, ideas, or communities" (Siemens, 2004). Even the most basic Blackboard or WebCT learning environment, which does not deal with much more than transmitting data, utilizes discussion boards. Discussion forums fall into the connectivist realm because they create networks and connections between students and instructors.

Perhaps this is the reason for increases in online course enrollment. In 2010, approximately 30% of all higher education enrollments were online (Seaman & Allen, 2010; Moloney, 2010). Online education creates diversity in student populations because it creates opportunities for working parents, adult learners, and return students that traditional classrooms were unable to offer. The relationship between education and the national agenda is also an important part of virtual education's future. President Obama has two goals for education, which include increasing the number of college graduates to 60% by 2020 and closing the achievement gap to ensure student success in college and careers (U.S. Dept. of Ed, 2010). Online learning in higher education is also influenced by "digital natives" currently enrolled in basic education. Research purports that one in four preschool age children are learning skills through the use of the internet that some adults have not mastered. This suggests that online education is important for the success of today's students, and universities should improve programs to meet the needs of learners (Moloney, 2010, p. 66).

For online courses to be successful, instructors and students must change their roles in the learning environment and develop expectations that coincide with the tectonics of blended, traditional, and online courses. Instructors' roles often become more complex and time consuming, while students roles become more flexible and independent (Hoskins, 2011). Instructors spend a significant amount of time building courses, as well as reading and responding to emails, discussion posts, and grading papers. Contact with students is consistent and daily. However, research informs that online instructors display a lower burn-out rate because they have more time for research, the convenience of working from home, and less travel time to and from the work place (McCann, 2009).

Accessibility and cost also must be given consideration in the appeal on online education. Students who favor online schooling do so because it allows them to focus more closely on knowledge acquisition and course work and less on the logistics associated with bricks and mortar options such as parking, travel time, and fuel costs (Chau, 2010). Online schooling reduces these factors, and also has the ability to significantly decrease costs for students. Two-year and four-year universities also find online education beneficial from a fiscal perspective. Universities are able to offer more courses to an increasing population of students without needing classroom space, parking availability, or the management of the use of other facilities. In addition to these cost saving measures an increasing amount of adjunct faculty members are utilized to teach distance courses, thus saving costs associated with more highly educated and tenured faculty such as higher salary and health care costs. It is suggested that the cost efficiency benefits both the student and the university. Allen and Seaman (2008) suggest that higher education benefits students in difficult economic downturns because it provides them with a way to improve skills and increase training without the burdensome costs that can be associated with attending bricks and mortar universities.

While online courses allow students to fit college into busy work schedules, they also require students to display more self-reliance. In many circumstances students must learn content without the assistance of face-to-face instruction, and must keep track of weekly assignments through the use of virtual tools. This is seen as a benefit of online education because it adds to students' toolbox of 21st century skills. They are learning the power of clear and concise written communication, and developing the skills to collaborate with peers and instructors in a different type of environment.

One cogent factor in online learning is the effect of time on student success. Some evidence suggests that learners in online environments spend more time on task than students attending traditional courses in bricks and mortar environments (Jaschik, 2009). This can be attributed to the fact that students are able to determine when and where to complete coursework at their convenience. Both instructors and students benefit from the "anytime, anywhere" appeal of online learning (Mayadas, 2009, p. 52). They are no longer synchronously tied to specific buildings and schedules, and have the freedom of mobility and fluidity offered by the online educational experiences.

In order for online learning to be widely adopted active learning and new roles for teachers and students are necessary (Zhu, 2010, p. 147). Some research suggests that teacher perspectives and perceptions are also linked to values and culture, and that this could influence the way instructors view their roles. Cultural connections that are most compatible with online learning are individualist cultures and constructivist learning theories (Zhu, 2010, p. 148). Institutes of higher education considering adoption should spend a significant amount of time also examining the demographic of their student population to determine what type of virtual environment would meet learner's needs. Zhu suggest that certain cultures respond differently to what he refers to as the "power distance." According to Zhu, the power distance is the willingness of less powerful people in society to accept inequality. To express this, Zhu (2010) discusses that in Chinese

culture students may not be willing to participate in communicative activities with peers because it does not involve interaction with authority. Another consideration that Zhu (2010) mentions is competitiveness of some cultures such as Chinese culture, interfering with the collaborative aspect of online learning (p. 148). These concepts should be given significant consideration in the development of new online learning environments, which seek to improve the student experience.

Since online environments are prevalent in the everyday lives of learners research suggests new literacies that can easily be developed through virtual course work such as attention, participation, collaboration, network awareness, and critical consumption (Rheingold, 2010). Network awareness and critical consumption are among the most noteworthy of these literacies because of the impact that they have on student's real world activities and future career endeavors. Being able to engage with new interfaces and learn foreign technologies is imperative to success in the 21st century workforce. Critical consumption can best be defined as the ability to discover resources when needed, but also to decipher valuable information from useless information. While students come to degree programs with the ability to use social media, they often do not possess the skills necessary to use them optimally (Rheingold, 2010). An instructional strategy that has the potential to work well in a virtual setting and could improve students' critical consumption skills is merging technology and assessment. This would permit students to build learning communities where interpretation and analysis could be developed through multiple perspectives (Edwards, 2010). Collaboration and authenticity are key components in successful online courses.

Courses that offer authentic assessments, which are applicable to life after the institution have become an integral part of online learning adoption in higher education. Reductionism was common among online course programs in their infancy and continues to occur today. Courses are reduced to the bare bones of standalone tests and assignments, much like one would see in a university survey course. While many online programs still favor "reductionism," a paradigm shift is beginning to occur, which favors complex course design where authentic tasks are embedded in the curriculum (Herrington, Reeves, & Oliver, 2006, p. 233). Authentic tasks have real-world relevance to the careers that students wish to pursue after completion of college coursework, and serve as practice and preparation for professional employment (Herrington, Reeves, & Oliver, 2006, p. 236).

Online educational experiences have the potential to be interactive environments where students are able to collaborate, communicate, share, and discuss. Growing enrollment suggests that there is a need for universities to revisit current curricula and design of online courses as challenging and rigorous environments in which students can prepare for the future. National educational goals demand that students are ready for the workforce after completion of college. Virtual learning environments, whether as supplements to traditional courses or completely online, should give students opportunities to take part in authentic assessment, real life experiences with relevance, and critical development of new literacies. The paradigm shift has begun and both students and instructors should be cognizant of their changing roles as learners and teachers.

Online learning as a dissociative process

Educational technology is at the forefront of educational debates in both K-12 and higher education. A crucial aspect of the controversy surrounding how technologies are used is the suggestion that online learning environments deviate in purpose from the relationships and connections made in traditional classroom environments. Central to this argument is the concept that "educational technology participates in the cultural context and is as much a part of the learning problem to be identified as it is of the solutions implemented" (Belisle, 2001, p. 25). Saugstad (2002), when examining the work of Aristotle, suggested that knowledge is primarily seen as a product rather than a competence which suggests human activity (p. 378). This

suggestion contributes to a critical view of the relevance of current online educational systems. Separating the person from the product does not contribute to learning. It also suggests that online learning in higher education, in its current state, does not sufficiently provide students with engaging and rigorous material in preparation for professional careers. Belisle (2001) suggests that this perspective of learning is flat and focused on learning as opposed to teaching (p. 15). Hamilton (2004) refers to the delivery of knowledge style systems as fast knowledge or 'McKnowledge' (p. 844). Although the socio-cultural aspect of online education has been realized it has not come to fruition.

Online education began in the 1990s when the Internet became widely utilized (Ribsaman, 2000). Ribsaman (2000) implies that there is a difference between distance education and online education, and this difference lies in transmission and interaction. Distance education promotes the autonomy of learners, but online education actually involves students in active participation, socialization, and interaction (Ribsaman, 2000). Distance education is offered by 56% of all 2-4 year institutions, while online education is primarily offered through public institutions and large universities (Waits and Lewis, 2003; Allen & Seaman, 2006). For this reason certain aspects of online learning environments have come into question with the primary concern being the impact on students. This view returns to the concept of dissociation, in that students in fully online programs miss out on campus experiences that would connect them with instructors and students (Bejenaro, 2008, p. 411). This suggests that online atmospheres are devoid of community experiences where students are able to build relationships and construct knowledge based on differences in perspective and background.

Critics of online education argue that virtual environments are not able to provide students with the same quality and caliber of education that traditional, face-to-face courses can. They also suggested that online environments expect too much from students in terms of self-discipline and this expectation can lead new students and students unfamiliar with online education to failure (Bejenaro, 2008, p.412). In order for students to be successful in online courses a different level on support is needed from instructors. It is easy for students to become isolated and completely independent in online course without the facilitation of the instructor/ Instructors should provide students with "corrective feedback, encouragement, and motivation" to ensure success in learning (Young, 2006, p. 73).

Sub-standard work is another common concern of critics. Since there are not national standards which outline minimum knowledge and skills that students should develop, it is difficult to discern whether online institutes are diploma mills or legitimate institutions of higher education (Pina, 2010, p. 122). Although there are initiatives to establish commonality among college coursework statewide and nationwide, plans have not been finalized (Pina, 2010, p.122). Although there are many legitimately accredited online educational programs, a lack of sufficient standards, confusing accreditation processes, and research studies makes them difficult to decipher from substandard institutions (Pina 2010, p. 123).

Another debilitating perspective that surfaces in dialogue pertaining to online learning is students as receptacles for knowledge. Participation is removed from autonomous, distance education courses, in which students receive transmitted materials, interpret those materials, and create a product in isolation from peers. Sfard (1998) describes this process in a way that views the student as a consumer of knowledge, rather than a creator of knowledge (p. 5). Another concept related to the idea of consuming knowledge is "dialogue in teaching," which includes the socio-cultural backgrounds of unique learners (Bakhtin, 2011, p. 1111). Some critics would argue that communication and dialogue between students and instructors is missing from distance education, and for this reason higher education administrators should consider a paradigm shift to more participatory and active online educational systems.

Some institutions, such as the University of Phoenix, have been successful in delivering what is considered to be quality online education to a diverse population of students, among them working adults. However, admission requirements have also been examined, and can explain why enrollment in some online colleges is so prevalent. The University of Phoenix requires nothing more than a high school diploma for admittance, and is also willing to transfer work experience into course credits. While this may seem like an acceptable alternative to traditional schooling, some argue that it perpetuates conditions that are not conducive to academia (Chau, 2010, p. 179). Online education critics concede that institutions like the University of Phoenix constitute a multi-million dollar industry, but are merely diploma mills that cause damage to more reputable institutions. This claim lies in the fact that online universities sometimes have what is considered sub-standard admissions requirements, faculty that have not earned doctoral degrees, or lack consistent accreditation. While tuition at an online university may be more cost effective for both the student and instructor researchers question who is paying the price (Pina, 2010).

A theory which seeks to explain education's leniency towards educational technology is the concept of memes or 'actively contagious ideas.' Technology is a consistent presence in our lives, which is constantly changing and evolving. Memes in educational technology might include using wikis, youtube videos, social networking, and Skype in the classroom, because it is the current contagion (Lynch, 1996). Although traditional education is working as usual, technology must have its place because it is the active contagion of the twenty-first century. Contingent with this notion is the idea that future professionals will not possess the proper skills necessary for success in society because students are receiving only enough information to be productive and efficient (Lynch, 2006). This concept directly correlates with critics who see higher education's lean towards capitalism through new management regimes and privatization. Chau (2010) argues that schools should not be run or seen as businesses because ultimately the main goal of universities is to educate students and create valuable citizens (p. 183). If education does take on a completely corporate business model dissociation is inevitable because "teaching and learning will become decontextualized, simplistic, and mechanistic" (Grineski, 2000, p.22).

Faculty also suffer in the wake of online course development and creation. Instructional design and curricula building may not be worthy of promotion or tenure, but are required components of teaching online courses. Giving full attention to instruction of both face-to-face and online classes can become a stressful challenge (Chau, 2010, p. 185). A misconception that students sometimes place on instructors is the constant availability. A shift to asynchronous learning outside of bricks and mortar time constraints creates an expectation in some students that faculty should be readily available at all times (Chau, 2010, p.185). Faculty may also experience loss of ownership of the course they created if it becomes property of the institution. This phenomenon contributes to the commodity, production, and consumer schema (Chau, 2010, p. 185).

The present push towards online learning is fueled by the theory that in order for students to compete professionally and globally they must have access to digital literacies and technological advantages. However, the digital divide is not considered in this argument. While access to technology is marketed as the key to the future, consumers in lower socioeconomic circumstances are begin left out of the equation (Chau, 2010, p. 185). If online learning is designed to reach a wider audience of learners and increase accessibility, flexibility, and mobility why is digital division ignored? Another implication is the element of privilege suggested to become prevalent as "diploma mills" grow and legitimate institutions struggle to compete. Quality education may become a commodity reserved only for wealthier populations (Chau, 2010; Pina, 2010).

Online education and distance education programs are becoming prevalent in institutions of higher education. However, instructors, administrators and students should approach online learning environments with caution. Online education, by definition, suggests that students and instructors interact, collaborate, discuss, and share. It also implies that students are critically

thinking, problem solving, engaging, and participating. Unfortunately, research suggests that online education, while the wave of the future, is not the norm. Distance education, which favors consumerism, capitalism, and dissociation is common practice. By these ramifications, success in an online program does not necessarily lead to excellence in a profession or career, or sensational citizenry, but productivity and efficiency.

Conclusions

For online education to be meaningful for students and instructors universities must question the mission of the institution. Are we in the business of producing workers or educating students? Can both of these activities occur simultaneously? Most importantly, is technology an effective way to teach and learn? These questions must be considered when evaluating the validity and success of online education.

While online education lacks the substance needed for higher education, distance education has the potential to set a high standard for valuable learning experiences in virtual environments. Instructors and students can benefit from an environment that is rich in communication, collaboration, and community. Courses without these imperative components have no true educational value, and should be redesigned to engage students in meaningful learning experiences. Students should be active participants who take part in constructing knowledge and making meaning (Green, 2010).

Distance education also embraces Gardner's (1983) theory of multiple intelligences and creates a space for digital literacies to be developed (Adams, 2004). If done correctly, distance education can meet the unique needs of all students through innovative technological tools such as Elluminate, Echo 360, and advanced options in blackboard. Schools that choose to add these components to their online courses can create opportunities for synchronous and asynchronous instruction, real-time communication, and collaborative course work. Tools that enable these opportunities include live and recorded lessons by instructors, chatting abilities, journal writing and peer review, and wiki sites. All of these learning options have the potential to reach diverse populations of students.

Online education promotes learning that is one-dimensional in scope and does not have the ability to effectively provide students with what they need. It is difficult to learn without the assistance of face-to-face instruction and meeting times, and often times students do not reach out to instructors when they are struggling in an online course. The dissociative process is in full effect when students do not feel that they have a communicable relationship with their instructors (Hamilton, 2004). Unfortunately, this version of online education is prevalent because it is cost effective and promotes a corporate model which creates productive employees for the work force (Chau, 2010).

While it may seem that technology is the wave of the future, some researchers caution this shift. Before distance education is adopted in either a fully virtual or blended environment universities should first decide if this option is an effective way to teach. The term "digital native" is used frequently in research that views educational technology favorably. However, there is has been a limited amount of research conducted, which suggests that students truly are digital natives and that technology is the best way for students to learn (Lea & Jones, 2011).

An essential discussion which needs to be at the center of distance education adoption is its relationship to best practice. Although twenty-first century students might be savvy in the technical uses of smart phones and tablets, are they skilled in using these tools for learning? It seems that technology in education is utilized because of its ability to lower costs for students and institutions, and to increase accessibility, but the conversation on efficacy remains unspoken.

References

Adams, N. B. (2004). Digital intelligence fostered by technology. *Journal of Technology Studies*, 30(2), 93-97.

Allen, I. E., & Seaman, J. (2006). *Making the grade: Online education in the United States*. Needham, MA: Sloan Consortium.

Allen, I. E. & Seaman, J. (2008). Annual report: The power of learning. Retrieved from http://www.apollogrp.edu/Annual-Reports/2008%20Annual%20Report.pdf.

Bakhtin, M. (1981). *The dialogic imagination*. Austin, TX: University of Texas Press.

Belisle, C., Rawlings, A. & Van Seventer, C. (2001) T*he educational multimedia taskforce 1995-2001: Integrated research effort on multimedia education and training.* Luxembourg: European Commission. Retrieved from: http://www.cordis.lu/ist/ka3/eat/training_publ.htm.

Bell, F. (2011). Connectivism: Its place in theory-informed research and innovation in technology-enabled learning. *International Review of Research in Open and Distance Learning*, 12(3), 98-118.

Bejerano, A. R. (2008). The Genesis and Evolution of Online Degree Programs: Who are they for and what have we lost along the way?. *Communication Education*, 57(3), 408-414.

Chau, P. (2010). Online higher education commodity. *Journal of Computing in Higher Education*, 22(3), 177-191.

Edwards, S. (2010, April). Teaching through assessment: The merging of technology and assessment in teacher education learning contexts. Paper presented at ACE2010, Digital Diversity Conference, Melbourne.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Green, N. C., Edwards, H., Wolodko, B., Stewart, C., Brooks, M., & Littledyke, R. (2010). Reconceptualising higher education pedagogy in online learning. *Distance Education*, 31(3), 257-273.

Grineski, S (2000). I've a feeling we're not in Kansas anymore: The commercialization and commodification of teaching and learning in higher education. *Bulletin of Science, Technology & Society*, 20 (1), 19-28.

Hamilton, D., Dahigren, E., Hult, A., Roos, B., & Söderström, T. (2004). When performance is the product: problems in the analysis of online distance education. *British Educational Research Journal*, 30(6), 842-855.

Herrington, J., Reeves, T. C., & Oliver, R. (2006). Authentic tasks online: A synergy among learner, task, and technology. *Distance Education*, 27(2), 233-247.

Hoskins, B. (2011). Demand, growth, and evolution. *Journal of Continuing Higher Education*, 59(1), 57-60.

Jaschik, S. (2009, June). The evidence on online education. Inside Higher Ed. Retrieved from http://www.insidehighered.com/news/2009/06/29/online

Lea, M. R. & Jones, S. (2011). Digital literacies in higher education: Exploring textual and technological practice. *Studies in Higher Education,* 36(4), 377-393.

Lynch, A. (1996). Thought contagion. New York: Basic Books.

Lynch, K, (2006). Neo-liberalism and marketisation: The implication for higher education. European Educational Research Journal, 5(1), 1-17.

Mayadas, A., Bourne, J., & Bacsich, P. (2009). Online education today. *Journal of Asynchronous Learning Networks*, 13(2), 49-56.

McCann, J., & Holt, R. (2009). An exploration of burnout among online university professors. *Journal of Distance Education*, 23(3), 97-110.

Moloney, J. F., & Oakley, B. (2010). Scaling online education: Increasing access to higher education. *Journal of Asynchronous Learning Networks*, 14(1), 55-70.

Pina, A. A. (2010). Online Diploma Mills: Implications for Legitimate Distance Education. *Distance Education*, 31(1), 121-126.

Ribsaman, M. (2000). What is distance education? Defining the concepts and terms which have characterized the field. Retrieved from http://www.distance-educator.Com/index1a101600.phtml.

Rheingold, H. (2010, September/October). Attention and other 21st century social media literacies. *Educause Review,* 45(5), 14-24.

Saugstad, T. (2002) Educational theory and practice in an Aristotelian perspective. *Scandinavian Journal of Educational Research*, 46(4), 373-390.

Seaman, I.E. & Allen, J. (2010). Class differences, online education in the United States, 2010. Babson Park, MA: Babson Survey Research Group.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher.* 27(2), 4-13.

Siemens, G. (2004). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*. Retrieved from http://www.elearnspace.org/Articles/connectivism.htm

U.S. Department of Education Office of Educational Technology. (2010). Transforming American education--Learning powered by technology. Retrieved from http://www.ed.gov/sites/default/files/netp2010-execsumm.pdf.

Waites, T., & Lewis, L. (2003). Distance education at degree-granting postsecondary institutions 2000-2001. Washington, DC: U.S. Department of Education, National Center for Educational Statistics.

Young, S. (2006). Student views of effective online teaching in higher education. *American Journal of Distance Education*, 20(2), 65-77.

Zhu, C., Valcke, M., & Schellens, T. (2010). A cross-cultural study of teacher perspectives on teacher roles and adoption of online collaborative learning in higher education. *European Journal of Teacher Education*, 33(2), 147-165.

About the Author

Sasha Reese is a doctoral candidate in Indiana University of Pennsylvania's Curriculum and Instruction program. She is a Learning Support Teacher and Student Support Team Leader at Commonwealth Connections Academy, a cyber charter-school. Sasha serves as an adjunct English distance education for Burlington County College in New Jersey. Her research interests include technology in special education, inclusive learning environments, transition to careers for students with disabilities, virtual learning, literacy, and global education.

Email: s.a.connors@iup.edu

Return to Table of Contents

Return to Table of Contents

Editor's Note: Computer Assisted Language Learning (CALL) advances teaching and learning of foreign languages with a wide variety of interactive audio and visual learning experiences in easy to use media formats. Whether for learning basic vocabulary or creating rich learning environments for language teaching in context, it has become the technology of choice for teachers and learners. Thus we can move to higher levels of learning and have context and meaning integrated into the language learning experience.

The effect of CALL program on expanding lexical knowledge of Iranian EFL learners

Leila Babapour Azam
Iran

Abstract

This study was designed to explore the effects of the Computer Assisted Language Learning (CALL) program on expanding lexical knowledge of Iranian EFL intermediate learners: (a) effect of CALL program on the long-term retention in vocabulary learning, and (b) impact of this program on contextualized vocabulary learning. The study was conducted with 40 learners studying in Marefat English Language Institute in Miandoab city. From the participants, those who had access to personal computer at home were selected as the experimental group and the others served as the control group. They were called CALL group and non-CALL group respectively. Although teachers' instructional approach and the material were the same for both groups at class, the experimental group exploited computerized facilities at home to find meanings and definitions of newly taught words, but the control group used desktop dictionaries and made a bilingual list of new words to memorize them. To measure two groups' long-term retention of new lexical items, three tests were administered. Based on the mean scores and standard deviations of the groups, CALL group had an intensive mental processing which resulted in long-term recall of words. An independent samples t-test was run to compare the participants' performances on contextualized vocabulary learning. Similarly, CALL group produced better results than non-CALL group. Teachers can integrate CALL in their teaching to improve learners' lexical knowledge and give them more autonomy in language learning.

Keywords: lexical knowledge, CALL, Non-CALL, EFL learners, Intermediate Level Introduction

Introduction

Vocabulary knowledge is one of the language sub-skills crucial for fluent language use (Nation, 1993). Lexical knowledge is an indicator of how well the second language (L2) learners can perform academic language skills such as, reading, listening and writing (Treiman & Casar, 1996). According to Ellis (1997), vocabulary knowledge is a predicator of learners' discourse comprehension, which allows grammatical rules to be patterned in the learners' mind. Having inadequate vocabulary hampers learners' language comprehension and production in a way that makes it more likely the learners will face difficulties in the path of academic achievement (Al Farsi, 2008).

However, the advent of Computer-Assisted Language Learning (CALL) seems to provide a new outlook for language teaching and learning as well as vocabulary acquisition. With computer and network becoming more and more popular, numerous CALL programs and on-line materials have flooded the field of language teaching and learning, providing learners with a variety of learning activities which have been a dream even decades ago. Accordingly, various forms of technology are being integrated into the teaching and learning of L2 vocabulary, particularly CALL programs. Rahimpour (2000) pointed out that recent years have witnessed an explosion of interest in using computers for teaching and learning English as a second language. A couple of years

ago, the use of computers in language teaching and learning language was of the concern only to a small number of language teachers who were familiar with computers. But recently, computer-assisted language learning has received a great deal of attention of many English and foreign language instructors and second language acquisition (SLA) researchers.

Due to its flexibility, learners can use CALL inside and outside classroom. They can use it as drilland-practice software. They can use it as a reinforcement or remediation tool to strengthen their English lexical knowledge and competency (Almekhlafi, 2006).

Ghapanchi and Anbarestani (2008) pointed out that CALL programs provide a stimulus to which the learner must respond. The stimulus may be illustrated in any combination of text, still images, sound, and motion video. The computer analyze the learner's response to indicate whether their response is right or wrong and then offers feedback and try to describe the learner's response and to determine exactly their errors.

One of the most recent efforts researchers have been pursuing is to enhance the process of vocabulary acquisition in more effective ways. Vocabulary knowledge has long been deemed a major determinant influencing English as a second language/English as a foreign language (ESL/EFL) (Laufer, 1997). Given the close relationship between ESL/EFL learners' vocabulary command and their ability to understand English spoken and written language, researchers have been searching for ways to effectively enhance students' acquisition and retention of new vocabulary knowledge. Many techniques of direct vocabulary instruction have been examined through such attempts, including glosses (Hulstijn, 1992; Watanabe, 1997) and mnemonic devices (Beaton, Gruneberg & Ellis, 1995). Despite scholarly interests (Carter, 1998; Nation, 1994), direct vocabulary instruction appears to be inadequate to prepare learners in ESL/EFL classrooms for the basic vocabulary size demanded by most English textbooks used at EFL intermediate level in Iranian universities.

Consequently, there are various kinds of CALL programs for improving vocabulary knowledge. Computer-Assisted Vocabulary Acquisition (CAVOCA) is an example of computer programs for vocabulary acquisition; it is a program for vocabulary acquisition in a foreign language. The CAVOCA program performs that by accelerating the acquisition process; it exposes learners to carefully selected L2 materials and take them systematically through the various stages and illustrates the outstanding features of the new L2 word and/or the differences between the L2 word and its nearest L1 equivalent or counterpart (Groot, 2000).

The prominent role of vocabulary knowledge in EFL learning has been increasingly investigated. The words to be learned may be presented in isolation or in context. Learning lists of vocabularies in bilingual words seems an attractive shortcut and in comparison with contextual presentation it doesn't take a great deal of time and also causes excellent short term results (Ghabanchi & Anbarestani, 2008). It is essential that the bulk of words be learned in a short time at the intermediate and advanced stage of language achievement (Groot, 2000).

As such, vocabulary learning and teaching is a central activity in the L2 classroom. One way in which vocabulary learning can be fostered is through the use of computer-assisted language learning (AL Farsi, 2008). The computer technology, particularly CALL programs and software, have the potential to play a significant role in vocabulary teaching and learning. Therefore, this study was intended to explore this relationship and shed more light on the nature of this interconnectivity between CALL and lexical knowledge.

Literature Review

According to Kawauchi (2005), it is self-evident that vocabulary knowledge and skills are important for successful communication in a second language. Words are the units of meaning. Sentences, paragraphs, and whole texts are formed from words. People's language ability is often determined by the number of words that they know. Kawauchi (2005) pointed out that many second or foreign language learners regard learning vocabulary as an essential matter in second language acquisition, so they devote most of their time and energy to memorize lists of L2 words and their bilingual dictionary is regarded as a basic communicative resource for them.

Most theorists (e.g., Krashen, 1985; Schmitt, 1997) in the field of second language acquisition believe that formal L2 instruction is not successful that much because learners do not receive enough input in the target language (Cummins, as cited in Blake, 1999). It is argued that technology can play a significant role in increasing all L2 learners' contact with the target language, and it may provide an important way for contact with the target language (Blake, 1999).

> The role of computers in language teaching has changed significantly in the last 30 years. We have come a long way. There is still a long way a head to go. In the past, utilization of computers were limited to text and only simple simulations and exercises, primarily gap-filling and multiple-choice drills were used. Technological and pedagogical developments now allow us to more fully integrate computer technology into the language learning process (Rahimpour, 2000).

Computer-assisted language learning is defined as "the research for and study of applications of the computer in language teaching and learning" (Levy, 1997: 1). The main aim of CALL is to find ways for using computers for the purposes of teaching and learning the language (Levy, 1997). More specifically, CALL is the use of computer technologies that promote educational learning, including word processing, presentation packages, guided drill and practice, simulation and internet application such as e-mail, chat and the World Wide Web (WWW) for language learning purposes (Levy, 1997).

In the area of Computer-Assisted Vocabulary Teaching and Learning, researchers and practitioners have been trying to find out how to link CALL with vocabulary acquisition and how to better make use of CALL in vocabulary instruction. To date, plenty of research in this regard has been conducted (e.g., De Ridder, 2002; Ellis, 1995; Groot, 2000; Yoshii, 2006). Studies on computer assisted vocabulary learning have touched upon various aspects of vocabulary learning, among which a line of research is to examine the effects of electronic or online dictionary use or the effects of look-up behavior or the click behavior on word retention. For example, Knight (1994) found that students who used computerized dictionary learned more words than those who did not. In the study conducted by Laufer and Hill (2000), it was reported that using multiple dictionaries reinforced retention.

De Ridder (2002) examined visible or invisible links to investigate whether the highlighting of hyperlinks affects incidental vocabulary leaning, text comprehension, and the reading process. The result showed highlighted links is clicked more often but without affecting speed, comprehension or learning of vocabulary.

Another line of research that has received great attention and stirred great interest among researchers is to examine the effectiveness of multimedia annotations or glossaries on vocabulary acquisition. For example, Yoshii (2006) examined the effectiveness of L1 and L2 glosses on incidental vocabulary learning in a multimedia environment. The result indicated that there was no significant difference between L1 and L2 glosses for definition-supply and recognition tasks but showed significant differences between picture and no-picture glosses for definition-supply

test only. In addition, the results revealed significant interaction effects between languages and tests and the effect of additional visual cues on vocabulary learning may rely on the nature of the tasks given.

Chun and Plass (1996) investigated effects of multimedia glosses on vocabulary on 160 college students learning German as a L2. The study indicated that the combination of text and picture glosses was more effective than text-only or text-plus-video glosses.

There is also another line of research which aims to investigate the effectiveness of CALL in comparison with traditional learning or teaching methods. For instance Tozcu and Coady (2004) examined the effect of direct vocabulary learning using CALL on vocabulary knowledge, reading comprehension, and speed of word recognition. The result showed that students who used tutorial CALL to learn highly frequent vocabulary did learn a significantly larger number of words than those in the control group. Although both groups showed increases in vocabulary gain and reading comprehension and a decrease in reaction time for frequent word recognition, the treatment students indicated significantly greater gains than the control group.

Groot (2000) described a Computer-Assisted Vocabulary Acquisition program (CAVOCA). The program was contrasted in a number of experimental settings with a paired associate method of learning new words in order to establish its efficiency. The results of the study revealed that there is no best method. The efficient method of L2 word learning depends very much on variables like degree of L1-L2 equivalence of the words to be learned, the intensity of processing, the age and cognitive level of the learner and the quantity and quality of rehearsal practice.

From the above-mentioned studies we can see the application of CALL to vocabulary learning has generally revealed positive results. Although these studies only explored limited aspects of CALL and vocabulary learning, some research results are also mixed or inconclusive. They contribute to better understanding of the nature of vocabulary learning and provide us with implications when we are trying to make an informed decision concerning vocabulary teaching. Consequently, this study was designed to examine the effect of CALL program on learners' lexical knowledge. To explore the existing relationship between CALL and learners' vocabulary knowledge, the following research questions and hypotheses were addressed.

Research Questions and Hypotheses

RQ1: What is the effect of CALL program on the long-term retention in vocabulary learning?

RQ2: What is the effect of CALL program on contextualized vocabulary learning compared to ordinary method of learning vocabulary in isolation through bilingual list?

RH1: The application of CALL program will affect the learners' long-term retention of vocabulary.

RH2: The application of CALL program will affect the learners' contextualized vocabulary learning.

Methodology

Participants

This study was conducted with 40 EFL students (male and female) studying in Marefat Institute in Miandoab. They were between the ages of 17 to 26 years old. Majority of the participants in the sample were female (twenty eight out of forty). Those who had access to personal computer at home were voluntarily selected and made the experimental group of the study called CALL group (n=20), and the others who didn't have access to computer made the control group called non-CALL group (n=20).

Instrumentation

Oxford Placement Test (OPT) was used to determine the students' level. Moreover, it was administered to assess students' knowledge of the key language as well as their receptive and productive skills.

A pre-test was administered to make sure about the homogeneity of the two groups in terms of their level of proficiency. The test consisted of 30 multiple choice items which were selected from the research course book Essential Words for TOEFL.

Two CDS were distributed to CALL group to work with, using their personal computers. One of them was Oxford Talking Dictionary (OTD) which provided students with English definitions of words, in addition to some synonyms, antonyms, contextualized example, visual and pictorial presentation of words. The second CD was Oxford Genie Dictionary (OGD). This one provided students with phonetic form of pronunciation of words and an audible pronunciation by a native speaker in both British and American accents. So that students could listen and repeat the pronunciation.

In order to assess the participants' long-term retention of newly learned words, three homogeneous tests were used in the study every four sessions: The first was a multiple-choice consisting of 20 items which were selected from first fifth lessons of the book. Every item was a sentence with a word missing as a blank. The second test was given in the eighth sessions which had the same content but with different questions. In the final test (session fifteen), the participants of the study were given a cloze test with 15 items from the last five lessons. Lastly, a cloze passage consisting of 30 items was administered to measure the participants' contextualized learning of new vocabulary items. It should be asserted that all the exams were given without any previous announcement.

Pilot Study

The purpose of the pilot study was to determine item characteristics and gain some insights about the test reliability and validity. All of the teacher-made tests for the study were administered to a sample of eight students - four from CALL group and four from non-CALL group - who were similar to the main population. Based on the obtained scores, it was revealed that these tests were reliable. The following Table indicates the reliability values for the tests used in the study.

Table 1
Test Reliability Values for the Pilot Study

	Test A	Test B	Test C	Posttest
Reliability	.81	.83	.75	.78

Procedure

Prior to the research, a placement test was given in order to select the participants are at the intermediate level. They were chosen to attend the intermediate level. Every session consisted of two phases. In the first phase, the teacher worked on the participants' lexical knowledge which was intended for the research purposes. In the second phase of the session, the teacher worked on the students' conversational fluency which did not play any role in the present study. Those participants who had access to personal computer, called CALL group, served as the experimental group and the other with no access to computer, called non-CALL group, served as the control group. The homogeneity of participants was determined by running the independent sample t-test (Table 2 and 3).

Table 2
Results of Descriptive Statistics for the Proficiency Test Scores of Two Groups

	N	Minimum	Maximum	Mean	Std. Deviation	Std. Error Mean
CALL group	32	48.00	70.00	58.31	7.79	1.37
Non-CALL group	32	47.00	70.00	57.96	7.88	1.39

Table 3
Independent Samples t-test for the Proficiency Test Scores of Two Groups

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	Df	Sig (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Scores	Equal variances assumed	.001	975	.175	62	.861	.34375	1.96077	-3.57576	4.26326
	Equal variances not assumed			.175	61.99	.861	.34375	1.96077	-3.57577	4.26327

As illustrated in Table 3, the F ratio for the means of the proficiency test scores of two groups (CALL group and Non-CALL group) proved to be non-significant at the 0.05 level. This shows that there are no statistically significant differences between proficiency test scores of two groups and thus, these groups are homogenous in terms of language proficiency.

Prior to the learning sessions, a pre-test was administered to check the level of proficiency about the course material, Essential Words for TOEFL. In order to check the homogeneity of the groups in terms of their prior knowledge of vocabulary items, an independent sample t-test on the pretest scores was used.

The participants of the study were required to attend the course, two sessions a week. Totally, they met 16 sessions for eight weeks, and the teachers' instructional approach was the same for both groups. Whereas the CALL group had access to the computerized facility for finding the meaning of newly taught words, the non-CALL group used desktop dictionaries for learning new words and memorized the bilingual lists. Therefore, they were following the traditional method for improving their lexical knowledge.

The CALL group was provided with two CDs; OTD and OGD. They were required to work with these CDs to learn new vocabulary items and expand their lexical knowledge. This program had four stages: at the first stage, participants were asked to insert OTD in the computer. They typed new words and pressed the enter key so that the information related to these words appeared on the screen.

First, students read contextualized examples and tried to guess the meaning of the new word. Then they could press the picture button on the screen to access its visual presentation (if any) and to try more in guessing the meaning. At the second stage, participants could read the definitions of the words. In this stage, participants could also obtain some synonyms and antonyms. The participants were required to eject OTD and insert OGD in the CD-ROM. They should have typed the words again. In the third stage participants should have worked on the pronunciation. The phonetic form of the pronunciation was screened in front of the word. Moreover, participants could hear the pronunciation in both British and American accent, by a native speaker. They might have repeated the words they liked to improve their pronunciation. In the final stage, participants read some idioms and expressions in which the related words were used. This section was beneficial for students not only to consolidate the words but also to expand their general knowledge of learning new idioms.

In contrast, non-CALL group was required to work with the desktop dictionary. In this situation, they were provided just with written phonetic form of pronunciation and had no access to any kind of idioms, synonyms or antonyms. They just read the translation or definitions of new words and could make bilingual word lists to memorize and learn them in isolation.

Every session, the teacher asked both groups orally some questions from previous lessons. In order to measure students' long-term retention, three achievement tests were administered every four sessions. Finally, one week after the end of the treatment, the participants were given another test as a posttest. In fact, through this test the researcher assessed the retrieval through a different instrument. This posttest was a cloze passage with 30 items, which was selected carefully from the course book.

Design

The design of the study was quasi-experimental intact group design including an experimental group and a control one. The former was provided with computerized facilities (CALL programs) beside the teachers' instruction, but the latter received the same instruction without any access to computerized programs. They followed a traditional method of learning new words through desktop dictionaries and bilingual word lists. In this study, the CALL program was considered as the independent variable and learners' lexical knowledge was regarded as the dependent one.

Data Analysis

In all experiments the effect of the two methods was measured twice: immediately after the learning session and four sessions later, to determine the long-term retention effect. Participants were not told about the delayed tests to prevent them from paying more than usual attention to the words after the learning session, which might invalidate the results. In addition to written exams, every session before starting the new lesson, the teacher asked students orally some of the previously taught words and asked them to tell the definitions, some synonyms and antonyms or even Persian equivalents of particular words.

To assess the long-term retention effect, descriptive statistics (means and standard deviations) were calculated. Later, a posttest was administered to compare the two groups' improvement in terms of lexical knowledge to see whether the two different lexical learning methodologies had any differing results. Since there were two groups in the study, the result of the posttest was submitted to SPSS software and an independent t-test was employed in order to compute the significance level of t-value.

After collecting the data, CALL group's and non-CALL group's lexical knowledge was analyzed.

RQ1: What is the effect of CALL program on the long-term retention in vocabulary learning?

Table 4
Results of Descriptive Statistics for Test A

	N	Minimum	Maximum	Mean	Std. Deviation
CALL group	20	10.00	19.00	14.60	2.79
Non-CALL group	20	13.00	18.00	15.55	1.50
Valid N	20				

As the descriptive data in Table 4 shows, the participants' mean score in CALL group in test A is 14.60 but the mean score of the participants in non-CALL group is 15.55. This implies that the non-CALL group outperformed CALL group. Therefore, the CALL group's lexical knowledge did not improve in comparison to non-CALL group's vocabulary learning as a result of computerized facilities.

Table 5
Results of Descriptive Statistics for Test B

	N	Minimum	Maximum	Mean	Std. Deviation
CALL group	20	13.00	18.00	15.80	1.70
Non-CALL group	20	10.00	19.00	14.30	2.12
Valid N	20				

By a brief look at the mean scores of test B for the participants of the two groups in Table 5, we notice that there is a significant improvement in the vocabulary learning of CALL group in comparison to their performance in test A. The mean score for CALL group is 15.80 but for non-CALL group is 14.30. In fact, the participants' access to the technological apparatus and computerized facilities did result in their improvements in vocabulary learning. However, non-CALL group did not improve in the learning of new lexical items as a result of making a bilingual list of new words to memorize them. This contrasts with their long-term retention outperformance in test A.

Table 6
Results of Descriptive Statistics for Test C

	N	Minimum	Maximum	Mean	Std. Deviation
CALL group	20	16.00	20.00	17.85	1.13
Non-CALL group	20	10.00	18.00	16.20	2.16
Valid N	20				

As the descriptive statistics in Table 6 indicates, CALL group had a higher mean and lower standard deviation in comparison with non-CALL group. This implies that in this test, CALL group did better than non-CALL group. In other words, the retention in CALL group is better than retention in non-CALL group. Furthermore, the participants' use of technological apparatus not only did result in their retention of newly learned vocabulary items but also their long-term retention of new lexis did improve in comparison with their performances in tests A and B. Although non-CALL groups' retention in test C is better than the one in test B, their performance in this test indicates a decrease in comparison with the mean score in test A. Therefore, CALL participants outperformed non-CALL ones in this test, and their access to computerized facilities did result in better retention in the two last tests than retention in test A.

To sum up, it was found that learning newly taught words via traditional system has better short term results, but learning via computerized facilities is more beneficial in long-term situation, and the rate of forgetting is much lower in technological vocabulary learning. In a word, participants of CALL group had less forgetting and more retention than participants in non-CALL group; and consequently, the first question of the study was responded positively.

RQ2: What is the effect of CALL program on contextualized vocabulary learning compared to ordinary method of learning vocabulary in isolation through bilingual list?

After answering the first research question positively, an attempt was made to provide a statistical analysis in order to answer the second research question. For this to happen, the collected data from posttest, which was a cloze passage consisting of 30 items, were submitted to statistical analysis. The analysis consisted of a descriptive statistics and an independent t-test to compare the overall performances of two groups in order to see which method results in a more fruitful and longer retention of words. From the scores obtained and according to means and standard deviations of test the following result was found. In this analysis the alpha was set at .05.

Table 7
Results of Descriptive Statistics for Posttest

	N	Minimum	Maximum	Mean	Std. Deviation	Std. Error Mean
CALL group	20	11.00	20.00	16.25	2.36	.52
Non-CALL group	20	9.00	18.00	13.85	2.32	.51
Valid N	20					

As indicated in Table 7, the participants in CALL group outperformed non-CALL group's participants.

The mean score of CALL group is 16.25 but the mean score of non-CALL group is 13.85.

Accordingly, the higher the figure the better the results. Therefore, the contextualized vocabulary learning of CALL group improved significantly as a result of their access to computerized facilities.

Table 8
Independent Samples t-test for Contextualized Vocabulary Learning

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	Df	Sig (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Scores	Equal variances assumed	.008	.927	3.242	38	.002	2.40000	.74038	.90119	3.89
	Equal variances not assumed			3.242	37.991	.002	.2.40000	.74038	.90117	3.89

As shown in Table 8, the difference between the participants' contextual vocabulary` learning in two groups was significant (p=.002). This implies that in contextual situation students who use CALL programs have a better learning than those who use desktop dictionaries.

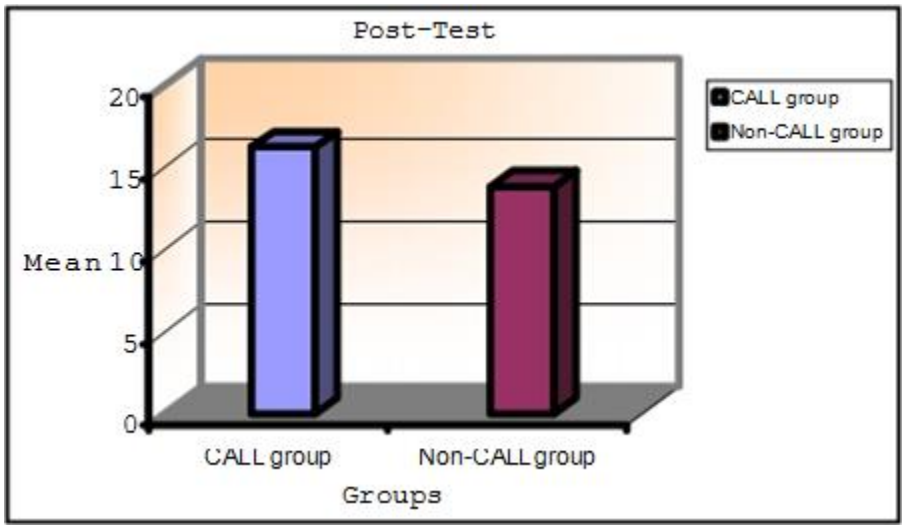

Figure 1. Comparison of means obtained in post-test by two groups

Figure 1 shows means for two participating groups on post-test. According to Figure 1, CALL group outperformed non-CALL group on post-test.

In sum, the second research question was responded positively in that CALL program does have a significant effect on contextualized vocabulary learning than ordinary method of learning vocabulary in isolation through bilingual list. It can be concluded that in contextualized situations, CALL is a better tool for learning lexicon, and it has a long-term influence which is much valuable in learning a new language.

Discussion

The findings of the study are in line with the results of Knight (1994), Laufer (2003), Yoshii (2006), Chun and Plass (1996) and Tozcu and Coady (2004). The studies argue that the application of computerized facilities to vocabulary teaching and learning does have positive results. To put it simply, through the application of CALL programs, learners' long-term retention and contextualized learning of newly taught vocabulary items are improved. However, in the study conducted by De Ridder (2002), CALL programs did not affect speed, comprehension or learning of vocabulary which is not consistent with the results of the study.

Conclusions

This study set out to explore the effect of CALL program on expanding lexical knowledge of EFL Iranian intermediate learners. The research was conducted with 40 students at intermediate level. Using computer technology and two CALL programs (OTD and OGD), the participants' performance in both CALL group and non-CALL group was measured based on the established criteria. The effect of CALL program on learners' lexical knowledge was determined by comparing the participants' performance in CALL and non-CALL group.

Based on the experiments, it was found that in both sorts of assessment those who had learned the words through CALL had less reduction of mean in delayed tests. It indicates that in using CALL program, learners have an intensive mental processing which results in long-term recall of words. By attending to the fact that in the cloze test users of CALL had better performance in both immediate and delayed tests, it can be concluded that CALL also produced better results in contextualized vocabulary learning than ordinary desktop dictionary method.

Previous research findings provide evidence supporting the effectiveness of CALL programs in improving learners' long-term retention and contextualized learning of vocabularies. The findings of the present study provided support for the close link between CALL programs as an effective method and learners' lexical expansion.

Pedagogical Implications of the study

Building on these facts and results, the following pedagogical implications can be concluded:

a) CALL can be an extremely powerful educational tool, offering individualized attention and allowing students to work at their own place.

b) CALL is considered a very useful tool for teaching EFL. It can be used as a classroom instructional tool, the effect of which was documented in research (e.g., Almekhlafi, 2006) or as an independent tool, the effect of which was documented in this study.

c) EFL teachers should be encouraged and given motivations to integrate CALL in their teaching, and hence improve their students' language proficiency.

d) Using CALL in and outside classroom can make teachers aware of individual differences in learning styles.

e) CALL can give the language teachers some changes in their roles. As certain activities are given to the students to work on at home, the teachers can act as a guide. This factor, also, may let the teachers have more free time in the classroom to work on other aspects.

Limitations and delimitations of the study

Certain limitations are imposed on this thesis. First, because of the small number of students, the number of the participants of the study was limited to 40. Second, due to the small number of the male students, majority of the students in the sample were female.

Furthermore, certain delimitations are imposed on this thesis. First, the study did not take into account different levels of proficiency. Only intermediate learners participated in the study. To examine the effects of CALL programs on lexical knowledge, various levels of proficiency should be taken into account. Second, among the language skills, lexical knowledge was investigated to explore its relationship with CALL programs. To explore the effects of these programs on language learning, different language skills (e.g., speaking and reading) can be used to see whether CALL programs have differing effects on learners' language skills.

Suggestions for Further Research

First, the present study did not take into account different levels of proficiency. Only intermediate learners participated in the study. To examine the effect of CALL programs on learners' lexical knowledge, different levels of proficiency should be included in the study.

Second, some other areas of language such as four skills are left to see whether CALL can be used in improving them, as was in vocabulary learning. In addition, finding how we can use CALL programs in learning the grammar of L2, are some unanswered questions left for further research in future.

Third, although the study included both male and female participants, no attempt was made in order to compare male and female performances in the form of independent groups. Thus, the role of gender was not considered in this study.

References

AL-Farsi, B. (2008). Morphological awareness and its relationship to vocabulary knowledge and morphological complexity among Omani EFL university students. Retrieved July 13, 2009, from http://www.asian-efl-journal.com/Thesis/Thesis-Al-Farsi.pdf/.

Almekhlafi, A. G. (2006). The effect of computer assisted language learning (CALL) on United Arab Emirates English as a foreign language (EFL) school students' achievement and attitude. Journal of Interactive Learning Research, 17(2), 14-40.

Beaton, A., Gruneberg, M., & Ellis, N. (1995). Retention of foreign vocabulary using the keywords method: A ten- year follow-up. Second Language Research, 11, 112-120.

Blake, R. J. (1999). Technology, multimedia, and second language learning. Retrieved May 28, 2008, from http://www.mhhe.com/socscience/Passport/spotlight.htm

Carter, R. (1998). Vocabulary: Applied linguistics perspectives (2nd ed.). New York: Routledge.

Chun, M. C., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. The Modern Language Journal, 80(2), 183-198.

De Ridder, I. (2002). Visible or invisible links: Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension, and the reading process? Language Learning & Technology, 6(1), 123-146.

Ellis, N. C. (1995). The psychology of foreign language vocabulary acquisition: Implications for CALL. Computer Assisted Language Learning, 8, 103-128.

Ellis, N. (1997). Vocabulary acquisition: word structure, collocation, word-class and meaning. In N. Schmitt & M. McCarthy (Eds.), Vocabulary: Description, acquisition and pedagogy (pp. 122-139). Cambridge, UK: Cambridge University Press.

Ghabanchi, Z., & Anbarestani, M. (2008). The effects of CALL program on expanding lexical knowledge of EFL Iranian intermediate learners. Retrieved September 13, 2009, from http://www.readingmatrix.com/articles/ghabanchin-anbarestani/article.pdf/.

Groot, P. J. M. (2000). Computer-assisted second language vocabulary acquisition. Language Learning and Technology, 4(1), 60-81.

Hulstjin, J. H. (1992). 'Retention of inferred and given word meanings: Experiments in incidental vocabulary learning'. In P. J. L. Arnaud & H. Bejoint (Eds.), Vocabulary and applied linguistics (pp.113-125). London: Macmillan.

Kawauchi, Ch. (2005). Proficiency differences in CALL-based vocabulary learning: The effectiveness of using "PowerWords". Retrieved August 27, 2000, from http://fleat5.byu.edu/_files/71Kawauchi.pdf

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. The Modern Language Journal, 78(3), 285-299.

Krashen, S. D. (1985). The input hypotheses: Issues and implications. New York: Longman

Laufer, B. (1997). "The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess". In J. Coady & T. Hulstjin, (Eds.), Second language vocabulary acquisition (pp. 20-34). Cambridge: Cambridge University Press.

Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. The Canadian Modern Language Review, 59(4), 567-587.

Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? Language Learning & Technology, 3(2), 58-76.

Levy, M. (1997). Computer-assisted language learning: Context and conceptualization. New York: Oxford University Press.

Nation, I. S. P. (1993). Vocabulary size, growth and use. In R. Schreuder & B. Weltens (Eds), The bilingual lexicon (pp. 115-134). Amsterdam: John Benjamins.

Nation, I. S. P. (1994). Morphology and language learning. In R. E. Asher (Ed.), The encyclopedia of language and linguistics (pp. 2582-2585). Oxford: Pergamon Press.

Rahimpour, M. (2000). Computer assisted language learning. A paper presented at FLEAT IV conference on foreign language education & technology. Kobe, Japan.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt, & M. McCarthy (Eds.), Vocabulary: Description, acquisition and pedagogy (pp. 199-228). Cambridge: Cambridge University Press.

Tozcu, A., & Coady, J. (2004). Successful learning of frequent vocabulary through CALL also benefits reading comprehension and speed. Computer Assisted Language Learning 17(5), 473-495. Treiman, R., & Casar, M. (1996). Effects of morphology on children's spelling offinal consonant clusters. Journal of Experimental Child Psychology, 63, 141-170.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. Studies in Second Language Acquisition, 19, 287-308.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. Language Learning & Technology, 10(3), 85-101.

About the Author

Leila Babapur Azam graduated with an MA in TEFL from the Islamic Azad University of Tabriz . She teaches in the Department of Teaching English as a Second Language. Ardabil Branch, Islamic Azad University, Ardabil, Iran.

Email: leyli_babapour87@yahoo.com

Return to Table of Contents

Editor's Note: The research question: "What is the most effective method of learning?" requires studies to be replicated in a variety of circumstances. Here is a study that could motivate a matrix of related studies to determine effectiveness for different grade levels and student characteristics.

Web-based instruction vs. text-based instruction and second language learners' grammar

Malahat Yousefzadeh
Iran

Abstract

This study investigated the effects of web-based instruction and text-based instruction with multiple-choice questions on English simple present tense. The purpose of this study is to assess the potential of web-based instruction for improving second language grammar. Thus our goal is to investigate whether web-based instruction has an advantage for the learning simple present tense by Iranian secondary school learners. Seventy participants were divided into two groups: web-based group and text-based group. The pre-test was administered before treatment and showed that there is no significant difference between participants. In treatment sessions, web-based group received instruction through online learning and text- based received instruction through traditional classroom instruction by teacher. The results (pre-test and post-test) were analyzed using t-test. It indicated the superiority of web-based instruction over text-based instruction.

Introduction

Pupils at secondary schools have difficulties with learning some grammatical concepts. The main problems are highlighted by second language learners are dealing with acquisition of tenses, especially simple present tense. They need to use the present simple tense a lot in English so it's really important to understand it well. Many students have problems with the form of this tense or how to make it, when they are in secondary school. They confuse in differentiating which subject use auxiliaries do and which ones use does. In relation to this, the researcher was interested in conducting a study which was focused on improving ability to use simple present tense. To address these problems we propose to create an English language teaching and learning based on World Wide Web. A good deal of research in the field of second language acquisition (SLA) supports the idea that second language learners benefit from web-based learning. The World Wide Web is one of the most exciting pedagogical resources in use today. It can create meaningful tasks and various materials for language learners.

Yousefzadeh (2012) said nowadays the dominant language teaching/learning debates is, using technology in second language teaching/learning. Using computers and multimedia programs have recently increased in language teaching and learning. Ragan, Boyce, Redwine, Savenye, and McMichael (1993) found that, in general, multimedia instruction reduces learning time by 30% compared to traditional instruction. Ewing (2000) also believes that students find chances for improvement in Computer Assisted Language Learning (CALL) environments that are unavailable in traditional second language (L2) classrooms. Learners can receive immediate feedback about their answers and correct their errors from the system. With a variety of hyperlinked multimedia documents and computer-mediated communication (CMC) tools, the Web can support language teachers to integrate Web resources into the language classroom (Son, 2007).

What makes the Web especially exciting as a resource for language teaching and learning is its possibilities for interactivity. Online language tutorials, exercises, and tests are available to anyone who has access to the Web. Web-based materials can be updated and distributed easily

and quickly, and feedback for many activities is instantaneous. Proponents of Web-based tutorials assert that such systems offer opportunities to increase student engagement and understanding of material; thereby, students have the opportunity to complete assignments and receive immediate feedback at any time (Cheng & Swanson, 2011). Mitchell and Jolley (1999) found significant positive correlation between students who used a self-guided, Web-based tutorial and exam performance. Maddux (1996) has stated that some unique characteristics of the Web include: (a) information on the WWW can be made interactive in nature; and (b) it often makes use of multimedia, including graphics, sound, and animation.

The Web provides more effective and efficient searching tools than traditional searches in libraries, and the pages retrieved from the web are more attractive and appealing than traditional printed media. Moreover, multimedia capabilities probably make the Web more attractive to many people. Dooly (2005, p. 8) has clarified that "innovative uses of the Internet and other ITC tools provide opportunities for collaborative language projects which focus on using the language to learn the language".

The aim of the study

The goal of this study is to find out whether learning simple present tense via web-based instruction will result in better learning than learning it using text-based instruction.

Research question:

Are there any significant differences between the groups of learners due to method of instruction (web-based instruction vs. text-based instruction)?

Alternative Hypothesis:

H1: There are significant differences between web-based instruction and text-based instruction in simple present tense learning of second language learners.

Independent and Dependent Variables

Independent variable was the instruction type (web-based and text-based) and dependent variable was students' scores measured by post-test.

Method

Subject: Subjects: 90 students in Ardabil secondary school were chosen randomly. They ranged in age 13-15. All of them were female. All of the 90 participants were at elementary level, based on the results of Longman Placement Test administered by researcher (Dawson, 2005). They had two years of experiences of studying English at school. Any participants who indicated even partial knowledge of simple present tense was excluded from experiment. Only the data from those participants who demonstrated absolutely no prior knowledge of the simple present tense on this study were included. After a pretest, 20 participants were excluded from the study.

Then they were divided into two groups. The two groups, in which there were (n=35) web-based group and (n=35) text-based group were placed in two treatment conditions. Both groups had the same English teacher. This controls for teacher effects on result of study.

Procedure The English simple present tense form was the target form chosen for this study because it is a form that the participants had not previously encountered. This was important because the purpose of this study was to investigate how learners learn unfamiliar forms through web-based learning. In order to verify that the participants had no prior knowledge of the English simple present tense, a pre-test was administered prior to the treatment in the experiment. A pre-test was used to measure the subjects' knowledge in simple present tense and to find out if there were any significant differences among the groups before and after the treatment, so that any

significant differences found at the time of the post-test will be due to the effect of the treatment. 20 multiple-choice questions were dedicated for pre-test.

Each question was followed by four choices. In treatment sessions, the participants in the web-based group was taken to a language laboratory in which there were 25 computers. Since the number of participants was more than number of computers, the web-based group was divided into two groups. It takes 2 times as much time to teach as web-based as a text-based for both students and teacher in web-based group. Before treatment sessions, the students in web-based group were given information about working on web-based learning and the teacher prepared the web-based group for treatment sessions before sessions began. The students were taught how to use the materials as independent learners. They could practice exercises, check the answers against the answer key and then read the relevant teaching points for which their answers were wrong. If any students required further help she then asked the teacher for clarification. The teacher would then explain to the students individually. This study was conducted online, so students accessed to the Internet in order to complete both pretest and posttest via the web-based lesson. The web-based group received web-based instruction on simple present tense in about 20 minutes a week (four sessions). The text- based group just received a placebo treatment on this field through the traditional text-based teaching.

At the end, a multiple-choice test including 40 items (each correct answer=0.5 point) was used as a posttest to depict the two groups' performance. For web-based materials, we used:

http://www.english-4u.de/grammar1.htm and for Text-based instruction, we used: English textbook for Iranian secondary school for instruction.

Data analysis

All 70 participants were homogenous based on a pre-test that was administered before starting the study. Results obtained by participants in the post-test were compared for the web-based and text-based in order to determine each of their effects of on simple present tense learning outcomes. A t-test was run to test the alternative hypothesis. The data were the score of two groups after the two types of learning condition (web-based and text-based).

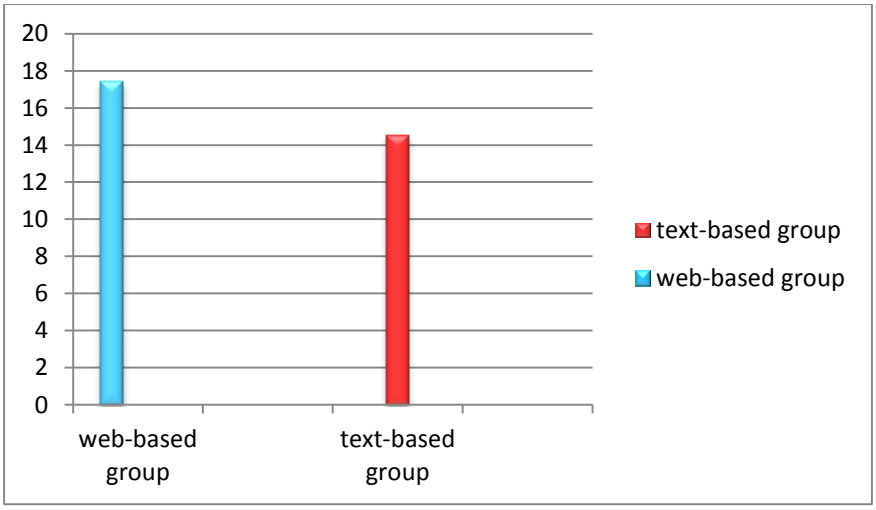

Figure1: Group's mean for post-test

The result of descriptive statistics shows that [Equation]= 17.45 and SD= 1.46 (web-based group) and [Equation]= 14.5 and SD= 2.47 (text-based group). Also the result shows that t(68)= -6.07,

p< 0.05, therefore web–based group is significantly different from text-based group, and we can support the alternative hypothesis. The results of the t-test shows significant differences between the students from the web-based group where web-assisted teaching as applied and that of the students in the text-based group where conventional teaching methods were applied on simple present tense learning.

The mean differences indicate the magnitude of the difference between the two groups.

Table 1
Means and Standard deviation obtained in post-test

	N	Mean	Std. Deviation
Web-based	35	17.45	1.46
Text-based	35	14.5	2.47

Table (1) shows group statistics. From this we can see that $\bar{x}$=17.45 and SD=1.46 (web-based group), and $\bar{x}$ =14.5 and SD=2.47 (text-based group).

Discussion and conclusion

This study was conducted in an attempt to determine whether the web-based learning method had any effect on the learning of simple present tense. At the end of treatments it was determined that there was an increase in learning levels of students in the web-based group and text-based group. However it was found the increase in the scores of the web-based group was higher than that of students in the text-based group. The result showed that web-based study support the development of simple present tense learning more effectively than traditional text-based instruction. The results of the current study are consistent with Tallmadge and Chitester (2010) who propose that Web-based tutorials assist students in compensating for insufficient background in any subject. Also this finding is in line with Nutta's study (1998) which showed significant differences in favor of the computer-based grammar instruction. It also lends support to the findings of Torlakovic and Deugo (2004).

To the contrary, Jarvis and Szymczyk (2010) found that despite the era of the digital native, our students have not, under certain circumstances, abandoned more traditional resources and it would be a mistake for practitioners and other resource providers to slavishly follow the digitalized medium route for everything. They also conclude that the tutorial CALL has a role but shows no sign of replacing paper- based materials. One of the important advantages of Web – based instruction is that it helps to learners' autonomy which is very important, because students always have positive attitudes toward learning grammar independently. Many students emphasized the fact that the websites, unlike any other material, provided them with a vast amount of activities which made grammar practice more interesting. This research should be conducted with different populations (e.g., high school students) and different types of skills.

References

Cheng, J., & Swanson, Z. (2011). An examination of the effects of Web-based tutorials on accounting student learning outcomes. Review of Higher Education and Self-Learning, 4(10), 14-28.

Dawson, N. (2005). New opportunities placement test. London: Longman.

Dooly, M. A. (2005, March/April). The Internet and language teaching: A sure way to Interculturality. The ESL Magazine: The Information Source for ESL/EFL Professionals Worldwide, 44, (8): 8.

Ewing, M. (2000). Conversations of Indonesian language students on computer-mediated projects: Linguistic responsibility and control. Computer Assisted Language Learning, 13(4), 333–356.

Jarvis, H. and Szymczyk, M. (2010). 'Student views on learning grammar with web and book-based materials.' English Language Teaching Journal. Vol. 61. Volume 64, Issue 1pp. 32-44.

Maddux, C. D. (1996). The State of the art in web-based learning. Computers in the Schools. 12 (4): 63-71.

Mitchell, M., & Jolley, J. (1999). The correlation: A self-guided tutorial. Teaching of Psychology, 26, 298-299.

Nutta, J. (1998). Is computer-based grammar instruction as effective as teacher-directed grammar instruction for teaching L2 structures? CALICO Journal, 16(1), 49-62.

Ragan, T., Boyce, M., Redwine, D., Savenye, W. C., & McMichael, J. (1993, January). Is multimedia worth it?: A review of the effectiveness of individualized multimedia instruction. Paper presented at the Association for Educational Communications and Technology Convention, New Orleans, LA.

Son, J. B. (2007). Learner experiences in Web-based language learning. Computer Assisted Language Learning, 20(1), 21-36.

Tallmadge, W., & Chitester, B. (2010). Integrating concepts using online tutorials in a freshman chemistry course. Transformative Dialogues: Teaching and Learning Journal, 4(2), 1-7.

Torlakovic, E. and Deugo, D. (2004). 'Application of a CALL system in the acquisition of adverbs in English'. Computer-Assisted Language Learning, 17 (2): 203–235.

Yousefzadeh, M. (2012).Mobile-based learning vs. Paper-based learning and collocation learning. WJEIS journal, 2(3), 216–220.

About the Author

Malahat Yousefzadeh graduated from Islamic Azad University of Ardabil with an M.A. degree in teaching English as a Foreign Language (TEFL). Her research interest is online learning.

E-mail: Yousefzadeh5351@yahoo.com

Return to Table of Contents

Return to Table of Contents

Editor's Note: The authors of this paper pose the problem of unasked questions to determine the meaning and value of student responses. Statistical validity can be enhanced by drilling deeper to find and quantify additional factors that influence student performance.

Students' perceptions of taking online physical education courses: a qualitative analysis

Lauren K. Cavanaugh and Timothy M. Sawicki

Abstract

From the inception of the online graduate physical education program at Canisius College in the fall of 2006, enrollment has grown 417% from the first year to the seventh year. Given the growth of online courses, extensive research is needed to determine best practices for online programs. The current study will qualitatively address the issue of likes and dislikes of 185 graduate students taking an online program through open ended questions analyzed using computer software called 'NVivo.' The most common likes reported by students taking online courses were "Flexibility" and that the students can continue to 'Work" while taking online classes, whereas the lack of interaction with classmates and lack of interaction with instructors were the commonly cited negative aspects of taking online classes.

Keywords: Online learning, student perceptions, graduate students, qualitative analysis, open ended questions, NVivo 10, likes and dislikes, physical education, flexibility, interaction

Introduction

The proliferation of taking online classes is increasing at a rate like no other in the history of higher education. Based on responses from more than 2,500 colleges and universities in the US, in the fall of 2010, there were 6,142,280 students enrolled in at least one online course. Nationwide, this number has increased by 560,000 students in one year and over 56% since 2006. Compared to the 1% overall growth in enrollment in the higher education student population, the growth rate of online enrollment has increased 10% (Allen & Seaman, 2011). The popularity of online enrollment is promising, with a dramatic increase in the preference of online education. From the inception of the online graduate physical education program at Canisius College in the fall of 2006, enrollment has grown 417% from the first year to the seventh year.

Greater satisfaction is also reported as students become more experienced taking online courses (Young & Norgard, 2006). Given the growth of online courses, extensive research is needed to determine best practices for online programs. The current study will address the issue of likes and dislikes of students taking an online program.

Moore and Wilson (2005) found "convenience" as a major factor influencing the students' decision to enroll in distance education. When taking face to face classes, not only is traveling to class an issue, but by taking online classes, students can use their time more efficiently while doing school work and not disrupting their social and working lives (Young & Norgard, 2006).

The positives of taking online classes are countered in the literature with negatives of having less student-instructor and student-student interaction. Student-course, student-instructor, and student-student are the various forms of interactions present while taking an online class (LaPointe & Reisetter, 2008; Yang & Durrington, 2010; Young & Norgard, 2006). Student-course interaction refers to interactions between the online students and the course content such as module question and answers. Student-instructor interaction refers to interaction between the course instructor and the students while student-student interaction refers to course requirements which have the students discuss course content with each other such as discussion board postings. In relation to

student-course, Song, Singleton, Hill, and Hwa Koh (2004) showed that course design, learner motivation, time management, and comfortableness with online technologies were the areas that greatly impacted online learning experiences. When taking online courses, students value instructor input through easy accessibility, course expertise, and caring of student progress. Lastly, interaction among students is needed for a sense of community within the online classroom. This can be a difficult task to achieve as synergy and communication between students can be lost if a course is not set up correctly (LaPointe & Reisetter, 2008).

When assessing online learning, many studies use a Likert scale to determine what students like and dislike about taking online classes. The limitations in using Likert scale surveys are that students must choose from a predetermined list of criteria. This limits their responses rather than providing a response that is more open ended to each individual. These predetermined surveys can miss important concepts related to online learning that were not included in the list of questions. With a 2-10% nationwide increase in online enrollment each year, there is a need to go beyond Likert scaled questions to determine student perceptions of online learning.

This study expands upon previous work by using open-ended questions. The students in the current study provided their own short answer responses to the questions: "What do you like about taking online classes?" and "What do you dislike most about online classes?" It is anticipated that these results will extend past research by specifically demonstrating what aspects students like and dislike about taking online classes through qualitative analysis.

Method

Participants

Participants included 185 graduate physical education students who were enrolled in the physical education Master's online program at a private institution in Western New York called Canisius College. Each student had previously completed an undergraduate physical education degree and is teacher certified. All participants in the study have taken at least one online class in the program prior to their participation.

Participants were made aware of data collection through an announcement notification during their online course. There was no penalty if they chose not to participate in answering the two questions.

Procedures

At the beginning of the *Motor Development* graduate course, participants were provided with two open ended questions: "What do you like about taking online classes?" and "What do you dislike most about online classes?" These questions were provided through their course introduction webpage on the Angel delivery system. Students voluntarily and anonymously completed their answers and submitted it to a secured drop box, where no one but the course instructor could view their results. Participants were given one week to complete these questions prior to the drop box closing.

Data were analyzed using NVivo 10 (qualitative analysis software) to determine the top five aspects students liked and disliked about taking online classes. Online research typically use Likert scales, which provide good indication on what students like and dislike about online classes, but participants have to choose from pre-existing choices.

NVivo 10 is the newest qualitative analysis software that allows researchers to complete sophisticated analysis of electronic text data contextually (Durian, 2002). NVivo 10 is the software that will be used to analyze the 185 student surveys. NVivo is a useful tool for symbolic or metaphorical trends in qualitative data down to the sentence level. NVivo searches particular

words or phrases in each paragraph and sentence and points to common trends in the qualitative data.

A graduate research student was granted access to the secured drop box to gather the data. The data, posted as a word document discussion post, was copied and pasted onto an Excel folder. Participants were given an identifying number and, on the subsequent column, a 1 or 2 indicating whether they were male or female, with 1=male and 2=female. Data were then transferred to NVivo 10 where further analysis occurred.

Data analysis

Inductive analysis was performed to analyze the open-ended short answer survey questions (Thomas, 2003). NVivo 10, a qualitative software package was used in developing the various themes and categories (QSR International, 2012) which made up the results.

Data was first gathered from electronic survey submissions and compiled into a word document. This document was uploaded to NVivo 10, where common words were searched to determine a word count for each word. These common words were then analyzed to determine like/synonym words that should be placed in the same category. For example, 'flexibility' was paired with 'freedom' because in every case they referred to the same aspect in meaning. After determining the common words, results were reviewed to establish if the meaning of the terms corresponded to the common word. For example, 'group work' was identified as a common word and in every case, 100% of the time, group work was a negative attribute.

Complete steps for analysis included:

Scanning all responses using NVivo 10 for most common words in each student response.

Viewing each common words for meaning and whether it referred to a like or dislike. For example, when communication was found as a common word, half the time it was a positive attribute and half the time it was a negative attribute.

Placing common words in categories where similar meaning was implied.

Therefore, we had categories of a common word(s) in like/dislike columns.

Interpreting the results and formulating conclusions/generalizations of the findings; and

Validating the data and results through member checking, disconfirming evidence, and extreme reviewers (Creswell & Plano Clark, 2011; Raeder, 2007).

Results and discussion

The online research from the 185 students enrolled in the physical education Master's online program at a private institution in Western New York has provided a number of likes and dislikes students have while taking online classes. It should be noted first that students were not limited to how many likes or dislikes they could list. The results showed an overwhelmingly higher number of likes regarding taking online courses than dislikes 633 to 242 (over a 2: 1 ratio). The results clearly showed when given the freedom to discuss taking online courses, students enjoy taking online courses (see Table 1.). This is consistent with other data from Cuthrell and Lyon (2007) whose students preferred independent modes of instruction due to their ease and convenience.

Table 1
Category and count of likes and dislikes of taking an online physical education Masters classes

Like	Count	Dislike	Count
Flexibility – flexible, freedom, pace, schedule(s), fit(s), convenience, family, pace	218	Lack of Classmate Interaction- relationship, meet, personal, classmates	87
Work – continue to work, coach, coaching, working, physically, busy	185	Lack of Instructor Interaction- – relationship, meet, personal, professor	68
Time – saves time, travel/sit, home, travelling/stay, own time, commute	177	Group Work – combined groups, groups, group papers	42
Communication – interaction, express ideas, communicate, conversation, can communicate with people	40	Communication – interaction, interacting, discussion, misunderstandings, talk, instructions, discussion	40
Organized – laid out, module(s)	13	Feedback – technology, professor response time	< 5
TOTAL :	633	TOTAL:	242

Responses from students were analyzed using Nivo10 and showed that students stated more likes with taking online classes than dislikes (72% to 28% respectively). The top likes of taking online courses were (in order of highest number responses): the courses allow "Flexibility" in completing assignment, students can continue to 'Work" while taking online classes, online courses save "Time", online courses promote "Communication," and online courses are "Organized" (see Table 1.). This is consistent with past literature from Young (2006) who found students appreciated the flexibility as it related to their own time management, allowing them to choose when and where to complete their online course work.

The top dislikes of taking online courses were (in order of highest number responses): students missing 'Interaction' with their peers, student missing 'Interaction' with their professors, students generally do not like 'Group Work," online courses do not promote "Communication," online courses in small numbers have issues with "Professors Response Times," and 'Technology."

Likes

The highest positive attribute to online courses is that students appreciate the 'Flexibility' that taking online courses provide. Many students stated this aspect and, in our groupings, 218 reports (1 out of every 3 students = 34% of total reports) under the category of 'Flexibility': Examples include: "The *convenience* of taking online courses is great. I like not having a commute. I also enjoy being able to work on things on my *schedule*." This student provided an example under the positive column with the common word 'Flexibility' by using the words convenience and schedule in their open ended response.

The second highest positive attribute to taking online courses is that students like the fact they can continue to 'Work,' receiving 185 likes. The following is a specific example taken from the data which illustrates this: "I like the fact that in online courses you can complete your assignments from home and it fits our *busy work* schedules. I love the freedom to work in my free

time; those with *families, jobs and coaching* know they have time for those activities." In today's tough economic times it is intuitive that an important benefit to students taking online courses would be the ability to continue their livelihood while going to school.

'Time' had the third common like count, with 177 positive reports. The students thoughts towards taking online courses is that it saves time in not having to travel to classes, saves times so they can teach and coach, and physically they do not have to leave the house to go to school. A student wrote the following: "My favorite things about taking online classes are that I can schedule the work on *my own time* and when it works best for me. And I do not have to *spend over an hour traveling* to the closest college for graduate classes in Physical Education."

Communication was shown to have 40 (6%) positive reports from students on likes of taking online courses. Positive communication centered on discussing topics with classmates and professors about assignments and content. "I really like how I can still *communicate with the professors and classmates* bounce and exchange ideas with each other." Online courses are found to allow for ideas to be shared and exchanged with classmates.

The fifth most stated like aspect of taking online courses is that the courses are 'Organized.' Thirteen students reported they liked online courses being 'organized' and they enjoyed the ease of completing modules. One student wrote the following: "I like having each *module organized* so that you can see exactly what you need to accomplish."

Dislikes

The negative aspects of taking online courses are far fewer in comparison to the likes in terms of numbers but there are several negative features that should be reported. The most prominent aspect students dislike about taking online courses is that there is a decreased "Interaction with Classmates" (87 responses = 36% of total negative responses) and decreased "Interaction with Instructor" (68 responses or 28% of total negative aspects). This is intuitive since few online courses have regular online interactions. An example of decreased 'Interaction' includes: "I prefer face-to-face classes primarily for the *interaction and discussion*, however online classes have been going great thus far." As seen by the previous statement, students miss interaction with classmates and the instructor but not enough to prevent them from enrolling in an online course (633 likes to 242 dislikes).

The third most cited dislike among online students is "Group Work." Many professors place students in groups to work together in an attempt to increase interaction between students. Unfortunately, students often cite 'Group Work' as a problem with online courses. "Without a doubt my biggest complaints about online classes have been *group work*. I feel it's tough to coordinate strong group work when you cannot sit around a table and discuss the issues and goals." Coordinating time with other online students, having equal work load for all students in the group, and communicating with other group members were reported 42 times under the heading of 'Group Work' in the dislikes of taking online courses.

Students reported the same number of negative (40) and positive (40) responses when it came to Communication. Students stated there is a 'lack of communication' while taking online courses, communication is difficult as it leads to misunderstandings of assignment requirements and discussions: "I miss discussions with classmates and asking professors questions in person because I am a people person."

A small number (less than 5 responses) of students reported technology as a dislike with taking online classes and another small number (less than 5 responses) listed professor response time as a dislike with taking online classes. "I don't like when you're doing your work and have a question and you have to wait for a response, when in the classroom you get the answer eight away."

In total 185 online Master's program students were asked the open ended question about what they like or dislike about taking online classes. Nivo10 reported 875 total student responses of positive and negative aspects of taking an online course. Students reported taking online classes as a like 633 times or 72% of total responses. Negative aspects of taking online classes had 242 responses or 28% of total responses.

Conclusion

The most common likes reported by students taking online courses were "Flexibility" and that the students can continue to 'Work" while taking online classes. This is consistent with previous research by Northrup (2002) and Young (2006), who both support flexibility as the most often cited advantage of taking an online course. The lack of interaction with classmates and lack of interaction with instructors were the commonly cited negative aspects of taking online classes. Similar to Cuthrell and Lyon (2007) students did not enjoy group work as it took longer to complete assignments and was difficult reaching their group members.

Future research

In the future, our goal is to refine the questionnaire to allow for participant demographics, and specific Likert scale questions based on the results of this research. More specifically, student demographics will include: gender, age, (whether they are single, married, kids…), technological resources, the number of online courses taken prior this point.

This survey will also be provided to faculty, both adjunct and full time faculty, to determine their perceptions on teaching in a fully online graduate program. More specifically, faculty will be provided with demographic data on gender, age, (whether they are single, married, kids…), how many online courses they teach, have taught, the technology used, how many years they have been teaching at least one course online, and what training they have to better prepare them to teach in an online environment.

These data will then be compared to determine the strengths, weaknesses, and overlap in student and instructor responses. Results will then be used to strengthen the online graduate physical education program.

References

Allen & Seaman, (2011). Going the distance: Online education in the United States, 2011. Retrieved from: http://sloanconsortium.org/publications/survey/going_distance_2011

Cresswell, J. W., & Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research* . Los Angeles: Sage.

Cuthrell, K. & Lyon, A. (2007). Instructional strategies: what do online students prefer? Retrieved April 23, 2013 from http://jolt.merlot.org/vol3no4/cuthrell

Durian, D. (2002). Corpus-based text analysis from a qualitative perspective: A closer look at NVivo. *Style, 38,* 738-742.

LaPointe, L. & Reisetter, M. (2008). Belonging online: Students' perceptions of the value and efficacy of an online learning community. *International Journal on ELearning, 7*, 641-665.

Moore, G., & Wilson, E. (2005). Perceptions of graduate students taking on-line and on-campus courses. *Journal of Agricultural Education, 46*, 23-35.

Northrup, P. T. (2002). Online learners' preferences for interaction. *The Quarterly Review of Distance Education, 3*, 219-226.

QSR International (2012). NVivo 10: Getting started. Retrieved from: http://download.qsrinternational.com/Document/NVivo10/NVivo10-Getting-Started-Guide.pdf

Raeder, S. (2007). Qualitative Methods. Retrieved from: http://peoplelearn.homestead.com/MEdHOME/QUALITATIVE/Coding.Example.pdf

Song, L., Singleton, E., Hill, J. R., & Hwa Koh, M. (2004). Improving online learning: Student perceptions of useful and challenging characteristics. *The Internet and Higher Education, 7*, 59-70.

Thomas, D. R. (2003). A general inductive approach for qualitative data analysis. Retrieved from: http://www.fmhs.auckland.ac.nz/soph/centres/hrmas/_docs/Inductive2003.pdf

Yang, Y. & Durrington, V. (2010). Investigation of students' perceptions of online course quality. *International Journal on ELearning, 9*, 341-361.

Young, A., & Norgard, C. (2006). Assessing the quality of online courses from the students' perspective. *Internet and higher education, 9*, 107-115.

Young, S. (2006). Student views of effective online teaching in higher education. *The American Journal of Distance Education, 20,* 65-77.

About the authors

Dr. Lauren K. Cavanaugh is Assistant Professor in the Department of Kinesiology at Canisius College. She teaches online and hybrid undergraduate and graduate courses. Her email is cavanau7@canisius.edu

Dr. Timothy M. Sawicki is Associate Professor in the Department of Kinesiology at Canisius College. He is the founder of the online Physical Education Masters program at Canisius College. His email is sawickit@canisius.edu

Return to Table of Contents