

Przemysław Krakowian  
University of Lodz, Lodz, Poland

# **Classroom Research and the Digital Learning Media**

## **1. Introduction**

It is commonly accepted that classroom research consists of some form of study of teaching and learning in an attempt to answer one or both of the following questions—how well students are learning and how effective teachers are in what they are doing. It often involves observation, occasionally gathering feedback from students concerning their learning and sometimes even experiments through which teachers can discover how students learn and how they respond to teaching approaches in terms of progress they make and the affect that accompanies different forms of classroom work. With the availability of the digital media, learning today may happen at least to a certain extent outside the class. Numerous educational institutions are opting for electronic learning environments as part of blended learning, with the driving idea behind such an approach that this additional exposure to learning opportunities is beneficial to students (Hollins and Robbins-Bell 2008). E-learning in various forms from programmed to collaborative learning, with the use of virtual learning environments, blogging and content management platforms, digital learning portfolios and even in its entirety in virtual environments such as Second Life® or New Victoria®, focuses on the *accumulation, organization, and delivery of content* (Gilroy 2001: para 5), as well as on *creating substitute experiences for the real thing* (Linser and Ip 2002: 606). Following the constructivist perspective, however, the significance of the digital media in learning is not so much that they provide better, faster and more pervasive access to learning resources, but rather that they create opportunities for interaction, communication and

joint thinking, and in this way foster the construction of meaning and knowledge (Garrison and Anderson 2003). One fundamental utility of the digital media, often neglected in debates on their effectiveness in teaching and learning, is the fact that students engaged in digitally mediated processes, leave a digital footprint every time they take any action. This article explores what kind of information can be derived from the digital media as an alternative to the traditional classroom research and how this information provides insight into student learning.

## **2. Virtual Learning Spaces**

Joint, shared and asynchronous access to resources via the Internet provides opportunities to dispense learning assets, materials and guidelines that have been seized from the very moment networking became available to the academia. While the early instances of creating learning spaces in the virtual environment of the Internet were, to say the least, cumbersome, they permitted access if only to those who knew what they were looking for, and most importantly, where to look for it. The cumbersome nature of access was mainly determined by the technology itself, which in the early stages of the development of the Internet consisted merely of the TCP/IP protocol, where lists of resources in the form of ftp archives resided in the root ftp directories and were available by e-mail requests. Unless the interested parties were given the exact location of the resource, identification of such assets involved trial and error process of locating materials that potentially could be of use.

The situation was somewhat improved with the advent of the Gopher protocol consisting of TCP/IP application layer meant for the distribution, searching, document retrieval over the Internet. The Gopher arrangement, the predecessor of the World Wide Web, automated the resources search process and presented the end user with summary lists of distributed assets, though still in a largely user unfriendly fashion of lengthy printouts with resource listings of sometimes very ambiguous names.

Gopher was eventually replaced by the Hypertext Transfer Protocol (HTTP) designed for distributed hypermedia information systems, which essentially consist of text rendered on a computer screen, or for that matter any other networked device, structured via references known as hyperlinks to other text in the same or a different location, which the user can access instantly, by a click of a mouse or, more commonly now, by touching the screen (Berners-Lee and Cailliau 1990). This arrangement, still very much in operation today, essentially permitted access to resources on a dispersed basis, with the traffic directed towards the resource, however; with no possibility to reciprocate (Attwell 2007).

With hypermedia, e-mail and the WWW pages, opportunities opened for the educators to create Learning Spaces, in essence arrangements of easily accessible

materials, instructions and user managed product oriented activities that lead to feedback information in the form of performed tasks available to the teaches at the completion of the process. One such arrangement was postulated by Dodge (1995), and in fact it is still used today.

Following Dodge (1995), a WebQuest is an inquiry-oriented, constructivist activity format where the learners retrieve the necessary information for the completion of their tasks from the Internet. WebQuests can be created with the use of commonly available programs, starting with word processors, to more complex and demanding software for authoring content. A WebQuest lesson consists of several distinct parts: introduction, the core task, processes involved, web resources, evaluation, and a conclusion stage concerning language reflection. While arrangements of this sort provide certain structure to the learning experience, apart from the final product there is no account, insight or monitoring of what and if anything the learners do in the virtual learning space. A natural development, therefore, of the virtual spaces was in the form of Content Management Systems (CMS) and more specifically learning oriented Learning Management Systems (LMS).

### **3. Learning Management Systems**

A Learning Management System is a scheme, an application or an application bundle for delivering, administrating, tracking, reporting, and records management of educational content, training and/or courses (Bailey 1993). Such schemes range from simple arrangements for user management and content delivery to complex course and training management provisions, with educational record-keeping and staff editing privilege registration, student and staff activity monitoring and course fee and payroll accounting. While individual configurations and components vary, most such systems will include a central administration module for user management, a content or course management module, activities module, internal communication and tracking system. What distinguishes a LMS from a Virtual Learning Environment is that it relies on the mechanisms of content deployment that are essentially those of virtual learning spaces, though augmented by management functions, while a VLE goes several steps further to ensure user interaction and collaborative learning experiences.

### **4. Virtual Learning Environments**

Virtual Learning Environments (VLE's) are a natural extension of the idea behind educational content management into the management of learner interaction, both with the learning materials and with each other that utilises Web 2.0 concepts resulting in creating reciprocal learning interaction (O'Reilly 2007). The concept of Web 2.0 was conceived in the late 1990's to denote not so much

a specific change in technology, as it is meant to denote web content that extends beyond the static web pages through the use of various mechanisms allowing the Internet users to interact with the content by selecting, probing and searching as well as collaborating with each other in a discourse that exhibits a social dimension, whose participants perform different roles and functions and act as architects of user-made content. This process is embedded in a community, which is primarily virtual in nature, though its participants may know each other from face-to-face meetings in actual classes in the real world of the school classroom. This virtual interaction remains a stark contrast to the earlier arrangement of learning spaces where the participants restrict their interactions with the content viewing and task completion. Activities of Web 2.0 nature implemented in VLE's incorporate social networking facilities and systems, user and learning blogs and videoblogs (vlogs), collaborative knowledge-sharing databases and wikis as well as Electronic Performance Support Systems (EPSS), various user-oriented web applications (webapps) including ePortfolios, mashups integrating webapps of differing provenience and folksonomies aiming at collaborative annotation and categorisation of user generated content (Peters 2009).

McAfee (2006) additionally categorises the various features and processes in Web 2.0 using the acronym SLATES, where he refers to information retrieval through keyword *search* from meaningful information ecosystems with the use of social tools that permit the creation of *links* to *authored* user content that is user augmented and updated through collaborative work of numerous individuals working together and extending beyond individual isolated web authors, where *tags* are created to manage and categorise content and *extensions* used to simplify and enable delivery of materials, of which the parties are informed through *signals* deployed via syndication technology such as RSS, mailing lists, timelines and noticeboards.

## 5. The Structure of the VLE's

Every CMS, LMS or VLE is built around the same core technology that includes four principal components comprising a web page server, a database programming and query language and an operating system, most commonly implemented in an Open Source architecture and referred to as LAMP, MAMP, WAMP or generically XAMP, where the first letter of the acronyms denotes the operating system implemented on the server—Linux, Mac OS X, or Windows respectively. The second Open Source component is responsible for displaying content in the form of web pages, and the letter A refers to the Apache web server. The M refers to the MySQL database with PHP, the P in the acronym, being the language for communicating between the web page server and the database. It is possible to use different, non-open proprietary software such as the Windows Web Server or MS Access, but most server administrators opt for more secure Open Source and Unix based systems.

Whatever the actual configuration of the core components may be, it is the database that permits operations revealing insights about the actions of the learner, no matter how trivial they may be. With every click of the mouse, with every stroke of the key on the keyboards, the database accumulates information that may be used in classroom research. What needs to be done is for the data to be displayed in a configuration that contains the information sought after.

## **6. Harvesting Data from the VLE Database**

A lot of the data present in the database is there for reasons that are determined by its structure which reflects the structure of the CMS/LMS/VLE system and the way in which information is stored in relational databases. Some of this information is displayed for the purpose of tracking student activity and progress in the form gradebooks, participation reports, summary reports and cumulative results on individual activities or groups of activities as well as for the individual student and clusters and groups of students. Most of the data that is displayed by the system as part of the VLE's grading and progress tracking functionalities can be harvested on a mass scale using the export facility or manually, when smaller student groups are concerned, through the cut and paste approach. It is conceivable that under certain circumstances, especially when particular information that is not available as a standard feature of the VLE is concerned, the researcher may be interested in creating additional code, in the form of independent queries, or as a plugin or an extension to an existing functionality that will extract the information in question. Krakowian (2013) describes how she tracks the emergence of vocabulary items in the writing done by students over the course of the semester making claims how they become part of their competence using a functionality that was created solely for that purpose, and which involves identifying and chronicling individual writing entries that were made by students on different occasions and in different tasks, activities and assignments. She also investigates interaction arrangements and networking links between the students and discovers that in activities involving collaborative work and a social format of tasks, help, assistance and collaboration is more likely to be sought from students who provide information on themselves, including a picture, a list of interest, blog entries, even if they are otherwise unknown to them and come from a different, group, class or even school. This conclusion is again possible owing to a tracking functionality written from scratch and not present in the original VLE.

A relatively straightforward modification of the structure of the database allows the system to store additional information regarding the choices made by the learners when reaching decisions in various tests that may be part of the assignments and assessment activities. While the system stores the information on the scores obtained in the activities, it disregards information that may be of

use in item performance analysis. With a straightforward modification to the code behind the activity displayed to the students, the VLE may be rendered to store that information, to later be used in item vetting, improvement and re-writing (Bachaman and Palmer 2002, Brown 1998, Krakowian 2010).

## **7. Volume of Interaction vs. Attainment Levels and Predictive Utility of Student Placement**

While some information the researcher may be seeking requires additional functionalities to be created, there is a substantial amount of information readily available in the form of logs of activity for individual students. IA Viability Report (2012) shows that there is a link between the volume of activity performed by students and attainment levels, which can be used in predictive validation of the internal placement test deployed for new students entering the language programme, which is conducted in a blended arrangement using a Moodle environment. The volume of activity is estimated using the information produced by Moodle in the form of activity logs with the VLE displaying actions undertaken by the participant in the form of hits on particular activity. A single hit may be anything from a simple click to view a resource to engaging in an activity that may altogether require complex actions. While this information is hardly a complete reflection of the whole commitment the participant is making, a simple correlation of the hits with the end of term achievement tests indicates that there is pronounced link between the two. This link additionally explains why in predictive validation of the entry placement test, which associates student placement results with attainment, the predictive utility goes beyond a simple correlation (Bachaman and Palmer 2002, Brown 1998).

## **8. WordSmith Tools in investigating student writing in VLE's**

Since a substantial portion of the information stored in the databases of various VLE's are various snippets of writing performed by the learners on different occasions and when participating in different activities, the collective texts produced by individuals over a selected period of time may be analysed looking for information. WordSmith Tools (WST) is a suite of corpus processing and analysis tools that are commonly used in researching language, mainly, though not exclusively, in lexical analysis. The suite comprises 3 main tools and collectively 11 functionalities for investigating concordances, wordlists and keywords in texts, both written and spoken. It has a long and well-documented track record of performance and has been used abundantly in research on language and learner language alike (McEnery and Wilson 2001, Botley, McEnery and Wilson 2000).

The standard features of WTS provide statistical indices concerning average word, sentence and paragraph length, all of which can be used as very rough estimates of language quality, with the assumption that with increased complexity of learner language, those indices rise as well. Apart from straightforward statistical information provided by WTS, an avenue worth investigating concerns the Type to Token ratios (TTR) in student writing, which may be used to ascertain the complexity and richness of vocabulary and how varied student expression may be. Tokens in a text are defined as individual words, and in WTS they can be equated with the number of running words in the text after numerical entries have been omitted (McEnery and Wilson 2001). Types, on the other hand refer to classes of words, where a class comprises any variant of the word or the word itself that is repeated in the text.

Tagging the text collections with information concerning their structure can reveal additional information on how learner language evolves over time. One common indicator pointed out by Biber, Johansson, Leech, Conrad and Finegan (1999) is how coordination and subordination are used to express meaning in texts. Indices involving the total number of coordinate and subordinate phrases in student writing can be easily computed by WTS, with the assumption that more advanced learner language will be characterised by a larger subordination index and lower coordination index.

## **9. Using the CPIDR: Propositional Density and Propositional Idea Density**

WTS statistical information is not an absolute measure of language quality, and conclusions drawn based solely in such indices may sometimes oversimplify the issue (Botley, McEnery and Wilson 2000). The notions of propositional density and idea density are an attractive premise in investigating language progress and language quality of both written and spoken texts and may be used alongside WTS indices. The CPIDR (Computerized Propositional Idea Density Rater, pronounced “spider”) is a computer program that allows the researcher to establish the propositional idea density of a written or transcribed spoken text without human intervention (Brown, Snodgrass, Kemper, Herman and Covington 2008). The authors of the program claim that it has been validated against human raters and the convergence is sufficiently high to lead to further applications in machine aided assessment (MAA).

Propositional density, also known as proposition density, or P-density, but sometimes referred to as propositional idea density, and understood as in Kintsch (1974) and Turner and Greene (1977), can be determined by the total number of content words such as verbs (but not auxiliaries), adjectives, adverbs, prepositions, and conjunctions against by the total number of words (Snowdon et al. 1996). In



a research study by Brown, Snodgrass, Covington, Herman, and Kemper (2008), a computer algorithm was conceived and perfected allowing the researchers to obtain accurate idea density measures. The implementation of this algorithm, the CPIDR was vetted against human raters and markers, and according to its creators, it agrees with them better than they agree with each other,  $r = 0.97$  vs.  $0.82$  respectively (Brown, Snodgrass, Kemper, Herman and Covington 2008:2).

Started by Kintsch and Keenan (1973) and Kintsch (1974), research into propositional density and idea density posits that propositions are the elements of the utterance involved in the process comprehending and recall of texts, both spoken and written. Following the Kintsch's paradigm (Kintsch and Keenan 1973; Kintsch 1974), with subsequent revisions of Turner and Greene (1977), the verb of the main clause alongside the subject, object, indirect object, and any other elements present form a single proposition. Additional descriptive elements such as modifiers in the form adjectives, adverbs which qualify the main verb, and qualifier phrases need to be seen as additional propositions.

The authors of the CPIDR program somewhat depart from Kintsch's ideas, as those differ from propositions in logic or logical semantics. The first, and probably most quantitatively important, point of departure concerns the fact that most of the information about the main verb in the main clause such as verb tense, aspect, and its modality is reduced in Kintsch's model of propositional density (Turner and Greene 1977). The second reason being that common nouns are not propositions in Kintsch's understanding of propositional density. As a result the model produces deflated measures of propositions, where paradoxically, the sentence expressing a more complex meaning would sometimes score lower on the measure of propositional density as it could express the same number of propositions in more words.

The measures in CPIDR are obtained based on the notion of idea density, which the authors of the program tend to use over the term propositional density, which is understood as the number of expressed propositions divided by the number of words. In terms of semantics, idea density constitutes a gauge of the extent to which the speaker is making claims or for that matter making requests rather than just referring to entities. Propositions here include the notions of verb tense, aspect, and its modality, as well as account for the common nouns. If the CPIDR algorithm could be integrated with any mainstream VLE, this would give the course administrator instantaneous access to information on the quality of the learner writing, making grading a less tedious process.

## 10. Conclusions

The information that can be derived from the digital media such as any arrangement of the Virtual Learning Environments or Learning Management Systems can



definitely be considered a viable alternative to traditional classroom research. As an added benefit, they offer data that are longitudinal in their nature without any added effort on the part of the researcher. While some of the data can be obtained relatively easily, without the need to construct procedures, applications or writing even a single line of code, investigating some aspects of learning in VLE's may require considerable forethought and the involvement of database management specialists. This second type of research, however, may offer insights that would not be possible to gain, or for which procuring data under normal circumstances would be at least problematic.

## References

- Attwell, G. 2007. The Personal Learning Environments—the future of eLearning? *eLearning Papers*, vol. 2 no. 1.
- Bachaman, L., Palmer, A.S. 2002. *Language Testing in Practice*. Oxford University Press.
- Bailey, G.D. 1993. Wanted: A Road Map for Understanding Integrated Learning Systems. In G.D. Bailey (Ed.), *Computer-based Integrated Learning Systems*. Englewood Cliffs.
- Berners-Lee, T., Cailliau, R. 1990. *WorldWideWeb: Proposal for a hypertexts Project*. Retrieved from <http://www.w3.org/Proposal.html>
- Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. 1999. *Longman grammar of spoken and written English*. Edinburgh Gate, Harlow, England: Pearson Education Limited.
- Botley, S., McEnery, T. and Wilson, A. 2000. *Multilingual Corpora in Teaching and Research*. Editions Rodopi.
- Brown C., Snodgrass T., Kemper S.J., Herman R. and Covington, M.A. 2008. Automatic Measurement of Propositional Idea Density from Part-of-Speech Tagging, *Behavioral Research Methods*; 40(2): 540-545.
- Brown, J.D. 1998. *Testing in Language Programs*. Prentice Hall.
- Dodge, B. 1995. WebQuests: A technique for Internet-based learning. *Distance Educator*, 1(2), 10-13.
- Garrison, R. Anderson, T. 2003. *E-learning in the 21st century: A framework for research and practice*. Routledge Falmer.
- Gilroy, K. 2001. Collaborative E-Learning: The Right Approach. *ArsDigita Systems Journal*.
- Hollins, P., Robbins-Bell, S. 2008. The Educational Affordances of Multi User Virtual Environments (MUVE). *Researching Learning In Virtual Environments* ReLIVE8: Open University Milton Keynes, 172-180.
- IA Viability Report 2012) (1) Unpublished Internal Report Paper at the Institute of English Studies at Łódź University.
- Kintsch W., Keenan J. 1973. Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*, 5: 257-274.
- Kintsch W. 1974. *The representation of meaning in memory*. Erlbaum.
- Kintsch W. 1998. *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Krakowian, K. 2013. *Constructing Meaning in Virtual Learning Environments: Towards a Cascading Model of Vocabulary Learning*. Unpublished PhD dissertation. University of Łódź.

- Krakowian, P. 2010. *Modern Test Theory Explained*. Scholar.
- Linser, R., Ip, A. 2002. Beyond the Current E-Learning paradigm: Applications of Role Play Simulations (RPS)—case studies. In *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2002*, 606-611.
- McAfee, A. 2006. Enterprise 2.0: The Dawn of Emergent Collaboration. *MIT Sloan Management review*. Vol. 47, No. 3, 21-28.
- McEnery, T., Wilson, A. 2001. *Corpus Linguistics: An Introduction*. Edinburgh University Press.
- O'Reilly, T. 2007. What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. In M. Hutter, F. Stephan, V. Vovk and T. Zeugmann (Eds.), *Design*, 65(65), 17-37.
- Peters, I. 2009. *Folksonomies. Indexing and Retrieval in Web 2.0*. De Gruyter.
- Snowdon D.A., Kemper S.J., Mortimer J.A., Greiner L.H., Wekstein D.R., Markesbery W.R. 1996. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the Nun Study. *JAMA* 528-532.
- Turner A., Greene E. 1977. *The construction and use of a propositional text base (Technical Report 63*. University of Colorado, Institute for the Study of Intellectual Behavior, Boulder.