Journal of Geography in Higher Education, Vol. 21, *No.1, 1997*

Evaluating the Effectiveness of Multimedia Computer Modules as Enrichment Exercises for Introductory Human Geography

JAMES D. PROCTOR & ANTHONY E. RICHARDSON, *University of California, USA*

ABSTRACT *Quantitative proof that multimedia enrichment activities are a positive benefit to lower-division undergraduate geography is an alluring though elusive goal. The results are presented of a careful experimental evaluation of two multimedia computer modules used as enrichment devices for an introductory human geography course at the University of California, Santa Barbara. The objectives were to determine their overall effectiveness, as well as the kinds of students and kinds of geographical knowledge and skills they best served. The rather disappointing results in respect of all three of these areas tend to corroborate one published allegation that quantitative evaluation of multimedia effectiveness is itself ineffective, due primarily to the inherent complexity of learning. The conclusion of this article, and of the study, is that an array of quantitative and qualitative evaluation methods will better serve the important objective of improving multimedia use at the university level.*

KEYWORDS Multimedia, university education, evaluation, human geography.

Background

Multimedia and Geography Education

The use of multimedia is now commonplace in education. Multimedia can be understood as "an evolving set of teaching and learning tools that, in their most sophisticated form, combine motion video images, sounds, text, and graphics in a computer-driven environment under the user's control" (Wilson & Tally, 1991, cited in National Education Association, 1994, p. 4). In the USA, multimedia has had a strong impact especially in pre-college level education, with the support of the National Education Association (1994) and organisations such as the National Educational Computing Association, which sponsors an annual symposium (National Educational Computing Association, 1994).

0309-8265/97/010041-15

To be sure, computer applications in geographic education are quite widespread (Le Gras, 1991; Unwin, 1991; Yun Hee, 1991; Antrop, 1992; Flowerdew & Lovett, 1992; Fitzpatrick, 1993; Gossette & Wheeler, 1993), especially in teaching areas such as GIS (Maguire, 1992; Raper & Green, 1992; Garner & Qiming, 1993; Merchant, 1993). Yet geography, with its central reliance on maps as visual representations of reality and its use of visually based concepts such as spatial diffusion or the landscape, appears ideally suited to adopt computer-based multimedia as an instructional aid, as a recent wave of publications has suggested (Kent, 1992; Lindholm, 1992; Carstensen *et ai., 1993;* Lueckenhoff, 1993; Foote, 1994; Svingen, 1994). One good overview of early and relatively recent uses of computers in geography education is found in Gold *et ai. (1991).*

This wave of enthusiasm for multimedia applications in geography and other disciplines has, however, not always been matched by careful evaluation of their effectiveness-a shortcoming which is especially true in higher education. As one measure, a recent search of the ERIC online database of educational publications revealed fully ten times as many citations devoted to evaluating the use of multimedia in pre-collegiate versus college settings (Educational Resources Information Center, 1995). Yet several decades of evaluation research at the university level have been performed (McLaughlin, 1973; Yeazell, 1974; McKillip & Baldwin, 1990; Mitra, 1994). The results of this research-that, for instance, multimedia may work better in conjunction with traditional methods than either alone (Young, 1972), or that multimedia enhancements to large lecture courses probably do not benefit all students equally (Garber, 1976)-have important implications for the development and application of multimedia at the college level, though courses designed on the basis of these research results (e.g. Liou, 1994) are probably in the minority.

The UCSB Geography 5 *Computer Modules*

The multimedia modules used to help teach introductory human geography at the University of California, Santa Barbara (UCSB) are part of an ambitious campus-wide programme of multimedia educational development (e.g. Fagan & Michaels, 1992; Chun & Plass, 1995; Prothero, 1995). The modules were primarily designed as supplements for Geography 5, the lower-division introductory human geography course taught at UCSB (Proctor *et ai.,* 1995). They have been developed with HyperCard software, a relatively widespread development medium for interactive multimedia in education (Ambron & Hooper, 1990). They operate in a Macintosh environment as a Macintosh lab facility exists on campus to support student work. Currently, we are weighing the merits of converting these modules to World Wide Web-accessible versions as a mode of distribution.

Our intent has been to develop a library of modules from which human geography instructors could select the most appropriate for their course. At present there are five including:

- an introduction to major conceptual approaches in human geography;
- a tutorial on the Huff extension to the gravity model;
- a discussion of errors and distortions in cognitive maps;
- an exploration of the use of models to project global population growth as a function of resource limits;
- a critical examination of the contrasting ideas of nature that arise in environmental disputes.

Sample multimedia elements include graphically enhanced sequential tutorial instructions, student-drawn electronic maps, digital movies and dynamic line charts showing how elements unfold and interact through time.

Fuller details on the two modules used in this study may serve as examples of the kinds of multimedia-based student interaction typical of the Geography 5 system. One module, entitled 'The Huff Model: Consumer Spatial Preference', introduces students to the Huff extension of the gravity model, which is applied to calculate the probability that they would patronise certain major shopping centres located in the vicinity of Santa Barbara as a function of student distance from these shopping centres, and of their 'attractiveness' to the student, defined in several possible ways. The module begins with a brief interactive tutorial in reading co-ordinate geometry information from maps and calculating Euclidean distance between two map points. It then proceeds to a step-by-step presentation of the calculations involved in determining Huff probabilities. Next, students are shown a map of five Santa Barbara retail centres as well as the location of their home, and asked to rank these centres in terms of their frequency of shopping there. Students then proceed step by step through the sequence of calculations required to arrive at Huff probabilities, with several steps being computer assisted to automate routine operations. When finished, students compare the results against their previous rankings. The final section consists of a series of discussion questions, in which students evaluate the adequacy of the Huff model and suggest improvements.

The second module, entitled 'Errors and Distortions in Cognitive Maps', introduces students to the nature of cognitive maps through a series of personal interactive experiments. The module begins with an introduction to behavioural geography followed by a description of cognitive maps and the methods used to study them. These methods are applied in a series of exercises requiring students to provide distance or direction estimates between geographical locations. For instance, students use a mouse to rotate an outline of South America to its correct orientation and align it to the appropriate position below North America. They are provided with feedback (in this case, an outline of South America in its correct position and orientation) and then presented with theory to explain their errors. The next exercise requires students to make different judgements, but this time they must apply recently learned principles to interpret their findings. The remainder of the module follows this pattern: students provide judgements, are introduced to theory and ultimately interpret novel data or scenarios based on that knowledge. Topics covered include errors of alignment and rotation, the hierarchical organisation of memory, theories of distance estimation and cognitive perspective.

The Geography 5 modules can be used in several ways. Each module can simply be completed by students as a roughly two-hour stand-alone exercise, in which they submit as a final product the printed output of their written and numerical module work. Alternatively, their module experience can be used as the basis for small-group student discussions, or for more integrative summary essays; the latter option is facilitated by means of a HyperCard-based anonymous peer review system linked to the Geography 5 modules in which students receive feedback on draft essays (Proctor *et al.,* 1995).

Of perhaps more fundamental importance, however, is the means by which the Geography 5 modules are integrated into the course. There are several viable options in this respect as well. At one extreme, the modules can be thought of as supplements, independent of lecture material. The instructor who, for example, does not have sufficient time to cover a particular topic in class could assign that module to students. At the other extreme, the modules can serve as enrichments to classroom material, giving students the opportunity to review and apply concepts introduced in lectures.

For the purposes of this study, we chose to evaluate the modules as enrichments to more difficult material presented during class lectures. In this sense, the Huff module served as an enrichment for lectures on spatial interaction (in particular, the gravity concept), and the cognitive maps module built from a general series of lectures in behavioural geography. We chose the enrichment mode of use owing to its relative ease of administration (students independently complete modules any time after attending the introductory lecture connected to them), as well as its potential in reinforcing elements of human geography included in the lecture syllabus. In the latter regard, it is always a challenge teaching introductory human geography-a course which could easily involve a year of student study-in the lO-week quarter we use. As enrichments, the multimedia modules can in theory allow instructors to move more rapidly through difficult course material during a lecture, knowing that students will get a better chance to digest this material when they encounter it in the computer module.

Objectives and Methodology

Evaluation of Effectiveness

Student feedback on the Geography 5 modules to date has provided us with a good deal of specific input that we have used to improve them, as well as general information on their perceived effectiveness (Proctor *et al.,* 1995). These surveys do little, however, in giving us some objective measure of the effectiveness of the modules as enrichment exercises. The question we are more frequently asked is one many developers of multimedia educational products hear: is multimedia really worth the effort on the part of students and instructors? Students may enjoy multimedia, and instructors may believe that their teaching effectiveness has improved because they are providing students with an additional learning resource. On the other hand, multimedia development is complex and risky, and takes a good deal of time and money.

Our intent in this study, therefore, was to perform a quantitative assessment of the effectiveness of the Geography 5 modules as enrichment exercises. Our particular interest concerned whether students learned the geography material better with than without the benefit of multimedia enrichments. This aim makes a good deal of intuitive sense to administrators who must decide whether to fund multimedia development, and to instructors deciding whether to invest their precious time in educational multimedia. If truly significant differences are found between students using multimedia enhancements and a control group who use only traditional educational approaches, this result may suggest that multimedia modules are worth the effort.

Yet this ostensibly simple and eminently practical research objective has been called into question recently by research specialists in educational evaluation. One researcher in particular, Thomas Reeves, has charged that any experimental attempt to demonstrate significant improvement in learning at the university level with-versus-without multimedia enhancements is ultimately flawed (Reeves, 1991). Reeves argues that learning is too complex a phenomenon to be reliably measured, citing as evidence a number of studies which surprisingly failed to show any statistically significant difference between students using multimedia and students following traditional educational methods. Our belief, however, was that a carefully designed experimental procedure that allows for disaggregation of different kinds of students and different kinds of knowledge and skills would result in measurable differences.

TABLE I. Student variables included in analysis.

General	Background	Module-specific	Quiz-specific
Class level	Computer familiarity	Module completed	Perceived difficulty of cognitive quiz
Gender	Mathematical ability	Level of enjoyment	Perceived difficulty of Huff quiz
SAT score	Map-reading ability	Perceived overall worth	Performance likely better if graded
		Perceived intellectual challenge	
		Experience of technical problems	
		Time taken to complete	

Major Questions

Three overarching questions guided our study:

(1) What is the overall effectiveness of the Geography 5 multimedia computer modules as enrichment exercises?
(2) What kinds of students benefit most by doing these module enrichments?
(3) What kinds of geographical knowledge and skills are best reinforced by these modules?

The first question is the most general: we wanted to know to what extent doing multimedia computer modules improved students' understanding of classroom material. The latter two questions disaggregate this overall effect with respect to the two key dimensions of students and course material. The student variables we were interested in are found in Table I; their values were determined by means of a questionnaire administered to all participants. The variables include (a) general characteristics such as gender and class level, (b) background (self-rated) in courses, concepts and skills relevant to the study, (c) variables related to student use of the modules, ranging from perceived value to student experience of technical problems, and (d) quiz-related variables such as perceived difficulty.

A variety of forms of geographical knowledge and skills is typically introduced in a lower-level human geography course. These can be organised as given in Table II with reference to the classic taxonomy of educational objectives of Benjamin Bloom (Bloom *et al.,* 1956). Bloom's taxonomy for the cognitive domain includes several forms of knowledge ranging from knowledge of specifics (e.g. terminology, specific facts) to knowledge of universals and abstractions (e.g. principles, theories). It also lists a number of desired skills ordered according to complexity, including comprehension, application, analysis, synthesis and evaluation. Bloom category numbers are listed next to each objective heading in Table II; these represent the major divisions only within Bloom's scheme.

Table II suggests how the two computer modules chosen for the study represented these different forms of knowledge and skills. The Huff module was designed to reinforce student knowledge of (in Bloom's terminology) "how to deal with specifics" by application of distance and gravity formulae to concrete situations, and to help students develop a range of skills from comprehension (understanding the purpose of the gravity model) and application (using the Huff model to calculate probabilities) ultimately to evaluation (assessing major limitations with the Huff model as a predictor of spatial behaviour). The cognitive maps module was designed to review students' knowledge of specifics (e.g. properties of cognitive maps) and abstractions (e.g. theory

TABLE II. Geographical knowledge and skills addressed in modules.

Bloom's taxonomy	Huff model module	Cognitive maps module
Knowledge		
Of specifics (1.10)		Recall definition of cognitive maps and principles of hierarchical organisation
Of how to deal with specifics (1.20)	Recall application of distance, gravity equations	
Of universals/abstractions (1.30)	Understand gravity concept in human geography	Understand rationale behind errors in cognitive maps
Skills and abilities		
Comprehension (2.00)	Understand components of gravity equation	Understand possible effects of errors and distortions on behaviour
Application (3.00)	Calculate distances Perform gravity, Huff model calculations	Interpret new data based on understanding of hierarchical organisation and errors of rotation Apply theory of alignment and cognitive perspective to concrete situations
Analysis (4.00)		Check consistency of hypotheses with given scenario
Synthesis (5.00)	Interpret results of gravity, Huff model calculations	
Evaluation (6.00)	Note limitations to gravity model	Discuss value of understanding cognitive maps

of errors in cognitive maps), and to apply this knowledge to develop skills such as comprehension (understanding effects errors and distortions may have on behaviour), analysis (checking the consistency of hypotheses with a given scenario), and evaluation (critically discussing the value of cognitive mapping research in terms of understanding human behaviour).

Procedure

The study proceeded in three major phases. The first phase consisted of introductory lectures on spatial interaction and behavioural geography, which were attended by all students (n = 100) participating in the study. In the second phase, students were randomly assigned to complete either the Huff model or cognitive maps module. Students taking one module thus simultaneously served as a control for students taking the other module. Some unplanned minor technical problems occurred during the administration of the Huff module; in addition, a small number of students in both modules reported occasional technical difficulties.

The final phase consisted of a quiz on relevant gravity model and cognitive maps material covered in lectures but developed further in the multimedia modules; all students answered questions in both areas. Our intent was that all questions could in theory be answered successfully by applying material from class lectures, though we presumed that completing the multimedia module would provide further preparation for doing well on the post-test. There was no pre-test included in the procedure: student participants had no previous college-level exposure to the subject areas included in this experiment, as verified by means of a questionnaire administered at its inception, and so we assumed that they had virtually no geographical knowledge and skills in these areas coming into the experiment.

The quiz was designed to assess student mastery of concepts and skills related to the gravity model and cognitive maps. Gravity model questions assessed students' abilities to (a) calculate Euclidean distances from maps, (b) comprehend the gravity model formula, (c) apply this formula to a simple example, interpret results and point out limitations, (d) perform a simple Huff model-type probabilistic determination of gravity behaviour, and (e) critically assess the meaning of results of probabilistic models. Cognitive map questions evaluated students' abilities to (a) understand and recall properties of cognitive maps, (b) interpret data based on their knowledge of common errors, (c) predict the effects of these errors on human behaviour, (d) analyse scenarios based on given assumptions, and (e) discuss the value of cognitive mapping research in terms of understanding human behaviour. Each of these two portions counted a maximum of 20 points.

The 100 students participating in the experiment were randomly assigned to complete the Huff module or the cognitive maps module. Participants worked at their own pace and at separate times in a university microcomputer laboratory. In addition to the assigned module, a general questionnaire including academic background and several module-based questions (Table I) was administered. Students answered module-based questions on a four-point Likert scale which ranged from 'disagree strongly' to 'agree strongly'. Additional questions determined whether students had attended relevant lectures, and these results were incorporated into the regression analysis.

Both quizzes were administered to the entire class during discussion sections (smaller group sessions linked to the large-group lectures), and time was limited to 45 minutes. The order of quiz completion was counterbalanced; no order effects were found in subsequent analysis. A subset of 12 participants was administered the quiz in a separate setting. These students were instructed to think aloud while answering the questions; this generated a protocol which was recorded and analysed separately.

Results

Separate regression models were constructed to predict the two quiz scores. Backward elimination was used to select the appropriate set of variables for the model. We started with a set of 15 student variables, and then eliminated the one with the lowest absolute t-value. As suggested in the analysis of the cognitive maps quiz presented in Table III, however, the vast majority of these variables were weak predictors.

The final model for predicting cognitive scores (Table IV) included SAT (the Scholastic Aptitude Test, taken by a majority of applicants to American universities) and 'module completed' (Huff vs cognitive) as strong predictors and 'perceived benefit to understanding' as marginally significant (removal of the latter variable resulted in a larger cross-validated standard error and smaller cross-validated *R-square).* In the Huff quiz model all variables except for SAT were eventually dropped. Although SAT was statistically significant in the Huff case, it accounted for only 21% of the variance in the model.

A comparison of mean Huff and cognitive quiz scores based on which module students completed is presented in Figure 1. Mean scores on the cognitive maps quiz were 9.7 for students completing the cognitive module, and 7.3 for students completing the Huff module, a statistically significant ($t(95) = 3.96$, $p < 0.001$), though not

TABLE **III.** Original model predicting cognitive map scores.

Independent variable	Coefficient	Std error	*T-stat*	p-value
Intercept	-3.6057	4.0682	-0.8863	0.3794
General				
Class level	-0.0914	0.3284	-0.2782	0.7819
Gender	-0.1656	0.6909	-0.2398	0.8114
SAT score	0.0105	0.0021	4.8957	0.0000
Background				
Computer familiarity	-0.2248	0.5286	-0.4253	0.6723
Mathematical ability	0.1364	0.5534	0.2464	0.8063
Map-reading ability	-0.1768	0.5755	-0.3072	0.7598
Module-specific				
Module completed	1.5830	0.8331	1.9002	0.0628
Level of enjoyment	0.2006	0.6333	0.3167	0.7527
Perceived overall worth	-0.5247	0.5790	-0.9062	0.3689
Perceived intellectual challenge	0.0923	0.4965	0.1859	0.8533
Perceived benefit to understanding	0.8689	0.5696	1.5255	0.1330
Experience of technical problems	-0.2384	0.3383	-0.7048	0.4840
Quiz-specific				
Perceived difficulty of cognitive quiz	0.0354	0.5339	0.0663	0.9474
Perceived difficulty of Huff quiz	0.3008	0.5029	0.5982	0.5522
Performance likely better if quiz graded	-0.3740	0.3844	-0.9730	0.3349

Notes: Residual standard error = 2.6169; multiple *R-square* = 0.4718; F-statistic = 3.2155 on 15 and 54 df; p-value < 0.001; cross-validated standard error: 3.4851; cross-validated *R-square: 0.0631.*

especially sizeable, difference. There was no statistically significant difference in **Huff** quiz scores between the two groups (Huff average = **11.4,** cognitive average = 10.7).

Analysis was performed on the top 20% and bottom 20% of scores to explore the possibility that certain students benefit from this form of instruction more than others (Table V); here, 'benefit' is defined operationally in terms of quiz score, assuming all students started at the same level of knowledge regarding the gravity model and cognitive maps. Although most tests were only marginally significant, the pattern of

TABLE IV. Final model predicting cognitive and Huff quiz scores.

Independent variable	Coefficient	Std error	*T-stat*	p-value
Cognitive map quiz				
Intercept	-5.2477	2.2950	-2.2866	0.0254
SAT score	0.0101	0.0019	5.4148	0.0000
Module completed	2.0928	0.6232	3.3579	0.0013
Perceived benefit to understanding	0.7362	0.4255	1.7301	0.0883
Huff model quiz				
Intercept	-0.4383	2.6298	-0.1667	0.8682
SAT score	0.0119	0.0025	4.7827	0.0000

Notes:
Cognitive map quiz: Residual standard error = 2.4357; multiple *R-square* = 0.4407; F-statistic = 17.334 on 3 and 66 df; p-value = 0; cross-validated standard error: 2.5880; cross-validated *R-square: 0.3685.*
Huff model quiz: Residual standard error = 3.1967; multiple *R-square* = 0.2603; F-statistic = 22.8744 on I and 65 df; p-value = 0; cross-validated standard error: 3.3008; cross-validated *R-square: 0.2114.*

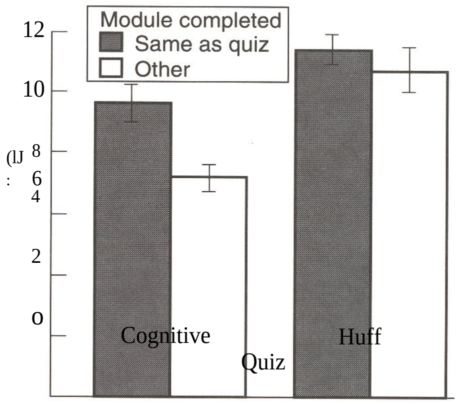

FIGURE 1. Comparison of mean cognitive map and Huff model quiz scores based on module completed. Error bars represent 95% confidence intervals.

results mirrored the total group findings for both high and low scorers, and for both the entire quiz and the disaggregated components. High and low scorers enjoyed the module just as much, found it similarly challenging, and believed it helped them understand the material to the same degree. Low scorers did, however, have a significantly lower mean SAT than high scorers (902 vs 1191), and spent more time on the module (98 vs 77 minutes).

Geographical knowledge and skills were partially disaggregated (in practice, they are often highly interrelated) by performing a separate analysis on subsets of each quiz. The cognitive maps quiz was fairly readily dividable into major Bloom categories (Table VI). Students who had completed the cognitive maps module scored significantly higher in lower-order realms, such as knowledge of specifics and comprehension. Huff quiz scores were divided according to the five major objectives of the quiz noted earlier, most of which serve multiple-though unique combinations of-Bloom cognitive categories. One test was marginally significant: students who had completed the Huff module scored higher than the cognitive maps group in calculating distances. All other tests showed no significant differences between the two groups.

TABLE V. Comparison of top and bottom 20% of student quiz scores.

Variable	Top 20% Mean	Std. error	Bottom 20% Mean	Std. error	*T-stat*	p-value
SAT score	1191.429	131.667	901.818	127.265	5.539	0.000
Time taken to complete	77.000	21.587	98.077	35.032	-1.990	0.057
Benefit to understanding	3.313	0.479	3.118	0.857	0.799	0.430
Intellectual challenge	2.983	0.854	2.765	0.831	0.589	0.560
Technical problems	2.375	1.204	2.471	1.007	-0.248	0.806
Perceived overall worth	2.688	0.946	2.765	0.903	-0.240	0.812
Level of enjoyment	3.00	0.730	2.941	0.748	0.228	0.821

TABLE VI. Disaggregation of quizzes into knowledge/skills categories.

Knowledge/skills	Cognitive group Mean	SD	Huff group Mean	SD	*T-stat*	p-value
Cognitive/skills:						
Cognitive domain;						
Knowledge of specifics	3.649	1.637	2.767	1.125	3.143	0.002
Knowledge of universals	0.730	0.450	0.730	0.462	0.311	0.757
Comprehension	2.568	1.55	1.317	0.833	5.154	0.000
Application	2.486	1.304	1.867	1.065	2.553	0.012
Analysis	0.946	0.780	0.900	0.706	0.299	0.765
Evaluation	0.730	0.769	0.533	0.724	1.267	0.208
Huff model quiz						
Objective:						
Calculate Euclidean distances	1.324	0.915	1.617	0.739	- 1.727	0.087
Comprehend gravity model	2.378	0.639	2.467	0.623	- 0.671	0.504
Apply/interpret gravity model	3.811	2.039	3.833	2.109	- 0.052	0.959
Apply/interpret Huff model	2.000	1.414	2.200	1.260	- 0.725	0.470
Critically assess results	0.270	0.508	0.350	0.547	- 0.716	0.476

Discussion

Effectiveness of Modules

We begin by applying these results to the three general questions we posed in our study.

(1) *What is the overall effectiveness of the Geography* 5 *multimedia computer modules as enrichment exercises?* The results above paint a remarkably unimpressive picture of the effectiveness of our modules. Test results suggest that the Huff module was entirely ineffective as an enrichment device for gravity model concepts in lower-division human geography, and the cognitive module was only moderately effective, with students who used it still giving correct answers for fewer than half the questions included in the cognitive portion of the test. The students' scholastic aptitude levels, as measured by their SAT scores, were much stronger predictors of their quiz performance than whether or not they had completed an enrichment exercise.

Protocol analysis revealed differences in the ways the test was taken by students who had completed modules and students from the control group. Those participants who had completed a computer module produced lengthier answers and generated more ideas on that portion of the exam, although their responses were not more likely to be correct. This result may suggest that the test did not adequately capture their knowledge; or it may mean that completing a module gave students a false sense of security that they understood the material. Our guess is that both factors may have been involved.

(2) What kinds of student benefit most by doing these module enrichments? As summarised above, there were few significant differences detected in the regression analysis and in comparing high and low scorers, other than that high-scoring students had higher levels of aptitude as measured by their SAT scores, and took less time to complete the modules. The overall lack of difference between these two groups may suggest that the modules did not discriminate against students with perceived low levels of computer familiarity, for instance, or students who found the material to be difficult.

(3) What kinds of geographical knowledge and skills are best reinforced by these modules? Disaggregated results for clusters of questions representing particular clusters of knowledge and skills did not yield expected results for both the Huff and cognitive modules. In the cognitive module case, the more significant higher-order domains such as knowledge of universals, analysis and evaluation were not markedly improved; in the Huff module case, disaggregation failed to improve the overall poor result, with the exception of distance calculation-more a mathematical than a geographical skill! These results were highly counterintuitive, since both modules offered students a very concrete opportunity to develop interpretation, synthesis and evaluation skills in comparison with the more abstract and introductory class lecture.

Value of Experimental Procedure

Surprisingly for us, our results largely corroborated Reeves's allegation raised above that no significant differences are usually found when performing comparative evaluation of multimedia effectiveness. On the face of it, this result is unusually negative: whereas, for example, we had thought that students doing the Huff module would gain valuable experience in applying the gravity model and would thus outperform the control group, the results seem to indicate that this is not so. Though the case of the cognitive module was stronger, it still did not produce the impressive results most people would consider necessary to justify continued multimedia development and use in undergraduate education.

Several hypotheses could explain these overall results. Perhaps the modules were so poorly designed and/or implemented that they truly had no beneficial impact on learning; yet they and the other Geography 5 modules were all carefully developed, pre-tested and revised several times prior to this experiment, all with significant qualitative student input, so we cannot put full weight on this hypothesis. A second possibility is that our principal evaluation tool (the post-test) was poorly constructed; this seems rather unlikely, however, given our experience at designing tests. Finally, student motivation to learn and perform well on the quiz could have been low, as our deliberate prohibition of students from doing both modules so as to provide control groups meant the quiz scores could not fairly count toward their final grade (though completing the modules and ~hequiz did; thus all students were motivated to participate). Yet we have little basis to believe that students did not try to do a reasonable job; indeed, many students took the experiment quite seriously, and when asked whether their quiz performance would likely have been better had it directly counted toward their final grades, an equal proportion agreed as disagreed. Each of these hypotheses, then, may be true in part, yet we do not feel that they can explain the results we obtained.

Perhaps it is worth revisiting Reeves's discussion to link it with our own experiences (Reeves, 1991). Reeves's main point is that learning (and certainly multimedia-based learning) is far too complex a phenomenon to capture adequately with statistical analysis of standard experimental evaluation techniques. Education, in Reeves's (though not necessarily our!) perspective, is still in its early stages as a science, and must necessarily devote itself to description rather than the sophisticated experimental techniques characteristic of more developed sciences such as physics. The fact that statistically significant differences were generally not found between the multimedia-enhanced and traditional education methods in our experiment would not, in Reeves's perspective, lead to any necessary conclusion regarding the educational merit of our Geography 5 modules. In fact, Reeves would probably argue that even if we were to find a greater preponderance of statistically significant results, this would have said more about the experimental environment we created for the evaluation than the real environment in which students learn geography.

What are we, then, to make of the effectiveness of our multimedia investment? Reeves does not reject evaluation altogether, only forms that rely exclusively on what appear to be rigidly controlled experiments and statistical analysis. In their place, Reeves suggests a broader style of evaluation based on Clifford Geertz's ethnographic style of 'thick description' (Geertz, 1973). This form of evaluation, which Reeves terms 'descriptive' evaluation, can include experimental studies and statistical analysis as components, but must be devoted primarily to formative assessment (evaluation focusing primarily on improving the multimedia product) rather than summative assessment (evaluation similar to comparative studies which attempts to indicate once and for all the value of a multimedia product). Reeves mentions methodologies such as interviews, observation and even computer modelling as superior modes of gaining thick descriptive understandings.

Lessons from Other Data

Thankfully, our evaluative procedure involved more than the post-test. It included a questionnaire as well, which provides additional information in conjunction with test results. For instance, analysis of responses suggests that student assessments of the worth of multimedia modules were fairly reliable indicators of their actual value in improving quiz performance. Specifically, the cognitive module, which as stated above improved students' quiz performance more than the Huff module, received higher reports of student enjoyment (3.3 vs 2.9, $t\,(84) = 3.8$, $p < 0.001$), perceived overall worth (2.8 vs 2.3, $t\,(84) = 2.33$, $p < 0.05$), and perceived benefit to understanding (3.3 vs 2.9, $t\,(84) = 2.21$, $p < 0.05$). These results may in part be due to technical difficulties encountered during the Huff module: Huff students reported a greater amount than cognitive maps students (3.0 vs 2.1, $t\,(84) = 3.84$, $p < 0.001$).

Qualitative information gained from student questionnaires and informal interviews conducted during the period of the experiment yielded further insight into possible benefits of the Geography 5 modules, and ways they can be improved. It is likely, for instance, that student enjoyment of computer modules may be as important in the long term as student 'learning'—at least as defined by improvement on tests—since enjoyment often translates into paying better attention to material presented in class, taking more geography courses later, and perhaps even considering geography as a major field of study. Other analyses of multimedia in higher education have essentially come to the same conclusion (e.g. Branson, 1971; Squizzero, 1976). With respect to this critical variable, students responded that they strongly preferred completing the Geography 5 computer modules as enrichments versus using more traditional methods (e.g. paper handouts). In particular, students mentioned animated portions of the modules (e.g. digital movies) as particularly enjoyable and helpful to their understanding of the material. In addition, survey results give a good basis to believe that any effort to minimise student frustration while completing modules (e.g. by eliminating possible technical malfunctions as suggested above, or by reducing the time necessary to complete each module) will result in more successful acquisition of module-based skills and knowledge. Also, we have learned that whereas some students are very open to computer modules, other students are highly resistant, owing to lack of computer familiarity, lack of interest, concern about time demands, or a combination thereof. It thus seems ideal to present computer-based multimedia enrichments as a resource that students can choose to use, rather than as exercises all students must complete, given the very high level of anxiety associated with computer-based learning among lower-division undergraduates.

These kinds of insights generally did not require quantitative measures to reveal, though they certainly did require attentive evaluation of student perspectives on the Geography 5 modules. The result is indeed more in the order of Geertzian thick description than demonstration of statistical significance, but our mosaic approach seems to perform far better in interpreting the complex interactions between students and geographical material via multimedia enrichments.

implications

Viewed broadly, evaluation has thus played a critical role in our continuing development of multimedia-based learning resources for lower-division human geography. The experimental procedure we followed in this study was valuable not least in that it demonstrated the elusiveness of quantitative measures of our success so far. As we are apparently not alone in this experience, we are not taking the results as negative indicators of the value of these modules. This overall assessment would not have been reached, however, if it were not for the qualitative input gained by student interviews and course evaluation surveys as mentioned above. If all measures we employed to assess the educational value of the Geography 5 computer modules revealed significant areas of concern, we would have had to arrive at a quite different conclusion from our quantitative results.

Evaluation will always be critical in multimedia applications in geography, as a means of assessing and improving their worth in assisting the student learning process. Given the more complex role of the instructor in a differentiated learning setting involving multimedia resources, evaluation provides information essential to optimize their use (Menges, 1994). A recent National Education Association publication offers the following perspective:

> With multimedia, students move to the center of the learning process. This does not mean that teachers move to the periphery. In using multimedia products, the teacher's role is to ensure that students are asking questions that lead in productive directions and that students are sufficiently challenged by their own inquiries. .. .As the educational process becomes more open-ended (as, for example, in classrooms where multimedia products are created as well as explored), the teacher's role in creating a suitable learning environment and structuring inquiry is even more critical (National Education Association, 1994, p. 16).

Experimental methods mayor may not provide instructors with the information they need to utilise multimedia effectively in the university geography curriculum, but evaluation in general-ideally following a number of coupled strategies-will always play an essential role in improving the quality of undergraduate education.

Acknowledgements

This research was kindly supported in part by a grant provided by the UCSB Office of Instructional Development. Dr Proctor also acknowledges support from National Science

Foundation Grant #SBR-9600985. We gratefully acknowledge the participation of the undergraduate students of Geography 5 at UC Santa Barbara during spring term 1995.

Correspondence: James D. Proctor, Department of Geography, University of California, Santa Barbara, CA 93106-4060 USA. Tel: 805893-8741. Fax: 805893-3146. Email:jproctor@geog.ucsb.edu

REFERENCES

AMBRON, S. & HOOPER, K. (Eds) (1990) *Learning with Interactive Multimedia: Developing and using multimedia tools in education* (Redmond, WA, Microsoft Press).

ANTROP, M. (1992) New information techniques in the teaching of geography, *Bulletin, Sodett! Geographique de Liege*, 28, pp. 57-63.

BLOOM, B.S., ENGLEHART, M.D., FURST, E.J., HILL, W.H. & KRATHWOHL, D.R (1956) *Taxonomy of Educational Objectives: the classification of educational goals. Handbook I: Cognitive domain* (New York, David McKay).

BRANSON, R.K (1971) Formative evaluation procedures used in designing a multi-media physics course, paper presented at the annual meeting of the American Educational Research Association, New York, ERIC Document ED 050140.

CARSTENSEN, L.W., Jf., SHAFFER, C.A., MORRILL, RW. & Fox, E.A. (1993) GeoSim: A GIS-based simulation laboratory for introductory geography, *Journal of Geography,* 92(5), pp. 217-222.

CHUN, D.M. & PLASS, J.L. (1995) Project CyberBuch: a hypermedia approach to computer-assisted language learning, *Journal of Educational Multimedia and Hypermedia,* 4(1), pp. 95-116.

EDUCATIONAL RESOURCES INFORMATION CENTER (1995) Online database search, 16 August.

FAGAN, B.M. & MICHAELS, G.H. (1992) Anthropology 3: an experiment in the multimedia teaching of introductory archaeology, *American Antiquity,* 57(3), pp. 458-466.

FITZPATRICK, C. (1993) Teaching geography with computers, *Journal of Geography,* 92(4), pp. 156-159.

FLOWERDEW, R. & LOVETT, A. (1992) Some applications of Murphy's Law: using computers for geography practical teaching, *Journal of Geography in Higher Education,* 16(1), pp. 37-44.

FOOTE, KE. (1994) The geographer's craft: a new approach to teaching geographical methods in the liberal arts curriculum, *Pennsylvania Geographer,* 32(1), pp. 3-25.

GARBER, H. (1976) Project evaluation of English composition methods, unpublished report, State University of New York, College at Oswego, ERIC Document ED 173798.

GARNER, BJ. & QIMING, Z. (1993) GIS education and training: an Australian perspective, *Computers, Environment and Urban Systems,* 17(1), pp. 61-71.

GEERTZ, C. (1973) *The Interpretation of Cultures* (New York, Basic Books).

GOLD, J.R, JENKINS, A., LEE, R., MONK, J., RILEY, J., SHEPHERD, I. & UNWIN, D. (1991) Computer assisted teaching and learning, in: *Teaching geography in Higher Education: a manual of good practice,* pp. 97-135 (Oxford, Basil Blackwell).

GOSSETTE, F. & WHEELER, J.D. (1993) Computer mapping in a regional geography course, *Journal of Geography,* 92(1), pp. 28-34.

KENT, W.A. (1992) The new technology and geographical education, *Geography and Education,* pp. 163-176.

LE GRAS, C. (1991) Educational software in the United Kingdom, *Cahiers Geographiques de Rauen, 36,* pp. 43-64.

LINDHOLM, M. (1992) The possibilities of hypermedia in geographical education, *Terra, 104(3),* pp. 169-180.

LIOU, H.-C. (1994) Practical considerations for multimedia courseware development: an EFL IVD experience, *CALICO Journal,* 11(3), pp. 47-74.

LUECKENHOFF, M. (1993) Teaching geography with Hypercard, *Journal of Geography,* 92(4), pp. 166-169.

MAGUIRE, DJ. (1992) Teaching earth and social science in higher education using computers: bad practice, poor prospects?, *Geographic Information 1992/3. Yearbook of the AGI,* pp. 313-317.

McKILLIP, J. & BALDWIN, K (1990) Evaluation of an STD education media campaign: a control construct design, *Evaluation Review,* 14(4), pp. 331-46.

McLAUGHLIN, G.H. (1973) Evaluation of a media-based credit course, 'Developmental Psychology', offered by Ryerson Open College (Toronto, Ontario Educational Communications Authority, Report Number 35, ERIC Document ED 115273).

MENGES, RJ. (1994) Teaching in the age of electronic information, in: *W.l.* MCKEACHIE (Ed.) *Teaching Tips: strategies, research, and theory for college and university teachers,* pp. 183-193 (Lexington, MA, D.C. Heath).

MERCHANT, *l.W.* (1993) Resource materials for teaching GIS, *Geo Info Systems,* 3(4), pp. 57-61, 68.

MITRA, A. (1994) 'Instructor-effect' in determining effectiveness and attitude towards technology-assisted teaching: report of a case study, *Journal of Instruction Delivery Systems,* 8(3), pp. 15-21.

NATIONAL EDUCATION ASSOCIATION (1994) *Multi-media: its promise and challenges for public education* (West Haven, CT, National Education Association, Research Division).

NATIONAL EDUCATIONAL COMPUTING ASSOCIATION (1994) *Recreating the Revolution: National Educational Computing Conference 1994* (Boston, MA, National Educational Computing Association).

PROCTOR, J.D., SUTTON, P.c. & MICHAELS, G.H. (1995) Multimedia guided writing modules for introductory human geography, *Journal of Geography,* 94(6), pp. 571-577.

PROTHERO, W.A. (1995) Using technology to tame the large oceanography class, *Journal of Geological Education* (forthcoming).

RAPER, J. & GREEN, N. (1992) Teaching the principles of GIS: lessons from the GISTutor project, *International Journal of Geographical Information Systems,* 6(4), pp. 279-290.

REEVES, T.C. (1991) Ten commandments for the evaluation of interactive multimedia in higher education, *Journal of Computing in Higher Education,* 2(2), pp. 84-113.

SQUIZZERO, W.E. (1976) A comparative study of a traditional approach and a multimedia approach to teaching the accounting cycle unpublished EdD dissertation (Nova University, ERIC Document ED 157585).

SVINGEN, B.E. (1994) New technologies in the geography classroom, *Journal of Geography, 93(4),* pp. 180-185.

UNWIN, *D.l.* (1991) Using computers to help students learn: computer assisted learning in geography, *Area,* 23(1), pp. 25-34.

WILSON, K. & TALLY, W. (1991) *Looking at Multimedia: design issues in several discovery-oriented programs* (New York, Center for Technology in Education, Bank Street College of Education Technical Report No. 13).

YEAZELL, M.l. (1974) A comparison of instructor-paced and self-paced groups in a criterion-referenced, multi-media educational psychology course, (paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, 15-19 April, ERIC Document ED 088956.

YOUNG, K. (1972) Validation of individualized instructional materials, ERIC Document ED TMOOI885.

YUN HEE, C. (1991) A study on the development of the coursework for geography education and its educational effects under the viewpoint of cognitive approach, *Journal of Geography Education, 25,* pp. 1-23.