Does Learning Occur through Gaming?

Author:
Sharon L. Gander, Senior Learning Analyst, Cerner Corporation, Cerner Virtual University, 2800 Rockcreek Parkway, W0721 Kansas City, MO 64117 (phone: 816–201–2623), (fax: 816–201–8623)
sgander@cerner.com

Author Biography

Ms. Gander is a Senior Learning Analyst with Cerner Corporation where she was Project Manager and Learning Analyst for the *HNAM DataQuest* Learning Game Development Team. She has her M.Ed. from Montana State University, Bozeman, MT. Ms. Gander has over 20 years experience providing education to both adults and youth and specializes in the development of alternative learning practices such as games, goal-based scenarios, and group-process learning events.

Abstract

Cerner Corporation's Cerner Virtual University (CVU) created a computer-based game, *HNAM DataQuest*: The Millennium Architecture Knowledge Adventure, to teach technical information systems concepts. This game was designed to match knowledge required with instructional and game design strategies chosen. CVU needed to determine whether or not the game taught the concepts intended was determined through pre/post testing. Therefore, if the game was properly designed, then an increase in delta of post-test minus pre-test scores would indicate the extent of new knowledge acquired by the learner-players. While not a broad-base scientific study, this report of CVU's results with computer-based gaming provides support for the use of games designed expressly to teach specific knowledge. The change in post-test minus pre-test scores indicates that only game playing could have produced ratios of this magnitude.

Introduction

This introductory section provides an overview on the development of Cerner Corporation's learning game, a review of the literature and the questions posed by the development of this product.

Development Background

In 1998, Cerner Virtual University, (CVU), developed a computer-based learning game to teach core processes and knowledge of Cerner's three-tiered client/server architecture and relational database. The resulting product, *HNAM DataQuest:* The Millennium Architecture Knowledge Adventure (a.k.a. *HNAM DataQuest*), was an experimental venture for CVU. As CVU worked through the development of this game, one of the reoccurring concerns was whether or not adult learners could actually acquire new knowledge through this medium. Therefore, a testing plan was developed and the results, based on the delta of post-test minus pre-test scores, were analysed.

Literature Review

A review of the educational literature provided limited data on whether or not learning occurs during gaming even though many educators firmly believe in and back gaming as an effective educational methodology. McMullen (1987) provided the only available data demonstrating that games were more effective than drill and practice CBT. However, his work was done with sixth-graders and not with adults. Dempsey and colleagues (1996) reviewed 100 games and found little substantive research on how to use computer games for educational purposes. Furthermore, Demsey proposed that games were designed to be entertaining rather than educational and that, consequently, learning was incidental. Wolf (1997) found only one case study existed that controlled learning outcomes produced by a business game (1975). Merrill (1999) indicated that the research is insufficient to prove that learning does in fact occur during gaming.

Therefore, a pre- and post-evaluation was needed to determine whether or not Cerner Corporation's target audience had in fact acquired the new and essential knowledge presented in the HNAM DataQuest. During beta testing for this product, CVU used a simple pre/post test and observation procedure as one element of its usability review. This testing was aimed at resolving whether or not gaming can be an effective learning strategy for adults learning highly technical content.

Questions Posed

CVU needed to determine whether or not the *experimental HNAM DataQuest* game was an effective learning tool and to what degree and where it was and was not effective. Likewise, CVU wanted to know whether *HNAM DataQuest* taught the concepts it purported to teach. The questions that needed answers were:

1. Do computer-based games teach adults?

2. Does *HNAM DataQuest* teach the Cerner-specific information systems concepts it purports to teach?

3. Can learning occur through a computer-based game whose format matches the learning content desired?

4. Is 'length (in minutes) of play' a factor that influences, whether or not, learning occurs?

Methods

This section describes the intended user audience, education design basis, and game strategy related to instructional design strategy. The final sub-section defines the match of audience, design and game strategy to the testing procedures.

Audience

This section reviews the match between the intended and tested learner-player audiences using three factors (roles, computer experience, and Cerner experience). Other factors were not specifically measured but may have influenced learning; these factors are identified. Overall the tested audience did match the intended user audience.

(a) Intended Audience.

The intended audience for the game was ‘anyone at Cerner who worked with Cerner's HNAM Millennium products’. The targeted range of role titles identified included:

(1) Application Developer;
(2) Application Specialist (a technical implementation consultant);
(3) Learning Analysts.
(4) Project Architect.
(5) Project Executive.
(6) Project Manager.
(7) Sales Consultant.
(8) Systems Analyst/Systems Engineer.
(9) Technical Writer.
(10)Tested Audience Roles.

Twenty-nine (29) individuals were tested. The test audience included the following roles and the number (#) of tested individuals in each role:

(a) Account Manager (2),
(b) Application Developer (2),
(c) Application Specialist (3),
(d) Call Centre Support Specialist (2),
(e) Certification Analyst (2),
(f) Functional Architect (1),
(g) Learning Analyst (1),
(h) Learning Coach (1),
(i) Multimedia Writer (1),
(j) Role not identified (2),
(k) On-line help Developer (2),
(l) Project Architect (5),
(m) Sr. Manager (1),
(n) Sr. Marketing Communications Specialist (1),
(o) Systems Analyst (2), and
(p) Technology Technical Consultant (1).

This tested audience included the needed roles and, in fact, represented a broader audience than required.

(b) Audience experience level.

The test-audience included both new associates with no Cerner experience and experienced Cerner associates as described in the Table A. The newest Cerner’s associates start employment at Cerner with a range of computer skills from no-previous computer experience to highly experienced information systems management backgrounds. Those who had information systems backgrounds would have general information systems knowledge and not specifics about Cerner’s architecture. Current Cerner associates should have experience both with computers and with Cerner’s specific information system. Therefore, in background experience with Cerner’s products and/or with general information systems experience the individuals tested adequately represented the typical Cerner audience – both experienced and newly hired.

The experience level of the 90 minutes-of-play group (Table A) was sufficiently different from other groups to impact their pre-test scores. This group started with more Cerner experience, more computer experience and more database experience than the other groups. Therefore, this group was expected to start with higher pre-test scores and show less change overall.

The test groups were arbitrarily assigned based on the date when individuals participated in testing. The first group tested were internal associates who had been with Cerner a number of years. Subsequent groups were new-hires, who had been with Cerner for only a few weeks even though some of them had worked with Cerner's products in previous occupations. All were volunteers willing to take from other work to assist in testing a product. Some of the new associates had information systems experience and others did not. All together, the test audience adequately matched the intended Cerner audience. (See Table A).

Table A: Average Audience Experience Levels by Length-of-play Groups and Composite-of-all-play Group

	90 minutes	60 minutes	30 minutes	Composite of all play
Number of Individuals	7	20	3	30
Years w/ Cerner	1.44	0.00	0.00	0.34
Computer experience	4.71	0.00	3.67	3.33
Database experience	2.57	0.00	1.67	2.13
Cerner products experience	1.71	0.00	1.00	1.03

(c) Tested audience gender balance.

The test audience was a random audience not selected for gender. However, it did contain both male and female learners in all length-of-play groups. There was a reasonable balance of men and women in each group in approximately the ratio that would be experienced throughout Cerner. However, gender was not a controlling factor, was not tracked, and has not been reported as a factor of results.

(d) Tested Audience Age Balance:

The test audience was a random audience not selected for age. Participants ranged in age from young college-graduates (approximately 21 years) to those over fifty years old. The mix of ages in each group was approximately the same balance as would be found throughout Cerner. However, age was not a control factor taken into account in the design of testing, was not tracked, and, therefore, has not been reported.

(e) Tested Audience Learning Styles/Thinking Preferences:

The test audience was a random, self-selecting audience not selected for specific learning styles or thinking preferences. Since this was not a variable for which the testing

controlled, and since the audience appears to have matched Cerner's typical associate audience, it was assumed that a wide range of learning styles and thinking preferences were represented. Some learner-players found gaming as a learning format easy, intuitive, and enjoyable while others found it difficult and less informative that other modes of presentation. These personal responses to the game format indicate that differences in preferred learning and thinking modes affect learners. However, the impact of match (or non-match) between preferred learning/thinking modes and learning was not a control factor taken into account in the design of testing, was not tracked, and, therefore, has not been reported.

(f) Match between tested and intended audiences:

The test audience effectively represented the intended market audience within Cerner in terms of roles, length of service within Cerner. Even though gender, age and experience in the information systems industry were not factors for which the study was controlled they were not noticeably out of alignment with Cerner's audience.

The Educational Design Basis

This section discusses the selected instructional design strategies intended to enhance and highlight the desired outcomes.

(a) Basic Design Principles

Key learning principles used by the HNAM DataQuest design team included: a] goal-based learning events with rich learning environments motivate learners to explore and facilitate learners discovery of underlying concepts and principles for themselves (Caine and Caine) (Shrank), (b) PEAs (process, entity, and attributes) provide motivating and memorable content (Merrill), (c) Just-in-time (JIT) learning support to create a rich environment controlled by the learner, and (d) experience (including mistakes) provides the basis for learning (Gagne, Briggs, & Wager.)

(b) Applying goal-based learning events

The game's learning environment was a resource rich situated learning environment similar to goal-based scenarios. Light-hearted, fun visuals and sound were added to increase the inherent motivation of problem-solving for goal. The rich environment provided many ways to solve the inherent problem and discover the underlying architecture's structure. Synthesis and review opportunities moved the tacit knowledge discovered in play into the explicit realm and reinforced learning.

(c) Applying Process, Entities, and Attributes (PEAs)

The 'play' of the game was based on two central Cerner process which together dramatized Cerner's three-tiered client/server architecture as it accesses its relational database. Conflict and motivational tension of play were created through the use of entities as game pieces. Each entity had: a] a unique light-hearted, fun, metaphorical image which became its game piece, b] distinct relationships to other entities, and c] unique attributes that must be understood to successfully play the game and build the correct relationships and processes.

Key processes were reflected in the connected path (between game pieces) on the playing area. The action of animated cartoons moving between entities reinforced relationships by dramatizing each entities input, output and relationship to other entities as they moved between game pieces. Depending on the state of the game the cartoon-reinforcement varied. When two game pieces had been connected correctly: a] the name of each entity appeared on-screen above the entity as it was touched by an animated cartoon, b] the entity/ game piece animated, c] an animated cartoon was generated as output, and d] that output-cartoon moved to the next entity/game piece as its input. However, where a given entity was incorrectly ordered, play stopped and error feedback appeared. The informational text in the error feedback reinforced the relationships and output of previously successfully encountered entity/game piece and provided hints and additional information about entities that might use the most recent successful output as their input. Error feedback did not provide a 'right answer' but encouraged continued exploration of relationships and attributes.

Attributes could be discovered many ways: through observation of the animation of game pieces, through use of the error feedback as previously described, or through use an attribute menu was available on each game piece. For each game piece the following attributes were available on the attribute menu:

(a) Who am I?
(b) Where do I live?
(c) What do I do?
(d) What is my input?
(e) What is my internal process?
(f) What is my output? and
(g) To whom do I relate?

Attributes described the essential functions of a given entity and its relationships to other entities.

(d Rich JIT Learning Support:

Just-in-time learning support was provided in order to strengthen learning and support a wide variety of learner-players' preferred learning and thinking modes. Learning support tools provided: a] Rules of road explaining how to play and what constituted 'winning', b] Tutorial on how to move and connect pieces, c] Hints (both automatic hints and hints-upon-user-request were available), d] Supporting documents describing Cerner-specific processes, events, and interactions, e] A summary video synthesis at the end of each round of play, and f] FAQ (Frequently Asked Questions) documents providing information about game play as well as information about Cerner processes and entities.

(e) Experience (including mistakes) provides the basis for learning**:**

Traditional adult education theory states that experience – active involvement – in learning is more valuable for adult learners than are more passive learning interventions (e.g. listening to a lecture). HNAM DataQuest was an active learning intervention, which provided a fun way for learner players to develop and test their understanding of Cerner's HNAM architecture. Mistakes provided opportunities to delve into the content at a deeper level. Successes were reinforced with light repetition, synthesis, metaphorical images and

opportunities to go to deeper and complex levels of understanding. Each round of game-play was another experience and another learning opportunity.

Game Strategy Related to Instruction Design Strategy

This section discusses the chosen game design strategies and their relationship to instructional design. Game strategies considered included: 'twitch' play, 'strategic' play, and the use of violence.

(a) Twitch Play vs. Strategic Play:

The design team defined 'Twitch' play as the video-like game strategy where speed and eye-hand coordination were key learner-player actions. Learner-players were scored on how fast they could accomplish the given task with the fewest errors. Building and organizing servers and databases were not deemed to be speed tasks or eye-hand coordination tasks. Twitch speed did not match the content or purpose embedded in PEAs. Therefore, 'twitch' video-like gaming was deemed an inappropriate strategy -- one that did not support the instructional design.

'Strategic' play was defined as gaming formats that focused on discovery of relationships and on uncovering the consequences when incorrect relationships are played. Organizing servers and selecting data fields from relational data has strategic relationships which learners needed to uncover. The relationships and consequences inherent in a strategy game format did match the PEA design and desired content. Therefore, a strategic game format was chosen.

(b) Violence in Play:

The design team considered the use of violence (killing off aliens or monsters or shooting down planes) to imply removing something from the play and equated that to destruction or loss of that entity. It was deemed inappropriate to imply that loss of information, servers, data or information services would be acceptable in an information system. Therefore, violence was not used in the design of *HNAM DataQuest*.

Testing Procedures

This section provides an overview of the test design, the testing procedures and describes how these were matched to the instructional design. The test design is less than scientific due to the nature and timing of the usability review in which it was embedded.

The specific pre- and post-tests and test-scoring procedures were designed to match the core knowledge outcomes required and to ensure that the *HNAM DataQuest* game was the only learning intervention that could have created the change in test scores. Three groups of test participants played *HNAM DataQuest* for different lengths of time. These three length-of-play groups provided the resulting data on the difference that length of play (in minutes) made in the amount of learning occurring through playing HNAM DataQuest.

Match of Game Design to Core Knowledge Outcomes:

The instructional design and game play format of *HNAM DataQuest* were specifically and intentionally selected in order to match the intended learning points and outcomes. Thus, learner-players were tested to determine whether or not they had learned the intended knowledge of processes, entities and their attributes. It was expected that *HNAM DataQuest* learner-players would show a positive change (delta) for their post-test scores minus their pre-test scores.

Testing Process:

A simple pre/post test was used to test change in knowledge. The test contained five multiple-choice questions and three short essays. The questions and their responses were the same on the both the pre- and post-tests. However, the sequence of questions and the order of each question's responses were re-ordered in the two tests to encourage test takers to think about their answers and to minimize rote memorization of questions and answers. The five multiple-choice questions were specifically focused on outcome knowledge needed. Each of the multiple-choice questions was designed to include at least one wrong answer, one answer that novices commonly use, at least one industry-standard answer that individuals trained in computer science would choose as well as the Cerner-specific answer that was preferred. Ratings on each answer indicated degree of preference for that response with 10 points being the preferred Cerner-specific response. The three essay questions were typical of questions that Cerner associates might be asked by a client about key technical concepts taught in the game; such as Cerner's three-tiered/client server architecture, the interaction between tables and keys in the relational data base, and the purpose and use of servers.

Length-of-Play Groups:

Testing done in groups where each group's length-of-play varied: 90, 60 and 30 minutes of play. There were slight variations in group's membership. Group membership was determined by the date on which the game was played and participants self-selected. This resulted in groups with uneven experience and backgrounds.

The 90 minutes-of-play group was composed of individuals who worked in an isolated testing station with no other players around them. In each case, one beta-test observer monitored both the test-taking and the game-playing. After completing the pre-test, these individuals used approximately 90 minutes to complete all three scenarios of the game. They were given the post-test immediately after completing their play. In total, this experience was a two-hour commitment to an intense process. Learner-players appeared mentally drained and exhausted at the end of this process.

Other learner-players tested and played the game while working in a classroom group situation with other learner-players around them. Each worked independently; they did not interact with each other. An observer monitored each person's test-taking and game-playing. Each person was given a pre-test, given time to play and then asked to complete the post-test. A sub-segment of this group was asked to take the post-test at 30 minutes-of-play and was then allowed to continue playing. This sub-group was not tested at 60 minutes-of-play. All other learner-players worked for one hour and were given the post-test after one-hour regardless of

how far they had progressed within the game. This 60 minutes-of-play group was the largest number of individuals tested.

Scoring Procedures

Test scores were tabulated using a simple spreadsheet and differences (delta) between pre- and post-test scores on each question were determined as follows. Preferred answers were specific to Cerner's information systems, while other answer were less preferred even though correct within the information systems industry generally. Some answers were not at all acceptable. Therefore, the multiple choice questions received 10, 5, or 0 points for the correct, industry accepted, or all other responses, respectively.

Essay question answers were evaluated for accuracy, clarity, and brevity. Rambling answers were considered to show a lack of understanding and an inability to effectively describe to a client the point of the question. Therefore, an answer could have been technically accurate and may not have received full points. Points given were: 10, for clear, accurate and concise answers; 8 for technically correct but wandering answers; 5 for correct on at least half of the concept, 3 for partial answers less than half correct and 0 for every other answer.

The results were analysed both on a one-tailed regression analysis by the time spent and finally by delta, change in response to specific questions related to which elements of the game had been completed during their play.

Testing Controls

Testing was not controlled for:

(a) Gender differences.
(b) Age differences.
(c) Learning style/ thinking preference.
(d) Impact of violence (or lack thereof).
(e) Choice of playing format — strategy versus twitch.

Testing was controlled for content that matched instructional design (PEAs) and for length of play. Testing was controlled for length-of-game-play, where length-of-game-play was the only learning intervention available between pre- and post-tests.

Results

This section reports the test data as it is related to each of the following four questions:

1. Do computer-based games teach adults?
2. Does HNAM DataQuest teach the Cerner-specific information systems concepts it purports to teach?
3. Can learning occur through a computer-based game whose format matches the learning content desired?
4. Is 'length (in minutes) of play' a factor that influences, or not, whether learning occurs?

Do Computer Games Teach Adults?

This section considers whether or not learning occurred across all groups of learner-players tested regardless of length of play. Test results are reported in Tables B (pre-test), C (post-test) and D (Delta Mean and Significance).

Table B: Table B: Pre-Test Range, Median, Mode, Mean and Deviation For Each of Three Length-of-Play Groups and the Composite-Of-All Play Groups

	90 minutes	60 minutes	30 minutes	Composite of all play
Pre-test Minimum Score	15.00	0.00	14.00	0.00
Pre-test Maximum Score	70.00	56.00	40.00	70.00
Pre-test Median	55.00	22.50	35.00	30.00
Pre-test Mode	#N/A	0.00	#N/A	0.00
Pre-test Mean Score	48.29	23.60	29.67	29.97
Pre-Test Mean Percentage	60%	30%	37%	37%
Pre-Test Standard Deviation	19.12	18.20	13.80	20.37

Table C: Table C: Post-Test Range, Median, Mode, Mean and Deviation For Each of Three Length-of-Play Groups and the Composite-Of-All Play Group

	90 minutes	60 minutes	30 minutes	Composite of all play
Post-test Minimum Score	35.00	15.00	25.00	15.00
Post -test Maximum Score	75.00	60.00	35.00	75.00
Post -test Median	70.00	40.00	33.00	40.50
Post -test Mode	#N/A	48.00	#N/A	30.00
Post -test Mean Score	62.86	38.10	31.00	43.27
Post -Test Mean Percentage	79%	48%	39%	54%
Post -Test Standard Deviation	14.14	12.00	5.29	16.14

Table D: Table D: Significance of Change for Each of Three Length-of-Play Groups and the Composite-Of-All-Play Group

	90 minutes	60 minutes	30 minutes	Composite of all play
Delta in Minimum Score	20.00	15.00	6.00	15.00
Delta in Maximum Score	5.00	4.00	1.00	5.00
Delta in Median	15.00	17.50	4.00	11.00
Delta in Mode	#N/A	48.00	#N/A	35.00
Delta in Mean	14.57	14.50	3.67	13.43
Delta in Mean Percentage Correct	18%	18%	5%	17%
Delta in Standard Deviation	(4.98)	(6.21)	(2.35)	(3.91)
Significance	<.01	<.01	>.05	<.01

As seen in the comparison of Tables B, C and D, the mean change delta of post-test minus pre-test scores for each length-of-play group and for the composite-of-all-play group indicated that learning did occur regardless of length of play. That is, the scores on the post-test questions increased an average of 13.43 points or 17% for most individuals regardless of how long they spent playing the game. Learning was directly related to the *HNAM DataQuest* game intervention as is indicated by a significance of <.01 for all groups. This positive delta indicates that learning of new concepts or refining of industry-standard concepts to fit the Cerner-specific model did occur.

The smallest deltas were shown for the 30 minutes-of-play group, which had completed only one round of play. The size of this group (three participants) made it difficult to arrive at any valid conclusions. With a significance greater than 05 (>.05) for the 30 minutes-of-play group, learning probably was directly related to game intervention even though the significance of this change can not be validated from this test-group. Other indicators of learning include specific responses to questions which learner-players had touch during play versus those they not touched (see Question 2 and Table E.)

The deltas and significance of changes (<.01) shown on Table D were strong for both the 60 minutes-of-play and 90 minutes-of-play groups. Learning was definitely happening in these two groups and, since the only intervention was the game, learning was related to *HNAM DataQuest*.

All together changes in test responses indicated that learner-players gained knowledge through playing HNAM DataQuest. Learning was happening generally across groups even in the 30 minutes-of-play group where change was the weakest. 2. Does *HNAM DataQuest* teach the Cerner-specific information systems concepts it purports to teach?

This section considers whether or not learning occurred for Cerner's specific information systems concepts. Test results reported in Table E provides mean change in responses for each of the eight (8) questions tested.

Overall the delta for most questions was positive and the mean overall change in scores averages a 13.43 point increase or 17% change per test. Change appeared to be occurring across all length-of-play groups for all questions. In particular, questions demonstrating change were those for which the majority of leaner-players had come into contact with their concepts during play. That is, the 30 minute-length-of-play group had not yet experienced the request server and may not have experienced the relational database while most or all members of the other two groups had experienced these concepts at 60 or 90 minutes-of-play.

Table E: Table E: Pre- to Post-Test Mean Change (Delta) In Scores By Topic/ Question for Each of Three Length-of-Time In Play Groups and Composite of All Play Group

	90 minutes	60 minutes	30 minutes	Composite of All play
3-tiered Client/Server architecture	2.86	5.00	1.33	4.13
Conversation	(1.43)	1.00	(1.67)	0.17
Relational database	2.86	1.25	0.67	1.57
Functions of a request server (a specific Cerner server)	0.71	3.20	(0.67)	2.23
General functions of servers as a class	0.00	(0.50)	1.33	(0.20)
Describe conversation	2.57	2.30	1.00	2.23
Describe tables and keys	4.00	1.10	0.00	1.67
Describe Purpose of servers	3.00	1.15	1.67	1.63
Mean Change In Pre-Post Test Scores	14.57	14.50	3.67	13.43
Overall Increased % Correct	18%	18%	5%	17%

Deltas did decrease for some questions. A decreased delta could have been indicative of 'unlearning'—the change that occurs when one discovers that their current model is incorrect and they are groping towards an improved model but have not yet discovered it. This was most likely to be happening for individuals with some industry standard knowledge who discovered that Cerner's specific model differs a little from theirs.

In the version of *HNAM DataQuest* tested by learner-players (the beta version), players were required to build the conversation three times–once for every round of play. They complained that this was tedious and frustrating. This may have been a factor in the small delta all three groups received for the multiple-choice question on the conversation. The decreased delta for

this question may also have been indicative of a poorly designed question. Regardless of fluctuations on the multiple-choice question, learner-players steadily improved their essay responses for the conversation as length-in-play increased (1.00, 2.30 and 2.57 respectively). Therefore, individuals in all three length-of-play groups became more capable of explaining the conversation clearly and precisely as they played this game longer. Learner-players' increased their ability to describe Cerner's implementation of the conversation as they spent more time playing *HNAM DataQuest*.

Also, mean scores improved with increased length-of-play for the relational database factual recall question–0.67, 1.25, and 2.85 for 30, 60, and 90 minutes of play respectively. They also improved for the relational database essay question. Learning was happening on the relational database questions.

Mean change in responses to server questions, both multiple choice and essay questions, were more variable. Knowledge of servers was the last concept introduced in the final round of play. Therefore, this variability in deltas may have been indicative of the stage in which players were when they were tested – exhaustion, concentration, and/or unlearning. The variability here may have been indicative of the need for additional rounds of play or, for the 30 and 60 minutes-of-play groups, it may mean that some portion of the learner-players had not yet reached the more advanced level of play which introduced the concept. At 60 minutes-of-play some players had completed all three rounds with exhilaration while others were just finishing their second round, were experiencing stress, and had not yet delved into servers. This variability of stages may have been reflected in the test scores for this concept.

In addition, groups that played longer increased the mean number correct for all three essay questions and not just for the key concept of the level that they had most recently played. In other words, they demonstrated the ability to remember and reuse concepts learned in the first round of play even at later stages of play. Overall change in responses to specific questions appeared to be directly related to the segments of play that each learner-player had experienced. Therefore, *HNAM DataQuest* taught the concepts that it purported to teach.

Can Learning Occur Through a Computer-based Game Whose Format Matches the Learning Content Desired?

This section considers whether it is possible to create a learning game that teaches specific content delivered in the game format, which has been designed to specifically teach instructional outcomes. This is a review of the match between design intentions and results.

HNAM DataQuest's chosen game format (strategic game play with consequences for failure) and chosen instructional design strategies intentionally matched the desired knowledge outcomes relative to Cerner's information system. The pre- and post-tests were designed to test both factual recall knowledge through multiple-choice questions and application of broad understanding through essay questions about Cerner's specific information system knowledge (the learning outcomes). Testing was controlled such that only game play could have been the learning intervention causing change in responses on the pre- and post-test. Differences between pre- and post-test scores indicated that learning did occur.

Therefore, *HNAM DataQuest* provided the learning intended. If this game could be designed to teach specific skills and knowledge then other games can also be designed to teach content-

specific skills and knowledge. Therefore, learning can occur through computer-based games whose game format and instructional strategies match the outcomes desired.

Is 'length (In Minutes) of Play' a Factor That Influences, Whether or not Learning Occurs?

This section considers whether or not length (in minutes) of play increases or decreases the amount of learning (the delta) that occurs through game play. Since length-of-play and date of play were the only variable between all three groups and since game-play was controlled and measured by pre-post testing and observation, any difference between groups would indicate whether length-of-play effected learning or not.

Players in the 30 minutes-of-play group were tested after playing one scenario or round of HNAM DataQuest. The first round of play for all learner-players was essentially the concept of Cerner's three-tiered client/server architecture as it uses servers in the conversation between the user's data request and the database in providing the requested data. On Table E, the 30 minutes-of-play group showed some improvement in both factual knowledge and essay responses of the Cerner's three-tiered client/server architecture and an increased understanding of servers. However, the change was not strong.

In the 60 minutes-of-play group, there were changes indicative of learning occurring in all conceptual areas that individual learner-players had touched. Results were somewhat mixed as both learning and unlearning appear to be still occurring. In addition, at 60 minutes of play some individuals had not yet touched the more detailed information about server functions and had only played the database portion of the game once while other had completed all three rounds. However, change did appear both in the factual recall on multiple-choice questions and in application of broad-based knowledge through the essay questions. Overall, changes were more than two times larger than found in the 30 minutes-of-play group.

In the 90 minutes-of-play group, all players had completed all three scenarios or rounds of play. After completing three scenarios, learner-players showed a stronger and deeper understanding of the architecture as shown increased delta indicative of stronger Cerner-specific responses on the essay questions. Responses to essay questions were much stronger in this group indicating an increased confidence in their ability to explain Cerner's specific information systems concepts.

For the 90 minutes-of-play group, there was still some confusion in some areas. As previously discussed these variations may have been attributable to many causes including

(a) exhaustion after 90 minutes of intense play;
(b) unlearning, and
(c) frustration with a design flaw requiring learner-players to rebuild one aspect of the game three times (once for each round of play).

Subsequent releases modified the design flaw.

Overall, the deltas for changes in minimum score, maximum score, median, mean score, mean percentage correct, and standard deviation all improved as time-in-play increased. Each increased dramatically between 30 and 60 minutes-of-play and then levelled off slightly between 60 and 90 minutes-of-play. The slight levelling may be due to several factors:

(a) The experience level of the 90 minute of play group was higher and their pre-test scores were higher than other groups were, which would indicate that the 90 minutes-of-play group may have been fine-tuning knowledge rather than developing new knowledge as the 60 minutes-of-play group with its less experienced associate base would have needed to, or
(b) The exhaustion factor for the 90 minutes-of-play group, or
(c) Frustration due to the design flaw previously mentioned, and/or
(d) the possibility that 60 minutes of game play is an 'ideal' length for learning from gaming.

As indicted on Table D, the significance of the change in scores increased as time in play increased. This significant change related to length-of-play indicated that length-of-play impacts the amount of learning acquired through game-play. The longer a learner-player played *HNAM DataQuest* the stronger their learning was.

Further Study Needed:

This section considers future studies needed to strengthen the body of knowledge around game playing as an instructional methodology. Suggestions are provided for re-creating these findings in more scientific studies and for studying factors not controlled in this study.

As indicated, literature about scientific studies and research on learning through game-play are limited even though educational practitioners value game-play as an effective methodology. Some studies focused on children as learners. Other studies were literature reviews. Still others used marketed games and self-reporting without controlling for or measuring changes in specific knowledge for each game.

This study controlled for change in knowledge over three different lengths-of-play for an audience of adult learners who demographically reflected the intended audience within one company, Cerner Corporation. The game design was intentionally created to deliver specific knowledge of Cerner's unique information system concepts. Testing was controlled to ensure that the game intervention was the only learning intervention that could have caused the change. In addition, length-of-play was controlled in order to determine whether knowledge increased with increased length of play.

Recreating these conditions with a different audience and/or different game in scientifically controlled situations would provide needed reinforcement that gaming can truly create learning. Without additional studies that provide similar controls and more rigorous audience controls, the results of this study are merely suggestive that learning does occur.

Additional testing needs to occur for game playing where such tests control for other factors such as gender, age, difference in learning and thinking modes, and use of violence in play. It would also be valuable to study different types of game formats ('twitch' versus 'strategy') where those formats match the intended learning points as closely as HNAM DataQuest's design matched its intended outcomes.

In addition, the difference between 60 and 90 minutes-of-play were not as large as those between 30 and 60 minutes-of-play. Therefore, it would be useful to determine whether there is an 'ideal' length of play that garners the greatest learning.

References

McMullen, D. (1987). *Drills vs. Games - Any Differences*? A Pilot Study. ERIC #ED335355.

Dempsey, J.V., Lucassen, B. A., Haynes, L.L. and Casey, M.S., (1996). *Instructional Applications of Computer Games*, paper presented at 1996 annual meeting of the American Educational Research Associate, ERIC #394500.

Wolfe, J. (1997). *The Effectiveness of Business Games in Strategic Management Course Work.* Simulation and Gaming Journal, v28 (4): p360-76, Dec. 1997.

Merrill, M. David, (1999). *What Motivates the MTV Generation*? Some comments on motivation, http://www.coe.usu.edu/it/id2.

Caine, R. M. and Caine, G. (1004). *Making Connections: Teaching and the Human Brain*, Alexandria, VA: ASCD. 1994. 0201490889

Merrill, M. David. Instructional Transaction Theory (ITT): *Instructional Design Based on Knowledge Objects, Instructional-Design Theories and Models: A New Paradigm of Instructional Theory*. Mahwah, NJ: Lawrence Erlbaum Associates and http://www.coe.usu.edu/it/id2.

Gagne, R., Briggs, L. Wager, W. (1992). Principles of Instructional Design (4th Ed.). Fort Worth, EX: HBJ College Publishers.