Journal of Information Systems Education

Volume 32
Issue 2
Spring 2021

Teaching Tip

Emerge2Maturity: A Simulation Game for Data Warehouse Maturity Concepts

Michael V. Mannino, Mohammed Khojah, and Dawn G. Gregg

Recommended Citation: Mannino, M. V., Khojah, M., & Gregg, D. G. (2021). Teaching Tip: Emerge2Maturity: A Simulation Game for Data Warehouse Maturity Concepts. *Journal of Information Systems Education*, 32(2), 77-91.

Article Link: https://jise.org/Volume32/n2/JISE2021v32n2pp77-91.html

Initial Submission:	3 March 2020
Accepted:	28 September 2020
Abstract Posted Online:	13 March 2021
Published:	9 July 2021

ISSN: 2574-3872 (Online) 1055-3096 (Print)

Teaching Tip
Emerge2Maturity: A Simulation Game for Data Warehouse Maturity Concepts

Michael V. Mannino
The Business School
University of Colorado Denver
Denver, CO 80217, USA
Michael.Mannino@ucdenver.edu

Mohammed Khojah
Faculty of Economics and Administration
King Abdulaziz University
Jeddah, Saudi Arabia Kingdom
mmkhojah@kau.edu.sa

Dawn G. Gregg
The Business School
University of Colorado Denver
Denver, CO 80217, USA
Dawn.Gregg@ucdenver.edu

ABSTRACT

This paper describes an innovative approach for teaching the challenges in the management of data warehouse development. The approach contains lecture material providing conceptual background about the management of data warehouse development, a simulation game supporting experiential learning, and a post-play debriefing to support synthesis of conceptual material and experiential learning. The simulation game, Emerge2Maturity, addresses learning challenges faced by students as they experience development over time, determine capabilities to balance costs and benefits for consistency with an organization's strategy, observe organizational learning effects on costs and benefits, and gain awareness of the impact of external events. To support decision-making by players and address these learning challenges, Emerge2Maturity uses two novel models: the Capability Assessment Model for choices about data sources subject to budget and resource constraints and the Configuration Model for transition among decision-making phases involving constraint levels, learning effects, and external events. Simulation in each phase and phase summaries provide opportunities for players to reflect about their progress in developing a data warehouse. Initial evaluation of Emerge2Maturity in a data warehouse course demonstrated the potential to improve instruction about maturity concepts pertinent to data warehouse development in organizations.

Keywords: Game-based learning, Data warehouse, Maturity model, Simulation

1. INTRODUCTION

Business simulation games have a long history in management education. Since the Beer Game (Anderson and Morrice, 2000), instructors have realized the benefits of using games to deliver knowledge and skills to students. Simulation games are effective in teaching students about business processes and the impact of information technology (IT) (Monk and Lycett, 2016). Simulation games facilitate learning about strategy, collaboration, integration, and development maturity (Leger, 2006), processes difficult to grasp using traditional learning methods without practice and experience. As evidence of the growing importance of games in business, Harvard Business Publishing (https://hbsp.harvard.edu/simulations-feature/) features many business simulation games, and Business News Daily reviewed the best business simulation games in 2019

(https://www.businessnewsdaily.com/10717-best-business-simulation-games.html).

As critical infrastructure for business intelligence, data warehouse projects involve large expenditures and high risk. A typical data warehouse project involves large capital investment, typically more than $1 million in just the first year (AbuAli and Abu-Addose, 2010). Business intelligence projects usually take 12-36 months to complete (Amin and Arefin, 2010). However, these projects fail at considerably high rates: 70-80 percent according to Inmon (2001), 90 percent according to Conning (2000), and 60-85 percent according to Patrizio (2019). Typical reasons for failure are lack of management support, poor data quality (especially data integration), ambitious project scope and deadlines, and lack of plans for long term maintenance (Merrick, 2014; Mitchell, 2017).

Although many university courses cover management of data warehouse development, traditional learning approaches fail to capture the complexity and challenges that occur in real situations. A 2019 review of syllabi of 20 data warehouse courses showed no active or experiential learning for management of data warehouses. Courses covered the conceptual background about data warehouse architectures, lifecycles, success factors, project team capabilities, and sample data warehouse designs. Course syllabi also omit coverage of important academic theories of management of data warehouses such as architecture selection factors (Choudhary, 2010) and maturity models (Sen, Sinha, and Ramamurthy, 2006; Sen, Ramamurthy, and Sinha, 2012).

This paper presents an innovative educational approach to address shortcomings in instruction about the management of data warehouse development. The three-part approach involves the conceptual background about key management topics, a business simulation game for experiential learning, and a post-play discussion and survey for reflection about the conceptual background and game-play. As a business strategy game, Emerge2Maturity involves alignment of an organization's business intelligence strategy with its data warehouse capabilities. Emerge2Maturity uses the Capability Assessment Model, a novel decision model, to evaluate cost-benefit tradeoffs among player choices for resources. Emerge2Maturity employs the Configuration Model to revise constraint levels and resource coefficients based on architecture evolution and the occurrence of events in each decision phase. Simulation in each phase and phase summaries provide opportunities for players to reflect about progress in developing a data warehouse. Initial evaluation of the student surveys for post-play in a data warehouse course provided evidence of satisfaction of learning objectives and suggestions to improve instruction about the management of data warehouse development.

This teaching approach, available from the primary author, contributes to both practice and theory. Emerge2Maturity is the first simulation game to address learning difficulties about the management of data warehouse development. Game-play of Emerge2Maturity, encapsulated in course materials to provide background and critical thinking exercises, provides students an enhanced learning experience. The novel analytical engine of Emerge2Maturity provides a concrete approach to help players evaluate tradeoffs among resource levels and strategy as an organization evolves to a mature state.

This paper continues as follows. The second section covers the first part of the instructional approach, background about maturity models, architecture selection, and learning curves presented as a lecture. The third section presents the design of Emerge2Maturity with an emphasis on models and game flow. The fourth section covers the remaining parts of the instructional approach with screen snapshots showing a student's experience when playing a game, a debriefing discussion, and a survey with student responses about the match between learning objectives and the design of Emerge2Maturity. The fifth section summarizes the paper and identifies future extensions.

2. LESSON PLAN AND CONCEPTUAL BACKGROUND

The lesson plan about the management of data warehouse development involves a module (second or fifth module in a 16-week semester) in a data warehouse course. The lesson plan can also be used in a business intelligence course with coverage of management of data warehouse development. The lesson plan contains a lecture about management of data warehouse development, usage of Emerge2Maturity after a demonstration about its features and design, and a debriefing discussion to help students synthesize conceptual material and game-play. The last two parts of the lesson plan provide the simulated experience with development decisions as well as a reflection about the lecture and game-play.

The lecture part of the lesson plan covers development challenges, data warehouse architecture selection, and maturity models as summarized in Table 1. The first part of the lecture provides background on the challenges of data warehouse development and the high failure rates and emphasizes learning curve theory to explain difficulties in data warehouse development. The lecture emphasizes factors identified in Figure 1 as unique difficulties for data warehouse development. Adelman (2012) identifies intangible benefits contributing to the difficulty to justify data warehouse investments until an organization learns to measure intangible benefits after some years. High reported rates of failure for data warehouse projects provide evidence about the learning difficulties that organizations face.

Area	Topics	References
Development challenges	Reasons for development difficulties, Project failures, Learning curve background	Deng, 2005; Plaza, Ngwenyama, and Rohlf 2010; Lapre and Nembhard, 2011
Business architecture choices and selection	Business factors, Common architectures, Architecture selection	Ariyachandra and Watson, 2010; Choudhary. 2010; Kimball and Ross, 2013
Maturity models	Stages of growth, Capability assessment	Paulk et al., 1995; Sen, Sinha, and Ramamurthy, 2006; Becker, Knackstedt, and Pöppelbuß, 2009; Sen, Ramamurthy, and Sinha, 2012; Spruit and Sacu, 2015; Carvalho et. al., 2019

Table 1. Summary of Lecture Areas

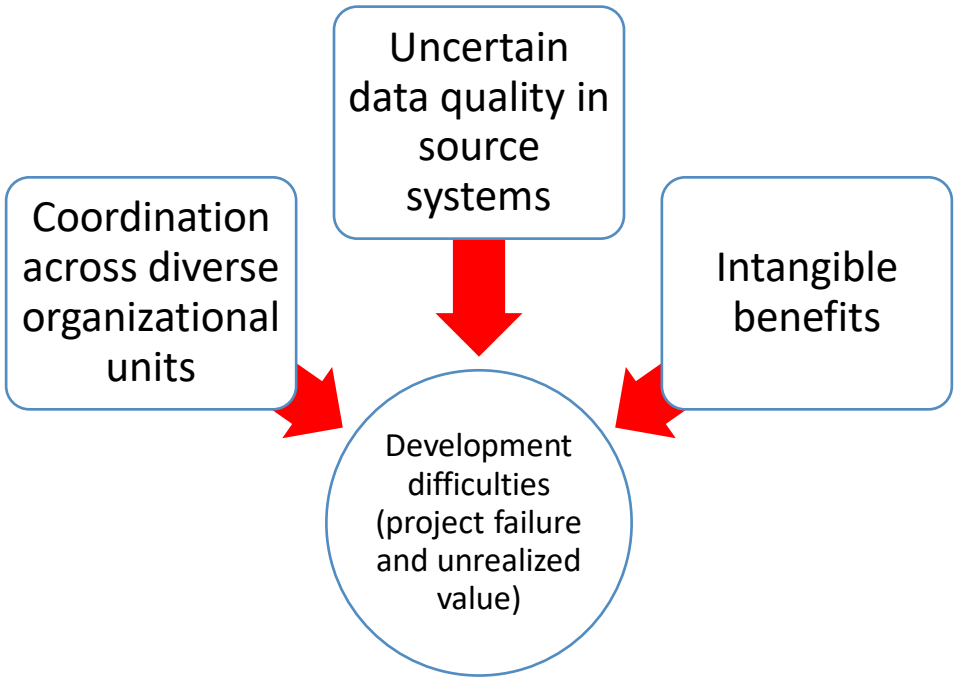

Figure 1. Factors Influencing Difficulties of Data Warehouse Development

Learning curves provide insight to understand intangible benefits and high costs due to uncertain data quality and coordination efforts in data warehouse projects. The lecture summarizes pioneering work by Wright (1936) on production costs in the aircraft industry as well as information technology applications on ERP systems (Plaza, Ngwenyama, and Rohlf, 2010), software development (Pendharkar and Subramanian, 2004), and help desk support (Deng, 2005). Kimball and Ross (2013) indicate that data warehouse projects have a steep learning curve. Merrick (2014) and Lindsey and Frolick (2003) provide several reasons that building a data warehouse may involve learning challenges for an organization.

The lecture hypothesizes two learning curves to explain the difficulties of data warehouse development. The business value learning curve (Figure 2a) shows a hypothetical relationship between deployment time of a data warehouse in an organization and the business value derived from its usage. The key insight from the business value learning curve is the initial difficulty to create high value from combining data sources. The data transformation learning curve (Figure 2b) shows a hypothetical relationship between data warehouse deployment time and transformation cost to resolve data quality problems. The curve provides insight about high fixed costs to discover and resolve unknown data quality problems during the initial period of usage.

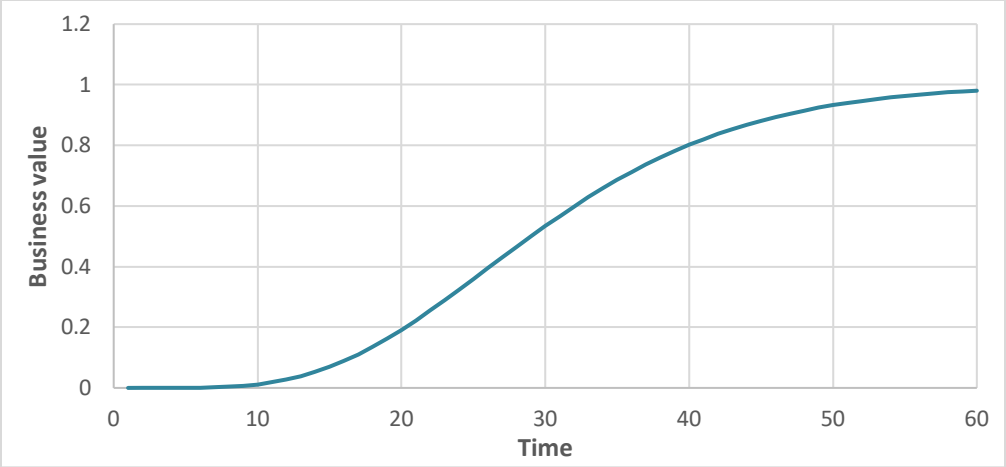

Figure 2a. Business Value Learning Curve

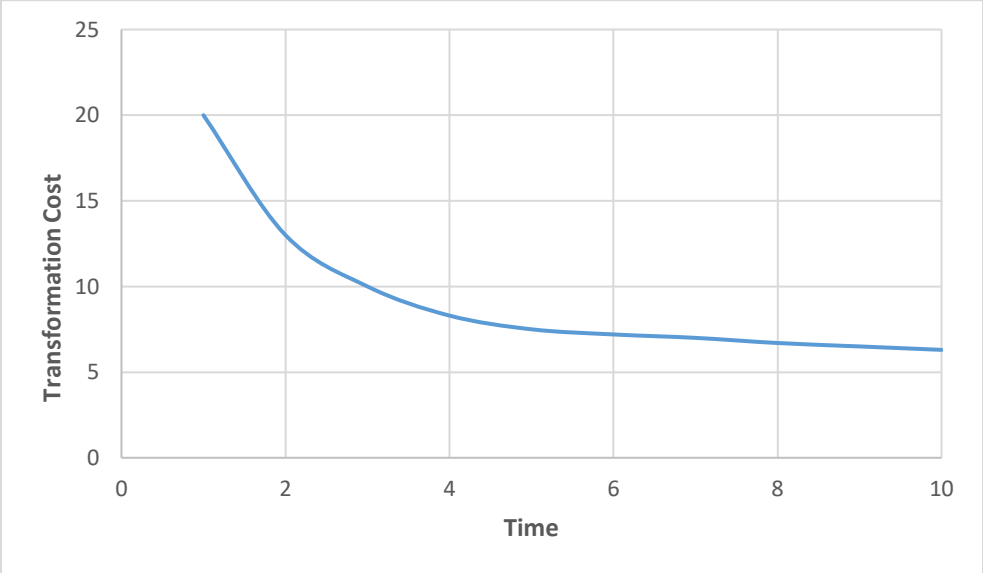

Figure 2b. Data Transformation Learning Curve

The remainder of the lecture emphasizes two management approaches for data warehouse development: architecture selection and maturity models. For architecture selection, the lecture presents four prominent business architectures (enterprise data warehouse, data mart, data mart bus (Kimball and Ross, 2013), and federated) along with a summary of influencing factors from studies about architecture selection. The lecture summarizes the factors indicated by Choudhary (2010) (resource constraints, information technology skills of staff, need for data integration, and perception of data as a strategic resource) and Ariyachandra and Watson (2010) (information interdependence, task routineness, and the level of sponsorship).

The lecture finishes with coverage of maturity models to evaluate progress over time for technology capabilities and deployment (Carvalho et. al., 2019). Following Becker, Knackstedt, and Pöppelbuß (2009), the lecture indicates that a maturity model contains a sequence of levels for the evolution of objects developed in discrete stages. For data warehouses, the lecture summarizes capability maturity models for data warehouses proposed by Sen, Sinha, and Ramamurthy (2006), Sen, Ramamurthy, and Sinha (2012), and Spruit and Sacu (2015) extending the Capability Maturity Model (Paulk et al., 1995), a well-known maturity model for software development. Each model applies the five levels of the Capability Maturity Model but develops different key process areas and features, unique for development of data warehouses.

3. DESIGN OF EMERGE2MATURITY

Data warehouse development is a complex process involving several related factors and extended periods to reach a stable solution. Organizations need to align capabilities with architecture selection and balance benefits and costs to operate a data warehouse. To address these complexities, Emerge2Maturity decomposes data warehouse development into decision-making phases using common factors across organizations. The game helps players grasp tradeoffs between

costs and benefits of acquiring capabilities while observing an organization's strategy for data warehouse development.

This section presents the design of Emerge2Maturity with an emphasis on design decisions and models. The design decisions indicate the scope of the game, while major models support a simplified representation of data warehouse development into game-play.

3.1 Game Design Decisions

The foundation of Emerge2Maturity involves decisions in four design areas as summarized in Table 2. The design of Emerge2Maturity followed an iterative process with simplification and decomposition of design alternatives.

Design Area	Design Decision
Organization specificity	Organization-independent using data source categories
Scope of decisions	Allocate resources for capabilities and observe strategy
Game duration	Multiple periods with learning effects and events
Cost benefit model	Demand-driven decisions about extraction, transformation, and integration

Table 2. Summary of Design Decisions

Business simulation games can be organization-specific or independent. Organization-specific games typically occupy a large part of a course, while organization-independent games support one or two modules of a course. Organization-specific games typically involve much complexity, while organization-independent games emphasize simplicity. Organization-specific games, such as the FinGame (Brooks, 2007), involve a hypothetical company with simulation and game features covering skills in financial management and analysis of financial statements. In contrast, the Beer Game (Sterman, 1989) is organization-independent, focusing on the BullWhip effect (Croson and Donohue, 2006), a symptom of coordination problems in managing a supply chain.

Since instruction about the management of data warehouse development typically involves a limited part of a data warehouse course, Emerge2Maturity is an organization-independent game with a simplified representation of data sources in categories. Data source categories involve common features with implications about costs and benefits of utilizing data sources in a data warehouse. Players focus on key aspects of data sources, making resource decisions with economic consequences for an organization.

To simplify player choices and model development, Emerge2Maturity involves capability assessment in resource decisions made by players. Players observe strategy elements related to capabilities as the game evolves. Strategy elements in Emerge2Maturity involve the number of phases in a game and the progression of constraints on budgets and resources as an organization matures in its deployment of a data warehouse. When transitioning among phases, Emerge2Maturity allows players to observe the impacts of learning difficulties and events. Because data warehouses typically mature over a long period, events (both internal and external) can affect budgets and resource limits.

Emerge2Maturity uses cost-benefit analysis, a simple economic model used in previous research about data sources. Rao and Osei-Bryson (2008) proposed a cost-benefit model that maps data sources to decision support views with estimated benefits to a firm and measurable quality levels. Ballou and Tayi (1999) developed a cost-benefit model to determine the quality level that provides maximum value. Emerge2Maturity uses a demand-driven approach that anticipates demand first and then selects related data sources (Winter and Strauch, 2002). This approach minimizes the risk of including data sources that might not be beneficial to decision-makers but also increases chances of missed opportunities.

In Emerge2Maturity, cost-benefit analysis applies to decisions about the extraction, transformation, and integration of data sources. Each capability adds value to queries and reports, but also involves fixed and variable costs. The first step toward benefiting from data sources involves extraction into temporary storage. The model in Ballou and Tayi (1999) selects data sources for relevance in decision-making. The second capability to increase the value of data sources involves data transformation to enhance data quality for accuracy, completeness, and timeliness (Ballou and Tayi, 1999). Transformation also resolves inconsistencies, applying business rules and summarization (Watson, Goodhue, and Wixom, 2002). The third capability, data integration, strongly influences the architecture of a data warehouse (Ramamurthy, Sen, and Sinha, 2008). Integration applies consistency rules and matches entities across data sources forming a single point of truth (Gulledge, 2006).

3.2 Game Flow and Model Development

Emerge2Maturity provides decision-making over multiple phases as depicted in Figure 3. In each phase, players make sequential or joint decisions about capabilities for extraction, transformation, and integration as represented in the middle box of Figure 3 (Manipulate Capability Decision Variables). Players attempt to maximize profit using details about costs, benefits, and constraints. The demand for information assets provided by capabilities is stochastic so players deal with uncertainty in assessing capabilities. After making choices for decision variables, a player can simulate and modify choices. After a limited number of choices, a player must commit choices for decision variables.

The game evolves over multiple phases representing budgeting or decision-making periods. The game controller makes the transition to the next phase of a game with learning effects and events as shown in Figure 3. The learning effect progresses over the phases, affecting coefficients for costs and benefits. Events influence coefficients and constraints on capabilities. A game terminates after a specified number of phases when an organization reaches its highest maturity level. The Capability Assessment Model (CAM) supports decision-making in each phase, while the Configuration Model (CM) provides details about phase transition.

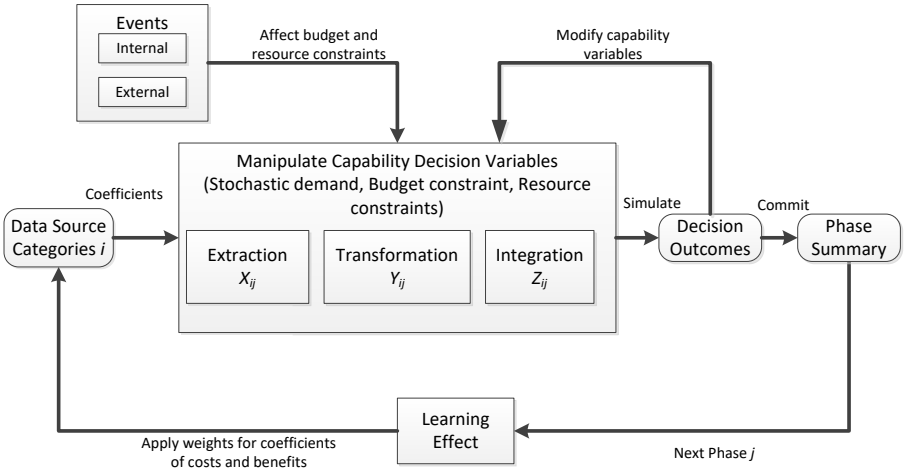

Figure 3. Game Flow in Emerge2Maturity

3.2.1 Data source categories. Emerge2Maturity uses features for technology, complexity, and size to define data source categories. Categories facilitate determination of cost and benefits of individual data sources as all data sources in a category share features. Technology level ranges from legacy systems to modern systems with the level determined by features of programming language, database management system, operating system, and hardware platform. Complexity involves the difficulty to transform diverse data for decision-making. Complex data requires extensive time and effort to analyze. Size involves the processing effort for data such as the number of rows. Larger data size requires additional storage and maintenance.

Categories determine coefficients for production, cost, benefit, and risk. Table 3 depicts the relationship of features to model components. The complexity and size of a data source determine the amount of production, variable cost, benefit, and risk. Technology and data size determine the fixed cost. Usage of data source categories and features is a definitive part of Emerge2Maturity.

3.2.2 Capability Assessment Model (CAM). The Capability Assessment Model (CAM) provides an optimization model for decision-making in Emerge2Maturity. The CAM is an educational model to demonstrate relationships among important variables of data warehouse capabilities. Figure 4 shows components of the CAM with decision variables, functions, and coefficients.

The CAM manipulates three decision variables (number of data sources *X*, transformation level *Y*, and integration level *Z*) used in processes for extraction, transformation, and integration affecting an organization's capabilities. Extraction involves selecting data sources and transporting data to include in a data warehouse. Transformation involves increasing data quality through operations on individual data sources. Integration involves combining data from different sources, matching, and consolidating common data. For each decision variable, Δ represents the incremental capabilities added in a phase.

The CAM uses stochastic demand, common in models in operations management (Miranda and Garrido, 2004; Schmitt, Snyder, and Shen, 2010) and econometrics (Browne and Zipkin, 1991; De Castro, Tabucanon, and Nagarur, 1997; Ben-Daya and Hariga, 2004). Demand is a function of production plus risk. Expected demand is the production level determined by values for decision variables and the uncertain risk or error term. Risk is modeled as a Normal distribution with a mean of 0 and standard deviation of *r*, a function of the features of a data source category.

The optimization model maximizes profit for each data source category subject to constraints on the budget for total costs, minimum capability levels (number of data sources, transformation level, and integration level) for each data source category, dependency of integration on transformation for each data source category, and maximum capability levels (number of data sources, transformation level, and integration level) for each data source category. Solving the model involves expected demand without the risk term.

	Model Component				
Feature	**Production (P)**	**Fixed Cost (FC)**	**Variable Cost (VC)**	**Benefit (B)**	**Risk (R)**
Technology		√			
Complexity	√		√	√	√
Size	√	√	√	√	√

Table 3. Mapping of Category Features to CAM Components

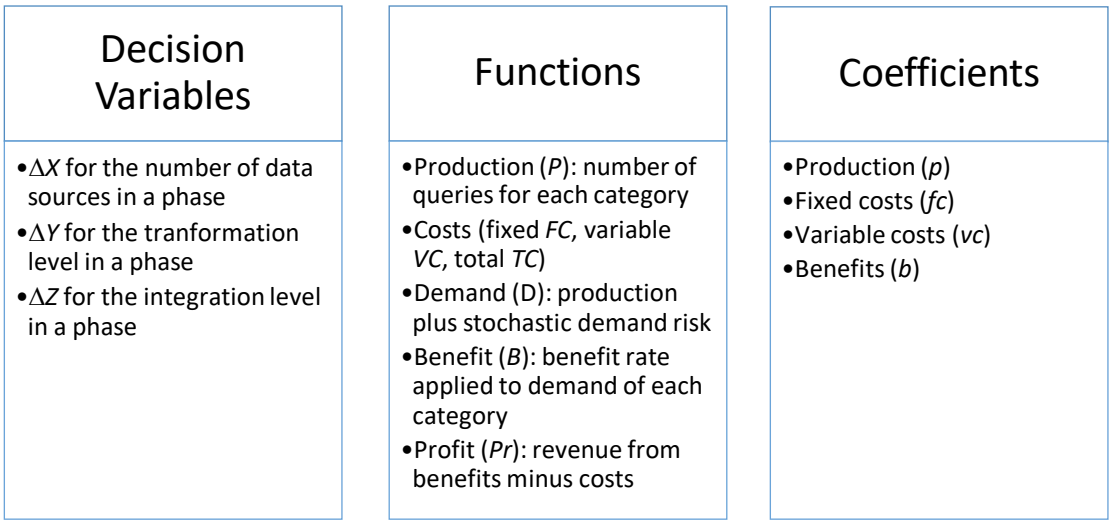

Figure 4. Elements of the Capability Assessment Model (CAM)

Source	Event	Impact
Internal	Minor acquisition	Increase number of data sources in selected categories
	Major merger	Add one or more data source categories
	Minor divestment	Decrease number of data sources from one or more categories.
	Major divestment	Drop a data source category and one or more data sources from some other categories
External	Recession	Decrease budget constraint
	Expansion	Increase budget constraint
	Minor regulation	Add a category with low benefits and a small number of required data sources
	Major regulation	Add a category with low benefits and a larger number of required data sources

Table 4. Impact of Events in Emerge2Maturity

The dependency on the number of data sources in the profit function adds considerable complexity for solving the model. To simplify a player's choices, Emerge2Maturity supports sequential choices for extraction, transformation, and integration. Initially, a player chooses the number of data sources satisfying relevant constraints. After a selection, a player selects the transformation level satisfying the relevant constraints using the selected number of data sources. After selecting the number of data sources and the transformation level, a player selects the integration level satisfying the relevant constraints.

3.2.3 Configuration Model (CM). Configuration of a phase involves revised levels for constraints about budgets and capabilities, revised weights applied to coefficients for costs and benefits, and random events that influence budget constraints. Constraint levels are determined dynamically based on organizational strategy and capabilities achieved in previous phases.

Coefficients for costs and benefits have base values. Cost and benefit coefficients change during the game based on organizational learning. Weights are applied for capability costs and benefits to reflect learning effects. As an organization acquires capabilities, it becomes more efficient with decreasing costs and increasing benefits for data sources. Emerge2Maturity uses the power law function, adapted from Wright (1936), to adjust weights applied to coefficients.

Events involve actions with long-term consequences, initiated internally or externally. An internal event involves actions within an organization such as a merger or divestment. An external event involves actions in an organization's environment such as a recession, regulation, or litigation. An organization reacts to events by adjusting their strategic view or capabilities. Emerge2Maturity uses a small set of random events as shown in Table 4. If an event occurs, the Configuration Model randomly adjusts constraints for data sources in a category, the budget constraint, or the number of data source categories.

4. GAME DEMONSTRATION AND DEBRIEFING

This section demonstrates the game interface showing results from an actual game-play along with discussion topics and the survey used after students finish game-play. Emerge2Maturity provides a web interface on standard browsers. Although Emerge2Maturity supports both educational and competitive games, this section only demonstrates the interface for competitive play. The educational mode provides additional

assistance to prepare players for competitive play. After the game demonstration, this section presents the debriefing discussion and survey to help students connect game-play with concepts about the management of data warehouse development.

4.1 Game Overview and Demonstration

Before starting game-play, students receive an overview of the concepts underlying Emerge2Maturity followed by a demonstration of game-play. Table 5 lists the topics covered in the overview before students begin game-play. After this presentation, students should be able to discuss the purpose of Emerge2Maturity, the role of data source categories, and game flow supported by Emergy2Maturity.

To begin game-play, a player chooses a game and begins in Phase 1. Emerge2Maturity provides games with several skill levels based on the number of phases, constraint levels, and category features. At the beginning of a phase, Emerge2Maturity displays the constraint levels (budget and resource levels for data source categories) and the features of the data source categories. Figure 5 shows three categories with constraints limiting each category to a maximum of 5 data sources, 30% transformation level, and 30% integration level. The feature table, below the constraint table, shows levels of technology, complexity, and size for each category as well as the maximum number of data sources. For example, category 1 has high technology, medium complexity, high size, and 20 data sources.

In Phase Simulation for Extraction, a player selects the number of data sources for each category that maximizes expected profit (Figure 5). For each choice, a player uses the simulation button to observe potential results from an uncertain demand. A player has a small number of attempts with the simulation before committing to an answer. After committing to an answer, Emerge2Maturity displays costs (expected and optimal) and profit (optimal, expected, and simulated) in bar graphs. Figure 6 shows committed choices of five data sources for Category 1, three data sources for Category 2, and four data sources for Category 3. A player then continues to the transformation and integration decisions.

At the end of a phase, Emerge2Maturity saves a player's decisions and outcomes and then initiates the next phase. The Phase Summary page shows expected costs and profits based on a player's choices for capabilities for each data source category. As a reference, the Phase Summary page also shows the optimal costs and profits. Figure 7 shows a good result with expected profit from player choices as $21,348.05 compared to the optimal profit of $21,490.15. For more detail, the Phase Summary page decomposes costs and profits by category and decision, showing both expected results from choices and optimal results.

Topic Area	Items and Notes
Business strategy games	Brief review of other business strategy games and learning difficulties addressed by Emerge2Maturity
Game flow and player decisions	Player decisions made in phases and changes occurring in transition among phases
Role of data source categories	Features of data source categories and influence of features on development variables
Examples of learning effects and external events	Simple examples depicting transition in a two-phase game with coefficient changes and external events

Table 5. Summary of Emerge2Maturity Topics

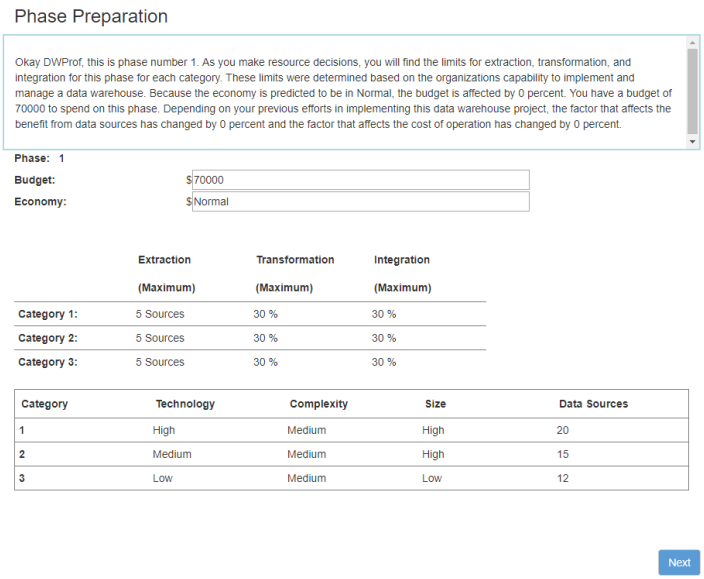

Figure 5. Phase 1 Preparation in Emerge2Maturity

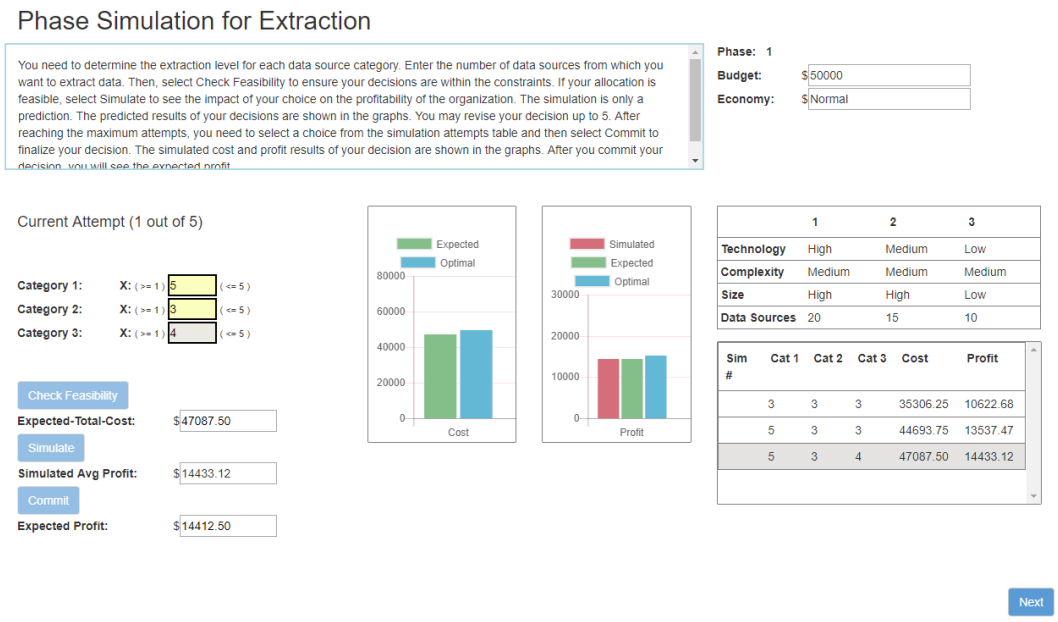

Figure 6. Phase Simulation for Extraction Decisions in Phase 1

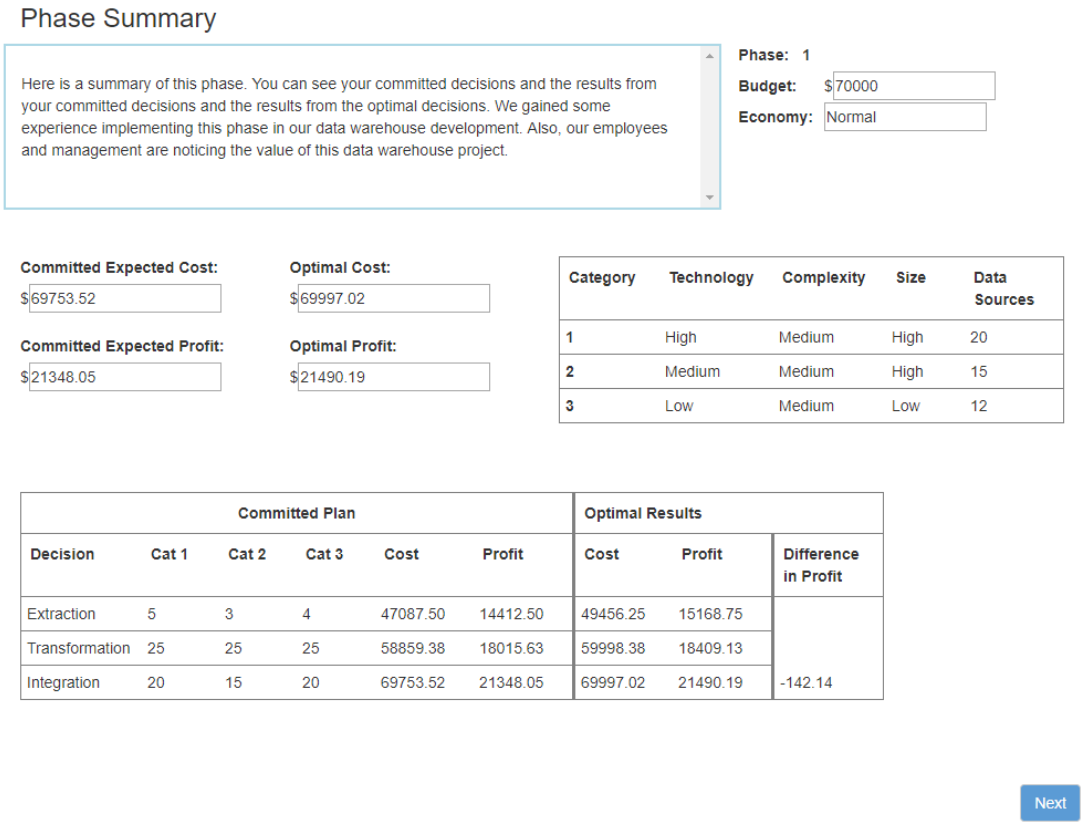

Figure 7. Phase Summary for Decisions in Phase 1

At the end of a game, Emerge2Maturity calculates a numeric score based on the difference between a player's total profit and the optimal total profit. Emerge2Maturity converts the profit difference to a qualitative score displayed on a five-

point scale as shown in Figure 8. In addition, Emerge2Maturity ranks players by score and displays the highest scores in a leaderboard. Emerge2Maturity uses points and a leaderboard to reward players for their accomplishments and encourage additional play.

As this demonstration indicates, Emerge2Mature provides a simulated, educational experience about the management of data warehouse development. Players focus on data sources grouped by important features for technology, complexity, and size. For data source categories, players manipulate capabilities for three related decisions in data warehouse development (extraction, transformation, and integration). Simulation allows players to observe the impacts of a limited number of choices. Phase results compare player choices for capabilities with optimal choices. When transitioning among phases, players observe a learning effect, strategy changes for capability and budget constraints, and impact of external events. A simple point system and leaderboard provide incentives to improve and compete with other players.

4.2 Debriefing Discussion

The debriefing discussion helps students connect game-play to conceptual material about the management of data warehouse development. The debriefing discussion can be done in a classroom or online setting. Before starting the debriefing session, each student plays Emerge2Maturity three times with at least one play using the educational mode and two competitive plays with at least one play having three or more phases. Table 6 summarizes the topics in the debriefing discussion.

The debriefing discussion begins with a summary of game-play. Students are provided several graphs and summary statistics about game-play. Students then discuss strategies employed to increase profits, mistakes made when poor profits occurred, and improvements made as play progressed in later phases.

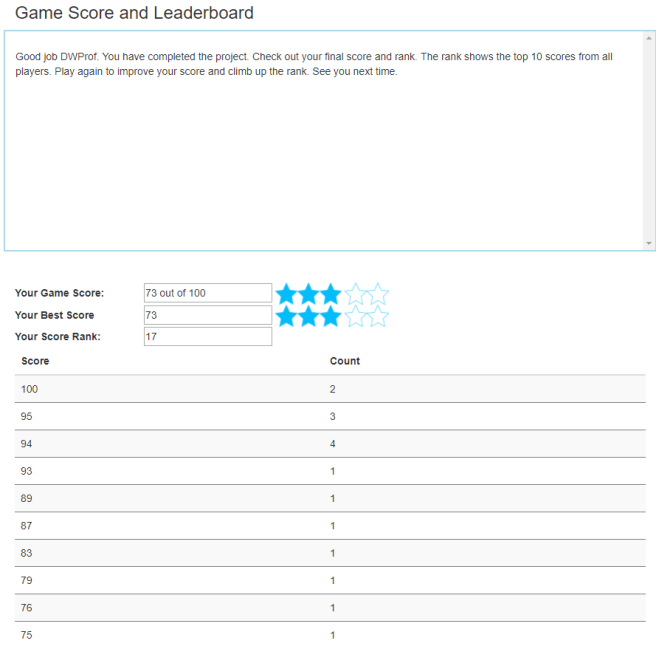

Figure 8. Game Score and Leaderboard

	Discussion Items
Results of game-play	Strategies to increase profits; Mistakes for poor profits; Revised play as phases progress
Student reactions	Insights and difficulties of game-play; Confidence about management data warehouse development projects
Game decisions	Strategy decisions observed in game-play for architecture selection, project funding, and resource constraints; Capability decisions made in game-play for data sources, technology, and personnel
Game elements	Importance of data source features; Measurement of costs and benefits; Impact of external events; Learning effect on data warehouse development
Data warehouse failures	Failures in game-play; Failures in real data warehouse development projects

Table 6. Summary of Discussion Topics

After the summary of game-play, students respond to questions about insights and difficulties encountered in game-play. Students indicate if they have developed more confidence about managing data warehouse development because of game-play.

Students then discuss decisions made in game-play and the game elements. Students distinguish strategy decisions provided in a game versus capability decisions made by players. Students reflect on the realism of game elements, especially the features of data sources along with the measurement of costs and benefits, the impact of external events, and the learning effect on data warehouse development.

The debriefing discussion concludes with questions about data warehouse failures. Students discuss the relationship of poor play strategies to data warehouse failures. To close the loop, students then discuss failures in data warehouse development projects not addressed by game-play.

4.3 Debriefing Survey

After the post-play discussion, students complete a survey to provide feedback for improvement of the learning objectives embedded in Emerge2Maturity. The learning objectives involve decision-making across phases; decision variables for extraction, transformation, and integration; cost-benefit trade-offs; and organizational learning. Table 7 lists the Emerge2Maturity learning objectives and design elements to address these learning objectives.

Gagne's (1970) learning events provide a useful framework to evaluate learning objectives. Table 8 shows Gagne's learning events and associated support in Emerge2Maturity. The survey evaluated the first seven events only ("Provide objectives" through "Feedback"), as these are the events associated with the design of the game. A future study will evaluate the other objectives ("Assess performance" and "Retain learning outcomes") using systematized learning analytics (Serrano-Laguna et al., 2017).

Most survey questions evaluate aspects of the game using a Likert-scale while some questions ask for detailed feedback. Tables 9 and 10 indicate 15 items in the survey corresponding to Gagne events. Several items are open-ended questions such as "Emerge2Maturity provides clear summaries at the end of each phase to help you gain insight about activities performed in the phase" and "Emerge2Maturity makes benefits tangible by calculating profits after each capability decision."

Ninety-nine students enrolled in a data warehouse course over three semesters completed the survey. The data warehouse course has a prerequisite of a first course in databases covering query formulation and database design.

Table 11 shows Gagne's learning events with summaries of responses from students. Students indicated a need to reduce the amount of text in the game, use more graphics, and provide additional details in certain parts of the game. Students suggested graphics and video effects to improve engagement and reduce reading of text. Students also indicated the importance of linking the learning objectives from the beginning of the game with actions in the game. Students wanted more details about the underlying organization and industry.

Learning Objective	Design Element
Use important features to explain costs and benefits of data sources	Features for technology, complexity, and size with impact on costs and benefits
Explain grouping of data sources into categories using common features	Group data sources into categories based on levels of features
Depict complexity of data warehouse development	Decompose a project into phases with standard decisions
Explain common strategy factors in data warehouse development	Use common factors such as budget, phases, and constraints on resources
Apply common capability decisions in data warehouse development	Make decisions about extraction, transformation, and integration efforts
Explain the relationship between strategy and capability	Show change in constraints as game progresses across phases
Understand intangible benefits in data warehouse development	Quantify benefits as a total profit made by the organization
Explain learning effects with increased benefit rates and decreased cost rates over time	Show reduction of costs and increase of benefits with efforts made in previous phases

Table 7. Learning Objectives and Design Elements in Emerge2Maturity

Gagne Event	Support in Emerge2Maturity
Provide objectives (Gne1)	Provide text-based objectives at the beginning of a game
Gain attention (Gne2)	Present the story of a company, create challenges to build a better data warehouse
Link to previous (Gne3)	Provide summaries at the end of each phase and game
Present content (Gne4)	Read about terminology, show the effect of decisions, and assess benefits and risks
Give guidance (Gne5)	Provide help documentation to play the game and give instructions about decisions in a game
Practice opportunity (Gne6)	Simulate their decisions before committing
Feedback (Gne7)	Show results of simulation attempts and committed decisions
Assess performance (Gne8)	Assess knowledge acquisition and skills gained by using pre-test/post-test (effect evaluation)
Retain learning outcomes (Gne9)	Follow up assessment (effect evaluation)

Table 8. Support of Gagne Events in Emerge2Maturity

Item Code	Item Statement
Gne1	Emerge2Maturity provides clear learning objectives in the Welcome page
Gne2	In the Game Preparation page, Emerge2Maturity presents a realistic business situation and creates a challenge to gain learner's attention
Gne3	Emerge2Maturity provides clear summaries at the end of each phase to help you gain insight about activities performed in the phase
Gne5	Emerge2Maturity provides useful help documentation and adequate instructions about playing the game
Gne6	Emerge2Maturity provides a useful simulation feature showing the effect of capability decisions before committing actual decisions
Gne7	Emerge2Maturity provides a useful summary of simulation attempts before committing to a capability decision

Table 9. Gagne's Items except Learning Objectives

Item Code	Item Statement
Gne4-1	Emerge2Maturity provides important features (technology, complexity, and size) that can explain costs and benefits of data sources
Gne4-2	Emerge2Maturity groups data sources into categories using common features
Gne4-3	Emerge2Maturity decomposes the complexity of data warehouse development into a sequence of standard phases
Gne4-4	Emerge2Maturity provides common strategy factors (budgets, phases, and resource constraints) needed in data warehouse development
Gne4-5	Emerge2Maturity provides common capability decisions (levels of extraction, transformation, and integration of data sources) made in data warehouse development
Gne4-6	Emerge2Maturity combines aspects of strategy and capability to help learners understand the relationship between them
Gne4-7	Emerge2Maturity makes benefits tangible by calculating profits after each capability decision
Gne4-8	Emerge2Maturity shows learning effects with increased benefit rates and decreased cost rates over time
Gne4-9	Emerge2Maturity shows the impact of events, such as change in the economy, on the budget for data warehouse development

Table 10. Gagne's Items for Learning Objectives

Gagne Event	Student Perspective
Gne1: provide objective	89 / 99 Agree Separate educational material from game material. Reduce text and use bullet points.
Gne2: gain attention	89 / 99 Agree Provide more details about the company and industry.
Gne3: link to previous	84 / 99 Agree Provide more details and statistical tools.
Gne4: learning objectives	90 / 99 Agree Provide more details that links game actions and the learning objectives. Indicate learning objectives in each corresponding part of the game.
Gne5: give guidance	82 / 99 Agree Provide detailed and short instructions.
Gne6: practice opportunity	92 / 99 Agree Provide more details on costs, benefits, simulated, expected, and optimal values.
Gne7: feedback	91 / 99 Agree Provide more details in graphs or additional text and statistical tools for analysis.

Table 11. Summary of Findings

Almost all of the individual learning objectives received some criticism by not being clearly addressed in the game. Some learning objectives seem difficult to observe. For example, the effect of events on business strategy happens randomly in the game. To maximize the benefits from using the game, players must play several times to observe the effect. However, students wanted more elaboration of learning objectives during the game, not just in the game concept page. Students also indicated that the relationship between costs and benefits and all its related aspects needs improvements throughout the game. Improvements can provide other factors that may affect the costs and benefits, such as data quality, elaborating the link between features and categories, and increasing the number of phases. Adding more text and phases to the game might increase the chances for a better learning experience, but the length of the game and the complexity of the text might also decrease the motivation to play.

5. CONCLUSION AND FUTURE WORK

This paper presented an instructional module to improve student learning about the management of data warehouse development. The module contains a lecture about the conceptual background, game-play with the simulation game, Emerge2Maturity, and a post-play evaluation of learning. The lecture part of the module covers development challenges, data warehouse architecture selection, and maturity models to provide the conceptual background for students. As the most innovative and important part of the module, Emerge2Maturity provides a simulated experience of capability decisions for data warehouse development using the Capability Assessment Model and the Configuration Model. The post-play discussion and survey provide an opportunity for students to reflect on the conceptual material and the simulated experience with capability decisions.

Extensions of Emerge2Maturity involve gamification and the Strategy Assessment Model. Although the current version of Emerge2Maturity includes gamification elements, such as the story narratives and a leader board, additional features will be beneficial to add. Gamification elements will follow suggestions by Deterding et al. (2011), Zichermann and Cunningham (2011), Kankanhalli et al. (2012), and Werbach and Hunter (2012). Examples include avatars for the player's profile, badges for high achievers, and level unlocking. After adding gamification features, a study will evaluate learning outcomes of Emerge2Maturity using a combination of a survey and an experiment. The Strategy Assessment Model will allow players to determine strategy elements of constraints and the number of phases using factors identified in data warehouse research, including information interdependence, task routineness, and level of sponsorship factors. The strategy extensions will support the learning objectives about strategy selection and adaptation as an organization matures in its data warehouse development.

6. REFERENCES

AbuAli, A. & Abu-Addose, H. (2010). Data Warehouse Critical Success Factors. *European Journal of Scientific Research*, 42(2), 326-335.

Adelman, S. (2012). Measuring Data Warehouse Return on Investment White Paper, *Data Management Association*, Phoenix, Arizona.

Amin, R. & Arefin, T. (2010). The Empirical Study on the Factors Affecting Data Warehousing Success. *International Journal of Latest Trends in Computing*, 1(2), 138-142.

Anderson, E. & Morrice, D. (2000). A Simulation Game for Teaching Service-Oriented Supply Chain Management: Does Information Sharing Help Managers with Service Capacity Decisions? *Production and Operations Management*, 9(1), 40-55.

Ariyachandra, T. & Watson, H. (2010). Key Organizational Factors in Data Warehouse Architecture Selection. *Decision Support Systems*, 49(2), 200–212.

Ballou, D. & Tayi, G. (1999). Enhancing Data Quality in Data Warehouse Environments. *Communications of the ACM*, 42(1), 73-78.

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management. *Business & Information Systems Engineering*, 1(3), 213–222.

Ben-Daya, M. & Hariga, M. (2004). Integrated Single Vendor Single Buyer Model with Stochastic Demand and Variable Lead Time. *International Journal of Production Economics*, 92(1), 75–80.

Brooks, L. (2007). *FinGame Online 5.0*. McGraw-Hill Education.

Browne, S. & Zipkin, P. (1991). Inventory Models with Continuous, Stochastic Demands. *The Annals of Applied Probability*, 1(3), 419–435.

Carvalho, J., Rocha, A., Vasconcelos, J., and Abreu, A. (2019). A Health Data Analytics Maturity Model for Hospitals Information Systems. *International Journal of Information Management*, 46, 278-285.

Choudhary, R. (2010). Key Organizational Factors in Data Warehouse Architecture Selection. *Vivekanada Journal of Research*, 24, 24–32.

Conning, P. (2000). *Data Warehousing and Data Mining in the Insurance Industry: Floods of Information, Fountains of Knowledge*. Conning & Company Insurance Research and Publications.

Croson, R. & Donohue, K. 2006. Behavioral Causes of the Bullwhip Effect and the Observed Value of Inventory Information. *Management Science*, 52(3), 323–336.

De Castro, E., Tabucanon, M., & Nagarur, N. (1997). A Production Order Quantity Model with Stochastic Demand for a Chocolate Milk Manufacturer. *International Journal of Production Economics*, 49(2), 145-156.

Deng, X. (2005). Differentiating the Effect of Cumulative Experience and Learning: A Field Study of Help Desk Support. In *Proceedings of the Americas Conference on Information Systems 2005*, 1775–1782.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining 'Gamification.' In *Proceedings of the 2011 Annual Conference on Extended Abstracts on Human Factors in Computing Systems - CHI EA '11*, 2425.

Gagne, R. (1970). *The Conditions of Learning, 2nd ed*. Oxford, England: Holt, Rinehart, and Winston.

Gulledge, T. (2006). What is Integration? *Industrial Management & Data Systems*, 106(1), 5–20.

Inmon, W. (2001). Information Management: Charting the Course: Little White Lies. *DM Review*.

Kankanhalli, A., Taher, M., Cavusoglu, H., & Kim, S. (2012). Gamification: A New Paradigm for Online User Engagement. In *Proceedings of the International Conference on Information Systems*, 3573-3582, Orlando, Florida.

Kimball, R. & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Wiley.

Lapre, M. & Nembhard, I. (2011). Inside the Organizational Learning Curve: Understanding the Organizational Learning Process. *Foundations and Trends in Technology, Information and Operations Management*, 4(1), 1-103.

Leger, P. (2006). Using a Simulation Game Approach to Teach Enterprise Resource Planning Concepts. *Journal of Information Systems Education*, 17(4), 441-447.

Lindsey, K. & Frolick, M. (2003). Critical Factors for Data Warehouse Failure. *Journal of Data Warehousing*, 8(1), 48-54.

Merrick, C. (2014). "9 Reasons Data Warehouse Projects Fail," R*JMetrics*. Retrieved April 1, 2021, from https://blog.rjmetrics.com/2014/12/04/10-common-mistakes-when-building-a-data-warehouse.

Miranda, P. & Garrido, R. (2004). Incorporating Inventory Control Decisions into a Strategic Distribution Network Design Model with Stochastic Demand. *Transportation Research Part E: Logistics and Transportation Review*, 40(3), 183-207.

Mitchell, T. (2017). Why Data Warehouse Projects Fail. Retrieved April 1, 2021, from https://www.timmitchell.net/post/2017/01/10/why-data-warehouse-projects-fail/.

Monk, E. & Lycett, M. (2016). Measuring Business Process Learning with Enterprise Resource Planning Systems to Improve the Value of Education. *Education and Information Technologies*, 21(4), 747-768.

Patrizio, A. (2019). 4 Reasons Big Data Projects Fail – And 4 Ways to Succeed, *InfoWorld.* Retrieved April 1, 2021, from https://www.infoworld.com/article/3393467/.

Paulk, M., Weber, C., Curtis, B., & Chrissis, M. (1995). *The Capability Maturity Model: Guidelines for Improving the Software Process*. Boston, Massachusetts: Addison-Wesley.

Pendharkar, P. & Subramanian, G. (2004). An Empirical Study of Learning Curve Theory's Application to the Software Development Effort-Experience Relationship in Software Engineering and Project Management. In *Proceedings of the Americas Conference on Information Systems*, New York, New York, 767-772.

Plaza, M., Ngwenyama, O., & Rohlf, K. (2010). A Comparative Analysis of Learning Curves: Implications for New Technology Implementation Management. *European Journal of Operational Research*, 200(2), 518-528.

Ramamurthy, K., Sen, A., & Sinha, A. (2008). An Empirical Investigation of the Key Determinants of Data Warehouse Adoption. *Decision Support Systems*, 44(4), 817–841.

Rao, L. & Osei-Bryson, K. (2008). An Approach for Incorporating Quality-Based Cost-Benefit Analysis in Data Warehouse Design. *Information Systems Frontiers*, 10(3), 361–373.

Schmitt, A., Snyder, L., & Shen, Z. (2010). Inventory Systems with Stochastic Demand and Supply: Properties and Approximations. *European Journal of Operational Research*, 206(2), 313-328.

Sen, A., Sinha, A., & Ramamurthy, K. (2006). Data Warehousing Process Maturity: An Exploratory Study of Factors Influencing User Perceptions. *IEEE Transactions on Engineering Management*, 53(3), 440–455.

Sen, A., Ramamurthy, K., & Sinha, A. (2012). A Model of Data Warehousing Process Maturity. *IEEE Transactions on Software Engineering*, 38(2), 336–353.

Serrano-Laguna, A., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2017). Applying Standards to Systematize Learning Analytics in Serious Games. *Computer Standards & Interfaces*, 50, 116-123.

Spruit, M. & Sacu, C. (2015). DWCMM: The Data Warehouse Capability Maturity Model. *Journal of Universal Computer Science*, 21(11), 1508–1534.

Sterman, J. (1989). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science*, 35(3), 321–339.

Watson, H., Goodhue, D., & Wixom, B. (2002). The Benefits of Data Warehousing: Why Some Organizations Realize Exceptional Payoffs. *Information and Management*, 39(6), 491–502.

Werbach, K. & Hunter, D. (2012). *For the Win: How Game Thinking Can Revolutionize Your Business*. Philadelphia, Pennsylvania: Wharton Digital Press.

Winter, R. & Strauch, B. (2002). A Method for Demand-Driven Information Requirements Analysis in Data Warehousing Project. In *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03)*, 9.

Wright, T. (1936). Factors Affecting the Costs of Airplanes. *Journal of Aeronautical Science*, 3, 122-128.

Zichermann, G. & Cunningham, C. (2011). *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. Sebastopol, California: O'Reilly Media.

AUTHOR BIOGRAPHIES

Michael V. Mannino is an associate professor in the business school of the University of Colorado Denver. Previously he was on the faculty at the University of Florida, University of Texas at Austin, and University of Washington. He has been active in research about database management, data warehouses, data mining, and organizational impacts of technology. He has published articles in major journals of the IEEE (*Transactions on Knowledge and Data Engineering* and *Transactions on Software Engineering*), ACM (*Transactions on Management Information Systems*, *Communications* and *Computing Surveys*), INFORMS (*Informs Journal on Computing* and *Information Systems Research*), and Elsevier (*Decision Support Systems*). His research includes several popular survey and tutorial articles as well as many papers describing original research. He is the author of the textbook, *Database Design, Application Development, and Administration*, in its seventh edition. He also developed four courses in the Coursera specialization, *Data Warehouses for Business Intelligence*, with total course completions exceeding 13,000 in October 2020.

Mohammed Khojah is an assistant professor with a demonstrated history of working in the higher education industry. He has skills in problem-solving, analytical skills, ITIL, and information systems. He holds a Ph.D. from the University of Colorado Denver focused on computer science and information systems. Research interests are in design science, computer-based learning, data analytics, and business strategies.

Dawn G. Gregg is the director of the information systems discipline and a professor at the University of Colorado Denver. She also serves as the director of assurance of learning for the business school. Prior to her current role, she has served as the Associate Dean of Programs and as the Administrative Director of the Entrepreneurship Center. She received her Ph.D. in computer information systems and her M.S. in information management from Arizona State University, her M.B.A. from Arizona State University West, and her B.S. in mechanical engineering from the University of California Irvine. While at UC Denver, she founded and was CEO of Developing Minds Software. Prior to her doctoral studies, she was employed for nine years as a research and development engineer. Her current research focuses on how to organize and maintain web-based content so that it can be used to better meet business and special education needs. Her work has been published in numerous journals, such as *MIS Quarterly*, the *International Journal of Electronic Commerce*, *Information & Management*, *Communications of the ACM*, and *Decision Support Systems*.

Information Systems & Computing
Academic Professionals

STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the Journal of Information Systems Education have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN 2574-3872