Process-Oriented Evaluation of User Interactions in Integrated System Analysis Tools

Chaiwoo Lee, Paul T. Grogan, Olivier L. de Weck

Engineering Systems Division
Massachusetts Institute of Technology
Cambridge, MA, United States

Abstract— **When computer-based tools are used for analysis of complex systems, the design of user interactions and interfaces becomes an essential part of development that determines the overall quality. The objective of this study is to investigate the processes and results of user interactions with integrated analysis tools to synthesize design implications for future tool development. In this study, two space exploration logistics tools are compared in a controlled user experiment. Through a comparative usability analysis, this study evaluated user performance and perception to provide design implications for future integrated analysis tools. For a comprehensive evaluation, multiple methods were used for data collection, including observation, questionnaire and interview. In addition to a result-oriented performance analysis, a process-oriented approach was used for analyzing patterns in user behaviors and errors. Results are presented with reference to the related features embedded in the interfaces of the two tools. Based on the comparative results, synthesized design insights for hierarchical structure, model transparency, automation, and visualization and feedback are discussed for integrated analysis tools in general.**

Keywords- Complex systems; software development; usability testing; space logistics; analysis tools

I. INTRODUCTION

Design of complex systems increasingly relies on computer-based analysis tools to gain insights of potential performance. An integrated analysis tool operates at the system level to identify and evaluate emergent properties including non-functional properties (commonly referred to as "ilities"). For example, evaluating the reliability of an automobile is an aggregate function of many components and does not solely depend on the selection of engine, transmission, or battery.

Integrated analysis tools incorporate many system components described using abstracted representations or models. Although a computer may manage computation and memory at large scales, effectiveness of an integrated analysis tool ultimately relies on the user interface. A human analyst must be able to leverage the tool for understanding the interrelationships contributing to system performance.

Thus, integrated analysis tools should not only be complete and high-quality, but also usable. The goal of this study is to investigate how users interact with integrated analysis tools. By comparing user performance and perception across multiple interfaces, insights for design implications are identified to inform development of future analysis tools.

An important property related to user interactions is usability. Although it has been defined in various ways to capture different system aspects, purposes or contexts of use, the concept of usability is often described to encompass effectiveness, efficiency and satisfaction. For example, ISO 9241-11 defines usability as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [1]. These three factors are not always important for all types of systems, but have been commonly used in previous studies. According to ISO 9241-11, effectiveness considers task accuracy and completeness, efficiency looks at resources needed to achieve goals, and satisfaction is defined as positive attitudes and feelings toward the system [1]. Previous studies have also used more specific factors and attributes for evaluation of usability, including learnability, productivity, adaptiveness, cognitive workload, rememberability, throughput, operability and attractiveness [2-8].

Usability evaluation is used to ensure if an interactive system is effective, efficient, and satisfying for users to achieve their goals. When a system is at the early stages of design, expert-based usability inspection is often used [9-10]. On the other hand, user-based testing can be done to effectively gather inputs from potential users at the later stages [11-12]. Traditional usability testing is limited in that latent needs are not expressed clearly or in full detail. Thus, a process-oriented method such as user observation is necessary to analyze user interactions and behavior in a greater detail [13-14]. Such investigation of process is important not only in usability evaluation, but also in analyzing other complex systems as well, especially when human interactions are involved [15-16].

This study aims to comparatively observe and analyze user interactions in two integrated analysis tools. The tools target space exploration logistics as a representative complex system. Space exploration involves careful consideration of physical components in hierarchical structures, human elements, and has critical temporal and spatial implications for logistics planning. Due to its complexity and intricacy, modeling and analyzing is aided by software tools with abstractions and simplifications. The two tools under investigation, a discrete event simulation application and a spreadsheet tool, represent diversity in interface designs. Since both tools are fully developed, this study employs a user-based testing method to comparatively analyze usability as well as patterns of use.

This paper is structured as follows. Section II provides an overview of the two space logistics tools and discusses the experimental design. Section III analyzes the results from the user evaluation using statistical evaluation and process-oriented methods. Section IV discusses the results and explains limitations of the study. Finally, Section V concludes with synthesized design implications for integrated analysis tools.

II. User evaluation

The study was designed as a comparative analysis of user interactions in using two space logistics analysis tools. A usability testing approach was employed to involve potential users in the evaluation process. The participants were given a set of tasks, and data were collected using multiple methods.

A. Space Logistics Analysis Tools

Two analysis tools, with similar capabilities but different underlying engines and user interfaces, were tested. The SpaceNet tool is a discrete event simulation software with a Java Swing graphical user interface.[1] Its immediate target users include students, researchers and analysts researching conceptual human space exploration. SpaceNet defines a network, resource, element, and event models, all of which can be accessed by users to define, simulate, visualize, and evaluate a mission. Simulated behaviors include propellant consumption during transportation and resource demands by crew members during an exploration. Figure 1 shows an input editor and a simulation result in SpaceNet.
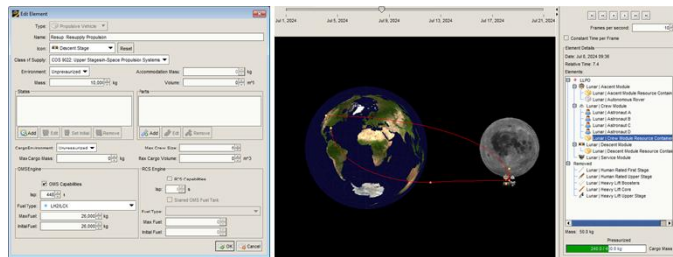

Figure 1. Examples of SpaceNet user interface

Figure 2. Example of spreadsheet user interface

The spreadsheet tool uses Microsoft Excel, as shown in Figure 2, to evaluate governing equations for resource consumption such as the rocket equation for propellant during transportation. In practice such a tool would be developed ad-hoc in the absence of a formal tool. For this study, a spreadsheet tool was developed with similar terminology, structure, and input options used by SpaceNet. Simulation is done with calculation, and results are visualized and evaluated

with changes in cell values and error indications.

B. Participants

Twelve MIT graduate students (nine males and three females) were purposively selected to represent the range of potential user population. The sample included eight people with background in space exploration and four people with no such background. The participants' ages ranged from 22 to 32.

C. Scenarios and User Tasks

Two mission scenarios, Scenario C and D, were designed with the same overall structure but differences in specific details. Both scenarios targeted lunar exploration; however Scenario C uses conceptual vehicles from the Constellation Program while Scenario D uses Apollo-era vehicles.

Both scenarios comprised of two parts, illustrated in time-location event diagrams shown in Figure 3. In Part 1, users were asked to create a simple model using a given tool of a detailed scenario comprised of launch stack assembly, launch, and transfer to and arrival in lunar orbit. In Part 2, users were given a baseline model and were asked to modify the input values to maximize the mission's effectiveness under a time restriction of fifteen minutes. Relative Exploration Capability (REC), a valuation of research taking place during an exploration, was used as measure of effectiveness.
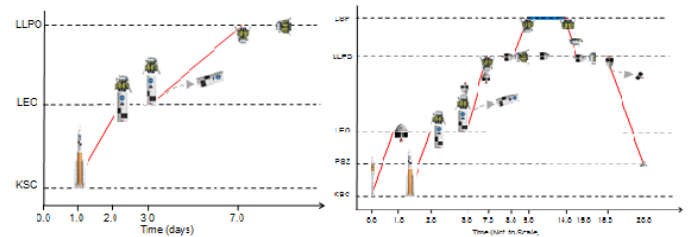

Figure 3. Scenario C time-location diagram (left: Part 1, right: Part 2)

Before testing, users were given a tutorial for both tools. Within the tutorial, step-by-step instructions were provided to design and modify missions based on a scenario different from Scenario C and D, but having a similar overall structure.

D. Study design

An orthogonal design was implemented to allow participants to evaluate the two tools and scenarios without interaction or ordering effects. Each participant completed two sessions of non-consecutive testing on different scenario-tool combinations. The study design is summarized in Table 1.

TABLE 1. SUMMARY OF STUDY DESIGN

Subject ID	1st Scenario	1st Tool	2nd Scenario	2nd Tool
5, 9, 13	C	SpaceNet	D	Spreadsheet
6, 10, 14	C	Spreadsheet	D	SpaceNet
7, 11, 15	D	SpaceNet	C	Spreadsheet
8, 12, 16	D	Spreadsheet	C	SpaceNet

E. Data collection methods

Multiple methods were employed to see various aspects of

user interactions and experiences with the two tools. Quantitative data were captured to analyze task performance and self-rated perception. Qualitative data were gathered for a process-oriented analysis and capturing detailed feedback.

1) *Observation and think-aloud:* User actions were observed real-time using a projected display, and recorded using Camtasia screen capture and audio recording software.[2] Participants were encouraged to talk aloud during testing, which gave richer description of the context and their thoughts.

2) *Questionnaire*: A one-page questionnaire was given at the end of testing sessions for the participants to rate and evaluate perceived usability on a seven-point Likert scale. The evaluation criteria are summarized in the next section.

3) *Interview*: An open-ended, semi-structured interview was conducted after the second session. During the interviews, which lasted around ten minutes for each participant, the participants spoke about study design, tool design, testing experience, and any other comments or suggestions.

III. RESULTS AND ANALYSIS

The results and analysis are presented in the following three sections. First, the results from comparative evaluation are analyzed using a statistical evaluation of the usability metrics. Second, results from questionnaires and interviews are presented to illustrate user perception. Finally, a process-oriented approach is presented to better understand errors created, detected, and corrected during the user tasks.

A. Data processing

The details of tasks and errors were logged for an in-depth analysis of user interactions. The video recordings from all testing sessions were coded by time, type of action and content of action. The type of action included start and end of task milestones, and creation, detection and correction of errors. The contents were coded to show the specific task milestones were being done, the specific types of errors, and the method of error detection and correction. For Part 2, the REC values were also recorded against a time dimension.
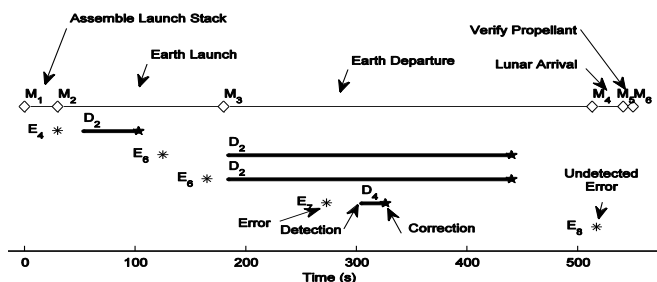

Figure 4. Task milestones and error analysis

For each participant, the coded data were first plotted and visualized along a timeline of milestones including the four tasks and verification as shown in Figure 4. Errors are identified as asterisks at the time of creation and stars when

---

2 Camtasia Studio: http://www.techsmith.com/camtasia.html

corrected. The duration of time in which the user is actively diagnosing and correcting an error is displayed as a solid line.

B. Performance-based usability evaluation

Metrics were defined to assess the degree to which the tools enable users to perform tasks easily. Since the user population consists of a homogeneous group and aims for usefulness and ease of use, effectiveness and efficiency are emphasized rather than emotional satisfaction. Also, since errors in modeling can often cause serious consequences, error prevention and tolerance, often a part of efficiency, was defined as a separate factor. Effectiveness metrics were mostly applied for analysis of Part 1 in both scenarios, where no time constraint was given. For Part 2, where users tried to arrive at the best possible result under a time restriction, efficiency metrics were applied the most. Specific metrics were defined for evaluating system usability in terms of performance as summarized in Table 2, and more detail can be found in [17].

TABLE 2. USABILITY METRICS FOR OBSERVATION ANALYSIS

Factor	Description	Metric
Effectiveness	Modeling missions completely with high research values	Completeness, outcome quality
Efficiency	The time and effort needed for users to achieve goals	Completion time, time in mode, time until event
Error tolerance and prevention	Making fewer errors, recovering quickly, and feeling a sense of control in usage	Error rate, recovery rate, recovery time

Quantitative data coded from the video recordings were used for the performance-based statistical usability analysis. A pairwise t-test was used for comparison between the two tools and two scenarios and an independent-group t-test was used for comparison between the two user groups. The results for Part 1 and 2 are summarized in Table 3 and 4, respectively.

TABLE 3. RESULTS FOR SCENARIO PART 1 (*: P<0.05, **: P<0.01)

Metric	Scenarios		Tools	
	C	D	SpaceNet	Spreadsheet
Completion time (s)	802.4	608.0	450.6	959.8
Time to correct task (s)	199.8	109.7	113.5	196.0
Time in task 1 (s)	75.9	55.0	72.8	58.2
Time in task 2 (s)	219.7	188.1	129.8**	278.0**
Time in task 3 (s)	85.4	93.0	70.5	108.3
Time in task 4 (s)	31.2	35.9	18.7**	48.4**
Tasks in 5 minutes (#)	1.50*	2.75*	2.58	1.67
Error rate (#)	3.08	2.00	1.83	3.25
Recovery rate (%)	70.6	80.6	73.1	78.0
Recovery time (%)	33.7	20.0	15.5*	38.2*

TABLE 4. RESULTS FOR SCENARIO PART 2 (*: P<0.05, **: P<0.01)

Metric	Scenarios		Tools	
	C	D	SpaceNet	Spreadsheet
Outcome quality (%)	255.4	223.8	170.0*	309.2*
Time to REC increase (s)	327.7	533.9	467.4	394.1
Error rate (#)	1.50	1.25	1.92*	0.83*
Recovery rate (%)	39.6	50.0	45.8	43.8
Recovery time (%)	14.8	16.5	20.6*	10.8*

In Part 1 of the two scenarios, SpaceNet was generally

more efficient, requiring significantly less time in task 2 (Earth launch) and task 4 (lunar arrival) and also consuming less time in recovery. The two scenarios were found to be comparable. In Part 2, the spreadsheet tool was better in terms of the overall effectiveness, significantly in improving REC, reducing errors, and spending less time recovering from errors. The two scenarios also found to have no significant difference in Part 2.

C. User perception

In addition to the performance-based objective evaluation, an evaluation of subjective user perceptions was completed. In the questionnaire, users were asked to rate how much they agree or disagree with statements of the metrics in Table 5.

TABLE 5. USABILITY METRICS FOR QUESTIONNAIRE

Factor	Metric
Effectiveness	Perception of outcomes
Efficiency	Mental effort, ease of use, complexity
Error tolerance and prevention	Annoyance, confidence, predictability, intuitiveness, familiarity

According to the questionnaire results as shown in Table 6, the subjects felt that the spreadsheet tool required significantly more mental effort than SpaceNet. Also, while not statistically significant, they perceived SpaceNet to be more convenient and easy to use, more capable of achieving high-quality outcomes, less complicated, less annoying, enables users to feel more confident, and more intuitive and familiar.

TABLE 6. QUESTIONNAIRE RESULTS (*: P<0.05, **: P<0.01)

Metric	Scenarios		Tools	
	C	D	SpaceNet	Spreadsheet
Mental effort	4.75	4.17	3.92*	5.00*
Convenience	4.67	4.50	5.17	4.00
Predictability	4.67	4.67	4.58	4.75
Outcome quality	3.58	4.41	4.25	3.75
Complexity	3.88	4.00	3.63	4.25
Annoyance	4.08	4.25	4.00	4.33
Confidence	4.42	4.42	4.67	4.17
Intuitiveness	4.75	5.08	5.25	4.58
Familiarity	4.50	5.00	5.00	4.50

User comments and qualitative feedback were gathered from the interviews. Most subjects agreed that SpaceNet was more graphical and visual. However, some of them said how SpaceNet seems "busier" and "less transparent", while it was "easier to play with inputs" using the spreadsheet. They also felt that SpaceNet was more helpful in error detection since it gave "pin-pointed messages," while they were unconfident of error detection with the spreadsheet. Subjects mentioned SpaceNet would be "more convenient in the long run", while the spreadsheet is "not intuitive" compared to their mental models and "not scalable and at some point chokes."

D. Process-oriented error analysis

The usability evaluation focused on how the tools enabled users to perform given tasks, and the degree to which the users felt the tools to be usable. While a few measures such as time

in mode looks at the process of interaction, usability testing alone is not sufficient to fully analyze the actions that occurred during various user interactions.

An error analysis was performed with a focus on the details of actions creating, discovering, and recovering from errors to better understand the processes taking place during the usability experiment. Data logged from the video recordings were gathered and coded by one of 11 error types described in Table 7. Analysis of frequency and association between error type and tools identified how many times various types of errors occurred, and how those errors were corrected.

TABLE 7. ERROR TYPES

Label	Description
Cargo nesting	Cargo is not properly nested inside parent element
Mission	Mission details such as location or time is incorrect
Propulsive burn	Error in completing a propulsive burn
Missing mass	Mass is missing from a calculation
Element staging	Element mass is not removed after being staged
Remaining delta-v	Remaining velocity incorrect after completing burn
Double-counted mass	Mass is double-counted in a calculation
Event timing	Time of event execution is incorrect
Propellant staging	Propellant mass is not removed after being staged
Available propellant	Incorrect propellant available for a burn
Specific impulse	Incorrect specific impulse for element being burned

The result from error frequency analysis is summarized in Figure 5. It can be seen that the types of errors differed by the tool used. In other words, SpaceNet and spreadsheet were susceptible to different types of errors. More specifically, while errors related to cargo nesting and propulsive burns were observed frequently in SpaceNet, these errors weren't found when using the spreadsheet. On the other hand, many users experienced errors related to verifying the remaining delta-v and propellant staging using the spreadsheet, and only a few such errors occurred when using SpaceNet.

Many of the errors created during a user experiment were ultimately detected and corrected. The modes of detection included feedback from element state (vehicle data outputs), propellant state (propellant mass outputs), clarification of task requirements, self-identification of errors by the participant, status or error messages within the tools, and written instructions provided with the tasks. Some errors were not detected but still resulted in a valid task completion.

The association between error types and detection methods is shown in Figure 6. The circle sizes represent the frequency that each type or error or detection method occurred. The lines show co-occurrence or association, where a certain type of error was detected with the connected method. The frequency of co-occurrence, or the strength of the association, is represented by the thickness of the lines. For example, the thick line connecting 'missing mass' and 'status message' for the spreadsheet tool means that the errors related to not adding mass during modeling was often detected by status messages displayed. It is observed that the detection methods and their associations with error types differed by the tools used. For example, many errors in SpaceNet were detected by checking

the element state, while many errors in the spreadsheet were detected using status messages and propellant state.
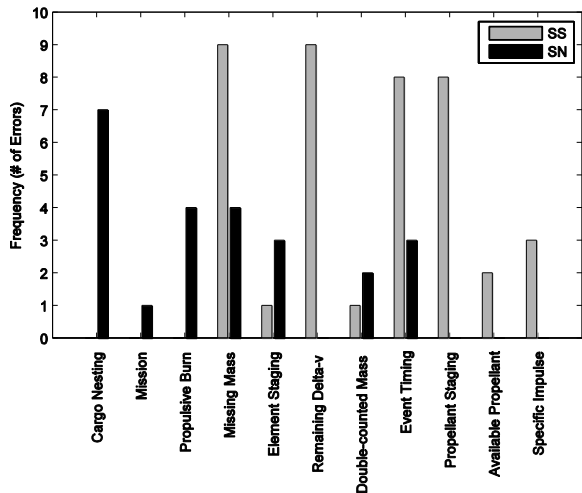

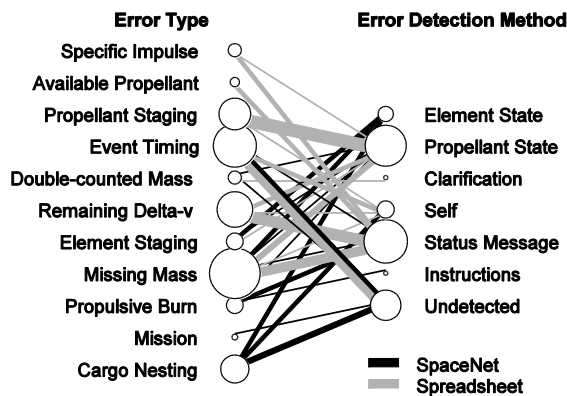Figure 5. Error frequency by tools (SN: SpaceNet, SS: Spreadsheet)

Figure 6. Association between error type and detection

IV. DISCUSSION

The results presented in the previous section indicate differences in usability across the two space logistics analysis tools. This section synthesizes four design insights from the observed differences and discusses limitations in generalizing the results to future analysis tool development.

A. Synthesized design insights

The first design insight studies the overall structure of the tool. The SpaceNet tool is generally hierarchical, with element models nested inside other elements and many of the user inputs presented as dialogs within other dialogs. In contrast, the spreadsheet tool is flat, using a two-dimensional grid of cells to contain all of the models and data used in the resulting analysis. Negative effects of hierarchical structure are observed in Part 2 of the study (more difficult to find and change inputs in SpaceNet resulting in lower effectiveness), and in the large number of cargo nesting errors in the SpaceNet tool. The user interviews identify both negative effects of hierarchical structure (SpaceNet seems "busier" and "easier to play with inputs" in the spreadsheet), and potential benefits (SpaceNet requiring less mental effort and being

"more convenient in the long run," while the spreadsheet being "not scalable and at some point chokes"). Thus, hierarchical structure is crucial for scalability, but a tool should provide methods to rapidly drill down into lower-level details. This discussion relates to the ecological interface design approach, where abstraction hierarchy is used [18].

The second insight concerns the transparency of underlying models' details. In the SpaceNet tool, details of propellant and resource consumption are hard-coded and treated as a black box during use. In contrast, the spreadsheet tool shows all equations and parameters in the cells where they are used. Positive effects of model transparency are observed in Part 2, where users were able to "play with" the inputs in the cells to improve the REC value. Model transparency appears beneficial, especially when coupled with a hierarchical structure, to provide a better understanding of the integrated models. This agrees with existing literature, where system visibility is recommended for user-centered design [19][20].

The third insight addresses the use of automation in calculation of intermediate inputs. The SpaceNet tool includes several automated processes supporting event timing and multi-burn transports, whereas the spreadsheet tool requires the user to maintain intermediate inputs during these processes. Automation in the SpaceNet tool likely contributed to the significant differences in Part 1 of the study, exemplified by the lower times for tasks 2 and 4 which both require multiple burning and staging of propulsive elements. In the spreadsheet tool, lack of automation directly contributed to the errors in intermediate inputs such as missing mass, remaining delta-v, element timing, propellant staging, and available propellant.

Finally, an insight to the use of visualizations and feedback is discussed. As a customized application, the SpaceNet tool includes plots and graphics representing the temporal and spatial nature of element states as well as error messages during simulation. The spreadsheet tool relies on cells to provide quantitative feedback and includes multiple checks for error conditions. Error detection depends on these feedback mechanisms, relying more on the element state display in the SpaceNet tool and the propellant level data and status messages in the spreadsheet tool. The user interviews indicate the graphical and visual feedback and pin-pointed messages in SpaceNet helped to identify errors, while the simple feedback in the spreadsheet did not build confidence of error detection. The importance of providing feedback and preventing errors is also discussed in existing principles and guidelines [19][20].

B. Study limitations

Several limitations reduce the strength of this study's conclusions. First, the study investigated two space logistics tools as cases of integrated analysis tools. The tool interfaces range from a graphical custom application to a simpler spreadsheet design. Tools in other domains likely have similar issues of representing large amounts of information in hierarchical formats, integrating smaller models of physical processes, and operating over spatial and temporal dimensions. Aside from conclusions particular to space logistics, the design implications can also contribute to other domains.

Second, the experimental design only considered relatively simple tasks. This does not address the effects of advanced analysis but rather simple conceptual analyses such as what may be performed as an introduction to space logistics. More complex scenarios would likely see larger benefits from hierarchical structure and may require more experience or domain knowledge for completion.

Finally, this study only explored students as potential users of the space logistics analysis tools. Although the sampled subjects represent a broad spectrum of potential users in the target population of these tools, other tools may have a different user population with specific expectations or modes of use. Due to the relatively small sample size (12), there are limitations in the quantitative conclusions drawn from the statistical results, however the qualitative results from the process-oriented approach are likely more robust.

V. CONCLUSION

This study sought to comparatively observe and evaluate user interactions with two tools for space logistics analysis. In addition to a result-oriented usability testing, a more process-oriented approach was used to analyze the user actions and errors. Through a human experimentation with various data collection and analysis methods, this study produced quantitative and qualitative results covering usability, user interaction processes, and user perceptions.

The usability testing results showed that SpaceNet was more efficient at parts of model creation, while the spreadsheet was more effective overall when modifying existing model. With a process-oriented approach, the specific user actions and errors were observed, recorded and analyzed. This enabled a rich set of data to be gathered, which generate important insights. Based on a frequency and associations analysis, it was found that the two tools differed largely in the types of errors generated and the ways that they were detected.

Four design insights for integrated analysis tools were synthesized from the study results. First, hierarchical structures enable the scalable study of large and complex systems, provided a tool allows users to quickly drill down into detailed data. Although the SpaceNet tool includes a strong hierarchy of model components, it was tedious to modify inputs nested in several layers of dialogs. Second, model transparency provides users the option of seeing details of the integrated models. Users of the spreadsheet tool benefitted from investigating the built-in formulas and relationships between cells. Third, automated processes reduce dependence on linking intermediate inputs. The automatic burn-stage process in SpaceNet contributed to efficient work and fewer errors compared to the spreadsheet tool. Finally, feedback and visualizations represent model state in richer formats than simple checks and numerical data. Graphical displays in SpaceNet help understand spatial and temporal effects and error messages help to identify and pin-point errors.

Future work seeks to implement the design insights to future versions of integrated space exploration logistics tools and evaluate the findings' applicability to other domains. In particular, terrestrial infrastructure systems are a potential use case for generalization. Infrastructure systems have similar processes of resource production, transportation, and consumption interacting with technical and human systems over spatial and temporal dimensions, and could build upon the findings of this study with a diverse user population.

REFERENCES

[1] ISO 9241-11, Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), Part 11: Guidance on Usability. Geneva: International Organization for Standardization, 1998.

[2] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, "Usability measurement and metrics: a consolidated model," Softw. Qual. J., vol. 14, no. 2, pp. 159-178, June 2006.

[3] M. Macleod, "Usability: practical methods for testing and improvement," Proc. Nor. Comput. Soc. Softw. Conf., February 1994.

[4] L. L. Constantine and L. A. D. Lockwood, Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design. New York, NY: Addison-Wesley, 1999.

[5] J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey, Human Computer Interaction. Wokingham: Addison-Wesley, 1994.

[6] ISO/IEC 9126-1, Software Engineering, Product Quality, Part 1: Quality Model. Geneva: International Organization for Standardization, 2001.

[7] Y. S. Lee, S. W. Hong, T. L. Smith-Jackson, M. A. Nussbaum, and K. Tomioka, "Systematic evaluation methodology for cell phone user interface," Interact. Comput., vol. 18, no.2, pp. 304-325, March 2006.

[8] B. Shackel, "Usability - context, framework, definition, design and evaluation," in Human Factors for Informatics Usability, B. Shackel and S. Richardson, Eds. Cambridge, UK: Cambridge University Press, 1991, pp. 21-31.

[9] R. Marchetti, "Using usability inspections to find usability problems early in the lifecycle," Pac. Northwest Softw. Qual. Conf., pp. 1-19, October 1994.

[10] J. S. Bradford, "Evaluating high-level design: synergistic use of inspection and usability methods for evaluating early software designs," in Usability inspection methods, J. Nielsen and R. L. Mack, Eds. New York, NY: Wiley, 1994.

[11] J. M. C. Bastien, "Usability testing: a review of some methodological and technical aspects of the method," Int. J. Med. Inform., vol. 79, no. 4, pp. e18-e23, April 2010.

[12] J. S. Dumas and J. C. Redish, A Practical Guide to Usability Testing. Portland, OR: Intellect Books, 1999.

[13] C. Lee, S. Bahn, G. W. Kim, and M. H. Yun, "Performance comparison of manual and touch interface using video-based behavior analysis," J. Ergon. Soc. Korea, vol. 29, no. 4, pp. 655-659, 2010.

[14] H. Jang, "A study on patterns in the use of hands when handling with screen-touch mobile multimedia devices," Master's Thesis. Kookmin University, 2008.

[15] J. Hense, W. C. Kriz, and J. Wolfe, "Putting theory-oriented evaluation into practice," Simul. Gaming, vol. 40, no. 1, pp. 110-133, February 2009.

[16] P. Todd and I. Benbasat, "Process tracing methods in decision support systems research: exploring the black box," MIS Q., vol. 11, no. 4, pp. 493-512, December 1987.

[17] P. T. Grogan, C. Lee, and O. L. de Weck, "Comparative usability study of two space logistics analysis tools," AIAA Space 2011 Conference & Exposition, Long Beach, CA, 2011.

[18] J. Rasmussen, "The role of hierarchical knowledge representation in decision making and system management," IEEE Trans. Syst. Man Cybern., vol. 15, pp. 234-243, 1985.

[19] D. A. Norman, The Design of Everyday Things. New York, NY: Basic Books, 2002.

[20] J. Nielsen, Usability Engineering. San Diego, CA: Academic Press, 1994.