Evaluation of the Advanced Collaboration Enterprise Services (ACES)

Funded by

Assistant Secretary of Defense
for Research and Engineering,
Deputy Assistant Secretary of Defense (DASD),
Emerging Capability and Prototyping (EC&P),
Joint Capability Technology Demonstration (JCTD) Office

31 August, 2016

*This report is funded through a task order against the University Affiliated Research Center (UARC) Indefinite-Delivery-Indefinite-Quantity (IDIQ) contract with University of Nebraska's National Strategic Research Institute (NU/NSRI) (NU/NSRI UARC Contract no. FA4600-12-D-9000, task order 56).

About This Report

The authors of this report are Dr. Douglas Derrick, Dr. Gina Ligon, and Mr. Joel Elson, Research Associate at the University of Nebraska at Omaha. Questions about this report should be directed to Dr. Douglas C. Derrick, dcderrick@unomaha.edu.

Multiple COCOMs IPLs, Joint Staff, and DoD CIO state the need for real-time multi-domain collaboration in garrison and on the move supporting quicker decision making. ACES is a joint venture that will purports to address the lack of efficient and secure information sharing across mission and classification boundaries. ACES is a fully-integrated operating system application to ensure secure, uninterrupted, discreet digital information exchanges across different classification levels and networks.

About NSRI

The National Strategic Research Institute (NSRI) at the University of Nebraska is one of 13 University Affiliated Research Centers (UARCs) in the nation. Established in 2012, NSRI is engaged in a long-term, strategic partnership with our Department of Defense (DoD) sponsor, United States Strategic Command (USSTRATCOM). The NSRI provides mission-essential research and development capabilities for USSTRATCOM as well as other DoD components and federal agencies focused on combating weapons of mass destruction. We maintain five core competencies for the CWMD mission:

- Nuclear Detection and Forensics
- Detection of Chemical and Biological Weapons
- Passive Defense Against Weapons of Mass Destruction
- Consequence Management
- Space, Cyber and Telecommunications Law

NSRI receives funding from project sponsors through contracts generated from our sole-source, Indefinite Delivery Indefinite Quantity contract with USSTRATCOM, as well as other direct contract vehicles and grants.

Citations

To cite this report, please use this format:

Derrick, D. C., Ligon, G. S., and Elson, J. S., (2016). "*Evaluation of the Advanced Collaboration Enterprise Service (ACES)*," Report prepared for the Assistant Secretary of Defense for Research and Engineering, Deputy Assistant Secretary of Defense (DASD), Emerging Capability and Prototyping (EC&P), Joint Capability Technology Demonstration (JCTD) Office.

Contents

List of Figures

List of Tables

Executive Summary

We conducted five separate assessments of the Advanced Collaboration Enterprise Services (ACES). These assessments involved multiple types of stakeholders including: military and support personnel (during a National Level Event exercise), government contractors with advanced education, industry users, and graduate / undergraduate students. In total, we had over 80 people participate in the experiments and evaluations. We evaluated ACES in a series of operational and laboratory experiments with various problem types and outcome measures. The primary problem categories were convergent problems (i.e., a single best solution) and divergent problems (i.e., ill-defined problems with no clear solution). The outcome measures varied across assessment, and included: time for solution development, novelty of solution, quality of the solution, elegance of the solution, willingness to collaborate, creation of shared mental models, information sharing, process satisfaction, outcome satisfaction, and perceived value.

It appears that ACES offers significant value, especially for geographically-separated, multi-team systems, and for ill-defined, divergent collaborative processes. Co-located teams with a convergent process did not benefit from ACES in our experiments. A summary of each assessment is below.

Assessment	Primary Users	Description	Result
National Event	Army and Navy Active Duty and Civilians	Support to a high consequence exercise for a convergent task. Examine potential operational benefits	Users reported high-engagement and high value. SME ratings indicated that ACES generated more novel and higher quality solutions.
Intelligence Support	Post-graduate Contractor SMEs	Create an intelligence support document. Assess the creation of shared mental models.	Significant and rapid increase of shared mental models among experts with disparate expertise.
Industry Usability Test	Industry Professionals and Normal Users	Link a development to market research team. Examine multi-team system.	The information sharing, process satisfaction, and resulting mental models were superb. Technical improvements identified.
Convergent Lab Experiment	Graduate and Undergraduate Students	Determine how ACES supports convergent collaboration tasks.	ACES teams performed slower and less effectively in co-located, convergent tasks. Technical improvements identified. However, willingness to collaborate dramatically increased and user satisfaction measures were high.
Divergent Lab Experiment	Graduate and Undergraduate Students	Determine how ACES supports divergent collaboration tasks.	ACES teams outperformed video conferencing, shared-text-space teams on all measures (novelty, elegance, and quality).

We recommend that ACES moves forward for a broader, operational JCTD.

Introduction

Developing and maintaining capabilities for the Nation requires long-term commitment. The objective of this project was to configure and demonstrate a technology suite that is purported to help national leadership to accomplish three things: allow for more interactive engagement with content displayed on current Department of Defense (DoD) Operations Centers' knowledge walls; enable enhanced in-room and between-site collaboration capabilities supporting content integration for decision-making; and provide a pliable platform upon which media-rich, multiple intelligence (MULTI-INT) constructs could further shape technological advancements.

Background

In order to determine the potential of this platform, we conducted multiple technical assessments in operational and controlled laboratory environments in order to provide holistic capability measurements and performance metrics. These metrics related to speed of convergence (e.g., decision-making, shared mental model), quality of the collaboration (e.g., users' process satisfaction), novelty (e.g., uniqueness of solutions generated), and quality of the work products developed. The purpose of the study was to provide feedback and assessments that the JCTD office can leverage to determine if this platform can provide optimized visualization, collaboration, and decision support in the digital operating environment.

For this effort, the ACES platform was installed and configured in four disparate locations. These included two at the University of Nebraska at Omaha, one at Dam Neck Annex, VA and one in Fort Bragg, SC (the description for these two installations are described in greater detail in Assessment 1). For the laboratory studies, ACES was installed in the Peter Kiewit Institute (PKI), room 335 and Mammel Hall, room 318. The rooms were outfitted with ACES/Photon installations to facilitate controlled, behavioral research efforts. These spaces were designed with the primary intention of allowing layout flexibility to support a diverse array of cognition experiments. The main spaces contain an existing Planar 2x4 touch screen array, which was integrated into the ACES-OE and was supplemented with a horizontal pair of 75" 4k resolution screens on an adjacent wall. This space also houses two L-shaped analyst workstations with two screens (one analyst for each workstation), and a dual analyst workstation with a 3 shared screen configuration. All workstations are on casters and able to move about the room, supporting multiple configurations. This allows collaboration experiments that utilize the entire space, as well as soft segregation between the touch screen array and 75" screen array as needed. Additionally, the "Nook" in the rear of the main space offers an area for small team collaboration with the use of a 75" 4k main display supported by another dual analyst work station. This workstation is also mobile and will have the ability to be added to the main space for large team experiments. Figure 1 below shows the architecture for both rooms and figure 2 shows a photo of the installation.

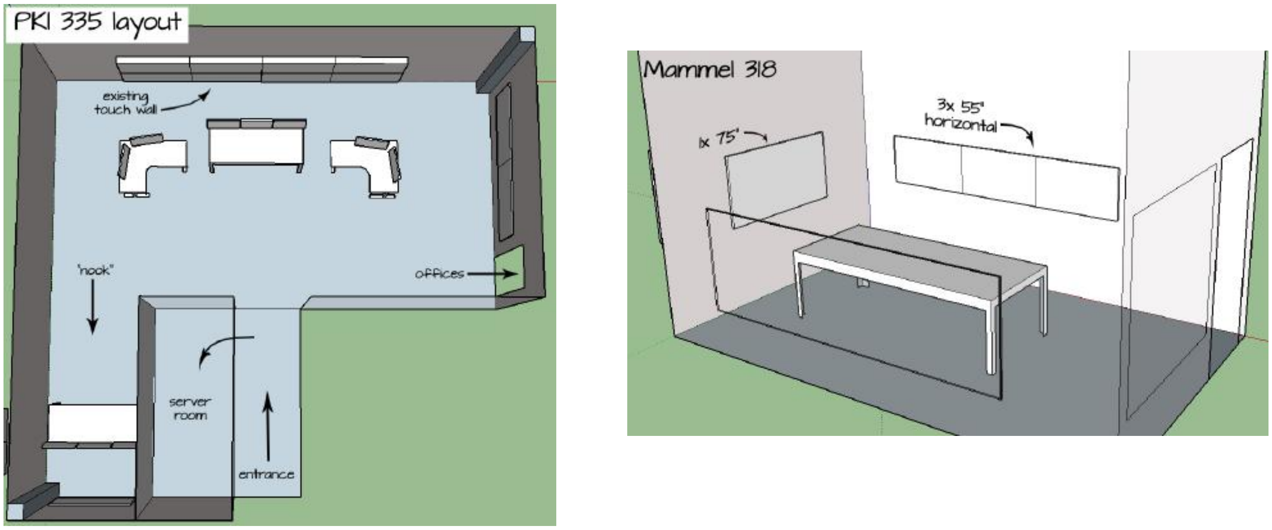

Figure 1. ACES Room Designs

Figure 2. ACES Installation in PKI, 335

Assessment Types

We conducted five separate assessments of the Advanced Collaboration Enterprise Services (ACES). These assessments involved multiple types of stakeholders including: military and support personnel (during a National Level Event exercise), government contractors with advanced education, industry users, and graduate / undergraduate students. In total, we had over 80 people participate in the experiments and evaluations.

Each of the assessments had a different purpose and will be described in detail in the following sections. A brief summary is below in order of most operational (i.e., most realistic/generalizable to DoD decision-makers) to least operational (i.e., most controlled / laboratory).

Assessment types:

1. **OPERATIONAL-LIKE ASSESSMENT.** National Level Event Exercise – the purpose was to deploy the system in an operational environment and compare its capabilities and potential benefits and/or shortcomings to existing processes as the exercise progressed. It was the closest evaluation to an operational test.

2. **GROUP COGNITION ASSESSMENT.** The context was a team developing an intelligence product and the purpose of the assessment was to evaluate the impact of the technology on the development of a shared mental model, and to determine how the technology impacted group processes and group satisfaction with the activity. A team of four subject matter experts with varying degrees of expertise had to produce an intelligence support product and used ACES to create the document.

3. **MULTI-TEAM SYSTEM ASSESSMENT.** The context was a large private firm conducting actual market research with people from target demographic. The purpose was to evaluate how the technology impacted a team-of-teams working in different locations with varying expertise on a similar problem. The teams evaluated the system on two occasions to allow development teams to collaborate in real-time with interviewers and end users.

4. **CONVERGENT SOLUTION EXPERIMENT.** In the laboratory environment, teams were asked to come to one correct solution to a well-defined, but complex problem. The teams solved four different problems, two using ACES, and two without ACES. The purpose was to determine speed, and quality of solution development.

5. **DIVERGENT SOLUTION EXPERIMENT.** In the laboratory environment, teams were asked to develop novel and new solutions to an ill-defined problem and complex problem. A control group with similar collaboration technology was also evaluated for comparison. The goal was to determine how ACES impacted novelty, quality, flexibility and fluency of the solutions.

Measurement

For group cognition, there are at least three theoretical constructs that need to be examined:

1. How does ACES affect the creation of team knowledge (shared mental models, team situation models)?
2. How does ACES affect group process satisfaction?
3. What is the relationship between group process engagement with ACES and operational outcomes?

Team Knowledge / Shared Mental Models

Team cognitive processes, such as situation assessment and coordination, rely on team knowledge. Team knowledge is critical to understanding team performance because it explains how members of effective teams interact with one another (Cannon-Bowers, et al, 2001) Team knowledge is also termed shared understanding, collective cognition, shared cognition, team mental models, shared knowledge, transactive memory, shared mental models, and so forth (Klimoski and Mohammed, 1994; Orasanu, 1990; Rentsch and Hall 1994). Team knowledge does not refer to a unitary concept; it refers to different types of knowledge that need to be shared in effective teams. In fact, researchers have proposed that teams may develop several different types of shared mental models including those related to equipment/technology, task requirements, team-member characteristics, and team interaction. (Klimoski and Mohammed, 1994; Cannon-Bowers, et al, 1993). Teams build knowledge about specific tasks (both declarative and procedural task-specific knowledge), items related to tasks (e.g. expectations of how teams operate), characteristics of teammates (e.g. strengths, preferences, weaknesses, tendencies of each individual), and attitudes and beliefs of teammates (Cannon-Bowers, et al, 1993). In addition to shared understanding of task requirements, it may be necessary for team members from different cultures to first develop shared models of team-member characteristics and of the individual and collective requirements for successful interactions.

Increased knowledge and understanding in any of these categories should lead to increased task performance. Team knowledge has been hypothesized to explain variance in team development, team performance, strategic problem definition, strategic decision-making, and organization performance (Klimoski and Mohammed, 1994).

Team knowledge is multifaceted and comprises relatively generic knowledge in the form of team mental models and more specific team situation models. Team knowledge features include type, homogeneity versus heterogeneity, and rate of knowledge change, etc. (Cooke, et al, 2007). There are multiple ways to measure the effectiveness of the creation and maintenance of team knowledge. For example, measurement features include knowledge elicitation method, team metric, and aggregation method. The hope is that ACES will speed the rate at which teams converge on a shared mental model and that it will also improve the fidelity of that team mental model (i.e., the model actually matches reality).

Group Satisfaction

Collaboration research indicates that people who feel dissatisfied with a technology-supported team process may discontinue use of such technology, even if it provides demonstrable benefits (Briggs, et al 2006). This is of concern because research has shown that people who find their experiences with a technology dissatisfying tend not to use it in the future (e.g., Hiltz & Johnson, 1990; Reinig, Briggs, Shepherd, Yen, & Nunamaker, 1996; Simon, Grover, Teng, & Whitcomb, 1996). We define satisfaction as an affective arousal with a positive valence on the part of an individual toward some object. In short, it is not only the outcome, but it is also the process of using ACES that matters. The evaluation was to determine if teams find using ACES more satisfying.

Multi-Team Systems

Multi-Team Systems (MTS), or “teams of teams” is highly related to the eventual use of ACES in the defense and security environments. As many of the same constructs that lead to greater performance within teams actually detract from MTS performance (Zaccaro et al.2012), ACES is in a unique position to generate knowledge for end-users about the ways that teams collaborate across locations, topics, and situations.

1. What team characteristics inhibit performance in a MTS environment in the context of ACES?
2. How can ACES facilitate group information sharing and problem solving in both ill-defined and concrete tasks?

Team versus Multi-Team Research

Defense environments require teams of diverse individuals working together on a problem set, but then also mandate that those teams come together to solve a common problem. For example, distributed teams work in crisis situations to solve problems where the nature of collaboration requires joint problem definition and information sharing, even when the teams have been socialized in markedly different environments (e.g., different branches of armed services) and have competing sub-goals under the broader effort (DeChurch et al., 2011). Given that ACES provides a shared workspace and common framework, it is likely that many of these dissimilarities will be ameliorated by the system. However, it is also possible that some will be exacerbated. Understanding what conditions facilitate teams working with teams from different host organizations and backgrounds is the intent of this research stream.

Nature of the Task

While some problems that MTSs face are more concrete and planned-for (i.e., have O-Plans in place), others are more ill-defined and ambiguous. It is critical to determine how ACES can facilitate information sharing in environments where the nature of the problem to be solved is complex and does not have a clear-cut set of solutions that have been a priori defined.

Assessment 1: National Level Exercise

Background

ACES was used to support participants in a National Level Event (NLE) Exercise in separate locations on 29 April 2016. The Exercise was time-bound, creating realistic conditions for decision making under crisis conditions. Participants were a mix of Army and Navy active duty and civilian DoD employees. The NLE Exercise took place using ACES at Fort Bragg and Dam Neck, and it required the participants in both locations to work together to reach a conclusion and coordinate a response between geographically dispersed operators and mission planners. The participants converged to produce a score sheet that was shared in order to provide mission operators the information needed to a plan an appropriate response. While participants in each location were familiar with the task and their teammates, participants had less familiarity with each other across locations. The exercise was classified, but it was an exercise global event that was "of great consequence."

Objective

The goal of this assessment was threefold. First, as participants were less familiar with the system, using ACES to support this particular exercise provided a realistic demonstration of the system utility as it related to collaboration outcomes such as performance, shared mental models, and user engagement. Second, because the exercise required participants who were distally located, this study provided some initial insights about how ACES functioned in a geographically dispersed environment. Finally, a third goal of this assessment was to gather data about the efficacy of ACES using pre and post measures from participants and expert raters.

Methodology

Six DoD participants in separate locations experienced minimal training on ACES system, and they completed an exercise highly relevant to their jobs using the system. Participants were given both pre and post-measures related to collaboration efficacy. Two Subject Matter Experts (SMEs), with a combined thirty years-experience with similar NLEs, provided ratings of the novelty, quality, and coordination of team performance. The exercise was a convergent task and lasted one day. The ACES nodes were installed, configured, and connected between Fort Bragg and Dam Neck. The teams in the two locations each were familiar with their own processes and each node was used by a different service (i.e., Army and Navy). Both teams had similar skill sets and expertise. Moreover, they had similar command guidance and organizational workflows.

ACES Training. An ACES facilitator conducted a brief training to familiarize participants with ACES features, operating norms, and interface components. This training was interactive and conducted in-person in both locations. After initial training, participants were able to ask their ACES facilitator for additional help if required during the exercise.

Pre-Assessments. A series of validated measures and open-ended survey items were administered via a paper-and-pencil instrument to assess: 1) confidence with the system, 2) mental models for other participant expertise/background, 3) problem definition. Each construct was assessed via a Likert scale (range of 1-5), and participants were asked to justify their ratings for each on open-ended items following each rating. On average, the six participants took 10-15 minutes to complete this questionnaire immediately after receiving training on the ACES system.

NLE Exercise. After completing the pre-assessments, participants worked on a Joint National Level Event (NLE) Exercise across locations. Elements of the exercise required the participants in both locations to work together to reach a convergent conclusion and take an action of great consequence.

Post-Assessments. Immediately following the Exercise, two types individuals completed post-assessments: participant and expert evaluators. Exercise participants completed a similar measure to the pre-assessment; this allowed for a within-participant comparison of changes in engagement, mental models, and problem definition. Expert evaluators, or subject matter experts (SMEs), assessed the performance results of the Exercise, as compared to other exercises they had observed in the past 5 years. These SMEs rated the Exercise team solution quality, novelty, comprehensiveness, and speed (i.e., time to solution convergence as compared to other exercises they have observed). Collectively SMEs had over 30 years of operational experience in support of these types of exercises. Table 1 shows the type and flow of the instruments used to assess this Exercise.

Table 1. Collaboration Instruments Used in Assessment 1.

	Motivation	Shared Mental Models	Engagement & Satisfaction	Performance	Self-Efficacy
Constructs	- To Collaborate - To Use ACES - For Divergent Thinking	- Problem Construction - Available Expertise of Others - Decision Making Gates	- Process - Outcome	- Novelty - Quality - Comprehensiveness - Speed	- To Use ACES - To Collaborate with Team in Future
Pre-Exercise	Individuals (Survey)	Individuals (Survey)			
During Exercise			Group (Observation)		

Post-Exercise		Individuals (Survey)	Individuals (Survey)	Group (SME ratings)	Individuals (Focus Group and/or Survey)

Analyses

Given the small sample size of Exercise participants (n=6) and SMEs (n=2), the survey ratings were assessed using descriptive statistics of mean comparisons and standard deviation of ratings. In addition, open-ended responses from both participants and SMEs were content coded by industrial and organizational psychologist familiar with collaboration, team mental models, and user engagement of expert systems.

Results

When asked to engage in a multi-location, time-bound NLE Exercise, the ACES environment allowed for innovative and rapid decisions, and was engaging to participants in both sites as assessed by both participants and SMEs. First, when assessed on a 5-point Likert scale (1 = low, 5 = high), pre-assessments indicated that the short training provided by the ACES facilitator resulted in an engaging session (x = 4.8/5), a perception that ACES would be valuable (x = 4.4/5), and that ACES would help the teams accomplish their goals (x = 4.4/5). Open-ended comments supported these findings:

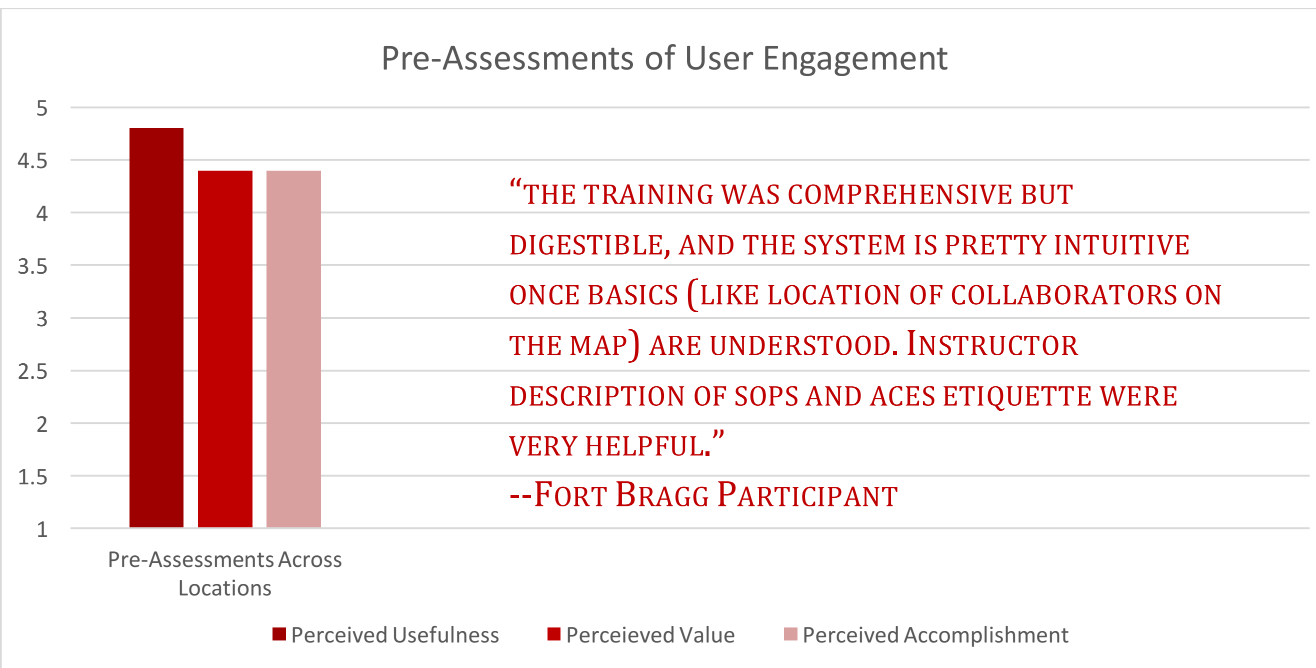

Figure 3. Pre-Assessment of User Engagement.

Because participants reported high inter-personal familiarity within and across locations, there were no detectable changes in shared mental models from the pre to post-assessments for this exercise.

Post assessments indicated that participants were engaged with the process (x = 5/5) and the outcome (x = 4.5/5) that resulted from use of ACES for the NLE Exercise. In open-ended responses, participants compared the ACES system process to previous exercises and reported elements such as joint authoring, situational awareness, and control over content flow as most engaging elements when compared to traditional collaboration approaches:

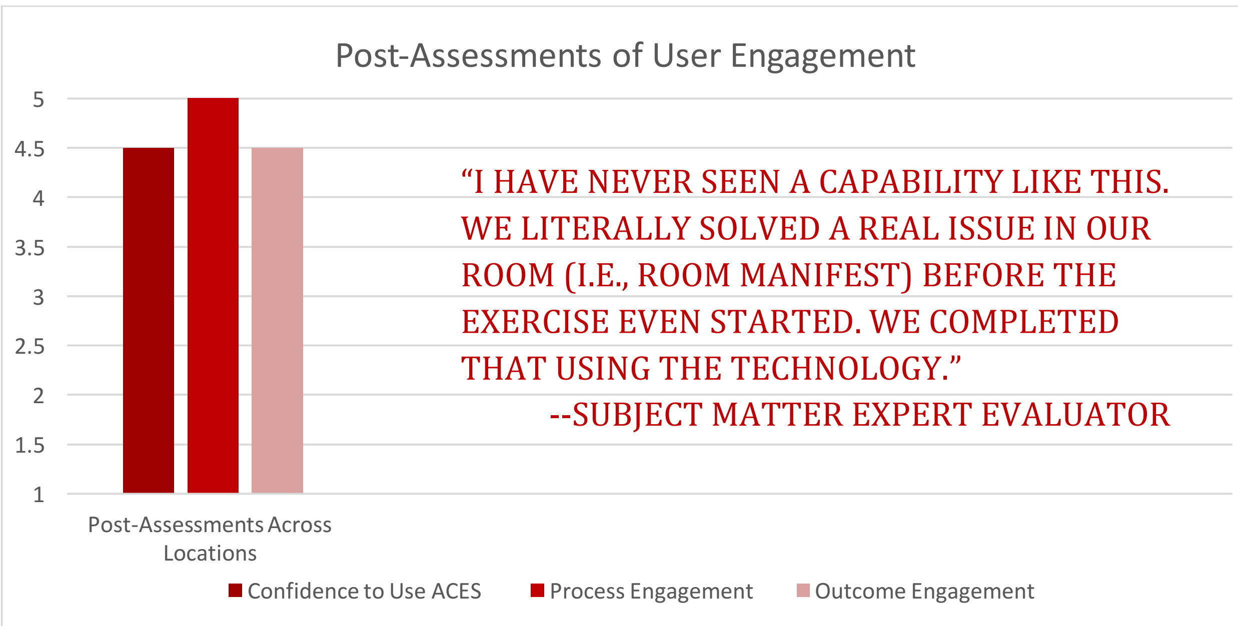

Figure 4. Post-Assessments of User Engagement

While user perceptions of engagement are critical for system adoption, the overall goal of ACES is to improve performance in collaborative settings. Thus, expert ratings of solution innovation were assessed via three indices: 1) quality (i.e., how effective each solution is for the problem at hand), 2) novelty (i.e., how unique the solution is compared to other solutions), and 3) comprehensiveness (i.e., how complete the solution is compared to the problem space it addressed). Each of these indices was assessed on a 5-point Likert scale from SMEs, and these raters then justified their assigned scores with written feedback.

Figure 5. Subject Matter Experts' Solution Ratings

In addition to the innovativeness of the solution, SMEs were asked to assess time to completion, particularly as compared to similar exercises they have observed in the past. The scale evaluators used ranged from 1 (much slower than other exercises observed in past five years) to 5 (much faster than other exercises observed in the past five years). SME Evaluators both rated exercise participants as "5s," or much faster at solution generation and convergence when compared to similar exercises in the past. One SME justified this rating, "ACES provided the ability to quickly assimilate all information from mission planning teams and operators in order to find, fix, and respond to events in the field faster and more deliberately." The other SME stated, "ACES system reduced decision making time 10x compared to existing information sharing systems."

Assessment 2: SME Intelligence Support Collaboration

Background

While Assessment 1 allowed for both participant and SME evaluations of how ACES supporting an existing team, Assessment 2 was designed to provide data on how ACES facilitated information sharing in a more nascent, newly formed team. Individuals are more likely to share information that is already known by all of the members instead of unique information held by the individual (Stasser & Stewart, 1992). When individuals withhold information from one another, whether intentional or not, the group has less information with which to make an effective decision. By sharing information and creating a shared mental model, or shared understanding of the necessary information, members of the decision-making group are better able to generate solutions (Arreola, Robinson-Morral, Crough, Wigert, Hullsiek, & Reiter-Palmon, 2011), evaluate ideas, and plan for implementation (Mumford, Feldman, Hein, & Nagao, 2001). Thus, processes and technology should be in place for individuals to share unique information relevant to problem set.

This is particularly relevant to newly formed teams, comprised of individuals with diverse expertise (Bell & Smith, 2010). For example, when individuals first work together in novel settings, the tendency is to find areas of overlap and agreement to facilitate more harmonious group norms. However, when individuals are brought together specifically to share different types of expertise to solve a complex, ambiguous problem, this type of group dynamic can be particularly problematic. Thus, a goal of Assessment 2 was to examine how ACES might facilitate mental model sharing across a set of Intelligence researchers with varying backgrounds and domain expertise.

Objective

Four researchers with doctoral degrees in clinical psychology, anthropology, political science, and organizational psychology were charged with developing a model to assess the effectiveness of strategic deterrence activities for the Defense Intelligence Agency and USSTRATCOM. In order to accomplish this, each researcher was to share his/her particular domain expertise to help frame different parts of the model. For example, the clinical psychologist, with expertise in threat assessment and measurement, was to identify indicators of State Leader Aggression that were difficult to detect without a systematic examination. The political scientist was to identify country-level differences that might lead to varying degrees of leader discretion in decision making authority.

This team was assembled based on their varying domain expertise; they had not worked together previously nor had all members met face-to-face. Thus, in March of 2016, members of the team traveled from their respective locations (Omaha, Lincoln, Ft. Leavenworth, Austin) to the Rosslyn ACES node to develop a product for USSTRATCOM J2 around the measurement of the "cognitive domain" of deterrence effectiveness. The goal of this Assessment of ACES was to determine how the system either facilitated or hindered collaboration for these individuals working on an ambiguous, ill-defined problem set.

Methodology

Participants spent 7.5 hours in the Rosslyn ACES node. At the beginning of the session, they completed a battery of pre-assessments to measure their pre-training motivation, their understanding of the problem set, and their understanding of each other's expertise. Participants then engaged in a 2-hour training session for the ACES system. Upon completion of training, participants were given a brief break and then reconvened to use the system in on a series of problem solving tasks related to developing a model to assess the effectiveness of deterrence. Upon completion of the session, the four participants immediately completed a set of post-measures meant to assess changes in a) motivation to use the system, and b) mental models for the problem and the team composition. In the following sections, we will detail each of these steps, the analyses conducted, and the results from this Assessment of ACES.

Pre-Assessments. A series of validated measures and open-ended survey items were administered via qualtrics to assess: 1) confidence with the system, 2) mental models for other participant expertise/background (i.e., open-ended items about both surface (e.g., degree) and deep-level (e.g., specific expertise) differences), 3) problem definition (i.e., open-ended items about requirements of the problem, ideas about successful completion).

Model Development Task. The task in this assessment can be described as a divergent, ill-defined problem. The overall goal of the participants was to develop a model to assess the cognitive domain of deterrence effectiveness. In other words, while we know a great deal about measuring an adversary's capability for strategic attacks, we know less about how to assess an adversary's intent to use that capability. Thus, the central task of the team was to develop a model for intelligence analysts to consider when analyzing data related to the cognitive space of deterrence. This required three related yet separate tasks. First, participants engaged in a problem construction task. Each member generated an individual construal of what the key problems of the model development were (e.g., measurement without direct access to the private cognition of an adversary). Next, participants engaged in a divergent problem solving task related to model refinement. Finally, participants engaged in an implementation-planning task to develop a way-ahead for the model completion. This three-stage approach was based on the

Mumford et al. (1995) creative problem solving model. Activities associated with elements of each of the problem solving stages were separated by a 20-minute break, resulting approximately 5 hours of concerted model development from the team.

Post-Assessments: Immediately following the session, the online qualtrics survey tool was again used to gather participant ratings and open-ended descriptions of engagement, mental models, and problem definition. The following section details the analyses conducted and the results from this process.

Analysis

Given the small sample size of Exercise participants (n=4), the survey ratings were assessed using descriptive statistics of mean comparisons and standard deviation of ratings. In addition, open-ended responses from participants were content coded by industrial and organizational psychologist familiar with collaboration, team mental models, and user engagement of expert systems.

Results

When asked to engage in a single-location, time-bound problem solving processes around the development of an intelligence product for deterrence, the ACES environment allowed for faster convergence of mental models about how to define the problem space. This was most illustrated in the participant understanding of what other expertise was available on the team prior to using ACES and immediately following the ACES session. Specifically, prior to using ACES, participants' descriptions of teammates focused on "surface characteristics" (e.g., job title, decision making authority, location/place of employment). However, at the conclusion of the ACES session, participants' descriptions of each other focused on "deep-level characteristics" (e.g., technical expertise, specific ways other members would contribute to the intelligence product). On a five-point Likert scale, clear differences emerged in the mental models of the team when assessed prior to and immediately following exposure to ACES:

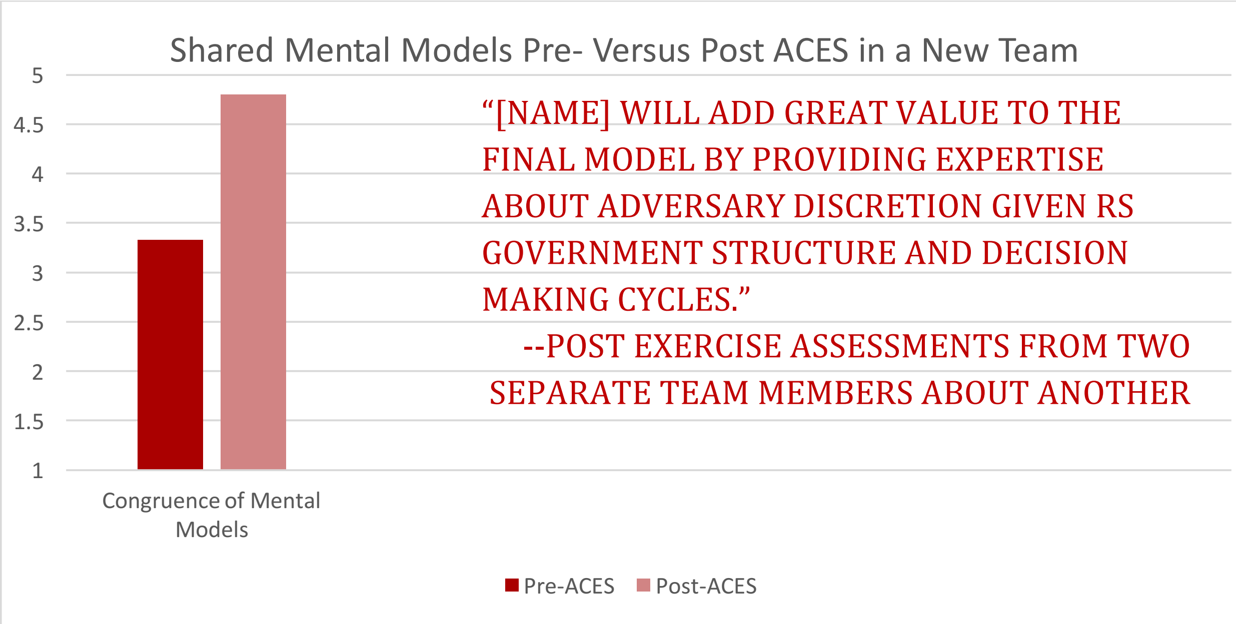

Figure 6. Shared Mental Models Pre-Versus Post ACES in a New Team

While greater congruence in shared mental models was achieved post-ACES, implementation planning was rated somewhat marginal by participants. Moreover, participants reported need for greater closure on the way-ahead. Thus, it seemed that in this Assessment, divergent processes were achieved via ACES, but convergent processes were less satisfactory to participants. In addition, participants described a need for more time with face-to-face interactions at the beginning of the session. This may be due to the lack of familiarity among the team. Thus, one recommendation could be to combine ACES with more conventional, face-to-face collaboration techniques in new teams.

Assessment 3: Industry Users for Real-Time Collaboration

Background

A large online retailer desired to test a redesigned mobile website and a customized native shopping application to enhance their user experience and increase their conversion rates. The company wished to have expert interface designers observe a usability study, coordinate the study in real time, and record all interactions for additional analysis. The goal was to help them better understand "what they may not know" from a shopper's perspective (e.g. is the customer looking for something that is not there, a feature that may be helpful, and why functionality may or may not be used).

ACES was used to connect these two teams with disparate, but related functions and to allow the development team to see in real-time the screen of the eye-tracking workstation that was monitoring the shopper at the other location.

Objective

From the company's perspective the objective of this project was to conduct usability testing with live eye tracking, participant observation, and video recordings of multiple usability sessions. For the assessment of ACES, observers in an "observation room" required multiple video angles of a proctor and participant located in a second "study room". The observers wished to be able to see both the screen of the mobile device with overlaid eye tracking to know where the participant was looking throughout the usability study. In addition, remote observers needed to be able to communicate with the interviewer seated with the participant in order to suggest a question to pose to the participant or issue directions to the proctor directly. The goal was to have the observers feel as if they were sitting in the room without the participants being physically located with the development team.

For the ACES evaluation, the primary functionality tested was screen and data sharing and collaboration with real consequences. The ACES functionality allowed a single resource, the infrared eye-tracker, to be shared in real-time with the two locations. Outcome measures were the quality of the information sharing, shared mental model of website user's experience, ACES users' satisfaction, value of the new capability, and proper functioning of the ACES system.

Methodology

The project looked to provide real time observation and coordination of a user experience analysis looking at a redesigned mobile website and native shopping application. A total of 16 study participants were involved in the project which took place over the course of two, non-consecutive days. The first day focused on the mobile website while another day was dedicated to the native mobile shopping application. Each day consisted of 8 usability sessions where a facilitator would guide a single participant through a series of scenarios and use-cases of the mobile interface being tested. These sessions lasted approximately 45 minutes, this allowed 15 minutes between sessions to reset before another participant was brought into the study room and a new session would begin. The interviewers and participant were located in the ACES room in Mammel Hall, 315 (see figure 3).

The development team located in the ACES room in PKI 335 would view the usability sessions in real time via the ACES-enabled conferencing system. The video conference was established at the beginning of the day. While a session was ongoing, the display located in the study room was turned off so the facilitator and participant would not see the observers. These observers had the capability to see an overlay indicating where on the mobile device's screen the participant was looking as well as see multiple angles inside the study room. The remote observers could send real time chat messages to the facilitator located in the study room if they had any requests or feedback for the facilitator while a session was ongoing.

Figure 7. Study room with mobile eye tracking rig and laptop (foreground center), ACES in Mammel Hall, 315

Participants were asked if they had ever done a focus group or a usability study before. They were then given an explanation of what to expect in this study and what was expected of them. For example, they were to provide feedback on what they liked and didn't like, or how something worked compared to how they expected it to work. In the start of the interview, initial information was solicited to better understand the participant. Questions asked addressed user familiarity with a smartphone, their general uses of their phone, and their experience with shopping on the smartphone, and other similar topics. Following these ice-breaker questions, the actual study began, consisting of a general exploration stage, a questioning stage, a directed task stage, and finally a conclusion and final comments stage. In the general exploration stage, users were asked to think allowed as they began to explore the app without any direction or task. After discovering the app on their own, the study proctor began to ask the user specific questions regarding the capabilities, layout, and features of the mobile interface being investigated. For example, what are the different sections and their purpose. In the third stage, users were then asked to complete a number of tasks while talking about their impressions and thought processes (i.e., a think-aloud protocol). In the final stage of the usability session users were asked concluding questions investigating their overall impression of the interface and any feedback about the interface being investigated.

Each session was recorded and included video of the study participant, and video of the eye tracking overlay of the mobile devices screen. Additionally, notes related to the session were logged and captured in Morae, a software allowing for real time annotation, time stamp marking and notetaking.

Assessment Tools

This study utilized hardware and software technologies from iMotions, Tobii, and the installed ACES systems. In the study room, a laptop running iMotions was used in combination with a web camera and infrared eye-tracking device from Tobii to capture and record participants use of an iPhone. The laptop was connected via ACES to the PKI room. The conferencing camera captured a front facing view of the participant and study proctor was also utilized. In addition, cameras from the ACES system were positioned to capture multiple angles of participants in the usability study room.

In the ACES observation room, observers viewed two large displays showing the multiple video feeds originating from the study room. The front facing video feed was run through a PC with HDMI input capture card to record this front facing angle. This front facing angle was also displayed on a screen in the observation room.

Communication between the two rooms was enabled by TechSmith's Morae software, allowing multiple observers to send instant text based messages to the facilitator in the study room. In addition, the Morae software allowed for time synchronized note taking and event logging of actions and observations regarding participants use of their mobile device.

The final measurement was with the two teams in semi-structured, post-assessment interviews. The purpose of the interviews was to determine their satisfaction with the process and the quality of the interaction. Additionally, we assessed if the system facilitated a shared understanding between the remote development team and the interviewer.

Analysis

For this project the participant was able to simulate a typical mobile shopping experience and provide the critical feedback that was beneficial for their system redesigns and upgrades. They were able to capture specific pieces of objective information utilizing heat maps and gaze patterns to determine customer preferences, likes, dislikes and missing information that could be utilized to make their website more interactive, efficient and productive.

ACES provided them a capability that they had never had before. It afforded them an opportunity to conduct an interactive study and complete their observation and analysis during the session from a remote location. The video capture enabled them to make notes and include markers to identify key points of information for more in depth analysis.

We analyzed the post interviews for themes about information sharing, the quality of process, and recommendations / dissatisfaction with the system. These themes and recommendations were consolidated by the research team.

Results

The company was very pleased with the information sharing – in their words, "it was unprecedented". ACES coupled with the eye-tracker provided the company with a capability that they had never had before. The ACES system was leveraged in both setup and execution of the study. Prior to the study beginning, ACES allowed for individuals setting up in the study room, to better understand how the observers would view the sessions by display the screens of the observation room in the study room itself. This ability enabled the setup to be adjusted and several layouts to be tested quickly and for both parties in the different locations to understand the setup benefits and drawbacks from the other locations perspective and the varying expertise. This pre-participant collaboration alone, created a more complete and improved usability test.

When setting up the study initially, the plan was to use a virtual machine hosted within ACES to enable a live screen share of the study rooms laptop. This setup worked, but did not provide a smooth video display capable of replicating the eye-tracking gaze points. Likely because the processing power on the laptop struggled to maintain the hardware supporting the eye-tracker and the ACES virtual machine. Ultimately it was decided to use the second input on the Cisco video conferencing system to directly feed the remote ACES room. This setup worked flawlessly.

Another finding related to the synchronization of video feeds from the multiple cameras located in the study room. While the cameras themselves provided the desired multi-angle look into the study room, each of the video feeds had slight delays. The decision was made to use a more direct (e.g., native Cisco) video feed as the main participant view, and to use the remaining cameras as the alternative angels into the study room. This latency will need to be improved in future iterations of ACES.

Related to the findings for the company, the usability study was a tremendous success and findings have had an immediate impact resulting from changes they have implemented. They stated they were able to immediately identify critical flaws in their website design. This insight allowed them to collaborate with their team members and develop immediate solutions. The resulting changes drove positive feedback from their customers, and a significant improvement in purchase conversion rates, sales and profits. This project was so successful that using the ACES environment is now their preferred method to capture information and they have immediate plans to use this for multiple design projects for both their mobile and desktop applications. In summary, the information sharing, process satisfaction, and resulting mental models were superb.

A number of recommendations were suggested for future iterations of the ACES-OE. First, a number of collaboration scenarios require the use of expert systems and devices that are brought into the ACES environment and needed to be accessed by users of the ACES system. Developing better integration technologies for these outside devices would greatly aid in the ability of ACES to meet a diverse range of use cases and scenarios. A second recommendation relates to the present implementation of audio/video equipment especially the synchronization of various video feeds from the IP cameras. As collaborators look to use these video feeds as information sources, their lack of synchronization negatively impacts their ability to be used as primary collaboration tools.

Overall, ACES significantly enhanced collaboration and provided a new capability. It did not fundamentally alter their process for usability test, but it did enhance it. The objective and subjective measurements indicated tremendous value.

Assessment 4: Convergent Solution Experimentation

For the final two assessments, we conducted four experiments that tested multiple conditions. The goal was to understand how ACES operated depending on different types of problems and task complexity. The idea of problem complexity has been studied for decades, and has led to numerous definitions of complexity, along with breakdowns of its various subcomponents. The complexity of a problem is often characterized by the number of resources and types of interactions of those resources required to solve the problem. For this assessment, we examined the structure of a problem.

Comparisons of ill-structured and well-structured problems have been around since research on problem solving from the 1970s (Simon1977). Well-structured problems are those for which all relevant information is known ahead of time, the problem state can be clearly defined and actions will result in predictable outcomes (Jonassen1997). Often well-structured problems have a single solution, and the challenge of problem solving is identifying the sequence of actions that will move the solver from the identified problem state to the solution.

A part of the complexity of ill-structured problems is the size of the set of possible solutions (Simon1971). Well-structured problems, as explained above, often have a single solution or small set of solutions. Ill-structured problems, however, may have a large set of solutions. For some problems, it is only important to arrive at a solution, without regard for solution quality. In these situations, a large set of solutions can make things easier, as it becomes a simple task to find a solution that resolves the problem. In other problems, it is important not only to come up with a solution, but to identify a **good solution**, or even the **best solution** to the given problem. In these situations, a large solution set can be a part of the problem, as comparing solutions and weighing alternatives becomes a challenge in itself.

To examine this, we created four experiments. This first experiment focused on well-defined, but challenging problems. This type of collaboration and decision-making is different than creating new solutions or options (Convergent Process Support Required). The second focused on ill-defined problems (i.e., only symptoms of problems). For this type of collaboration, there needs to be processes for data synthesis and conceptual combination (Divergent Process Support Required). Existing academic literature suggests benefits and challenges arise when moving from individual to group decision-making. Decision quality is enhanced when multiple perspectives are brought to weigh in on a particular issue (Amason, 1996). However, when adding more individuals to the process, they are likely to prefer one solution to another and take sides. Sheer power wins the battles of choice, meaning the more people you have supporting an idea, the more likely it is that the idea will be selected (Eisenhardt & Zbaracki, 1992). This can become more complicated when examining how individuals tend to choose sides. Natural separations, or fault lines, within the group can occur based on the status of the group members – not on issue or values (Eisenhardt & Bourgeois, 1988).

The goal was to determine how the ACES platform supported these two types of problems that required different types of collaboration support (e.g., Convergent and Divergent). Figure 4 below shows the experiment plan for the following two assessments.

		Problem Type	
		Divergent Task	Convergent Task
Collaboration Modality	Video Conference with Common Text Space	Team Creative Strategic Problem Solving	Not Planned
	ACES	Team Creative Strategic Problem Solving	Critical Thinking / Single Solution
	In Person	Not Planned	Critical Thinking / Single Solution

Figure 8. Experiment Plan for Divergent and Convergent Experimentation

Background

Convergent thinking is the type of thinking that focuses on coming up with a single answer to a problem (i.e., the collaborators "converge"). It is oriented toward deriving the single best, or most often correct answer to a question. Convergent thinking emphasizes speed, accuracy, and logic and focuses on recognizing the familiar, reapplying techniques, and accumulating stored information (Cropley 2006). In this view, answers are either right or wrong. The solution that is derived at the end of the convergent thinking process is the best possible answer the majority of the time.

The intent of this assessment was to determine how ACES performed / facilitated teams when the solution space didn't require divergent thinking, but rather generating a single, established solution. This type of process focuses mainly on case matching and recognition of similarly- or identically-structured problems. Team members likely engage in comparative analysis of expected and past cases and elicit potential responses to the situation based on isomorphic events where established solutions exist. Decisions are confined to generally using or producing established solutions and options.

Individuals are likely to compare what they know about a prior, similar situation and use that schema to code how this situation will unfold. Depending on how much knowledge individuals have of similar problems or solutions, the amount of information seeking required may vary. After benchmarking against pre-formed mental schema, individuals are likely to compare the sources of information, if more than one source exists. For instance, if you have infrared, intelligence, and satellite sources giving you similar or contradictory information, you may weigh each type differently when deciding which type of action should be taken.

Objective

Teams of students were assigned to complete various programming challenges. The goal of the assessment was to compare the time it takes to complete each challenge with the ACES environment and without the ACES environment in order to test the effects of collaboration using and being immersed into the ACES environment. The key measurement was time to complete the solution, quality of the solution generated, and the users' satisfaction with the system.

Methodology

We recruited five teams to participate in the study. All of them were college undergraduate students that have programming experience. The students were randomly assigned to a team and asked to complete a pre-survey questionnaire. The teams were then randomly assigned to either ACES first, or the in-room-only condition first. They were then given the following instructions, "In this study you and your partner will be completing four programming challenges in Java using the HackerRank website. Two of the challenges will be completed with the ACES system, and two in a normal environment. For the ACES portion of this study, you will be given a short presentation on the functionality of ACES before you begin. When you have finished each challenge, let us know so we can save your code, then advance to the next challenge. We will begin with the (ACES/non-ACES) portion of the experiment".

Each of the challenges is described in the table below.

Table 2. Challenge Descriptions

Challenge Number	Description	Justification
1	This challenge tests the user's knowledge of strings. In order to complete this challenge, the user will have to be able to compare each character of a string to determine if the string is an anagram or not. If the string is an anagram the output should say "anagram," if not, the output should say "not an anagram".	This challenge was chosen because it is easy enough to understand, but difficult enough to allow collaboration. The teams will be expected to work together on the challenge to come up with the most effective code that will output what is expected
2	This challenge tests the team's ability to comprehend strings. Each team will have to determine in a word is a palindrome, if it is-- the output of the code will be "palindrome." If it is not a palindrome, the output should say "Not a Palindrome." This challenge will be fairly simple and should not take a lot of time, especially since each team will be collaborating with each other to come up with the best solution	This challenge was chosen because it tests a basic concept that is taught in Java 1. We wanted to make sure that the concepts were easy to understand, but difficult enough finish alone within the time constraint. We are trying to inspire collaboration by choosing challenges that are easy to understand, but difficult to complete alone.
3	Challenge 3 tests the teams' ability to use a string and count how many words are in the string and display that number as the output. This challenge seeks basic knowledge of Java, but easy to understand capabilities. What is expected of the teams is to solve the challenge	This challenge was chosen for the ACES study because it will force collaboration inside of the teams. There is no right way of solving this problem, and we are hoping that the code will be more creative and unique

	using collaborative methods. This can be done by using IDE's that each team is familiar with and copying that code into the Hackerrank website.	inside of the ACES environment rather than in a different setting.
4	This challenge is asking more of the user's knowledge of programming. The biography informs the user of the methods that will need to be used in order to complete the challenge. Using what is given to the teams the challenge should be solely based upon collaboration.	This challenge was chosen because of this demand for collaboration. It will take discussion and planning to come to a conclusion that meets all of the requirements of this challenge. This challenge should prove that ACES environment does enhance creativity rather than an environment that is distracting and does not have the same capabilities that the ACES environment has.

All the teams were given and introductory training to the ACES system and had a chance to use it and get familiar with the technology prior to the experiment. At the end of the experiments, the uses were given a post-survey measure that gauged their satisfaction with ACES and asked open-ended questions about the system.

Assessment Tools

Process and outcome satisfaction were measured based on the instrument developed by Briggs, Reinig, and De Vreede (2006) to assess the affective arousal on the part of a participant with respect to procedures and tools used in a collaboration interaction. These two, four-item scales were meant to measure affective aspects of satisfaction, asking whether certain aspects of a collaborative interaction were valued (such as perceptions of process fairness). In the present study, we obtained appropriate scale reliability of α =.996, and thus aggregated items to obtain scale scores.

We also measured the time to converge and derive the solution, the completeness of the solution, the quality of the solution, and their perceptions of the ACES system.

Analysis

We began by analyzing the time that each team took to complete the various challenges. This data is summarized in table 3 below where each of the values is in seconds. The labels have the following mean: ACES1 is the first challenge that team completed using ACES and ACES2 is the second. NON1 is the first challenge not using ACES, etc.

Table 4 shows the percentage of the problem that was solved by each team in each condition. It is important to note that the time was capped at 30 minutes (1800 seconds). So, it is conceivable that the teams would have completed the exercise given more time.

Table 3. Summary of time to complete

Team ID	ACES1	ACES2	TEAM AVERAGE ACES	NON1	NON2	TEAM AVERAGE NON	DIFFERENCE (NON-ACES)
2	615	1800	1207.5	259	202	230.5	-977
3	1231	256	743.5	1572	1714	1643	899.5
5	1800	1800	1800	1800	600	1200	-600
4	1800	752	1276	1061	991	1026	-250
1	1626	1800	1713	927	818	872.5	-840.5
AVERAGE	1414.4	1281.6	1348	1123.8	865	994.4	-353.6
COMBINED AVERAGE	1348			994.4			

Table 4. Percentage of problem solved

Team ID	ACES1	ACES2	NON1	NON2
2	100%	60%	100%	100%
3	100%	100%	100%	100%
5	0%	0%	0%	100%
4	50%	100%	100%	100%
1	100%	0%	100%	100%
AVERAGE	61%		90%	

Results

Time to complete

Of note, the teams that did not use ACES performed faster on the convergent tasks by an average of 5:53.6 seconds and only one of the teams saw a speed improvement using ACES, while four teams performed much faster without it. In fact, when comparing with teams, team speed without ACES was faster from 2:10 to 16:27 minutes. There are multiple confounding variables that could explain this:

1. It could be that the personalities of the individuals and team dynamics played a role. This would need to be evaluated further to determine if certain personality types (e.g., introverts) have difficulty in this new flatter collaborative environment.

2. The fact that the technology is new. There may have been a learning curve associated with it. This learning curve may have "gotten in the way" of the benefits of working to solve the problem.

3. The problems varied in difficulty and not all team members' abilities were the same. However, random assignment should have washed out this variance and comparing within teams would not account for this anomaly. Further data collection is required to be certain.

4. The technology may not be well suited for convergent types of problems.

Completeness / Quality

Similar to time to complete, teams generally performed better when not using the ACES system for the convergent problems. On average the teams completed 90% of the solutions in the non-ACES condition, and only 61% with the ACES platform. Additionally, all of the teams improved or performed equally well without the ACES system. Four out of the five teams improved their performance and one team performed equally well. Similar factors as outlined in the section above may have been at play.

It should also be noted that at the start of the sessions (in the pre-survey), team members were asked, "Do you feel prepared to work collaboratively on this project?" and asked to reply on a scale from 1-5 with 5 being the most eager to collaborate. The average for all users was a remarkably low 1.2. This lack of initial willingness to collaborate may have been a large driver in the fact that benefits were not realized. It is also interesting to note that after the users had interacted with the system, they were asked a series of questions in a post-survey. One of these questions, was "How motivated are you to use the ACES system for future collaborations?". The average was 2.3, nearly double the initial question. It appears that ACES may have had an effect on willingness to collaborate at the individual level. It would be interesting to see if this trend would continue based on repeated exposure to and use of the system over time.

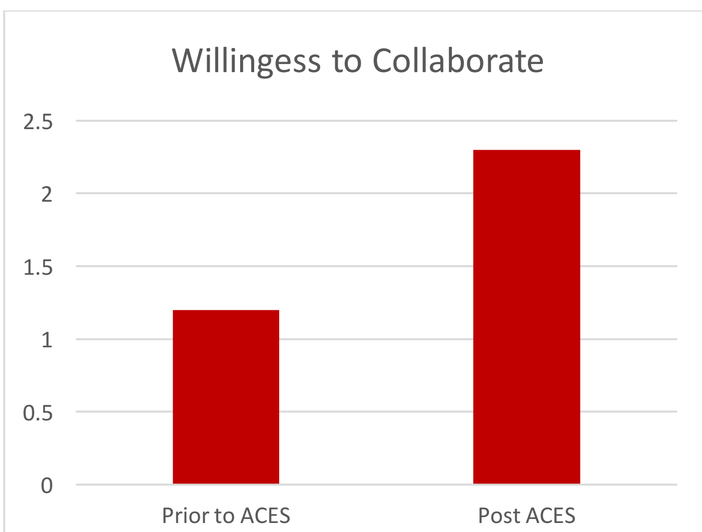

Figure 9. Willingness to collaborate prior to and post ACES interaction

User satisfaction measures

After the users' interaction with ACES, we asked them a 12-item survey that had 3 constructs, measured by 4-items each. The constructs of interest were: perceived value of the ACES system, process satisfaction, and outcome satisfaction.

In order to gauge perceived value, the following questions were asked on 5-point scale:

- The ACES work session was worth the effort I put into it.
- The things I accomplished in work session the ACES warranted my effort.
- The results of the ACES work session were worth the time I invested.
- The value I received from the ACES work session justified my effort.

In order to gauge process satisfaction, the following questions were asked on 5-point scale:

- I feel satisfied with the way in which the ACES work session was conducted.
- I feel good about the ACES work session process.
- I feel satisfied with how my team collaborated during the work session.
- I feel satisfied with the procedures used in the ACES work session.

In order to assess outcome satisfaction, the following questions were asked on 5-point scale:

- I liked the outcome of the ACES work session.
- When the ACES work session was finally over, I felt satisfied with the results.
- I am happy with the results of our team's collaboration using ACES.
- I feel satisfied with the things I achieved in the ACES work session.

In order to ensure scale validity, we calculated Cronbach's alpha. The results were as follows: Perceived-Value = .968, Process-Satisfaction = .885, and Outcome-Satisfaction = .937. So, each of the measures could be calculated using an average. The results were: Perceived-Value = 3.94, Process-Satisfaction = 3.58, and Outcome-Satisfaction = .3.80

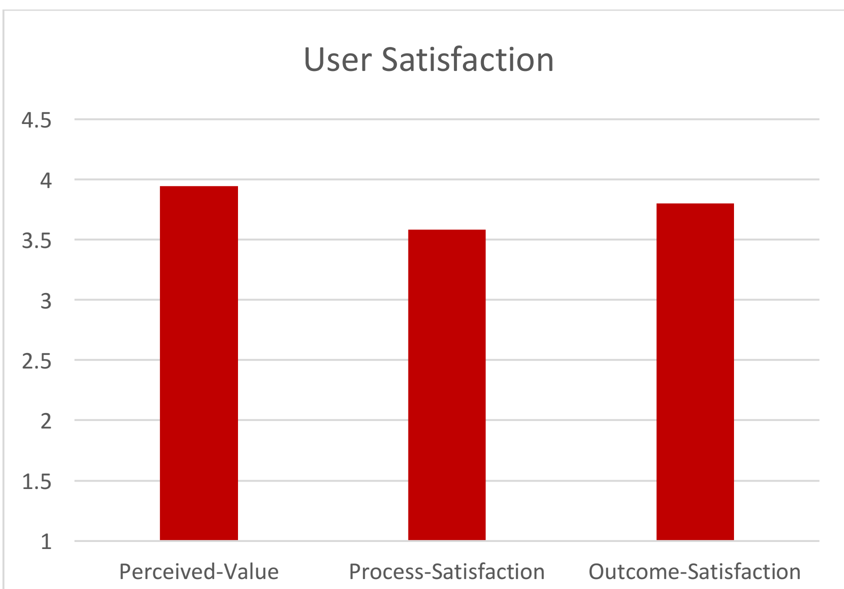

Figure 10. Measurement of User Satisfaction

The results are surprising because the teams performed slower and worse using the ACES system than without it, yet they all had high ratings for the perceived value and satisfaction measures. The indicates that perhaps learning the technology played a role in the degraded performance and that the users saw great value in the system and the processes associated with it.

User comments and feedback

After the experiment, we asked the users a series of open-ended questions, some of the most pertinent are summarized in the tables below.

Table 5. After today, what concerns do you have about the ACES system?

The system is far from easy to use. Sharing entire VMs seems like overkill, especially if I want to share only one application within that VM. There seems like very, very few ways in which you can share a screen, which slows me down -- completely opposite to what the system should be doing. Simply, put I have to think entirely too much in order to use the system.
How easy it would be to schedule time to use the facility for any projects that I would like to do in the future?
Bugs make it difficult to work sometimes
I noticed that while I was able to view my partner's screen, I was not able to directly change anything on his screen. This means that I needed to physically tell my partner some changes that may have helped his solution.
The system was a bit slow when using it, also make sure quality hardware is used if needed.
The front screen bars are kind of annoying because they make it hard to see content sometimes. My partner also seemed to have less control over the large screen vb than I did.
Passing of VMs can be problematic since one item moved to your area messes up the VM and you have to resize it. / 4K screens can make it hard to use some of the smaller stuff on the screen.
Sometimes it can get slow to respond

Table 6. What were your favorite features of the ACES system? Why?

The ability to easily manipulate the positioning and locations of the virtual machines. It makes moving work around to be viewed by other people easy.
ability to watch other persons screen made it easy to tell where we were doing it wrong
Being able to share my VM with my teammate was nice. We were able to see the different ways we were going about solving the problem
I like being able to send my screen to others because it makes it so much easier for them to see what I'm working on.
Easily "throwing" up my screen on the big screen so that everybody could see it.
I like being able to drag back and forth between team screens especially because of the way we were able to use it for syntax discovery combined with coding.
The ability to easily share screen with each other made it easier to figure out what to do.
You could pull up screen of your computer and show it on your partners screen

Table 7. What improvements would you like to see to current ACES system?

How you share and what you can share over the system feels it needs to be reworked completely.
There is a little bit of lag other PCs don’t have this that I work with
Just basic bug fixes and polish
Being able to directly input into another person's VM is a need. It may have been because I used the 3 screen setup, and we used my partner's VM in the shared screenspace in the middle, but I'm not sure. / / Another improvement would be the ability to resize VM's. I found myself struggling to correctly place my whole VM onto my screen
I would like to be able to have ACES auto-fit the screen I drag onto my screen. The resizing thing doesn't work for me for some reason.
A more reliable system, some problems happened and I did not know why or what was going on/causing them.
Just a little more seamless user experience. It was mostly fine but there were a few hiccups with moving things around. Automatically snapping to the size of the screen you moved your window to would be very helpful as well.
A better UI. A better way of handing window sizing of VMs / Better definition of left/right screen on the L desk.

Assessment 5: Divergent Solution Experimentation

Background

Problem solving research examines problems ranging from highly-structured to ambiguous. Creative problem solving is unique from other types of problem solving in that it focuses specifically on ill-defined or ambiguous problems (Anderson, 1983; Mumford, Whetzel, & Reiter-Palmon, 1997). Well-defined problems are highly structured and have only one correct solution. Consequently, only one pathway exists to solve these problems. Conversely, ill-defined problems have many possible solutions, are more complex, and require a less uniform problem-solving process than do well-defined problems (Mumford et al., 1991). Ill-defined problems are also highly ambiguous, making individual differences in creativity more apparent. As the differences between these two types of problems suggest, performance on well-defined problems has been shown to be independent of performance on ill-defined problems (Schraw, Dunkle, & Bendixen, 1995).

These divergent, creative processes focus on new solution generation, data analysis, conceptual combination and synthesis of ideas, rather than information dissemination. Decision-making is more complex as data and sources are considered and solutions are generated. The generation of alternative options is based on prior mental models, conceptual combination of existing solutions, and problem definition. Individuals must gather knowledge from past similar, but not identical, cases and from expert sources that will deliver new, useful information. Collecting new information happens continuously throughout the process. Once new information has surfaced, that information is combined with past cases. From the result of this conducted analysis, synthesis, and combination, individuals may decide to seek even more information to ensure that proper solutions are generated. Decision-makers should define the problem in multiple ways prior to generating solutions to assure that all parts of the event are being addressed and a complete solution may be generated. For divergent processes, the technology chosen to facilitate decision-making should focus on rapid data synthesis and situational sense-making.

Multifarious problems present an ongoing challenge for organizations. Indeed, many of the important strategic problems facing organizations demonstrate complexity, interdependence, and are ill-defined. The purpose of this study was to examine how teams with a divergent process support may or may not benefit from the ACES technology. To test this, we developed a strategic problem with twelve different formulations (i.e., problem symptoms). Strategic problem formulation is defined as a formalized causal representation of a given symptom or web of symptoms. Organizations often become aware of a potential "problem" through the observation of its symptoms.

The distinction between a problem and its symptom is worth noting. For example, a loss in market share represents a symptom while possible problem formulations range from new technologies, incoming competitors, governance misalignment, or supply chain inefficiencies. The ill-defined, complex, and interdependent nature of complex problems result in a wide range of possible problem formulations, effectively obscuring the emergence of a clear choice. Subsequently, multifaceted problems often result in disagreement and confusion within the top management as to the appropriate course of action.

Methodology

We first identified real organizations facing complex problems. Six problem descriptions across various industries were reviewed. Upon evaluation for appropriateness and complexity, a difficult problem facing a large consumer electronics retail firm was selected. A review of the information surrounding the problem indicated competing views among stakeholders regarding how to proceed, as well as the presence of multiple interrelated problem formulations. Subsequently, we adopted the underlying framework of the company to serve as a baseline in developing a problem that would be appropriate in an experimental setting. Details regarding the focal company's salient resources and capabilities were collected as well as information regarding the firm's history. Relevant problem formulations were also categorized and labeled. The complexity and approximate length of the problem formulations were increased or decreased to ensure equivalency across formulations.

The end result placed participants in the role of turn-around specialists at "Bordet Electronics," a fictitious consumer electronics firm facing declining sales and falling stock prices. Participants were first given a brief history of the company and were told that their task was to develop a unique and creative solution that would provide economic value and help the company compete. 33 students from a Midwestern university were recruited to participate in this study and were compensated with course credit and participants were randomly assigned to a virtual team consisting of three people and the teams were randomly assigned either to a video-conference with chat environment or to the ACES environment. The video-conference team environment is shown in figure 11.

The screen in the video-conference condition allowed the teams to edit the solution simultaneously and to see each of the teammates. It also provided an overview of the resources and capabilities of the company, the different formulations, and a button to review the entire case again. Each team was given twelve different formulations of the problems at Bordet and the categories of problem formulations included: the emergence of Internet, customer experience, product offerings, and the management of the businesses in which the company operates.

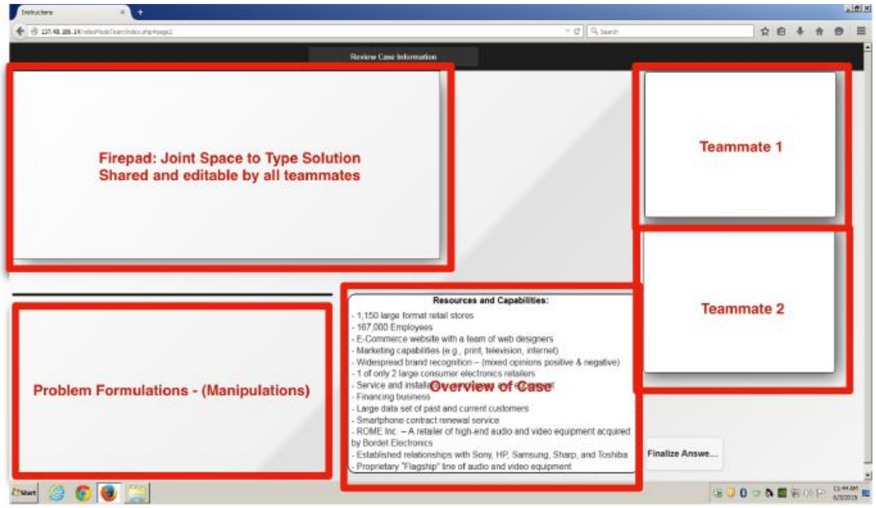

Figure 11. Screenshot of the User Interface for Video Conference

Teams were again instructed to virtually collaborate and develop a solution that was both novel and valuable. They did this by talking and working in the shared space to produce the solution. A screenshot of one of the combined team videos is shown in figure 12. Based on the random assignment, the sample included 5 teams in the video condition, and 6 teams in the ACES condition. Before the study in the ACES room began, participants were trained on how to use the Photon client to control the ACES environment.

Once the teammates were satisfied with their solution, they pushed the button to finalize and submit it. Finally, all participants completed a post-survey.

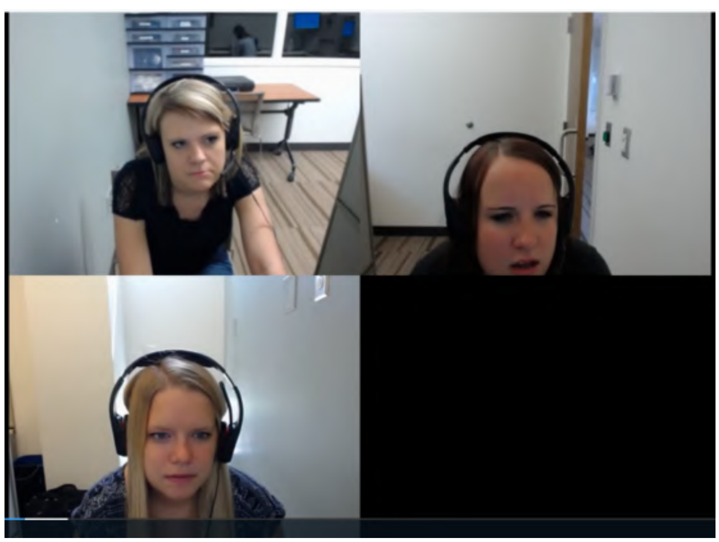

Figure 12. Sample Video of the Teams working on the Problems

Assessment Tools

Three research assistants that are familiar with evaluating performance on ill-defined problems developed benchmark ratings scales to evaluate the novelty, quality, and elegance of the solutions. Using procedures developed by Bessemer and O'Quinn (1995), raters then underwent 15 hours of calibration training and then independently rated all solutions. Raters used a modified Delphi method to reach agreement across ratings. Ratings were all in agreement within one scale point and the Intra-class correlation coefficient for each measurement was 0.792, 0.791, and 0.691 respectively (novelty, quality, elegance), showing excellent measurement reliability.

Raters evaluated the degree of novelty of the proposed solutions on a 5-point Likert-type scale, where 1 represents a common and or frequently proposed idea and 5 represents an uncommon and or infrequently proposed idea. The greatest discrepancies among raters were resolved through discussion.

Raters also evaluated quality by considering how many elements of the problem were effectively addressed by the solution. Solutions were then rated on a 5-point Likert-type scale, where 1 represents no elements addressed and 5 represents over five elements addressed. The greatest discrepancies among raters were again resolved through discussion.

Raters assessed elegance of the solution by determining how parsimoniously the problem was realistically solved. Parsimony was rated on the clarity, brevity and completeness of the solutions. Solutions were rated on a 5-point Likert scale where 1 represents a redundant and unrealistic solution and 5 represents a succinct and realistic solution.

Analysis

Because of the excellent inter-rater reliability, we were able to average the measurements from each of the raters across all teams for all conditions. The results are shown in table 8 below.

Table 8. Average data from video chat and ACES environments

Condition	Novelty	Elegance	Quality	Word Count
Video	2.2	3.0	3.0	194.8
ACES	2.6	3.6	3.8	239.7

Because of the small sample size, it is difficult to draw any definitive conclusions as to the differences. We simply do not have the statistically power to categorically state support for the hypotheses that ACES-supported groups generate more novel, more elegant, and higher quality solutions. However, even with the limited sample size, the statistics are suggestive that these hypotheses would be supported. We conducted Welch's two-sample t-test for each variable and found that even with only 11 teams, there was statistical support to show that the difference in elegance (t=2.0581, df = 6.1383, $p < .05$ one-tailed) was statistically significant. Similarly, quality (t=1.3525, df = 6.0635, $p < .1$ one-tailed) and novelty (t=0.8822, df = 7.7427, p = .2021 one-tailed) would likely be significant with a more robust sample.

Results

The teams that used ACES performed better in all three measures of the experiment. They generated more novel solutions and the creativity of the team was better. Moreover, their solutions were more elegant as measured by how parsimoniously and realistically the problem was solved. Parsimony was rated on the clarity, brevity and completeness of the solutions. Realism was rated based on how the applicability and how implementable the recommendations were. Finally, the ACES teams also generated higher quality solutions. This was determined by the number of issues that were addressed and the overall effectiveness in addressing the problem. However, it should be noted that the standard deviations of solution novelty, quality, and elegance indicate that a fair amount of variability exists within groups. Thus, it is likely that other factors may also contribute to solution creativity. Figure 13 below summarizes the results.

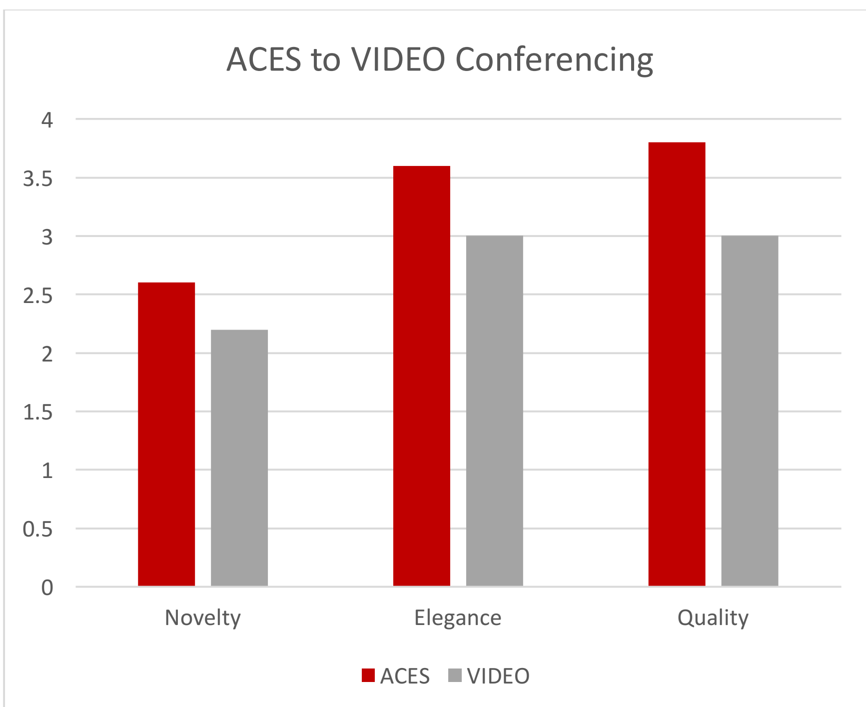

Figure 13. Performance of ACES to Video Teams

Findings and Conclusions

Table 9 summarizes the result of each assessment.

Table 9. Summary of all assessments

Assessment	Primary Users	Description	Result
National Event	Army and Navy Active Duty and Civilians	Support to a high consequence exercise for a convergent task. Examine potential operational benefits	Users reported high-engagement and high value. SME ratings indicated that ACES generated more novel and higher quality solutions.
Intelligence Support	Post-graduate Contractor SMEs	Create an intelligence support document. Assess the creation of shared mental models.	Significant and rapid increase of shared mental models among experts with disparate expertise.
Industry Usability Test	Industry Professionals and Normal Users	Link a development to market research team. Examine multi-team system.	The information sharing, process satisfaction, and resulting mental models were superb. Technical improvements identified.
Convergent Lab Experiment	Graduate and Undergraduate Students	Determine how ACES supports convergent collaboration tasks.	ACES teams performed slower and less effectively in the convergent tasks. Technical improvements identified. However, willingness to collaborate dramatically increased.
Divergent Lab Experiment	Graduate and Undergraduate Students	Determine how ACES supports divergent collaboration tasks.	ACES teams outperformed video conferencing, shared-text-space teams on all measures (novelty, elegance, and quality).

It appears that ACES offers significant promise, especially for geographically-separated, multi-team systems, and for ill-defined, divergent collaborative processes. Co-located teams with a convergent process did not benefit from ACES. The greatest benefits came to experts with diverse expertise, across geography.

Recommendations

ACES shows great promise. However, the technical recommendations specifically identified in Assessments 3 and 4 need to be considered and the interface needs to be improved. Additionally, most of the experiments serve as pilot data and because of small sample sizes, additional data collection is needed for conclusive statistical power.

We recommend that ACES move forward to a broader operational JCTD with both operational and experimental testing. We recommend that additional experiments also be conducted to validate the ACES value propositions.

Selected References

Amason, A. C. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Academy of Management Journal, 39*(1), 123-148.

Anderson, P.A. (1983). Decision making by objection and the Cuban missile crisis. *Administrative Science Quarterly, 28*, 201-222.

Arreola, N. J., Robinson-Morral, E., Crough, D. A. S., Wigert, B. G., Hullsiek, B., & Reiter-Palmon, R. (2011). *Creativity in teams: The role of shared mental models.* Poster presented at the 26th Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

Bell, B.S., & Kozlowski, S.W.J. (2002). A typology of virtual teams: Implications for effective leadership. *Group & Organization Management, 27*, 14-49.

Cannon-Bowers, J. A. and Salas, E., "Reflections on shared cognition," *Journal of Organizational Behavior,* vol. 22, pp. 195-202, 2001.

Cannon-Bowers, J. A., Salas, E., and Converse, S., "Shared mental models in expert team decision making," in *Individual and group decision making*, N. J. Castellan, Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1993, pp. 221-246.

Cropley, Arthur (2006). "In Praise of Convergent Thinking". *Creativity Research Journal*. 18 (3): 391–404.

Briggs, R.O., Reinig, B.A., and de Vreede, G.-J. Meeting Satisfaction for Technology-Supported Groups An Empirical Validation of a Goal-Attainment Model. *Small Group Research*, *37*, 6 (2006), 585–611.

Cooke, N. J., Salas, E., Cannon-Bowers, N. J., and Stout, R. "Measuring team knowledge," *Human Factors,* vol. 42, pp. 151-173, 2000.

DeChurch, L. A., Burke, C. S., Shuffler, M. L., Lyons, R., Doty, D., & Salas, E. (2011). A historiometric analysis of leadership in mission critical multiteam environments. *The Leadership Quarterly, 22(1),* 152-169.

Eisenhardt, K. M., & Bourgeois, L. J. III. (1988). Politics of strategic decision making in high-velocity environments: Toward a midrange theory. *Academy of Management Journal, 31*(4), 737-770.

Eisenhardt, K. M., & Zbaracki, M. J. (1992). Strategic decision making. *Strategic Management Journal, 13*, 17-37.

Hollingshead, A. B. (1996). Information suppression and status persistence in group decision making: The effects of communication media. *Human Communication Research, 23*(2), 193-219.

Klimoski, R., and Mohammed, S., "Team mental model: construct or metaphor?," *Journal of Management,* vol. 20, pp. 403-437, 1994.

Mumford, M. D., Feldman, J. M., Hein, M. B., & Nagao, D. J. (2001). Tradeoffs between ideas and structure: Individual versus group performance in creative problem solving. *The Journal of Creative Behavior, 35*(1), 1- 23.

Mumford, M. D., Mobley, M. I., Reiter-Palmon, R., Uhlman, C. E., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal, 4*(2), 91-122.

Orasanu, J. "Shared mental models and crew decision making," Cognitive Sciences Laboratory, Princeton University, Princeton, NJ Technical Report No. 46, 1990.

Reinig, B., Briggs, B. Shepherd, M.., Yen, J. and Nunamaker, J., “Developing and Validating an Instrument to Measure the Impact of Group Support Technology on Affective Reward.” HICSS 95.

Rentsch, J. R., and Hall, R.J., "Members of great teams think alike: a model of team effectiveness and schema similarity among team members," *Advances in Interdisciplinary Studies of Work Teams,* vol. 1, pp. 223-261, 1994.

Stasser, G., & Stewart, D. (1992). Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgment. *Journal of Personality and Social Psychology, 63*(3), 426.

Zaccaro, S. J., Marks, M. A., & DeChurch, L. A. (2012). Multiteam systems: An introduction. *Multiteam systems: An organization form for dynamic and complex environments*, 3-32.