Journal of Universal Computer Science, vol. 17, no. 12 (2011), 1638-1658
submitted: 15/10/10, accepted: 28/1/11, appeared: 1/8/11 *© J.UCS*

Investigating Collaborative Innovation in a Virtual World Task

Philip Vahey
(SRI International, Menlo Park, California, United States
philip.vahey@sri.com)

John Brecht
(SRI International, Menlo Park, California, United States
john.brecht@sri.com)

Charles Patton
(SRI International, Menlo Park, California, United States
charles.patton@sri.com)

Ken Rafanan
(SRI International, Menlo Park, California, United States
ken.rafanan@sri.com)

Britte Haugan Cheng
(SRI International, Menlo Park, California, United States
britte.cheng@sri.com)

Abstract: While much has been written about the importance of innovation, there is still much to learn about the specific behaviours that lead to innovation among groups. In this paper we introduce a framework of innovation based on behaviours identified as being conducive to collaborative innovation. We also report on a study of a task designed to elicit innovation supportive behaviours in a virtual world environment. The task resulted in a variety of solutions and a range of participant behaviours, and specific behaviours were correlated with innovative solutions. Multiple forms of analysis provided unique insights into participant behaviour, and the combined set of analyses led to a richer understanding of participant behaviour than found through any individual analysis. The paper also presents implications for how organizations may scaffold group interactions to increase the chances of successful collaborative innovation.

Keywords: Innovation, Collaboration, Virtual Worlds, Log File Analysis, Word-Space Models
Categories: K.3.1, L.3.0, L.3.0, L.6.1, L.6.2

1 Introduction

Productive innovation has been identified as a core skill that is crucial to economic prosperity, individual and national competitiveness, and economic development [Andrews, 04; Carlson, 06; Fagerberg, 04]. While innovation has been relatively well studied from an organizational and management perspective [e.g., Tidd, 94], relatively little is known about the sociocognitive processes of innovation [Shunn, 06], and even less is known about collaborative innovation. In this paper we introduce a framework for understanding behaviour associated with productive collaborative innovation. We discuss the development of a virtual world task designed to allow the close study of collaborative innovation and the study of analytic methods and findings.

1.1 Collaborative Innovation

Laboratory studies of innovation have been limited to studying a small number of relevant behaviours. This research includes studies of creative idea generation [e.g., Fink, 07], functional fixedness (in which subjects limit their use of an object to preconceived notions: a cardboard box that contains thumbtacks may be seen only as a device for holding, and not as a shelf that can be tacked to a wall) [e.g., Frank, 03; Wiley, 98], and creative design [e.g., Goel, 92]. These lines of research have shown the importance of creative imagery, cognitive flexibility, and the consideration of multiple perspectives in considering the potential space of problem solutions (what we call *problem frames*). Furthermore, diversity and similarity can act synergistically in group creativity [Miura, 04], and tasks that minimize conflict [Straus 94] and that have aligned incentives [Birnholtz 07] lead to greater productivity.

Post-hoc studies of innovation in real-world settings have considered a wide range of behaviours and contexts. While these post-hoc studies do not allow the same close examination of behaviours that is possible in laboratory studies, they have been particularly effective in showing that collaboration is more productive when participants come from diverse backgrounds with diverse areas of expertise [Page 07]. However, this diversity may also lead to an increased overhead in communication. So, understanding and harnessing the group collaborative behaviours associated with innovation takes on great importance. While the research to-date has made great strides in understanding collaborative innovation, we still do not have a thorough understanding of how specific team behaviours influence the process of innovation.

In this study we take a first step in moving beyond the limitations of existing studies, focusing closely on a constellation of behaviours that have initial evidence for being important in collaborative innovation: we call these “innovation supportive behaviours.” It is not our intent to imply that these behaviours are the only behaviours, or even the most important behaviours, for productive collaborative innovation. Instead, these behaviours were chosen due to initial evidence of their effectiveness in supporting innovation. We focus on these behaviours through the examination of a collaborative task in which multiple aspects of innovation supportive behaviours are necessary in the creation of an effective solution.

We are able to focus on a group of behaviours in a complex collaborative task because our task does not require that actions and behaviours conform to the limitations of the physical world. Instead, our task takes place in a highly instrumented virtual world, allowing us to capture and analyze participant behaviours while they engaged in complex collaborative tasks.

1.2 Virtual worlds and the study of innovation

We chose a virtual world for our task because the study of innovation supportive behaviours requires that we go beyond traditional laboratory settings. While earlier research tended to focus on one set of behaviours, such as idea generation, it was our goal to provide a task that required participants to engage in the types of rich thinking and deep communication that are required in most real-world innovation tasks. These include spatial and conceptual reasoning and communication of complex ideas, all deeply situated in a particular context. Studies of innovation may also be aided by eliciting and observing interactions that require significant real-world overhead, such as the creation of specialized equipment or the development/modification of large physical spaces. Virtual worlds have the potential to provide such environments [Dalgarno, 10].

The use of a virtual world also allows us to go beyond some of the limitations of post-hoc analysis of "real world" innovation. By careful instrumentation of the environment, virtual worlds allow for the capture of detailed real-time information in a non-obtrusive way. For instance, video and audio recordings in real world settings can capture participant actions and talk, but these methods are often obtrusive, and the cost of analyzing the resulting data can be prohibitive. In addition, virtual worlds allow researchers to alter the social dynamics of environments to examine participant reactions [Bailenson, 08]. Virtual worlds may also shed light on the behaviours of people when they are not in the virtual world [Fox, 09; Yee, in press], possibly allowing for extrapolation beyond the virtual world.

These characteristics of virtual worlds make them particularly well suited for studying aspects of human interaction that are otherwise difficult to capture.

In this paper we report on a study with the dual goals of expanding the types of environments and, therefore, tasks that can be used to study collaborative innovation and expanding the forms of data capture and analysis that can be used to better study and analyze innovation supportive behaviours among groups. We discuss three significant and related efforts:

1. Creation of a framework of "innovation supportive behaviours;"
2. Design and implementation of a virtual worlds task whose features are conducive to eliciting innovation supportive behaviours; and
3. Instrumentation of the virtual worlds task and the resulting analysis of participant behaviours.

This study engaged three groups, each with three participants, in a task implemented in the Second Life (SL) virtual world environment. We conducted multiple analyses using real-time observation and data from log files. We found that our task resulted in a variety of solutions and a range of participant behaviours. We found multiple forms of analysis provided unique insights into patterns of participant behaviour, and that the combined set of analyses led to a richer understanding of participant behaviour than found through any individual analysis. The paper also presents implications for organizations that wish to scaffold group interactions to increase the chances of successful collaborative innovation. Given the type and amount of log data collected for each group and our goal of exploring the methodological links between our framework of behaviours, task design, and analytic approaches, we limited the study to a smaller group of participants. Based on the results of this exploratory study, future research to expand the scope of the work, including sample size, is warranted.

2 Defining "innovation supportive behaviours"

To meet the challenge of addressing important aspects of innovation, we conducted a review of the innovation literature to determine dimensions of behaviours that have been consistently identified as high value. Once these dimensions of innovation were captured, we engaged in a process of evidence-centered design [Mislevy, 06] that guided the task-creation process.

The framework was not intended to cover all possible aspects of group innovation, nor to determine if participants engaged in a canonical problem solving process [e.g. Bransford, 93; Sternberg, 99]. Instead, our goal was to identify core aspects of collaborative innovation that would: 1) "naturally" occur in group problem-solving environments; 2) be bounded enough so that we would "know it when we saw it;" 3) be generative enough that the project will inform the field of research on group innovation; and 4) reasonably test the instrumentation of a virtual world to capture collaborative behaviours.

The behaviours that form our framework focus on group behaviours and interactions. By focusing on the group, our analysis avoids issues of "reading the minds" of individuals, focusing instead on those behaviours that are most relevant to group collaboration and innovation. For instance, our analysis need not attempt to differentiate between a participant who discovers new knowledge and a participant who is reporting on something they already knew; instead, our analysis is based on the observable act of a participant sharing information with the group. The specifics of this framework are:

`Problem framing and reframing` [Page, 07; Wertheimer, 82]. Problem framing refers to the frameworks participants use to understand their task in the service of generating and considering solutions. Examples of problem framing include using polar or Cartesian coordinates; or analyzing a network infrastructure according to the number of nodes, the cost, or the average transmission time of a message. Our analysis investigates the problem frames a group constructs, and how the group changes the problem framing as they come to better understand the task.

`Knowledge construction and sharing` [Lee-Kelley, 05; Pavitt, 05; Wickramasinghe, 06]. Knowledge construction and sharing refers to the information and knowledge that is referenced in group communications. This information and knowledge includes strategies, heuristics, attempted solutions to a problem, and declarative and factual knowledge. Our analysis investigates the knowledge that participants share, and the strategies, heuristics, and solutions the group constructs.

`Testing hypotheses and solutions` [O'Sullivan, 08]. Testing hypotheses and solutions refers to participants analysing the effectiveness of their hypotheses or solutions. Testing hypotheses includes all interactions in which participants run a test to determine the accuracy of their hypothesis or the validity of their solution.

`Group participation` [Page, 07; Miura, 04]. Group participation refers to the ways group members engage other group members in the group's task. For the group to engage in other innovation supportive behaviours productively, the group must find ways to participate that leverage the contributions of each member so that the group makes decisions and builds knowledge productively. We analyze the ways the groups engage in behaviours that support group cohesion, productivity, and equity. We note that, for the sake of generality, we allowed group participation to emerge naturally, without an explicit assignment of roles.

3 Task Design

Innovation framework dimensions	Requirements of the virtual world task
Problem Framing and Reframing	• Client is flexible in accepted solutions • Multiple solutions are possible • Non-obvious features lead to new metrics
Knowledge Construction and Sharing	• Participants receive unique information • Multiple strategies are possible and can be tested
Testing Hypotheses and Solutions	• Superficial understanding leads to inefficient solutions • Participants can test partial solutions • Participants can compare solutions
Group Participation	• Client requires group reporting • Solutions are possible without collaboration • Efficient solutions require collaboration

Table 1: Mapping features of the virtual world task to our innovation framework

To investigate the effectiveness of using a virtual world environment to elicit and analyze behaviours consistent with our framework of innovation supportive behaviours, we created a set of a-priori task requirements, and then engaged in a process of evidence-centered design to tightly align our task with our framework (the evidence-centered design process is discussed in detail in Section 3.1). The a-priori requirements included the following:

- An effective solution should require collaboration among participants, but less effective solutions could be created with minimal collaboration.
- The task should leverage the 3D nature of the virtual world environment.
- The task should not require specialized knowledge of any particular field.
- The task should be suitable for adults.

We began with the idea of a playground design task [Roussou, 05], in which adult participants would create a virtual playground from pre-existing resources. However, early pilot testing showed that teams of potential participants often based their actions on knowledge of existing playgrounds, and did not generate the range of behaviours and solutions we had hoped. We found a more productive task to be the creation of a virtual wireless network, which is described in detail in sections 3.2 and 3.3.

3.1 Evidence-Centered Design

To increase the likelihood that the final task would result in a variety of innovation supportive behaviours, we engaged in an iterative evidence-centered design (ECD) process [Mislevy, 06]. Prototypes of evidence collection methods were developed in Second Life to determine the feasibility of collecting the data for the different aspects of our framework. The forms of evidence available were used to create observable variables, which provide the evidentiary base for determining the degree to which group behaviour is consistent with the literature on innovative behaviours.

The ECD approach was then used to generate a set of task requirements. Starting with the a-priori requirements, we continued to iterate and refine the requirements to ensure that the task was consistent with our framework and the observable variables. Example requirements are shown in Table 1.

3.2 Task Description

The task sets participants on an island in the Second Life virtual world platform. Each participant interacts with the environment and the other participants through an assigned avatar. Built-in Second Life capabilities provide each participant (via her/his avatar) with a "first person" view of the island, the capability to walk or fly around the island, a set of objects in a personal inventory, and the ability to communicate with other participants. The task engages three participants as a team in creating a virtual inworld communications network using wireless network transceivers that are called `munchkins`. Participants are directed by a client avatar (controlled by a member of the research team) to place munchkins to efficiently connect three locations on the island, denoted by colored flags, to a satellite dish on the hypothetical island resort (Figure 1). Participants communicate primarily through the use of typed text chat. Voice communication between participants was not permitted.

Figure 1: The island

Munchkins' behaviours are similar to those of real-world network transceivers, with a few exceptions. Munchkin connections are literally line-of-sight, meaning that if there is any obstacle in their way, including something as seemingly insignificant as the leaf of a tree, they cannot communicate (an exception is water, as the munchkins work perfectly underwater). Munchkins are either connected or not in a binary manner, based on the range of the munchkins: there is no gradual fading of the signal as found in most transceivers.

Each participant is responsible for placing munchkins of a specific type, differentiated by color. Participants place munchkins by dragging a munchkin template from their inventory to the intended place on the island, and the number of available munchkins is essentially unlimited.

Any munchkin can connect to any other munchkin or any flag. Different munchkins have different ranges and costs, which were provided to participants in a text document in their avatars' inventories (green munchkins have the shortest range and cost the least, red have the longest range and cost the most). However, green munchkins have the highest cost per meter, making red munchkins the most efficient over long ranges, and green or blue the most cost efficient when there are obstacles.

An implication of different range munchkins is connection asymmetry: the connection between two munchkins is dictated by the weaker signal strength. This can be seen in Figure 2, which represents munchkin ranges as circles of a specified radius: for two munchkins to connect, both centre points must be within the range of the weaker of the two munchkins. From pilot testing we expected this asymmetry in connection range to be non-obvious to participants.

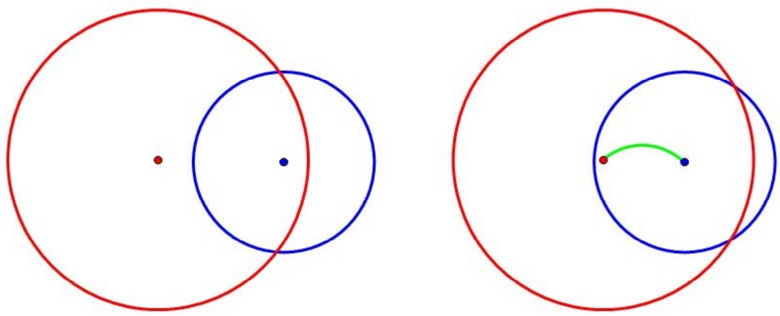

Figure 2: An example of unconnected Munchkins (left) and connected Munchkins (right). Note that the ranges are not visually explicit in the virtual world task.

While any munchkin can connect to any flag, the flags are placed such that the "obvious" solution of creating three single-colour munchkin networks to connect the flags (e.g. connect the green flag using only green munchkins) results in a particularly inefficient use of munchkins. The green and blue flags were placed such that a relatively straightforward placement of a sequence of munchkins could be used to connect to the satellite dish, with few obstacles. There were, however, significant obstacles between the red flag and the satellite dish. As a result, if participants implement the superficial solution of connecting each flag with a network consisting of solely the same colour munchkin, the primary benefit of each munchkin is not realized: the connections to the green and blue flags do not take advantage of the lower cost/meter of the red munchkins, and the connection to the red flag, which requires more munchkins due to obstacles, does not take advantage of the lower unit cost of the green and blue munchkins.

In addition to their main first-person view of the environment, participants have two map views of the island. The map that displays the munchkins and network connections is always oriented north-up and is located at the top left of Figure 3. Green arcs represent networks that connect to the satellite dish, and red arcs represent local networks that do not connect to the dish. Immediately below this map is a munchkin detector based on the colour of the assigned munchkin. When the user clicks the coloured square under the map, the display indicates the nearest in-range munchkin. The other map is located at the bottom right of Figure 3: this map is a feature of Second Life map and shows participant location and orientation on the island.

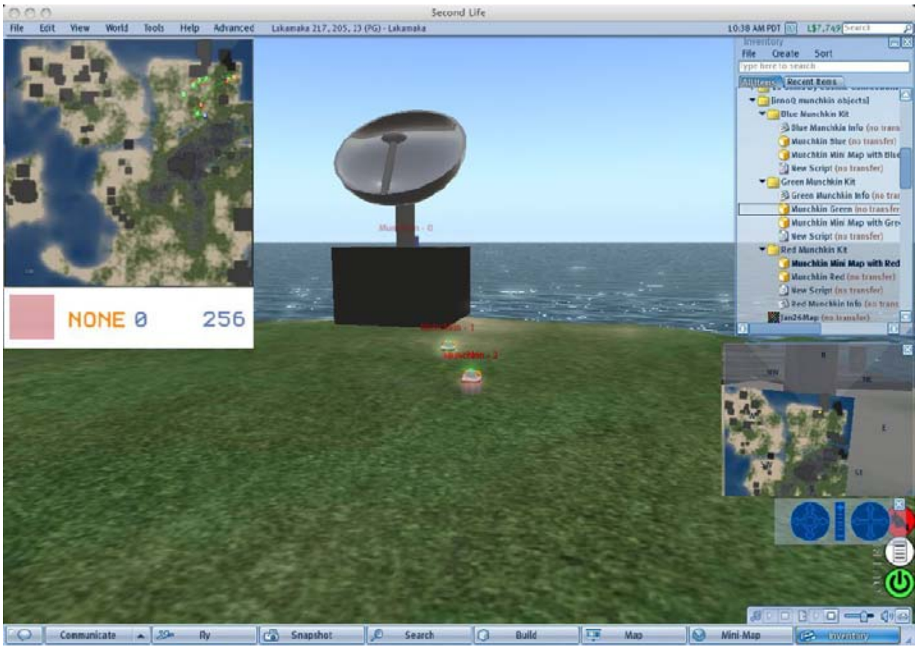

Figure 3: Participant view showing interface widgets, satellite dish, and munchkins

Key task features, mapped to task requirements, are shown in Table 2.

3.3 The Sessions: a short overview

After a brief introduction to Second Life, a group of three participants sign into the environment using predefined avatar names and passwords. The participants are brought to the Lakamaka Resort, where each is already outfitted with a particular munchkin kit (i.e. a red, green, or blue munchkin and its associated note are in the inventory, and the avatar is already wearing the appropriate map and munchkin detector). Munchkin assignment is as follows:

- Paula: Red munchkins
- Tam: Blue munchkins
- Jair: Green munchkins

Innovation dimensions	Requirements of the virtual world task	Task Features
Problem Framing/ Reframing	• Client is flexible in accepted solutions • Multiple solutions are possible • Non-obvious features lead to new metrics	• Client specifies only that all flags must be connected • Many configurations of munchkins connect all flags • Number of munchkins, total cost, and solution elegance are all reasonable metrics
Knowledge Construction/ Sharing	• Participants receive unique information • Multiple strategies are possible and can be tested	• Each participant has one colour munchkin and relevant technical information • Different strategies result in different and reasonable configurations of munchkins
Testing Hypotheses and Solutions	• Superficial understanding leads to inefficient solutions • Participants can compare solutions	• Straightforward networks result in costly solutions • Maps provide comparison of network solutions
Group Participation	• Client requires group reporting • Solutions are possible without collaboration • Efficient solutions require collaboration	• Shared text board • Each participant can create a network with one colour munchkin • Efficient solutions require using specialized capabilities of each munchkin

Table 2: Key features and task requirements

Upon signing in, the participants find themselves at a location on the resort where the client, named Dale, is waiting for them. Dale is an avatar controlled by a researcher, and provides the cover story for task administration. Once all participants are signed in, Dale introduces the task through a standard script that introduces participants to the session. He explains that he is in charge of setting up a wireless network at the Lakamaka resort, which is undergoing construction. He explains the role of the satellite dish, how the munchkins work and how to place them, and how to tell if a flag is connected. Dale explains the interface widgets, and gives the participants time to experiment (including placing munchkins, and then deleting unwanted munckins). Dale states that the participants should make a plan, write it on an inworld shared text board, and keep the text board updated with the current plan. He states that the munchkin kit in their inventory not only has the munchkin kit, but also a note that recaps the information in the introduction, and has more information on the munchkins (it is here that the technical information on the munchkin range and cost is found). Dale then takes questions from the participants.

Once the introduction is complete, the participants begin their attempt to solve the task. In general we saw that participants decide on a plan (an example plan is each participant attempts to connect "their" flags as defined by the assigned munchkin colours). Participants communicate via the use of text chat and actions as they complete the task. The participants place their munchkins, check connections, test hypotheses, and work with the goal of getting the flags connected. The sessions each took approximately two and a half hours.

The two initial sessions engaged participants new to SL and new to the task. Both teams completed the task, but the final solutions were very inefficient and did not take advantage of the unique capabilities of the different colour munchkins. The solutions used far more munchkins than necessary and were far more costly than necessary.

We then brought together three of the initial six participants to form a team of "veterans," where each participant had previously engaged in the activity. This was done to allow us to investigate group behaviour when participants already had deep knowledge of the task. In particular, we were interested in participant behaviour when (a) initial startup time learning the environment was not an issue and (b) the fact that there were differences between munchkins was hinted at more strongly, with the expectation that participants would discover the differences.

As before, the participants signed in using their assigned avatar (all participants used the avatar and munchkin kits they were previously assigned) and congregated near Dale. Dale again provided a set of instructions (modified to take advantage of their experience). This time Dale explicitly told the participants that prior solutions were inefficient and told them they should consult their technical guides for detailed munchkin information. Participants were also provided with maps of the prior solutions. In this final session the participants struggled with connecting munchkins with different ranges, until they figured out the asymmetry. They then undertook a more efficient solution, but still had difficulty due to line-of-sight and other technical issues. They did not have a final solution by the end of the session (that is, not all flags were connected), but the nearly-complete solution they did build was significantly more efficient than the prior solutions.

4 Data Sources and Analysis

4.1 Data Sources

The primary data source for this study was the collected log files. As participants engaged with the task, all actions (user actions and implicit munchkin actions, such as checking for connectivity) were transmitted to a remote server. This server computed the munchkin network state and logged all actions. The log files were then imported into a spreadsheet program and with a small amount of manipulation were used for human coding purposes. These files were also used as input to word-space analysis tools. Each log entry included a timestamp, name of the action source (either an object or participant), Second Life key (unique ID) of the action source, the URL query string passed to the servlet, and the result of the query as computed by the servlet. Logging the entire query directly allowed for replay of the server state at a later time (see Playback Tool below), while also allowing for a string that could be imported into a spreadsheet to create a well structured, human readable record of the action; for example, "COMMAND=PLACE & X=110.707800 & Y=58.225950 & TYPE=BLUE" is the command to place a new Blue munchkin at the coordinate 110.7, 58.2.

The research team conducted inworld observations and collected video files of the sessions using screen capture software, in case these were needed to shed light on any ambiguity in the log files. The team found the video capture data to be largely ineffective for our needs. The video necessarily focused on only one part of the island, and so was not useful when the participants were distributed across the island. In addition, video analysis is extremely time consuming. We found the log files, and the corresponding playback tool described below, to be adequate for our uses.

4.2 Data Analysis

We report on four primary results, each using a unique method of analyses.

Real-time inworld observations were used to gain ground-truth about participant behaviours, such as to determine the ways the participants were using the system, the types of interactions they were engaged in, and to determine if they appeared to be taking advantage of Second Life functions that we were not logging. The result of an analysis that uses the inworld observations as well as the log files described in the next section will be reported in Section 5.1.

The log files were the basis of a qualitative analysis, which compared the behaviours of different groups. Comparisons included temporal comparisons, in which we compared the order in which groups engaged in particular activities, quality comparisons, in which we compared the relative quality of the interactions, and quantitative comparisons, in which we compared the relative number of times groups engaged in particular activities. To conduct these analyses we created a playback tool, which precisely recreates the participant activity. A screenshot of the playback tool is found in Figure 4.

The left side of the tool shows a map of the island, upon which is rendered the evolving network graph resulting from participants' Munchkin placement. This map is rendered using the Java Universal Network/Graph Framework (JUNG) library, the same tool that renders the participants' map display. The top right of the tool shows the log entry parser and playback controls. The user can step through log entries and see the source of each action (a simulation object or a participant), the time the action occurred, the query/command sent to the server, and the result. Actions include placing or moving a munchkin, sending a text chat, writing on the text board, etc. When actions resulted in a change to the network graph, the map was updated appropriately. The bottom right shows the cumulative chat log.

The user may "play" the log, rather that simply step through it, in which case the actions are performed automatically in sequence, at a speed anywhere from real time to 300 times actual speed (assuming the computer can process the requests at that speed). We also associated each kind of action with a unique sound that is played when it occurs. The addition of sound to the playback enhanced the ability to notice patterns in behaviour. For instance, listening to the sounds during feedback made it eveident that activity often occurs in spurts of chat (planning), laying down munchkins (executing the plan), and then moving munchkins (refining). The result of this analysis is discussed in Section 5.2.

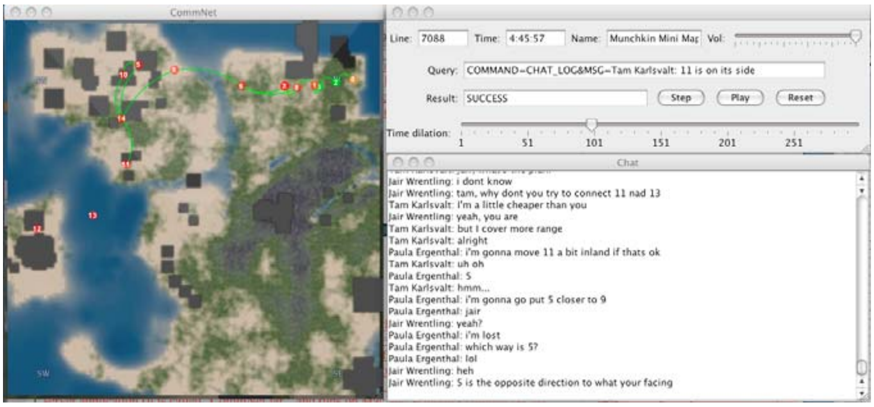

Figure 4: The playback tool, showing the updated island map and the participant chat

Word-space methods (see, e.g. Widdows 08) were used to analyze the amount of similarity in word use between participants, and investigate changes over time. For each participant, over a specified timeframe, we compute a vector in arbitrarily large n-dimensional space (each word is assigned an arbitrary vector in the n-dimensional space, and all word-vectors for a participant are concatenated). This process results in a vector for each participant in each timeframe. When participants use similar language, their vectors point in a more similar direction as compared to when participants use dissimilar language. This analysis can be used to compare two participants and to compare a participant to herself in two different timeframes (e.g., to ask the question: did the participant "change her tune": if so, we expect the angle between the two vectors to be relative large). The result of this analysis is discussed in Section 5.3.

Hidden Markov chain modelling was used to compare participant behaviour over time. Input to the models was based on human coding of the log files. Using the data logs generated by the system and the observable variables developed through our evidence-centered design process, the team coded an entire participant session. The result of this analysis is discussed in Section 5.4.

5 Results

5.1 Result 1: Teams using different strategies could experience success

Figure 5 shows the solution for each of the three sessions. In the first session the group created a network that connected the flags with munchkins of the same colour.

The solution created in the second session was similar, in that a line of blue munchkins connect the satellite dish to the blue flags, and green munchkins connect the satellite dish to the green flags, but green munchkins were also used to connect the red flags (in large part because the participant responsible for the red munchkins in this group had technical difficulties, and the participant responsible for the green munchkins continued to place munchkins to connect the red flag). We note key differences in these initial solutions. The first solution shows the relatively long range of the red munchkins (however, this was apparently not noticed by the participants). The second session's solution used far more munchkins but shows that flags can be connected with another colour munchkin.

The third solution shows the more sophisticated solution of strategically using different colored munchkins for different purposes. The red munchkins were used to cover long distances, and the blue and green munchkins were used when the terrain or obstacles made the use of the red munchkins difficult or inefficient. This group ran out of time before finishing, but they were on a path to an efficient solution, and we believe that they would have completed the task if given enough time.

This result is significant in that it shows that the task could elicit a wide range of solutions, including expected inefficient solutions as well as a fairly efficient solution.

Figure 5: Three final solutions (in chronological order of session)

5.2 Result 2: Qualitative comparisons show differences in innovation supportive behaviours between groups

While all sessions showed some similarities, there were marked differences between the first two sessions and the last session. Groups were most similar in group participation. They also engaged in a significant amount of testing of hypotheses and solutions, as well as knowledge construction and sharing. However, because only the participants in session 3 discovered that the munchkins had different ranges (and costs) the types of hypotheses tested and knowledge shared was qualitatively different in session 3 as compared to the other two sessions. The influence on problem framing and reframing was even greater, as only the participants in session 3 engaged in genuine problem reframing.

Group participation. All sessions showed positive group participation. Team members provided support for collecting different perspectives and ensuring that everyone participated. An example of this is shown in Excerpt 1, which was early in the first session: in this short exchange all participants had a chance to provide input into the group actions, and two statements were quite explicit in stating enthusiasm for the process (“Awesome!” and “Cool”).

```
Paula:  should we start building a path?
Tam:    yes. but if we have to have our own color trail
        back to the dish, should we just divide and
        conquer?
Jair:   ok team. yes. i was going to type that!
        Awesome!
Paula:  How about we try it?
Tam:    Cool
```

Excerpt 1. Session 1 shows an example of good group participation and initial knowledge sharing of a strategy (divide and conquer)

Testing hypotheses and solutions. All sessions also showed significant testing of hypotheses and solutions. Excerpts 2 and 3 show an example of hypothesis testing from sessions 2 and 3 respectively. Note, however, that the excerpt from session 2 includes the unchallenged assumption that all munchkins have the same, short range. This is in contrast to excerpt 3, in which Paula notices that the different solutions show different ranges. The team then decides to explicitly test the range of the munchkins. When that shows a surprising result they decide that the terrain must be getting in the way of the munchkins’ line of sight.

```
Paula:  I need some advice -- what are you doing to
        make that link? are you putting the munchkin at
        the flag? mine are always red
Jair:   I'm just dropping munchkins near the satellite,
        and they need to be close. like 4 or 5 feet
Paula:  OK, I'll head back to the dish
Tam:    so you have to start near the dish?
```

Excerpt 2. Session 2 shows Paula requesting knowledge and Jair sharing, based on Jair’s testing hypotheses about creating a connected network

```
Paula:  hey that's kinda weird, that in map 1 [map of
        solution from Session 1] the red muchkins are
        way farther than that, but the MK light up.
Jair:   that is weird
Paula:  can one of you put a munchkin right next to
        mine?
Tam:    Yes Jair, I was thinking the same thing
Paula:  whoever has the 30 m range
[munchkins are placed and they do not connect to
        Paula's]
Paula:  maybe it's the terrain. k then let's use cheap
        munchkins here
```

Excerpt 3. Session 3, Paula shares knowledge based on reading the map of an earlier solution, and the group testing hypotheses about munchkin ranges and line-of-sight

`Knowledge Construction and Sharing.` Excerpt 4 shows knowledge construction and sharing that occurred in session 1 immediately following the initial task introduction. Jair finds the technical information about her munchkin, and shares the cost with the team. It is of interest to note that the notecard specifically states the munchkin colour, cost, and range in the same sentence (e.g. "Green munchkins have a range of 12 meters and cost $30 each."). Jair does not relay all this information, and other team members do not check their own cards. This is in contrast to Excerpt 5 from session 3, in which, also immediately after Dale introduces the session, Paula checks the cost of her munchkin, and explicitly asks other to check the cost of their own munchkins. This comparison has an immediate effect on the group ("12 meters? Dannnggg") and sets the stage for sophisticated solution testing and a more efficient problem framing that uses cost per meter.

```
[Immediately after Dale's introduction, during which
        he introduces the information found in the
        kits.]
Dale:   OK, i'm going to leave you now so you can get
        to work, unless you have other questions
Jair:   One moment pls while i read the notecard. Do we
        have to purchase the Munchkins (mk)? It says
        that mk's cost $30. I have 0$
Dale:   you have as many as you need. but hopefully you
        can minimize costs. But no, it doesn't come out
        of your money here, we will add up later to see
        how much I will be charged to build the network
[Tam goes on to ask Dale about the size of the island]
```

Excerpt 4. Session 1, Jair shares knowledge of munchkin costs

```
Paula: do all the munchkins cost the same amount?
Tam:   Yes, I think so
Dale:  you all have info about the munchkins in your
       kits. you should find out there.
Paula: have either of you checked the cost of your
       munchkins? mine are $100 and have a range of 50
       m (I have red).
Tam:   at the very end of "Blue Munchkin Info.," it
       it says blue munchkins cost $60 (each)
Jair:  mine have a range of 12 meters and cost 30
       dollars each
Paula: 12 meters? Dannnggg. tam what is the range of
       yours?
Tam:   (and have a range of 25 meters)
```

Excerpt 5. Session 3, Team members share knowledge of munchkin costs and ranges

`Problem framing/reframing.` Excerpt 6 shows how participants in session 3 came to frame the problem as one of cost per meter. As a result, the team not only had a more efficient metric for judging solutions, but also was able to conduct tests that were more relevant to finding an efficient solution. As we see in Excerpt 6, they are having trouble connecting the red munchkins. However, with the knowledge that the red munchkins `can` connect over far ranges, and that it is efficient to do so, the participants decide to investigate areas of the island that may have fewer obstacles to line-of-sight ("maybe we should stay on the flat"), and are thus on their way to a highly efficient solution. This type of sophisticated reasoning was not available to the other groups, in large part because they did not realize the different costs and ranges of the munchkins, and so could not create an efficient metric.

```
Jair:  because green is the cheapest per meter, we
       should always use green MK unless we need to
       cross a gap
Paula: but isn't red the cheapest per meter? Red is
       $2/m. except it lies cause the MK actually has
       a crappy range.
Jair:  oh you're right. Duh. Hmmmmmmm. so our first
       strategy is the best, we just need to get the
       red MK to work right.
Paula: maybe we should stay on the flat. like look @
       the map. if we go a bit north to the edge of
       the island, the MK might have better range.
Jair:  that sounds good.
```

Excerpt 6. Session 4, Jair uses the problem frame of "cheapest per meter," then Jair and Paula hypothesize and share solutions

5.3 Result 3: Word-space analysis shows changes over time

Word-space analysis can provide a “gestalt” of how a particular group is behaving in terms of our framework. If a group begins the task using different problem framing and/or different information, one would expect their initial vectors to point in quite different directions. Then, if members engage in productive group participation that results in sharing knowledge and coming to a shared problem framing, the vectors would converge. Using data from session 3 (the only session that constructed a sophisticated problem framing) we see exactly this pattern. This result is at least consistent with our hypothesis that some key features of innovation-supportive behaviors would be detectable even with relatively automated analysis.

We analyzed session 3 by splitting the logs into four segments, each of approximately 20 minutes. After repairing participants’ orthographic slips, removing stop-words, and stemming the log data, we compared each participant to themselves over each timeframe, as well as to other participants. Messages were, by and large, brief, fragmentary, and infrequent, which limited the applicability of word-space analysis to anything but large time spans of the discourse. As a result we consolidated the timeframes into two segments roughly corresponding to the first half and second half of each session. This analysis shows that the vocabulary used in the first half of the activity is not particularly similar to the vocabulary used in the second half (timeframe 1-4 to timeframe 5-8). Individual participants changed from the first half of the session to the second half, and perhaps most importantly, two participants converged in the second half and one did not. Figure 6 shows this graphically, with the length of each line segment representing the angle (arccosine of the normalized dot product) between the corresponding vectors. The vertex P1to4, for example, represents the log of Paula’s chat across timeframes 1-4; the vertex J1to4 corresponds to the log of Jair’s chat across the same timeframe; the length of the line connecting these vertices corresponds to the angle between the chat vectors in word-space. In this case we found that the resulting distances could be reasonably represented in planar fashion (there is no reason to believe that a planar representation would be possible in general.) The obvious convergence of Paula’s and Jair’s chats and lack thereof with respect to Tam corresponds to a phenomena observed during the real-time observation, and one that was verified through an close reading of the log files. It was at about the halfway point that the team determined that an efficient solution would use red munchkins over long distances, and green munchkins for short ranges. This marginalized Tam, as there was not much use for blue munchkins in this strategy.

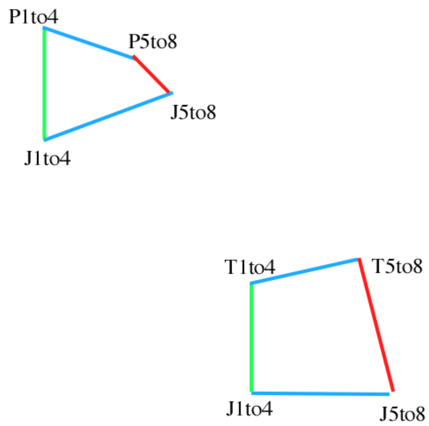

Figure 6: Vectors comparing participants in the first and second half of the session. Paula's and Jair's speech converged, as indicated by the relatively short line between P5to8 and J5to8; Tam did not converge with the others, as indicated by the relatively long lines from T5to8 and J5to8 and P5to8

Perhaps equally significantly, in the second half Paula and Jair drifted into creating a shared shorthand language (note the "MK" in the transcript fragment, above), a drift that shows up quite clearly in the word-space analysis.

5.4 Result 4: Hidden Markov chain models show differences over time

Hidden Markov chain modelling indicates that the first half of session 3 was spent primarily in planning through text chat, which corresponds to the observation that much of the first half of the session was spent trying to figure out how to get munchkins to connect over long distances. The second half of the activity was spent primarily in placing and moving munchkins, which corresponds to the observation that much of the second half was spent in actually implementing a solution to the problem. This analysis complements the word-space analysis, both by including participant actions in the analysis in addition to message content and by positing causal linkages among the types of activities participants engaged in over the different timeframes.

6 Implications and Significance of the study

This study advances our understanding of two important topics in collaborative innovation: the analysis of innovation supportive behaviours, and the use and instrumentation of virtual worlds for analyzing and understanding complex collaborative phenomena. Through our task design and analysis we have shown that virtual worlds can be used to study the important phenomena of innovation, and we have shown that virtual worlds can be instrumented to capture key participant behaviours.

With respect to analyzing innovation supportive behaviours, we have shown that some behaviours, while perhaps necessary, are not themselves sufficient to support collaborative inquiry. In particular, we note that all sessions showed significant group participation, hypothesis testing, and knowledge construction. However, despite these positive behaviours, two of the three groups did not discover the key piece of information that was key to an effective solution. Even more striking is that this information was discovered and shared by a member of one of the "unsuccessful" groups, but this information was not taken up by the group and didn't influence the framing of the problem. Furthermore, the primary representation provided to the participants, the map of connections, provides visual feedback showing different munchkin ranges, and yet the participants still did not discover the differences.

This finding may have implications for the structuring of collaborative teams. For instance, it may be productive to set organizational norms for key behaviours, such as checking information from multiple sources before accepting it as true. While this may seem obvious, social pressures may prevent individuals from checking on information presented by a participant, as such a check could be interpreted as a lack of trust. If checking information was an accepted practice for a collaborative team, it may allow for more effective teams.

With respect using instrumented virtual worlds for analyzing and understanding complex collaborative phenomena, we have shown that the properties in our framework can be elicited and analyzed in a virtual world environment. This initial finding suggests that future research may find there to be efficiencies in the use of virtual worlds for researching complex behaviours, such as those required for innovative collaboration.

A closer look at the analysis reveals an implication of our choice to limit the medium of direct communication to typed messages. As discussed above, the infrequency and terseness of typed messages limited the time resolution of word-space analysis, and we noticed saw that two participants in session 3 came to use a shared vocabulary in discussing their solutions.

In further analysis we found that these two participants "discovered" that they shared a powerful representation which they could use to effectively communicate about their solution: the real-time island map. By referencing the map they could express intended actions, difficulties, hypotheses, etc. At about the same time, one of them also "discovered" that they could move munchkins remotely (a non-obvious feature of SL in which one could view far-away part of the island, and manipulate objects that are in view). This action-at-a-distance was consonant with the emerging norm of using the map as the primary reference object in communication. Coincident with discoveries was an easily observable phase-transition-like change in both their discourse (more focused, common language as detected in the word-space analysis) and their actions (more frequent, more efficacious, as shown in the Hidden Markov Chain Analysis).

Would their discourse have continued to evolve over time, now that they had something they shared – and knew they shared – to talk about? Are there scaffolds that could be embedded to induce this kind of transition earlier, and more inclusively? The answer to these questions, especially the latter, is an issue for future research, and may have significant implications for structuring collaboration for innovation.

Acknowledgements

The authors would like to thank Valerie Crawford for her instrumental role in this project, as well as Yukie Toyama, Julie Remold and Geneva Haertel for their contributions to the project. We would also like to thank the participants who were willing to engage with the munchkin task. This manuscript is based upon work supported in part by the National Science Foundation under Grant No. 0745694. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[Andrews, 04] Andrews, P. (2004). Investing in biotech. Presented at Securities Institute and the Australian Stock Exchange. http://www.sd.qld.gov.au/innovation/sti/downloads/biotech.doc

[Birnholtz, 07] Birnholtz, J.P., Grossman, T., Mak, C., Balakrishnan, R. (2007). An Exploratory Study of Input Configuration and Group Process in a Negotiation Task Using a Large Display, Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 07), April 28 - May 3, San Jose, CA, pp. 91-100 .

[Bransford, 93] Bransford, J.D. & Stein, B.S. (1993). *The IDEAL problem solver (2nd ed.)*, New York: Freeman.

[Cagan, 07] Cagan, J. (2007). The cognition of engineering design—An opportunity of impact. Cognitive Science, 31, 193-195.

[Carlson, 06] Carlson, C., & Wilmon, W. (2006). Innovation: The Five Disciplines for Creating What Customers Want. Crown Business, New York, NY.

[Clark, 09] Clark, D., Nelson, B., Sengupta, P., & D'Angelo, C. (2009). Rethinking science learning through digital games and simulations: Genres, examples, and evidence. In National Research Council's (Eds.), Learning science: Computer games, simulations, and education. Available from http://www7.nationalacademies.org/bose/Clark_Gaming_CommissionedPaper.pdf

[Dalgarno, 10] Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? British Journal of Educational Technology, 41(1), 10-32.

[Fagerberg 04] Fagerberg, J. 2004. Innovation: A guide to the literature. In J Fagerberg, D., Mowery, & R. Nelson, (Eds.). 2004. The Oxford handbook of innovation. Oxford, UK: Oxford University Press.

[Fink, 07] Fink, A., Benedek, M., Grabner, R., Staudt, B., Neubauer, A. (2007) Creativity meets neuroscience: Experimental tasks for the neuroscientific study of creative thinking. Methods, Volume 42, Issue 1, Neurocognitive Mechanisms of Creativity: A Toolkit, Pages 68-76, ISSN 1046-2023, DOI: 10.1016/j.ymeth.2006.12.001.

[Frank, 03] Frank, M., and Michael Ramscar. "How do Presentation and Context Influence Representation for Functional Fixedness Tasks?" Proceedings of the 25th Annual Meeting of the Cognitive Science Society, 2003.

[Goel, 92] Goel, V. and Pirolli, P. (1992). The Structure of Design Problem Spaces. Cognitive Science 16: 395-429.

[Lee-Kelley, 05] Lee-Kelley, L., & Blackman, D. (2005). In addition to shared goals: The impact of mental models on team innovation and learning. International Journal of Innovation and Learning, 2(1) 11-25.

[Mislevy, 06] Mislevy, R. J. and Haertel, G. D. (2006), Implications of Evidence-Centered Design for Educational Testing. Educational Measurement: Issues and Practice, (25), 6–20.

[Miura, 04] Miura, A. and Hida, M. (2004) Synergy between diversity and similarity in group-idea generation. Small Group Research, 35(5), 540-564.

[O'Sullivan, 08] O'Sullivan, D. and Dooley, L. (2008). Applying Innovation, Sage Publishing, Thousand Oaks, California.

[Page, 07] Page, S. (2007). The difference: How the power of diversity creates better groups, firms, schools, and societies. Princeton University Press, NJ.

[Pavitt, 05] Pavitt, K. (2005). Innovation processes. In J. Fagerberg, D. Mowery, & R. Nelson, (Eds.). 2004. The Oxford handbook of innovation (pp. 86-114). Oxford, UK: Oxford University Press.

[Roussou, 05] Roussou, M., & Slater, M. (2005). A Virtual Playground for the Study of the Role of Interactivity in Virtual Learning Environments. In Proc. of PRESENCE 2005: The 8th Annual International Workshop on Presence, London, UK,

[Schunn, 06] Schunn, C., Paulus, P., Cagan, J., & Wood, K. (2006). Final report from the NSF Innovation and Discovery Workshop: The Scientific Basis of Individual and Team Innovation and Discovery. Washington, DC: National Science Foundation. Available at http://www.lrdc.pitt.edu/schunn/innov2006/talks/I+Dreport-final.pdf

[Sternberg, 99] Sternberg, R. (1999). *Cognitive Psychology. (2nd edition)*. New York: Harcourt Brace College Publishers.

[Straus, 94] Straus, S.G. & McGrath, J.E. (1994). Does the medium matter? The interaction of task type and technology on group performance and member reactions. Journal of Applied Psychology, 79, 87-97.

[Tidd, 94] Tidd, J., Bessant, J., Pavitt, K., 1994. Managing innovation: integrating technology, Market and Organisational Change.

[Wertheimer. 82] Wertheimer, Max. [1945] 1982. Productive Thinking. Chicago, IL: University of Chicago Press.

[Wickramasinghe, 06] Wickramasinghe, N. (2006). Knowledge creation: A meta-framework. International Journal of Innovation and Learning, 3(5), 558-573.

[Widdows, 08] Widdows, D. & Ferraro, K. (2008). Semantic vectors: A scalable open source package and online technology management application. *Sixth international conference on Language Resources and Evaluation (LREC 2008).* Marrakech, Morocco.

[Wiley, 98] Wiley, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. Memory & Cognition, 26(4), 716-730.