15th International Command and Control Research and Technology Symposium
June 2010, Santa Monica, CA

"The Evolution of C2: Where Have We Been? Where Are We Going?"

Topic 8: C2 Assessment Metrics and Tools

Interruption Interview: An approach to elicit situation assessment for an ill-defined task

Angela L.S. Tan, Shuli Yu, Boon Kee Soh

POC: Angela L.S. Tan
DSO National Laboratories
27 Medical Drive #11-01 Singapore 117510
+65-6485-7155 tlisin@dso.org.sg

Abstract

With the rise of asymmetric warfare, it has become more difficult to model and predict what adversaries do. Hence future commanders need to be able to adapt to changes and pre-empt eminent dangers. In this paper, we describe the Interruption Interview method which we developed for our research to elicit cognitive processes used in assessing an intelligent enemy. The technique is largely influenced by Klein's approach to naturalistic decision making and Endsley's theory of situation awareness. The cognitive dimensions developed by Klein were used to guide the interview and this was found to be generally sufficient for an appropriate appreciation of the task. The study was conducted in a warfighting microworld and participants were interrupted at semi-regular intervals. They were asked about what they would do next and these typically were counter-actions to their predictions about the enemy. This paper presents our methodology and findings associated with the methodology.

1. Introduction

With the rise in asymmetric warfare, soldiers are no longer constrained to the battlefield. They participate in Operations Other Than War, such as peacekeeping and relief operations, which can challenge their understanding of the situation (Holmquist & Goldberg, 2007). The difficulty ranges from differentiating combatants and civilians to interpreting adversarial situations with evolving goals, multiple intentions, and deception. Having an opponent that shares similar abilities implies a need for stronger adaptation techniques. They can be described as operating in the unknown-unknown cynefin space since they face problems which are ill-defined and they may not have thought about previously (Kurtz & Snowden, 2003).

The problem is exacerbated by the speed of technological evolution. We are not conditioned to modern day weapons and tools used by our human adversary – it is much easier for us to fear natural evolutionarily threats (e.g. leopards and spiders) than industrial revolutionarily threats (e.g. cars, knifes, and guns) (Platek, Keenan, & Shackelford, 2007). The challenge is greater when the decisions are made in a remote C2 environment. How do the commanders make sense of the situation and identify threats from the abstract information in the display?

Although there are existing efforts to build computer interceding models and tools to help military personnel predict adversarial behavior (Gilmour, Krause, Lehman, Santos, & Zhao, 2005; Kott & Ownby, 2005), the experienced human commander remains the most efficient engine to understand and counter the adversary (Lebiere & Best, 2009). However, there is limited empirical research that investigates how human assesses the enemy, especially in complex and ill-defined situations. Lebiere and Best (2009) highlighted that it is impractical to run such studies because they are resource intensive: the manpower required to observe, elicit and analyze how people think about the enemy is significant, and it is difficult to control the experimental protocol in adversarial scenarios which are fraught with complex and interdependent decisions.

The only traces of cognitive psychology research in our literature search used the Critical Decision Method and Think-Aloud Problem Solving method to elicit the battlefield mental model and the strategies in handling an adversary (Cohen, Adelman, Tolcott, Bresnick, & Marvin, 1993; Cohen, et al., 1995). These studies reported the factors commanders consider when making predictions in the battlefield and three general approaches commanders use to handle an adversarial situation: (1) reacting to the enemy; (2) predicting future enemy action and reacting to these predictions; and (3) proactively shaping the enemy.

The research by Cohen and associates is a welcome initial attempt to examine the interplay between two adversarial parties. We wish to extend their work by focusing on the cognitive processes of how military commanders assess the enemy, which includes but is not limited to:

1. Types of predictions: What do people try to predict about the enemy?
2. Basis of their judgment: Do they comprehend and predict the enemy based on a profile of the enemy or do they assume that the enemy would behave like themselves? If there is a profile of the enemy, what do they profile?
3. Situational considerations: What are the cues they look out for? How do they interpret cues which they have not thought of previously?
4. Order of recursion: Do they think about the enemy thinking about them thinking about the enemy? How many order of recursion do they use?

To achieve the project objectives, we utilized the arduous lab-based approach, as briefly described by Lebiere et al. (2009), and came up with the Interruption Interview. Despite the large amount of resources committed, we were still limited in our control of the study. In this paper, we outline our data collection approach, the challenges we faced, and how we could have done it better. We also explain why we felt the technique can be extended to studies that are targeted at eliciting the situation assessment in an ill-defined problem space. To maintain the focus on the methodology, only the results that (1) address the validity of the study and (2) demonstrate the effectiveness of the methodology are presented.

2. Development/Design of Interruption Interview

The Interruption Interview is designed to study situation assessment, which is also commonly termed as sensemaking (Klein, Moon, & Hoffman, 2006). The Interruption Interview methodology is a hybrid of Endsley's work in SA and Klein's work in sensemaking. Endsley (1988) defines Situation Awareness as the "perception [noticing] of the elements in the environment within a volume of time and space, and the (2) comprehension of their meaning, and the (3) projection of their status in the near future" (p. 97). She later narrowed the definition to refer to the knowledge state relevant to the three levels and described the process as situation assessment (Endsley, 1995b).

While the two terms are often used interchangeably, it is important to draw the distinction as it affects how the research would be conducted and the conclusions that can be drawn (Durso & Sethumadhavan, 2008). To clearly delineate the two concepts, we use the term "sensemaking" to refer to the process and "SA" to refer to the knowledge state. We started with an initial interest in the sensemaking of the adversarial environment and found the SA concept to be one of the better frameworks for guiding our probing and data analysis.

Here, we present the spectrum of experiment controls we considered and how we decide on the implementation of the experimental controls.

Nature of study: exploratory vs. confirmatory

While there were studies on how experienced commanders handled adversarial situations, ours was targeted at a cognitive level that is lower on the micro-macro spectrum (Cohen, et al., 1993; Cohen, et al., 1995). We found numerous psychological constructs, such as mindreading and inference processes, that could possibly affect one's SA about the enemy. However, we were unsure if we had a comprehensive list of all relevant processes and their relationship with one another.

Hence, this study was designed to be an exploratory study to identify the processes exhibited rather than a confirmatory study with specified hypotheses. An example of a confirmatory study would be one that tests the effect of a cognitive process (e.g. comparing options) on one's awareness of the adversary. However, the plan was not to impose upon our participants specific processes for assessing the enemy. Instead, for our exploratory study, we were looking for instances where mindreading took place, what it meant, how it happened, and how good it was.

The Critical Decision Method (CDM) surfaced as one of the better methods for studies in sensemaking (Klein, Calderwood, & MacGregor, 1989). CDM is a story-telling technique where subject-matter-experts are asked for an incident involving a challenging decision. The interviewer then probes deeper along pre-defined cognitive dimensions (e.g. goals, cues and expectations). It supports the elicitation of tasks in a naturalistic environment which is dynamic, complex, and ill-defined.

However, there may be bias in the details the participant's give since the interview is conducted post-hoc. The participants may have forgotten some details of the incident account. Other than the lack of ground truth, we were more cautious about post-hoc rationalization. As our study entailed the analysis of processes at a fine micro-level, we wanted to differentiate between reasoning about the situation and rationalizing what happened. With CDM, it was difficult to draw the distinction.

We decided that we would adopt the cognitive dimensions for specifying our questions. However, we would not ask for critical incidents. Instead, we would use a simulated environment and probed for the cognitive processes in the decisions made in the simulation. With the simulated environment, we could capture the decisions before they were made and compare them with the actual events that happen subsequently. This would improve the rigor of our study as it enabled us to better differentiate reasoning from rationalization. Our data would more accurately reflect one's sensemaking of the enemy.

Semi-controlled simulation

The use of simulated environment for sensemaking studies is not new. Dorner (Klein, et al., 1989) had used simulations that possess complex relationships and achieved notable discoveries

in how people manage resources over time. In one of his studies, participants implemented measures to promote the well-being of the inhabitants. He collected the thought patterns of the participants and one of his findings was that the good participants developed more measures for each goal they were trying to achieve. Similarly, our study was designed to elicit the cognitive processes and patterns that would be effective for handling adversaries.

One challenge in the use of simulation for our study was that the opponent had to be adaptive. A fully controlled simulation with pre-determined cues and situations would not possess these qualities as the opponent would not learn and adapt to the evolving situation. The scenario, thus, would not call for the rich cognitive thoughts used for assessing an enemy. To achieve the required level of fidelity, the opponent had to be relatively intelligent.

Although Artificial Intelligence (AI) may be used to simulate the opponent, it was not within the scope of the project to develop an intelligent AI adversary. As we did not have the resources to develop an adaptive AI that learned from the behavior of the participant, we had an experimenter play the role of an intelligent opponent. In order to maintain some measure of consistency in the game scenario across participants, the experimenter executed a semi-controlled script that consisted of a series of opening moves, high-level objectives, and strategies. The experimenter had the flexibility to tweak his low-level actions or tactics to adapt to the changing enemy picture.

Form of probing

As we progressed with the design of the simulation study, we also struggled with how the interview should be implemented. There were three known approaches. The first was the real-time probing which involved interviewing participants as they played the game. This was not ideal as the pace of the simulation was very fast and there were too many probes to implement. Although it was possible to slow the pace of the game, asking questions as participants compete against the adversary was extremely intrusive and would disrupt their concentration.

The second was retrospective interview, where participants would be shown a video recording of the simulation after each game. The experimenter would pause the video periodically during the playback and interview the participants at each event. This was also not ideal as it faces the same issue of rationalization as the CDM. While watching the replay of the game, participants can pick up events that they miss during the game and unknowingly think that they notice those events.

The last method was the Interruption Interview. We would freeze the simulation at semi-regular intervals and interview participants. We would not show them any playback and instead, depend on the participants' memory of what happened. The interval between the freezes would be carefully controlled such that they would be able to remember most of the events.

3. Interruption Interview procedures

The Interruption Interview can be simply define as an interview that interrupts a participant's task in a simulated environment for the purpose of understanding how one comprehends percepts in the environment and projects these percepts into the future. Its procedures, however, should extend to include the pre-task planning and post-task comprehension. The pre-task planning is where most of the projections are collected while the post-task comprehension is lighter and concludes the study.

The following sub-sections details the methodology from the preparation needed before conducting the Interruption Interview (Stage 0) to the Post-Task Comprehension (Stage 4). We also include a short recommendation for studies where participants undergo multiple Interruption Interviews.

Step 0: Preparation

1. A suitable simulator for the problem space of interest is acquired and/or configured for the study.
2. The roles played by confederate experimenters are defined. Minimally, an interviewer and a note-taker are required. As an example, for our context of studying how one reacts to an adaptive enemy, an experimenter is recruited to play the role of the opponent (hereby known as the "gamer"). If participants are working professionals and it is difficult to reschedule them, a larger team size (with reserved experimenters) is preferred to ensure that the team of experimenters is always ready for the study.
3. All experimenters are briefed about the protocol. This is especially important when the team size is large and there are multiple experimenter roles. As an example, one of the rules for the gamer is that he should not meet the participant until the study is completed. More challenging rule includes pseudo-doctrinal rules such as the relative combat power needed before executing an attack.
4. An event logging software (e.g. MORAE) is acquired. As a large volume of data (events and their associated mental processes) is to be expected, effort should be taken to keep the data organized as they are being collected. An efficient event logging tool can help ease the data collection, processing, and analysis.
5. Participants are recruited. The rules for our selection and recruitment are typical of most studies. We recommend "Working minds: A practitioner's guide to cognitive task analysis" as a reference for recruiting participants for sensemaking studies (1996).
6. Like all other lab-based studies, pilot studies should be conducted to test the material and protocol. This is very important for Interruption Interview since the experimenter team size is larger than most other experimental studies. All the experimenters need to be familiar and coordinated to ensure that the protocol implemented is consistent across trials.

Step 1: Pre-task planning

We recognize that people are not always thrust into complex and ill-defined tasks with no prior warning, especially in adversarial situations; instead, they often have some time to plan and prepare for embarking on their mission. To account for this in our study, we allocate participants a limited amount of time (ten minutes) to plan their strategy before each adversarial game scenario begins. The pre-task planning phase gives participants an opportunity to think through the complexity of the situations and develop mitigating strategies and contingencies against their opponent.

As participants would have thought about the enemy and predicted enemy action during this planning stage, we interview the participants to elicit the concerns and considerations they have, and probe them to discover if their thoughts about the enemy have impacted the plan. The considerations elicited using the probes generally contains elements of what the participants think would happen in the simulation. The interviewer can run through the series of events which the participants think will happen, and ask if there are additional events that the participants would like to add. The interviewer should try to elicit as many predictions as possible.

In generating the first set of questions, we suggest following closely the cognitive dimensions suggested by Hoffman, Crandall, & Shadbolt (1998) and Crandall, Klein, & Hoffman (2006, p. 79). The questions should be modified to suit the context of the study. We found these dimensions to be generally sufficient and that such an approach would lead to minimal refinement during the pilot study. Table 1 presents our final set of questions.

Table 1: Interview probes in the pre-task planning session

Theme	Interview Questions
Lead-in	Why did you decide to [perform this action]?
	How did you come up with this [plan]?
Own Goals	What is your goal of performing [this action]?
	How did you decide on this goal?
	Why will the enemy fall for your plan?
Cues	How will you decide when to execute this action?
	How did you come up with this cue?
	How important is this knowledge in your decision? Why?
Expectation	What do you expect to happen?
	Why do you think [this action] will happen?
	How did you know that [this action] will happen?
	Why do you think the enemy will [do this] again? (if there was an earlier experience that led to the expectation)
Enemy's Goal	Do you know what the enemy was trying to achieve? If yes, what was it? How did you learn the en goal?
	How do you know that the enemy would (not) be [at this location]?

Step 2: Simulation and freezing

After the planning interview, the participants execute their plan. The participants are informed that the simulation would freeze and they would be interviewed but they are not given information on when the freeze would take place. The time of freezing is semi-regular since having a regular time is too predictable and would affect the participant's concentration on the game. It is also not fixed since the task is dynamically changing and it is difficult to determine an appropriate time to pause the game in advance.

As different experimenters play different roles, there are different dwell times for each experimenter. As an example, the gamer's dwell time was during the interruption when almost all other experimenters are busy. As such, rules as to when the simulation should be frozen, has to be set and agreed upon by all experimenters. The entire session is orchestrated to ensure that the experimenters have a good situation awareness of the experiment and are ready for the interruption, as well as the resumption of the game.

As the study is a hybrid of lab-based and naturalistic studies, there are many confounding factors that can impact the findings. We recognize that any missing link in our chain of conducting the study would further skew the results of the study. As such, we try to document as detailed as possible the protocol of the study. In the course of our study, we generated many rules for different aspects of the experiment. The following is the list of protocols which can be extended to other studies:

1. Like a commander in the battlefield, a single experimenter is responsible for calling and ending the interruption. In our case, it was the interviewer.
2. The game is interrupted at approximately every five minutes. This is about eight to twelve events in each game. Each event is typically a cue or a countermove. We find this number of events to be manageable for recalling the events.
3. The games are not interrupted when there is intense action. In our study, the intense actions were times of engagements.
4. At the conclusion of the interruption, there is a signature question that the interviewer would ask (what do you think will happen next?) and this would alert all experimenters that the game is resuming soon.

Step 3: Interruption

During the scenario, the interviewer makes a record of events as they unfold, which serves as a checklist in the Interruption Interview. When the scenario is frozen, participants are first asked to recall the events that occur in the game so that the interviewer may verify the participants' recollection. If the participants have no recollection of the event, the interviewer does not remind them or probe them for their thoughts about those events. Participants may have

missed noticing the event or the event was not significant. In either case, probing participants about these events may lead participants to rationalize the events and can possibly impact the subsequent game segments in the simulation.

For events which they have a recollection, they are asked to explain how they assessed the situation and arrived at their decisions. The probes used in the planning should be modified to fit the context, see Table 2. As there are many possible situations that can arise during the simulation, some of the earlier probes are not applicable for all situations. For example, the probes for a failed plan differ from the probes used when the participants spot the enemy.

Table 2: Probes for the interruption sessions

	Failed plan	Enemy Encounter	Change in plan	New action
(1) Recalling the event				
What was your reaction? What were you thinking about when [you saw this]?	x	x		
Why did you change your plan? Where did the idea come from?			x	
Why did you decide to this? How did you decide on this?				x
(2) Participant's goals				
What was your objective? How did you decide on this goal?		x	x	x
(3) Cues used in decision making				
Did you spot anything that was different from your plan?	x			
Did something in the game trigger the change? What was it?			x	
Did something in the game trigger this action? What was it?				x
How did you spot [this]? How did you use [this] to come up with this action? Have you seen this before? If yes, how was your response related to the previous encounter?	x	x	x	x
(4) Expectations about the game				
Were you surprised to see this? What did you think was going to happen? Why did you think so?	x	x		
What did you expect to happen? Why did you think [this possible action] would happen?			x	x
How did you know that the enemy would (not) be here?		x	x	x
(5) Understanding the enemy				
Did you know what the enemy was trying to achieve? If yes, what was it? How did you know that the enemy was going to do this?	x	x	x	x

In our pilot study, four broad categories of situations were identified: (1) plans that were executed and failed, (2) changes in plan after encountering the enemy, (3) unexplained changes in plan, (4) new action that were not discussed previously. For each situation, we then questioned the need to probe in each of the cognitive dimensions. As an example, it was unnecessary to ask about the goals since it would have been discussed in the earlier interview. For probes that investigate how the cues were processed, we asked different sets of questions for the four situations.

At the conclusion of the interruption, participants are asked what they think will happen next and what they plan to do when the game resumes. The simulation will resume until the next Interruption Interview, where they will be probed with this set of questions again. The iterations of playing the game scenario and conducting the Interruption Interview continue until the game scenario ends, in this case, when the participant either wins or loses the game.

Step 4: Post-task comprehension

A final interview is conducted when the simulation ends. This interview only focuses on the comprehension of the events in the last segment of the experiment task. The interviewer can also ask other optional questions relating to lessons learnt from the experiment and thoughts about the experiment.

For studies where participants undergo multiple game simulations, we recommend that the participants draw a summary of the events after each simulation. The summary serves two purposes: First, participants have the opportunity to recall the events that have happened and gain a deeper appreciation of the task. Second, the summaries can be pinned on the wall in the next experimental session to help the participants recall of the previous sessions.

4. Case Study

The study revealed many interesting findings. Here, we selected two findings which are relevant and appropriate in conveying the value of the Interruption Interview methodology. First, we demonstrate the effectiveness of the methodology by addressing one of the many research questions we had about adversarial decision-making.

Second, we addressed the validity of the findings with the intervention of the Interruption Interview. While SA and sensemaking are increasingly gaining attention and recognition as viable constructs for decision-making, the methodology and measurements are under great scrutiny (Crandall, et al., 2006p. 79). There are criticisms that the SA construct is not coherent with basic human information processing constructs such as long term memory. There is also disagreement within the community in the way SA is measured and many advocates for multiple SA measures to be taken (Parasuraman, Sheridan, & Wickens, 2008; Wickens, 2008). More

specifically, the freeze technique in the Situation Awareness Global Assessment Tool (SAGAT) is criticized for being intrusive as the measurement may prime participants to watch out for parameters, which appear in the questioning, more regularly than they do normally to achieve a high SA score. Endsley, however, found no significant performance difference between trials with and without the freezes (Breton & Rousseau, 2003; Edgar, Edgar, & Curry, 2003; Salmon, Stanton, Walker, & Green, 2006). We recognized that our methodology faces the same scrutiny.

Methodology

There were ten participants in this study. Each participant was tasked to strategize a series of seven real-time strategy games played between the Blue and Red Force. They command the Blue Force, which includes a city, key installations, land, and air forces. Both Red and Blue Forces have equal resources (types and number of troops) and the same objective of capturing the opponent's city. They reside physically in different rooms and the commanders do not meet during the span of the experiment.

In the first four games, the Red Force (played by a confederate experimenter) exhibited the same initial moves to prime the participant into believing that the enemy always does the same act. In the subsequent three games, the experimenter changed his strategy and performed a second set of initial moves. While doing so, the experimenter also has to exhibit a substantial amount of intelligence and adaptiveness. The game situation had to unfold naturally and the game outcome should not be controlled. To achieve this effect, the experimenter thought through and documented his rules of engagement, contingencies, and tactics during the experiment design phase and tried his best to adhere to the list during the experiment.

The Interruption Interview was performed in the first, third, fifth, and seventh game. This was designed to capture how the participant builds a profile of the enemy (over the seven trials) and how well he adapted to the change (Trial 5 to 7). In these trials, participants were stopped and interviewed at semi-regular intervals for their assessment and projection of the situation and the enemy. These projections were graded as correct, wrong, partially-correct, or undefined based on the intentions of the experimenter (who played the Red Force), as well as the outcome of the subsequent game play. The verbal protocol analysis provided two types of quantitative measurements for prediction performance: frequency of prediction and accuracy of prediction.

The scores for the seven games were also collected and were used as the measurement for task performance. These scores allowed us to test for correlation between the prediction performance and task performance. We also compared the mean of task performance between trials with and without interview to study the effects of the Interruption Interview.

Results

A total of 463 predictions were collected from the forty Interruption Interview trials (10 participants x 4 trials each). The correlation analysis was performed at the trial level (N = 40) and at the participant level (N = 10). At the participant level, the data was the mean score of the four trials that each participant had.

The first objective was to demonstrate the effectiveness of the Interruption Interview by addressing one of the many research questions in our study. Specifically, the interest is in how prediction performance affects task performance in an adversarial environment. We found a correlation between the frequency of prediction and game performance both at the participant level ($r(8) = .691$, p=.027) and at the trial level ($r(38) = .318$, p=.045). There was no significant correlation between prediction accuracy and game performance both at the participant level ($r(8) = .074$, p=.840) and at the trial level ($r(38) = -.127$, p=.434).

The second objective is to study the effect of the Interruption Interview on the task performance in our within-subject study. We asked some of our participants if they find the interviews disruptive to their thought processes. Of those respondents, all felt that the interviews provided them with additional time to think about the problems and that they had a better appreciation of the situation during the trials with the Interruption Interview.

However, results indicate no significant difference in game performance between trials without the Interruption Interview ($M = 8986$, $SD = 1541$) and trials with the Interruption Interview ($M = 9125$, $SD = 1554$), $t(9) = .843$. This indicates that the Interruption Interview did not have an effect on their performance.

Discussion

The result in this case study suggests that the accuracy of predicting an enemy does not affect the outcome of the game. One does not have to be able to very accurate in his prediction of the enemy but he should try to make as many predictions about the enemy as possible. This was suggested by the significant positive correlation between the number of predictions made and the game performance. A possible explanation could be that when participants thought about the enemy, they were mentally prepared for the case and had contingencies for both the case and the alternatives. When the events unfolded and it did not conform to their prediction, they were able to quickly modify their plan to handle the situation. As such, there was better game performance when there was higher frequency of prediction but not with the accuracy of the prediction.

One factor that may confound this finding is with the evaluation of the predictions. The evaluation was judged by comparing the prediction (level 3 SA) with what happened in the game

and what the opponent was trying to do. The problem is that there are a handful of cases where we faced difficulty in evaluating the predictions. They can be generalized into three categories:

1. The SA cannot be observed from the playback of the game. Examples include enemy's goals, emotions, and cognition. These SA have to be rated by the opponent as quickly as possible. This is difficult since it takes time for the conversation to be transcribed and the confederate experimenter may not fully recall his intention for that specific point.
2. The SA was dependent on a hypothetical event. This is difficult when the SA was true (i.e. the opponent would behave under the hypothetical circumstance) but the hypothetical event did not happen.
3. The SA was partially correct. We gave a score of 1 for correct SA, 0.5 for a partially correct SA, and 0 for incorrect SA. While 0.5 was not the best way to rate the partially correct cases since there is a variation of correctness, we had to simplify the process.

It is very unlikely that the difficulty in judging the predictions would have confound the findings as this difficulty was experienced in all the trials. Future studies can look into how these difficulties can be resolved and test if the frequency of predictions has an impact on the outcome of the game.

The second hypothesis examines the effect of the interruption on task performance. Participants reported that they had more time to think about the enemy with the interruption though we do not find a difference in their game performance between games with and without the Interruption Interviews. Even if the interruption had affected the game performance, our primary aim of eliciting the sensemaking processes was not really affected by the interruption method.

We recognized that other than the interruption, there were many other factors affecting the perception and processing of the adversary. Examples include the pace of the fight, fatigue and stress. We did not control for these factors as it was not the interest of this study to examine their effects. Rather, we were interested in how people think about the adversary. As such, the deeper their analysis, the more valuable the findings were. Although the Interruption Interview appears to disrupt their thought processes and possibly aid their memory recall, we were not too concerned with its effects since this is an exploratory and not a confirmatory study. The improved memory recall was not a concern as the fidelity of the set-up was low to begin with.

5. Conclusion and future research

The Interruption Interview approach outlined in this article enabled us to perform an in-depth analysis of the interplay between various psychological constructs related to adversarial thinking. However, the technique is resource-intensive and inefficient. One should pick up the

technique only if the there is a genuine interest in an in-depth analysis and the cognitive task is sufficiently complex and ill-defined.

Despite these difficulties, there are still merits in the technique especially if one is trying to gain an in-depth insight into the sensemaking in an ill-defined problem space. With the data collected using the Interruption Interview protocol, we could study the sensemaking mechanism that binds our perception, comprehension, and prediction. Though we do not show our full analysis protocol and results in this paper, we demonstrated the possibility of quantifying an insight that we obtained from the study. From our data, we were also able to build the ontology of the concepts used in the task. We understood how an enemy's awareness is built from both situational factors such as the terrain and enemy's profile such as the enemy's rules of engagement. In addition, subtle but frequently used constructs such as the SA of the enemy (note that this is different from SA about the enemy) also surfaced in our analysis. We suspect that this construct is also applicable to the study of team SA.

The constructs obtained in an Interruption Interview can be used for the development of an exhaustive set of SAGAT questions. We compared the constructs we uncover in our study with the SAGAT questions developed for Military Operations on Urban Terrain (Endsley, 1995a). The dimensions reported in these studies were mainly physical. Our study found that participants assessed three other dimensions: (1) abstract dimensions (e.g. enemy's goals), (2) enemy's perception (what the enemy has seen and can see), and (3) enemy's cognition (e.g. enemy's attention, experience). Though it is still early to conclude the usefulness of these dimensions, having an all-rounded set of SAGAT questions would definitely lead to research studies that are more robust.

The large pool of literature in SA suggest that our findings can potentially guide us in personnel selection, design of training and equipment that lead to better SA, as well as the evaluation of these designs (Lampton, et al., 2005; Matthews, Pleban, Endsley, & Strater, 2000). We can use our findings for the development of Situation Awareness Global Assessment Tool (SAGAT) for enemy awareness. Our ability to measure one's awareness of his adversaries can inform us about the factors that impact our sensemaking, as well as the effect of each level on task performance.

From this exploratory study, we have identified the key processes responsible for a good SA about the enemy. We are currently building a set of material for training these processes and testing its effectiveness. This would be tested in a controlled setting where the sample size is larger and the results are more measurable and conclusive.

While there are existing methods, such as Hierarchical Task Analysis (HTA), for the development of SAGAT questionnaires, our Interruption Interview is more detailed. To quantify, we found that people project the attention and emotion of their enemy and use the information in their decisions. While we were aware that people make such assumptions (before this research is

reported), we did not find accounts for such nuances of considerations in HTA-alike tools. For HTA, the experimenter needs to know the context very well in order to develop the measurements and injects. The definition of measurement is dependent on Subject-Matter-Expert and therefore, lacks objectivity. Unlike Endsley's measurement of situation awareness, questions in the Interruption Interview were semi-structured; the interviewer only probed for information that the participants know.

The Interruption Interview is, however, very time-consuming and we would only recommend it for ill-defined tasks where other methodologies do not work and there is sufficient interest to look deeper at the decision tasks. If the tasks are well-defined and comprehensible, HTA should suffice. Otherwise, the Critical Decision Method is a parallel method to the Interruption Interview to study sensemaking.

One of the challenges in conducting an Interruption Interview is getting the team of experimenters to come together and coordinate the effort. As the scope has to be focused, it may be difficult to get experimenters who have common interests. In addition, some of the experimenter roles require multi-disciplinary and specialized skills. In our case, the experimenter, who plays the role of the opponent, needs to (1) play the game well, and (2) be aware of the importance of experimental control. Even with both skill sets, he needs to be able to control his emotions and meta-cognition. With a big team, this can become very effortful and difficult to control.

Like other verbal techniques, such as the Think-Aloud Method and Critical Decision Method, one challenge in using the Interruption Interview is the participants' capacity to externalize their thoughts. Participants may not be aware of their own internal thinking or do not have the linguistic ability to share their thought process. Probing these participants using simpler jargons, such as those applied in the critical decision method, may solve the problem. A more difficult class of participants is those who are not comfortable with sharing their thoughts. It is difficult for experimenters to identify these participants a priori and reject their participation. One strategy would be to recruit more than the required number of participants and terminate the study as soon as the experimenter assessed that the participant is not appropriate.

6. References

Breton, R., & Rousseau, R. (2003). *Situation awareness: A review of the concept and its measurement* (No. TR-2001-220): Defence Research Establishment Valcartier.

Cohen, M. S., Adelman, L., Tolcott, M. A., Bresnick, T. A., & Marvin, F. (1993). *A Cognitive Framework for Battlefield Commanders' Situation Assessment*. Arlington, VA: Cognitive Technologies, Inc.

Cohen, M. S., Thompson, B. B., Adelman, L., Bresnick, T. A., Tolcott, M. A., & Freeman, J. T. (1995). *Rapid Capturing of Battlefield Mental Models.* (No. Technical Report 95-3). Arlington, VA: Cognitive Technologies, Inc.

Crandall, B., Klein, G., & Hoffman, R. R. (2006). Working minds: A practitioner's guide to cogntitive task analysis

Dorner, D. (1996). *The logic of failure: Recognizing and avoiding error in complex situations*: Basic Books.

Durso, F. T., & Sethumadhavan, A. (2008). Situation Awareness: Understanding Dynamic Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*, 442-448.

Edgar, G. K., Edgar, H. E., & Curry, M. B. (2003). *Using Signal Detection Theory to Measure Situation Awareness in Command and Control.* Paper presented at the Human Factors and Ergonomics Society 47th Annual Meeting, Denver, CO.

Endsley, M. R. (1988). *Design and evaluation for situation awareness enhancement.* Paper presented at the Human Factors and Ergonomics Society 32nd Annual Meeting, Anaheim, CA.

Endsley, M. R. (1995a). Measurement of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*, 65-84.

Endsley, M. R. (1995b). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*, 32-64.

Gilmour, D. A., Krause, L. S., Lehman, L. A., Santos, E., Jr., & Zhao, Q. (2005). *Intent Driven Adversarial Modeling*. Paper presented at the 10th International Command and Control Research and Technology Symposium: The Future of C2.

Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998). Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 40*(2), 254-276.

Holmquist, J. P., & Goldberg, S. L. (2007). *Dynamic situations: The soldiers situation awareness* (No. RTO-TR-HFM-121-Part-II): NATO Research and Technology Organisation.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE intelligent systems, 21*(4).

Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man and Cybernetics Part A, 19*(3), 462-472.

Kott, A., & Ownby, M. (2005). *Tools for Real-Time Anticipation of Enemy Actions in Tactical Ground Operations*. Paper presented at the International Command and Control Research and Technology Symposium.

Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Systems Journal, 42*(3), 462.

Lampton, D. R., Cohn, J. V., Endsley, M. R., Freeman, J. T., Gately, M. T., & Martin, G. A. (2005). *Measuring situation awareness for dismounted infantry soldiers*. Paper presented at the Interservice/Industry Training, Simulation, and Education Conference.

Lebiere, C., & Best, B. J. (2009). From Microcognition to Macrocognition: Architectural Support for Adversarial Behavior. *Journal of Cognitive Engineering and Decision Making, 3*, 176-193.

Matthews, M. D., Pleban, R. J., Endsley, M. R., & Strater, L. D. (2000). *Measures of infantry situation awareness for a virtual MOUT environment.* Paper presented at the Human Performance, Situation Awareness and Automation: User Centered Design for the New Millennium Conference, Savannah, GA.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs. *Journal of Cognitive Engineering and Decision Making, 2*, 140-160.

Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics, 37*(2), 225-238.

Wickens, C. D. (2008). Situation Awareness: Review of Mica Endsley's 1995 Articles on Situation Awareness Theory and Measurement. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*, 397-403.