JDDM
Journal of Dynamic Decision Making

Open Access

Original Research

Strategies, tactics, and errors in dynamic decision making in an Asian sample

C. Dominik Güss[1], Ma. Teresa Tuason[2], and Lloyd V. Orduña[3]

[1]Department of Psychology, University of North Florida, [2]Department of Public Health, University of North Florida, and [3]Research & Development Center, University of Baguio

The current study had three goals: (1) to investigate strategies, tactics, and errors as predictors of success and failure under uncertainty following the dynamic decision making (DDM) and complex problem solving (CPS) framework; (2) to use observation and to examine its reliability and potential as a data collection method when using microworlds; and (3) to investigate the applicability and validity of a microworld developed in the West, to an Asian sample. One hundred three participants in the Philippines took the role of fire chief in the microworld WINFIRE (Gerdes, Dörner, & Pfeiffer, 1993). Their strategies, tactics, and errors were observed and coded by experimenters as they worked individually on the simulation twice. Results showed that (1) DDM strategies, tactics, and errors predicted success and failure in WINFIRE, and strategies and tactics that led to success increased while errors decreased over time; (2) strategies, tactics, and errors can be validly assessed through observation by experimenters, specifically that two types of decision makers were identified: the active, flexible, and big picture planners and the slow or cautious, and single-focused decision makers; (3) these findings together with participants' survey ratings speak for the applicability of the microworld in an East Asian sample and for its validity. Findings are potentially relevant for experts and for training programs, highlighting the benefits of virtual environments during DDM.

Keywords: strategy, tactic, dynamic decision making, complex problem solving, naturalistic decision making, errors, success, cognitive biases, microworld, uncertainty, virtual environment, culture

The fire chief has a problem. Fires are spreading in the forest near the city, and with the current wind strength and direction, they will most likely move toward the city. This is not the only problem, however. North of the city, in the forest, there are several small fires starting due to the extreme summer heat. What can the fire chief do? What strategy will the fire chief choose? Should the fire chief distribute the firefighting units and give them commands to quickly extinguish the fires, or order them to clear an area close to the city to stop the fires from spreading? Should the fire chief focus on the most urgent fire first, or try to deal with all the fires at the same time? We presented participants with this situation simulated as a virtual environment in order to observe their strategies, tactics, and errors.

Developing and selecting successful strategies and tactics and avoiding errors are important to daily life. A strategy provides a general framework for action, for example: *Focus on the main fire first.* Strategies can be defined as broad directions to meet long-term goals (Güss, 2000) and provide answers to the questions: "What do I want to accomplish and why?" Tactics, which are more specific, address short-term goals by constructing and executing short-term plans to implement a strategy (Bonissone, Dutta, & Wood, 1994), for example, *Since the fire is strong, send three fire trucks to the fire.* Tactics answer the question: "How can I get things done under the given circumstances?" Both levels, strategic and tactical, are crucial for successful action: tactics are used to react quickly to environment changes, whereas strategic plans provide more continuity.

Strategies play an important role in complex real life decision making, yet it has been problematic to link them to success or failure, or to describe them as errors or biases. One reason for the difficulty in analyzing errors and success of strategies is related to methodological problems. First, many analyses of errors in naturalistic contexts are retrospective (e.g., Reason, 1990). One starts with the negative unintended outcome and tries to identify errors that led to this outcome. Such post-accident analyses can fall for the hindsight bias, i.e., in hindsight it seems easy to identify the strategy that led to failure, when at the time the situation actually occurred, the outcome was not easily foreseeable (Fischhoff, 1975). Further, it is possible for bad outcomes to happen as a consequence of sound decision strategies. Also, given that human beings are part of, and act within complex, uncertain systems, it is difficult to identify one specific cause of error (Lipshitz, 2005). Some researchers argue that errors cannot easily be detected in dynamic environments due to unclear links of causation in the system (e.g., Flach, 1999; Rasmussen, 1990) or that it is difficult to define what the right decision strategy is in an ill-defined situation. As Klein (1999, p. 97) put it, "when we move to natural settings where we don't have tight control of the stimulus conditions we may have difficulty identifying biases and errors."

This and other methodological problems have been recognized by Montgomery, Lipshitz, and Brehmer (2005, p. 10), who summarized the state of naturalistic decision making NDM methodology and argued, "NDM still lacks generally accepted criteria of rigor that are suitable for its particular nature. Such criteria are essential for guiding the design of good research and the evaluation of research outcomes." However, this linking of strategies to success or failure is an important task. For instance, survey results

Corresponding author: C. Dominik Güss, Department of Psychology, University of North Florida, Jacksonville, FL 32225. e-mail: dguess@unf.edu

have shown that over 70% of accidents in medicine, aviation, and various industries – and this is a conservative estimate – can be attributed to human error (Cook & Woods, 1994). Therefore, the ability to identify successful strategies, tactics, and errors in decision making is key. Exactly this can be done with complex and dynamic computer-simulated problem scenarios used in the field of complex problem solving CPS and dynamic decision making DDM (Dörner, 1996; O'Neil & Perez, 2008). The pitfalls of naturalistic error and strategy analysis such as being retrospective, hindsight bias, lack of control, and questionable links of causation can be avoided with standardized computer-simulated problem scenarios. Thus, the main goal of this article is to assess strategies, tactics, and errors empirically in a laboratory situation. Like a few researchers in the past, our study investigated strategies, tactics, and errors over time, related them to performance, but unlike those researchers, we used observation as a method.

Strategies in Dynamic Decision Making (DDM) and Complex Problem Solving (CPS)

There has been recent interest in using virtual environments (also called microworlds or simulations, these terms will be used synonymously) to study strategies in the fields of DDM and CPS (e.g., Baker & Delacruz, 2008; Brehmer, 2004; Funke, 2003; Güss, Tuason, & Gerhard, 2010). Schmid, Ragni, Gonzalez, and Funke (2011) distinguished between the DDM and CPS research traditions. According to the authors, CPS refers to research mainly conducted in Europe (Broadbent, 1977; Dörner, 1980; Frensch & Funke, 1995) where the primary goal is to investigate how naive subjects deal with complex, nontransparent, and dynamic problems simulated as microworlds. The term DDM is more frequently used in the North American tradition (e.g., Gonzalez, 2005), but is similar to CPS in that it also focuses on environments consisting of connected variables that change over time. Additionally, the DDM tradition might use simpler simulated problems. This does not mean, however, that it is easy for participants to solve them. In fact participants often do very poorly (e.g., Funke, 1991; Strohschneider & Güss, 1999).

Using computerized simulations or microworlds in the study of CPS and DDM, it is possible to develop complex theories of human thought and behavior (e.g., Dörner, 1999; Dörner & Güss, 2013) and to adequately identify which strategies in a specific situation will be more likely to lead to success and which to failure. In microworlds (Brehmer & Dörner, 1993; Frensch & Funke, 1995; Funke, 2003; Gonzalez, 2005; Schmid, Ragni, Gonzalez, & Funke, 2011), a participant is confronted with a complex, uncertain, and dynamic problem situation and performance is assessed by automatically recording and saving to a computer file each decision the participant makes, along with changes in the system state. Predictors of individual differences in DDM and CPS and their relation to performance have been widely studied, such as intelligence (e.g., Rigas & Brehmer, 1999; Strohschneider, 1991; Wenke, Frensch, & Funke, 2005), personality variables (e.g., Schaub, 2001), or problem-relevant domain-specific knowledge (e.g., Putz-Osterloh & Lemme, 1987). Such studies have certain limitations, as solely focusing on performance outcomes will not reveal details about the DDM process (Spector, 2008).

Therefore, some DDM and CPS researchers have suggested using these computer protocols to examine behavior patterns (Fischer, Greiff, & Funke, 2012; Strohschneider & Güss, 1999), or strategies (Schoppek & Putz-Osterloh, 2003; Wüstenberg, Greiff, Molnár, & Funke, 2014), or common tactical failures (e.g., Dörner, 1996; Ramnarayan, Strohschneider, & Schaub, 1997). Thus, the goal of the current study is to address the need expressed by researchers to study strategies and errors systematically, specifically their role in DDM (e.g., Qudrat-Ullah, 2008; Schoppek & Putz-Osterloh, 2003). With particular focus on the process of DDM (as opposed to the relationship of individual differences to outcomes), the current study investigated strategies, tactics, and errors leading to success or failure by observing participants as they worked on a microworld. We will discuss our research using the term DDM, although the microworld WINFIRE has been used in both the CPS and DDM tradition.

It was postulated that strategies and tactics would be adapted to the situation over time. Successful strategies and tactics would increase in frequency and errors would decrease in frequency. Before discussing specific strategies, tactics, and errors, we will describe the WINFIRE simulation in detail, to illustrate its demands and constraints.

Demands of the WINFIRE Simulation

In the microworld WINFIRE (Gerdes et al., 1993), a participant takes the role of a fire-fighting commander who tries to protect cities from approaching fires, put out existing fires, and save the forest. Similar fire simulations, e.g., NewFire or the Firechief simulation in Australia (Omodei, McLennan, Elliott, Wearing, & Clancy, 2005) or C3Fire in Sweden (Granlund, 2003; Johansson, Trnka, Granlund, & Götmar, 2010) have been used by other researchers focusing more on the results and the influence of outside variables (such as personality traits or intelligence) or simulation variations than on decision-making process analyses.

Simpler versions of the FIRE scenario have been used in previous research to investigate DDM strategies. Schoppek (1991), for example, administered the scenario five times to 22 German participants, analyzed strategies in the logfiles (e.g., duration of scenario, number of commands and other, more sophisticated indices, such as initial distribution of fire-fighting units) and showed that a variety of problem-solving operations correlated with the use of multiple learning strategies to prepare for exams. Dörner and Pfeifer in Germany (1991) compared 20 participants who worked under the stress of sound to 20 participants who worked without stress on five versions of FIRE. They found that although the two groups did not significantly differ in performance, they showed different strategic approaches, such that stressed participants, for instance, gave more goal commands to individual units and fewer goal commands to several units. In their study, the relationship between different strategic behaviors and performance, however, was not examined.

The design of the current WINFIRE scenario requires participants to distribute their resources, puts the decision maker under time pressure and requires quick and decisive actions. According to the microworld criteria discussed previously, WINFIRE is moderate in complexity (consists of many variables), high in dynamics (develops in a nonlinear way), and moderate in transparency (unknown new fire locations and unforeseen wind changes). Specifically, the complexity results from having to select from four main (and three other) command options for each of 12 units at any given time. WINFIRE is highly dynamic because the situation changes even without any intervention from the participant; in 11 minutes, 15 fires break out at programmed times and changes in wind direction and strength are programmed as well. WINFIRE is moderate in transparency because fires start at staggered unknown times and locations.

The presumed DDM core strategies in WINFIRE consist of assessing situations quickly and identifying crucial situations, prioritizing, flexible planning that considers resource allocation and situational demands, and quick decision making which takes long-term effects into account to avoid further escalation of the problems.

Participant Observation

Extant research describing dynamic decisions in a FIRE simulation (Dörner and Pfeifer, 1991; Schoppek, 1991), referred to decisions of participants automatically saved in log files. These log files accurately reflect each decision of every participant. One disadvantage of these log files is that they mostly indicate tactics and system data (e.g. number of commands, duration of a simulation), as opposed to strategies.

The current study attempted to use observation in order to assess tactics and strategies (like some in Schoppek's log-file analysis, 1991) using a bigger sample size. Using observation, one can assess a decision in relation to a specific context as shown on the computer screen. Using one helicopter can be "good" or "bad". It depends on the context, for example, of how big a fire is or how distant the helicopter is from the fire and if there are other nearby units that would be available. This context is sometimes missing if one solely analyzes the log files. Whereas observation is one of the key methods used in analyzing naturalistic decision making NDM in the field (e.g., Lipshitz, 2005), it has been rarely used to assess strategies and tactics in the study of microworlds in the laboratory. Thus another goal of the current study is to assess the applicability of participant observation as a method of data collection using the WINFIRE scenario.

The Selection of DDM Strategies

The key demands of WINFIRE are related to situation assessment, planning, and decision making under time pressure and ever-changing conditions. Both planning and decision making are probably the key steps in the DDM process (e.g., Dörner, 1996; Funke, 2010). Planning can be defined as the development of a course of action aimed at achieving a goal (Güss, 2000; Hayes-Roth & Hayes-Roth, 1979). Decision making, then, refers to the selection and execution of a course of action. Participants have to take the described constraints in WINFIRE (e.g., slow moving trucks, wind strength and wind direction, lack of water) into consideration and plan how to distribute and allocate their resources to mitigate forest fires.

To identify potentially successful and unsuccessful strategies for the current study, we used both top-down (theoretically guided) and bottom-up (empirically guided) approaches; top-down, because theoretically they are based on the planning and decision-making steps of DDM; bottom-up because strategies, tactics, and errors are triggered by the context of WINFIRE and were operationalized and identified within this context. The strategies, tactics, and errors were observed when participants worked on the simulation in pilot studies.

The following set of four strategies and tactics are related to planning because they involve decisions based on predictions about future developments of the system environment – whether more proactive general strategic planning or more reactive tactical planning in specific circumstances. The planning strategies and tactics were expected to correlate positively with success (i.e., percent of remaining forest area at the end of the simulation).

The second set consists of four errors because they involve immediate decisions that are not adapted adequately to the current state of the environment. The errors were expected to correlate negatively with success. One exception was the strategy "Effect control and flexible strategy." We could not make a clear prediction because a sudden change in behavior pattern can be either positively or negatively associated with performance depending upon the change.

Planning Strategies and Tactics

- Proactive strategic planning: Active distribution of units, operationalized as distribution of fire-fighting units before the first fire starts in cycle 4 and use of patrol command that makes units patrol a specified area independently throughout the simulation. This strategy is expected to correlate positively with performance as it takes time for units to travel from one location to another. If units are not well distributed to cover the entire terrain, they may be too far from an emerging fire, allowing the fire to spread before they get there and making it harder to contain. The patrol command can be used to assign areas for units to patrol independently and search for fires, which can minimize potential travel time to emerging fires.
- Flexible strategy: Sudden change in behavior pattern, for example, suddenly selecting several units at once and giving them the same command. (Previous research has shown how stress can reduce flexibility in problem solving, e.g., Renner & Beversdorf, 2010.) A flexible strategy usually indicates the realization that an executed problem-solving approach did not show the expected result and a new problem-solving strategy has been executed to address the problem situation.
- Tactical Planning: Number of helicopters sent to a fire in the first 10 seconds after it starts. When a fire starts, a participant needs to plan. The participant has to detect the fire, have the goal to extinguish the fire, predict how long it takes a unit to reach the goal, predict spreading of fires due to wind. The participant must select one or more units (depending on their location and the strength of a specific fire), give them the goal command, and indicate where they need to move. The participant must consider the length of units' travel times, the wind direction, and spreading of the fires, and decide upon the appropriate number of units to send to a fire to control it without having too many units in one place, leaving areas vulnerable to other emerging fires.
- Tactical Planning: Number of trucks sent to a fire in the first 10 seconds after the fire starts.

Decision-Making Errors

- Single-focus strategy and lack of multi-tasking: Number of decisions related to one fire only and not to other fires burning more than 5 seconds (counted for each of the 9 observation cycles where new fires started, minimum 0, maximum 9). (Multitasking has been described as crucial for many dynamic tasks, e.g., Hunt & Joslyn, 2000; Salvucci & Taatgen, 2008.) While focusing on one fire, one cannot neglect the other burning fires.

- Incorrect order or function of commands: Number of tactical decision-making errors, such as sending a unit to a fire and forgetting to click the extinguish command, or mixing up commands such as clicking the patrol command instead of the extinguish command when attempting to extinguish a fire, or clicking on extinguish without selecting a specific fire-fighting unit to receive the command. (These errors are related to slips and lapses as described by Reason, 1990; slips are actions that are carried out but not in the intended way, and lapses are omissions in a sequence of actions.) Commands have to be given in a certain form and sequence. Tactical mistakes in executing commands will not lead to the intended results.
- Perceptual inaccuracy leading to inappropriate action: Number of units sent to a burned field. (Once a forest area has burned, it will remain grey, thus, sending units to burned areas instead of burning areas is a waste of resources and will have no effect.)
- Lack of foresight and impulsivity: Number of distant units sent to a fire instead of nearby units. (Other research has shown that impulsivity was related to lower performance and higher response frequencies in a dynamic task, Quiroga, Martínez-Molina, Lozano, & Santacreu, 2011.) As previously mentioned, it takes units a relatively long time to travel to a fire location. We observed participants who, perhaps due to stress, sent distant units to a fire when they could have sent a unit nearby. By the time the unit reached the fire, the fire had spread.

Cultural Applicability

Although several cross-cultural studies in the field of DDM exist (e.g., Güss, Tuason, & Gerhard, 2010; Strohschneider & Güss, 1999), most of the research has been conducted in Western Europe and the United States. For psychology as a science, it is imperative, however, to study individuals across the world and to test theories and methods in various cultural and ethnic groups. Using the WINFIRE simulation, research has shown that already the problem perception differs among participants from different countries (Güss, Glencross, Tuason, Summerlin, & Richard, 2004). Brazilian participants, for example, perceived FIRE as more complex and more difficult compared to Indian and US participants.

The current investigation of DDM strategies and tactics and errors, was conducted with an Asian sample, specifically in the Philippines. The applicability and validity of a microworld in the Filipino sample will be investigated. Besides gathering WINFIRE data and analyzing observation data, survey questions about the simulation will be distributed to assess what participants think of the simulation.

Goals of the Current Study

In summary, the current study has three goals. First, the current study addresses a need to systematically investigate strategies, tactics, and errors in DDM over time and as predictors of performance. The four planning strategies and tactics are expected to correlate positively with performance and the four decision-making errors are expected to correlate negatively with performance. Second, as observation is used as a method to assess strategies, tactics, and errors in their context, the study aims to examine the reliability of observation, and its potential as a data collection method. Third, to increase generalizability of the findings, the applicability and validity of the microworld WINFIRE developed in the West, is tested in an Asian sample.

Method

Participants

Participants were 103 undergraduate students with various majors of study, from two universities in the northern Philippines. Fifty-nine percent of the participants were female, 41% were male. Their ages ranged from 18 to 35 years ($M = 20.0$, $SD = 2.59$). Computer simulation data from seven participants were incomplete (i.e., from 6 participants we have only one of two Fire data sets, from one participant none of the two Fire data sets were saved). We have observation data from all, but one participant. When initially performance was regressed on strategies and tactics, three outliers in Fire A and three in Fire B were deleted due to their Mahalanobis and Cook's distances. Thus the final data set was 96 participants in Fire A and Fire B.

Instruments

WINFIRE. WINFIRE (Gerdes et al., 1993) was used as research instrument and was already briefly described in the introduction. On the computer screen in WINFIRE, participants see the forest, cities, red fire-fighting trucks, yellow helicopters, lakes, and a black stony area, and the command options. If fire-fighting units can extinguish a burning field in time, the field will remain green. If they come too late, the field will be completely burned and will turn grey. Wind strength and direction determine how fast fires will spread and in which direction.

Participants are given eight command options. By selecting a unit and clicking the *goal* command, participants can send a unit to a specific location, for example, to a fire, or to a lake to fill up water. When units receive the *search* command, they independently search for fires in their vicinity. Units that receive the *extinguish* command extinguish fires, and units that receive the clear command cut trees to create a barrier to prevent spreading of fires. The performance criterion in WINFIRE is proportion of saved forest at the end of the game. At the end of each time step, the percentage of unburned forest is determined and saved in a computer file.

The 11-minute WINFIRE simulation consists of 111 time steps (trials), each lasting 6 seconds, after which the proportion of burning forest is updated. For the participants, the situation seems to develop continuously. During the first 2 minutes of the simulation, no fires start, giving participants time to familiarize themselves with the situation and to distribute fire trucks and helicopters strategically. Then a few small fires start, which can be contained easily because there is little wind. After 4 minutes of the simulation, fires start simultaneously and spread because of wind strength. In total, 15 fires (some starting on neighboring fields) are programmed to start at different times.

The same simulation was played twice by each participant because we were interested in seeing whether and how strategies, tactics, and errors would change between WINFIRE A and B. WINFIRE A and WINFIRE B were exactly the same simulation; due to its difficulty, participants performed very poorly in WINFIRE A. Without any intervention, 45.18% of the forest would be saved at the end of the simulation. Overall, participants saved 50.06% ($SD = 8.89$) of the forest in WINFIRE A. If we excluded the 4 outliers whose performance scores were more than 2 standard deviations above the mean during WINFIRE A, the proportion of saved forest at the end of WINFIRE A would be only 48.44 ($SD = 4.00$). Thus, the simulation was very difficult for the participants, even after reading the instructions and practicing in a test game. Despite the simulation's difficulty, participants' performance improved in the second simulation. They saved 63.18% ($SD = 17.83$) of the forest at the end of the WINFIRE B game. Thus, using the same simulation twice was necessary to see how well participants performed and whether they tried the same or different strategies or tactics, or made the same or different errors.

Coding system and Training. Experimenters were trained to identify the 4 planning strategies and tactics and 4 decision-making errors, and to code them accurately. The two authors trained the three Filipino experimenters on the coding system. The training, which took a period of 1 week consisted of doing practice coding together as a team, and then coding individually and assessing the reliability of the coding. For each participant, 2 experimenters seated behind them when they worked on WINFIRE and used the coding system to code each decision participants made and certain predefined resulting events. The 11-minute WINFIRE time was divided into 22 cycles, each lasting 30 seconds (a time step in WINFIRE lasted 6 seconds). At cycles 4, 5, 7, 9, 10, 12, 13, 14, and 15, fires started. The strategies and tactics coded were described in the introduction.

Interrater reliability was always calculated for the two experimenters present at each observation experiment. Percent of agreement and Kappa were calculated. According to Fleiss (1981), a Kappa between .40 and .60 is fair, between .60 and .75 is good, higher than .75 is excellent. Overall, the Kappas ranged from fair to excellent (see Table 1).

Procedure

Initially, each participant received a 3-page instruction sheet explaining the context of WINFIRE and the command options, including a screenshot. Participants kept the sheet for the experiment duration. After reading the instructions, participants played a 10-minute test version of the WINFIRE simulation to familiarize themselves with the commands and screen. Two experimenters were seated about 3 feet slightly behind the participant and were instructed to explain the commands in the test game and to answer questions participants might have regarding the nature of the commands. However, experimenters were not allowed to give any form of advice on how to work or succeed in the simulation. Then, each participant worked on exactly the same WINFIRE simulation twice. While participants worked on WINFIRE, experimenters were instructed to observe participants' mouse clicks and actions on the screen and to fill out the coding sheet.

Two Kappas raised questions: Kappa .39 for *Sudden change in strategy* and Kappa -.02 for *Sending unit to burned field*. It is important to consider, however, that these two strategies were defined and coded as yes/no compared to the other categories, which were numerical counts. For *Sudden change in strategy*, the interrater agreement was actually 96%. So 96 times out of 100, both raters agreed with "no" and 1 time they agreed with "yes." Only 3 times out of 100 did one rater code "yes" and the other "no." A similar pattern was found for *Sending unit to burned field* (97% agreement). Thus, despite the Kappas, the interrater agreement for these two strategies can be considered very high. For such a complex coding system, the Kappas and the agreement percentages can be regarded as highly satisfactory, probably due to the long and intensive training given to the experimenters.

Results

Before we describe strategies, tactics, and performance, we would like to briefly provide information about participants' perception of the WINFIRE simulation. The objective characterization of the WINFIRE problem situation as moderate in complexity, high in dynamics, and moderate in transparency has been validated by participants' subjective perceptions assessed through rating scales and by participants' performance. After completing both versions of WINFIRE, participants answered a 28-question survey using a 7-point Likert scale (1-yes to 7-no) regarding the characteristics of WINFIRE. Referring only to questions with extreme mean value ratings above 6 and below 2, participants perceived that: small actions had big consequences ($M = 1.65$, $SD = 1.09$), the simulation was close to real life ($M = 1.80$, $SD = 0.95$), a single command could have a huge impact ($M = 1.92$, $SD = 1.30$), the simulation described a realistic situation ($M = 1.96$, $SD = 1.14$), many factors came together and influenced each other ($M = 1.73$, SD

Table 1. Planning strategies and tactics and errors and their operationalizations and interrater reliability.

	Percent Agreement	Kappa
Planning strategies and tactics		
Proactive strategic planning: Active distribution of units and use of patrol command	95.6 %	0.71
Effect control and flexible strategy: Sudden change in behavior pattern	97.0 %	0.39
Tactical planning: Number of helicopters sent to a starting fire	90.0 %	0.46
Tactical planning: Number of trucks sent to a starting fire	91.7 %	0.78
Errors		
Single-focus strategy and lack of multi-tasking: Decisions related to one fire only and not to multiple fires	95.6 %	0.90
Incorrect order or function of commands: Number of tactical decision-making errors	86.1 %	0.73
Perceptual inaccuracy leading to inappropriate action: Number of units sent to a burned field	96.7 %	-0.02
Lack of foresight and impulsivity: Number of more distant units sent to a fire instead of nearby units	93.3 %	0.42

$= 0.97$), and it was important to be successful in the game ($M = 1.67$, $SD = 1.04$).

Predictive Validity: Correlations between Strategies, Tactics, and Errors and Overall Performance in WINFIRE A and WINFIRE B

In order to measure the predictive validity of the assessed strategies, their frequencies were correlated with success in both versions of the WINFIRE simulation. Success was defined as percentage of forest saved at the end of the simulation. Success was better in WINFIRE B ($M = 63.18$, $SD = 17.83$) compared to WINFIRE A ($M = 50.06$, $SD = 8.89$), as a paired samples t-test showed, $t(91) = -7.42$, $p < .001$, with a large effect size, Cohen's $d = .68$. The correlation between performance in WINFIRE A and WINFIRE B was $r = .34$, $p < .001$.

We controlled for potential covariates such as computer skills, years using a PC, and problems with using the mouse. None of these correlated significantly with success. Therefore, they were not included as covariates in further analyses.

Table 2 shows the correlations between strategies and success in WINFIRE A and WINFIRE B. It was expected that the first four planning strategies and tactics would be successful and correlate positively with success in both WINFIRE versions and that the second set of four errors would be unsuccessful and correlate negatively with success in both WINFIRE versions. None of the expected correlations was significant for WINFIRE A (we excluded 4 extreme outliers for this analysis – participants UB02, UB17, UB23, and SLU23, whose performance was higher than 70% of saved forest area). The non-significant correlations are probably due to a floor effect with 70% of all participants in WINFIRE A having a performance between 41% and 49% of saved forest area. For WINFIRE B, 6 of the 8 correlations were significant; of these 6 correlations, 2 were medium and 3 were large (Cohen, 1988). The direction of all WINFIRE B correlations was as expected with the exception of *Sending more distant unit to a fire*. One would expect this to be a tactical mistake. It makes more sense and is more economical to send a unit that is close to a fire to the fire location. Further examination revealed that sending a more distant unit to a fire correlated positively with number of trucks ($r = .24$, $p = .02$) sent to a starting fire. These data indicate that participants corrected their mistake and sent other trucks with the distant unit to the fire. Although this is an error, in terms of performance, it was regarded as successful because ultimately, the participants were still sending units to a fire. They sent other closer units after sending the distant ones first and realizing that the far units are relatively slow.

Two strategies and tactics did not correlate significantly with success in either WINFIRE version: *Sudden change in strategy* and *Sending unit on burned field*, possibly due to their frequency, i.e., below 2 in both WINFIRE simulations (see Figure 1). Additionally, *Sudden change* can be either positively or negatively related to performance depending on the change made.

WINFIRE A: Cluster Analysis

In a next step we conducted cluster analyses for WINFIRE A and WINFIRE B to determine specific subgroups in our sample according to strategies and tactics used and to determine how well the selected strategies and tactics predicted success in WINFIRE A. As for the correlations before, we also excluded 4 extreme outliers for the cluster analysis – participants 2, 16, 22, and 54, whose performance was higher than 70% of the saved area. We used the two step cluster analysis approach, because we did not know the number of clusters in advance, because it is a relatively robust technique, and because it is a novel method that addresses weaknesses of the k-means clustering method and the hierarchical clustering method (Bacher, Wenzig, & Vogler, 2004). Although two step cluster analysis is a method that can also incorporate categorical independent variables, according to the authors, it performs especially well if all variables are continuous. In our study we used eight variables, all of them being continuous.

The cluster analysis for WINFIRE A results in 2 clusters. Cluster 1 comprises 37.4% ($N = 34$) of the participants and cluster 2 comprises 62.6% ($N = 57$) of the participants. The two most influential predictors were *Tactical planning: Number of trucks sent to a starting fire* (importance = 1.00) and *Proactive strategic planning: Active distribution of units and use of*

Table 2. Planning strategies and tactics and errors and their Pearson correlations with performance in WINFIRE A and WINFIRE B.

	Performance in WINFIRE A	Performance in WINFIRE A
Planning strategies and tactics		
Proactive strategic planning: Active distribution of units and use of patrol command	.16	.31**
Effect control and flexible strategy: Sudden change in behavior pattern	.05	.12
Tactical planning: Number of helicopters sent to a starting fire	.12	.58***
Tactical planning: Number of trucks sent to a starting fire	.07	.55***
Errors		
Single-focus strategy and lack of multi-tasking: Decisions related to one fire only and not to multiple fires	.02	-.47***
Incorrect order or function of commands: Number of tactical decision-making errors	-.13	-.25*
Perceptual inaccuracy leading to inappropriate action: Number of units sent to a burned field	.04	-.15
Lack of foresight and impulsivity: Number of more distant units sent to a fire instead of nearby units	−.15	.18†

Note. † $p < .09$. * $p < .05$. ** $p < .01$. *** $p < .001$.

patrol command (importance = 0.75). The means of the 4 strategies and tactics and 4 errors in the two cluster groups are represented in Table 3. We can characterize group 1 as active, flexible, and big picture planners and decision makers. Group 2, which consists of almost two thirds of the participants, can be characterized as slow or cautious, and single-focused decision makers. Comparing the performance of cluster group 1 ($M = 49.11$, $SD = 5.03$) and group 2 ($M = 48.07$, $SD = 3.23$) using an independent samples t-test, group 1 did not perform better than group 2, $t(88) = 1.19$, $p = .24$. With the exceptions of sending a distant unit to fire and sending unit on burned field, the means confirm the hypotheses regarding correlations of strategies, tactics, and errors with performance.

Looking at the results and means of the two cluster groups one could have the impression that action leads to success. Simple activity, however, defined as total number of commands given to units during the WINFIRE simulation, did not correlate significantly with success. This correlation was $r = .12$, $p = .23$ for WINFIRE A and $r = .11$, $p = .28$ for WINFIRE B, showing further that it was the specific strategies, tactics, and errors, and not simply sending many units around, without plan or focus, that predicted success.

WINFIRE B: Cluster Analysis

Another two step cluster analyses was conducted for WINFIRE B. The cluster analysis also resulted in 2 clusters. Cluster 1 comprises 36.8% ($N = 35$) of the participants and cluster 2 comprises 63.2% ($N = 60$) of the participants. The two most influential predictors were *Single-focus strategy and lack of multi-tasking: Decisions related to one fire only and not to multiple fires* (importance = 1.00) and *Incorrect order or function of commands: Number of tactical decision-making errors* (importance = 0.37). The means of the 4 strategies and tactics and 4 errors in the two cluster groups are also represented in Table 3. Group 1, which consists of about one third of the participants, can be characterized as slow or cautious, and single-focused decision makers (similar to group 2 in WINFIRE A). We can characterize group 2 (which is now the higher performing group as we will see, and similar to group 1 in WINFIRE A) as active, flexible, and big picture planners and decision makers. Comparing the performance of cluster group 1 ($M = 54.88$, $SD = 11.72$) and group 2 ($M = 68.35$, $SD = 19.03$) using an independent samples t-test showed that group 2 performed significantly better than group 1, $t(91.84) = -4.25$, $p < .001$. With the exception of sending distant unit to fire, all the means of the other 7 variables reflect the hypotheses on which strategies and tactics and which errors should correlate with performance.

Changes between WINFIRE A and WINFIRE B

Comparing WINFIRE A and WINFIRE B, an interesting question is to what extent participants stay in the respective successful and non-successful clusters, and if participants move from the non-successful cluster to the successful cluster. Results of a Chi-square test show marginal significant changes, χ^2 ($df = 1$, $N = 89$) $= 2.65$, $p = .10$. Most of the participants who were successful in WINFIRE A were also successful in WINFIRE B ($N = 24$ out of the 33, see Table 4). Nine out of the 33 were successful in WINFIRE A and weak in WINFIRE B. Less than one third of all participants performed poorly in WINFIRE A and also in WINFIRE B ($N = 25$). What is interesting to note is the group that performed poorly in WINFIRE A, but then performed well in WINFIRE B ($N = 31$). All in all, roughly one third of all participants performed poorly in both WINFIRE simulations, one third performed well in both WINFIRE simulations, and one third improved their performance, where they first performed poorly in WINFIRE A and then got better in WINFIRE B.

Strategies, Tactics, and Errors in WINFIRE A and WINFIRE B

The means of the strategies, tactics, and errors in the first and second WINFIRE simulation are presented in Figure 1 and Table 5. The four planning strategies and tactics to the left were expected to enhance success (with the exception of "Effect control and flexible strategy"), while the four errors to the right were expected to be detrimental to success.

Table 3. Means of the tactics and errors for the 2 Custer groups in WINFIRE A and WINFIRE B (2 separate cluster analyses). The two cluster groups in both WINFIRE A and WINFIRE B differ on all 8 strategies, tactics, and errors, except the ones in the grey underlined fields.

	WINFIRE A		WINFIRE B	
	Group 1	Group 2	Group 1	Group 2
Planning strategies and tactics	37.4 %	62.6 %	36.8 %	63.2 %
Proactive strategic planning: Active distribution of units and use of patrol command	1.94	0.30	2.26	2.60
Effect control and flexible strategy: Sudden change in behavior pattern	0.62	0.16	0.49	1.27
Tactical planning: Number of helicopters sent to a starting fire	1.85	0.56	1.11	2.92
Tactical planning: Number of trucks sent to a starting fire	5.62	1.70	3.06	6.65
Errors				
Single-focus strategy and lack of multi-tasking: Decisions related to one fire only and not to multiple fires	3.85	4.28	5.80	1.15
Incorrect order or function of commands: Number of tactical decision-making errors	14.79	20.98	15.20	6.97
Perceptual inaccuracy leading to inappropriate action: Number of units sent to a burned field	0.76	0.49	0.89	0.27
Lack of foresight and impulsivity: Number of more distant units sent to a fire instead of nearby units	1.06	0.09	0.43	1.07

Table 4. Distribution of poor and strong performance according to the two clusters in WINFIRE A and WINFIRE B.

	Cluster WINFIRE B		
	"weak" performance	"strong" performance	Total
Cluster WINFIRE A			
"strong" performance	9	24	33
"weak" performance	25	31	56
Total	34	55	89

Results showed that the four successful planning strategies and tactics to the left increased in WINFIRE B compared to WINFIRE A, and the two most frequent errors to the right decreased from WINFIRE A to WINFIRE B. The most frequent category was *Number of tactical errors*, for example, sending a unit to a fire and forgetting to click the extinguish command or clicking extinguish without selecting a specific firefighting unit to receive the command.

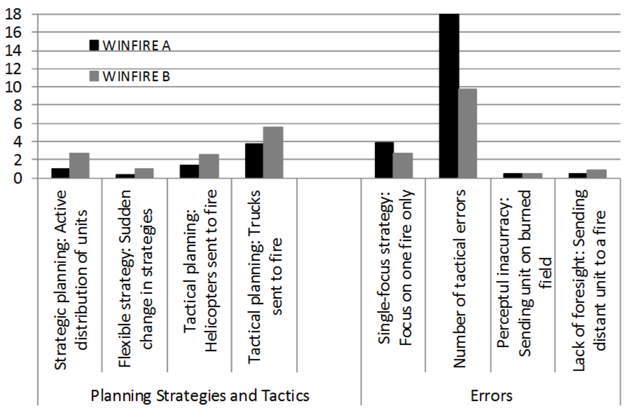

Figure 1. Means of the strategies, tactics, and errors in WINFIRE A and WINFIRE B.

Strategies, Tactics, and Errors in Different Cycles of WINFIRE A and WINFIRE B

To further investigate decision making during a dynamic task, researchers have shown that it is beneficial to scrutinize the task and look at specific parts and events (e.g., Sohn, Douglass, Chen, & Anderson, 2005). Figure 2 shows the development of successful planning strategies and tactics over time in both WINFIRE simulations. Figure 3 shows the development of errors over time in both WINFIRE simulations. We expected successful planning strategies and tactics to increase, i.e., to be higher in WINFIRE B than in WINFIRE A (Figure 2). Indeed, the three expected planning strategies and tactics increased in WINFIRE B. Comparing the frequency of planning strategies and tactics in both WINFIRE simulations (calculating repeated measures ANOVAs and reporting only the main effects for WINFIRE version), units were more actively distributed at the beginning of the simulation and received more patrolling commands, $F(1{,}188) = 30.23$, $p < .001$, $\eta_p^2 = .14$; tactical planning was more frequent, i.e., the use of helicopters, increased, $F(1{,}188) = 12.17$, $p = .001$, $\eta_p^2 = .06$, and more fire-fighting units were sent to fires, $F(1{,}188) = 11.34$, $p = .001$, $\eta_p^2 = .06$.

We expected errors to decrease, i.e., to be lower in WINFIRE B than in WINFIRE A. (Figure 3). Tactical errors decreased and focusing on only one fire also decreased comparing the first and second WINFIRE simulations. Comparing the development of errors in both WINFIRE simulations (calculating repeated measures ANOVAs and reporting only the main effect for WINFIRE version), the number of tactical errors declined, $F(1{,}188) = 31.48$, $p < .001$, $\eta_p^2 = .14$, and the focus on fighting several fires simultaneously instead of only one fire increased in WINFIRE B, $F(1{,}188) = 7.61$, $p = .006$, $\eta_p^2 = .04$.

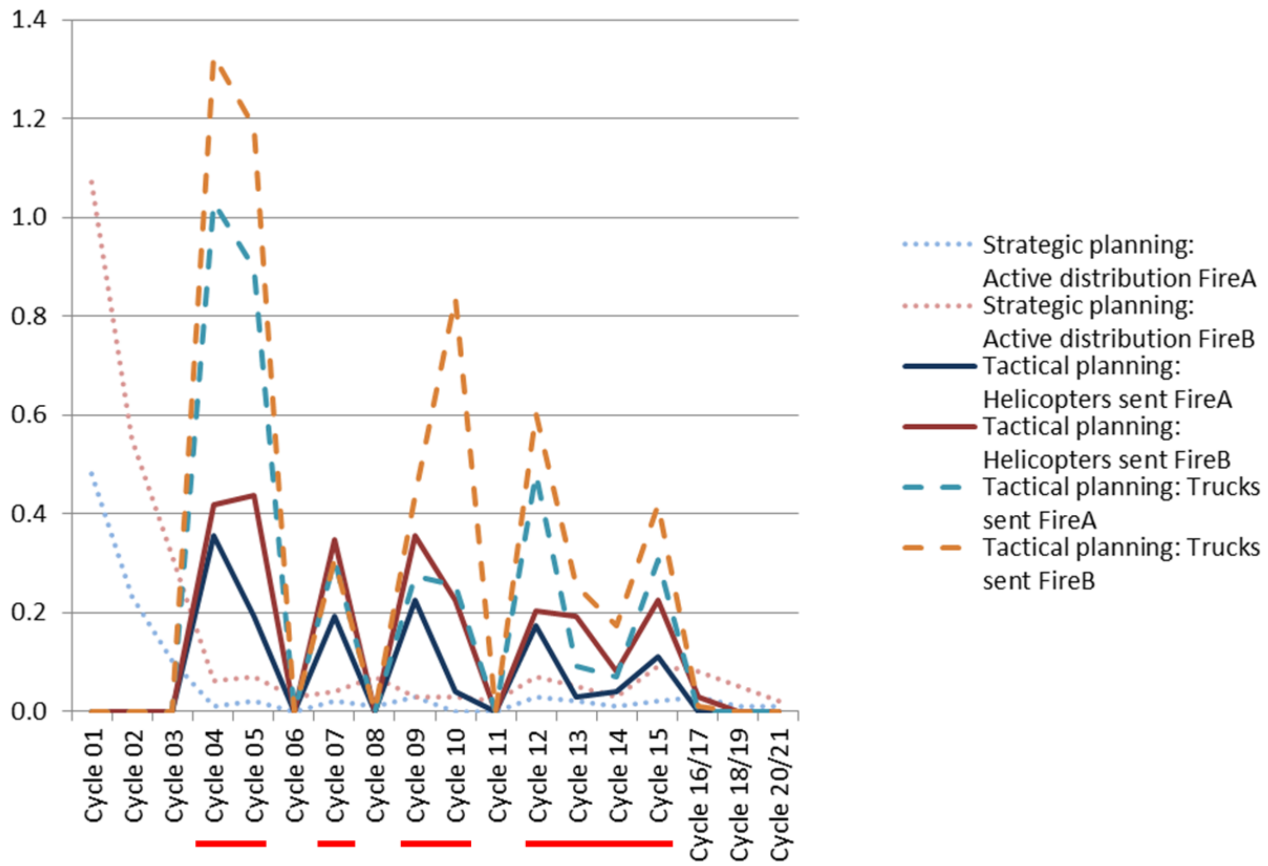

Figure 2. Means of the three most frequent successful planning strategies (solid lines) and tactics (dotted lines) in WINFIRE A (blue) and WINFIRE B (red). Cycles marked with a red bar indicate the start of new fires.

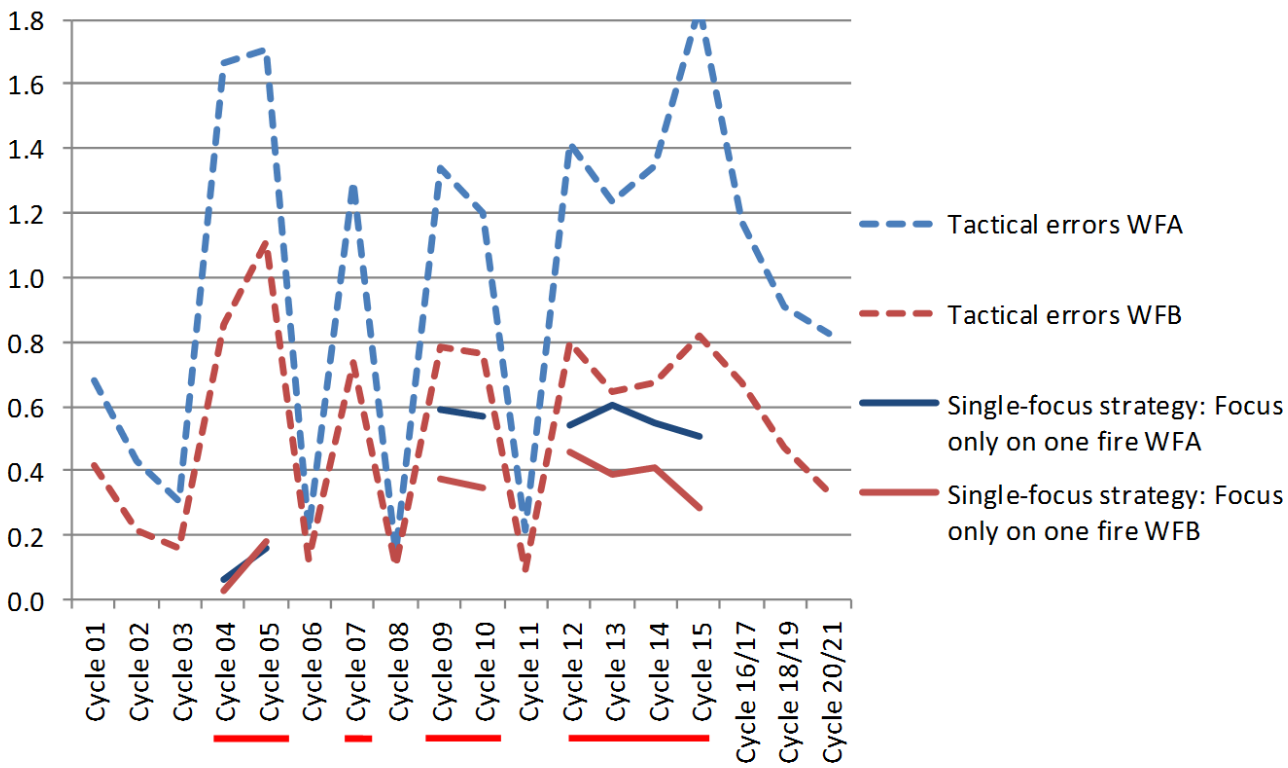

Figure 3. Means of the two most frequent errors in WINFIRE A (blue) and WINFIRE B (red). Cycles marked with a red bar indicate the start of new fires.

Discussion

The current study had three goals. First, it addressed a need for a systematic and precise investigation of strategies, tactics, and errors in DDM over time relating those to performance. Second, the potential of observation as a method in DDM was examined. Third, the applicability and validity of a microworld developed in the West was investigated in an Asian sample.

Strategies, Tactics, Errors, and Performance

Results showed interesting patterns in strategies, tactics, and errors specifically referring to planning and decision making. For example, tactical decision-making errors were quite frequent under this time pressure conditions and sudden change in strategy was rarely observed. Participants mostly stuck to what they did, and did not change their strategic approach, so that one third of all participants was consistent in their poor performance and another third was consistent in their good performance. This finding is consistent with other research highlighting mental sets, i.e., a fixed way to solve problems (e.g., Luchins & Luchins, 1959). Specifically for the third of participants who performed poorly both times, this shows the lack of strategic flexibility in adapting strategies to changes in environment (e.g., Cañas, Antolí, Fajardo, & Salmerón, 2005). The last third of our participants, the group whose performance improved, switched from a slow, cautious, and single-focused strategy to an active and flexible strategy. This confirms what Rasmussen (1990, p. 458) pointed out as a necessity, "dynamic shifting among alternative strategies is very important for skilled people as a means to resolve resource-demand conflicts."

The observed strategies, tactics, and errors correlated with performance in the expected direction in WINFIRE B and could explain variance in performance. Correlations and cluster analyses revealed that successful participants compared to unsuccessful participants showed more proactive planning, more active decisions, more multi-tasking, more flexibility, and fewer decision-making errors (i.e., single-focus strategy, impulsivity, incorrect order). It is important to highlight that decision-making activity alone did not correlate with performance. This clearly supports the notion that what is important is not the number of activities or responses to a problem per se, but the intentionality and directedness of the activity to solve the problem. Results showed the predictive validity of the postulated strategies, tactics, and errors and highlight the importance of strategies and tactics in the problem-solving process (see e.g., Anderson, 2005; Newell & Simon, 1972; Rieskamp, 2008). One can assume that the successful strategies in WINFIRE will also lead to success in other complex and dynamic problem situations. For example, multi-tasking strategy and tactical plans that lead to problem solving success were also identified in several other studies (e.g., Blackie & Maharg, 1998). We come back to this point when we discuss the applicability of our findings.

Observation as a Method in DDM

The second goal for this study was methodological, i.e., to investigate whether observation of strategies and tactics by experimenters can be a robust method in the field of DDM. Often times, researchers rely on the saved computer log files of participants to analyze their behavior (e.g., Güss & Dörner, 2011; Schoppek, 1991; Wüstenberg et al., 2014), and while doing so gives pertinent data on system and tactics, "bigger contextualized behaviors" and strategies can hardly be captured. Results of the participants' observations have shown that strategies and tactics can be identified and coded, and thus allow researchers to study DDM processes in their context as opposed to solely focusing on DDM outcome or the saved log files. Particularly for this study, inter-rater agreement was acceptable to high. Such credibility and reliability in the coding method is probably due to the time-intensive training received by experimenters.

Nowadays, although computer programs can be written to analyze these huge datasets and probably also some strategies, tactics, and errors as in our case (see also Wüstenberg et al., 2014), still many strategies, tactics, and errors refer to several variable combinations over time in a specific context shown on the computer screen, and are thereby difficult to program.

Applicability of WINFIRE in an Asian Sample

Most studies involving WINFIRE and similar simulations have been conducted in Australia, Germany, Sweden or the United States. To increase generalizability of research findings and to test the applicability of WINFIRE, we had a relatively big sample of students in the Philippines work on the simulation. One indicator of the applicability of WINFIRE in this sample is that strategies tactics, and errors correlated with performance in the expected direction. It was also possible to distinguish successful from unsuccessful participants in WINFIRE on the basis of their patterns in strategies, tactics, and errors. In WINFIRE A, the most important predictors of the cluster analysis were *Tactical planning* and *Proactive strategic planning*. In WINFIRE B, the most important predictors of the cluster analysis were *Multi-tasking* and *Fewer tactical decision-making errors*. In this Asian sample, two kinds of planners and decision makers stand out, those who are active, flexible, and who see the bigger picture, which therefore perform successfully, and another group who is slow, more cautious, and are fixated on a single focus, which then performs poorly.

Another indicator of the simulation's applicability is the survey data results. All of the objective criteria of WINFIRE are confirmed in the participants' subjective assessment of the simulation as describing a realistic situation, where small actions or a single action had great impact and consequences, and where many factors came together and influenced each other leading to success or failure.

Limitations and Recommendations for Future Research

Although the current study showed high inter-rater-reliability of strategy and tactic observation, it could well be that other operationalizations of strategies in other simulations could be more prone to inter-rater bias. Researchers working with microworlds can rely on observation, and as we did here, concurrently with saved log files of participants' decisions and saved changes in the system to analyze strategies (see e.g., Frensch & Funke, 1995; Gonzalez, 2005; Güss, 2011) or processes (e.g., Koop & Johnson, 2011), and performance. Although saved data might not always allow for investigations of more complex behavior patterns or strategies, they are certainly highly accurate and reliable and allow for different kinds of process analyses (see Baker & Delacruz, 2008).

We would like to acknowledge that several tactics that we have operationalized in this study could potentially be derived from log files provided the output files are programmed accordingly. Correlations among observations and log file data would add to the validity of future studies.

Another limitation related to the method of observation is that strategies, tactics, and errors related to one decision-making step are easier to observe compared to strategies, tactics, and errors related to another decision-making step. For example, goal definition can be inferred in WINFIRE by observing participants commanding units to go to certain locations. What cannot be observed, however, is how the participant came up with this particular goal and whether the participant had alternative goals in mind or goal conflicts. It is similarly hard to observe behavior related to information collection. For instance, eye-tracking methodology would be required to collect data on the specific aspects of the computer screen a participant focused upon, i.e., which city, which fire trucks, and which fires. In addition to the saved decision logs, one could also use thinking-aloud protocols to gain information about all decision-making steps (Güss et al., 2010). Thus, our focus on the planning and decision-making steps is partly due to the constraints of the observation methodology we chose.

Relevance and Applicability of Findings to Training and Other Domains

The primary focus of WINFIRE is not to simulate accurately all the details of fires and the physical environment, but to observe decision making under information overload and extreme time pressure (Omodei et al., 2005, using the Firechief simulation, argued the same). Thus, one can expect that strategies, tactics, and errors identified for the WINFIRE context can also be found and observed in a range of other time-pressured dynamic situations, such as military or plant operations: Tactical planning and being prepared for an emergency; having flexible and alternative plans of action in case the current approach is not successful; decisive and quick actions adjusted to the situational demands; multitasking and dealing with different problems at the same time; avoiding slips and lapses; information collection and perception of key problems. These DDM strategies represent domain general competencies (see also Greiff & Funke, 2009; Güss, 2010).

Especially because the WINFIRE simulation has been used as a powerful training instrument in management (Motowidlo, Dunnette, & Carter, 1990), in medicine and aviation (Thomas, 2009), and in team coordination among fire emergency responders (Toups, Kerne, Hamilton, & Shahzad, 2011), the findings we gathered specifically in identifying kinds of decision makers and planners that lead to success or failure may provide more information and insight in training. Specifically the findings show that successful performance is related to increasing proactive strategic planning, increasing tactics for carrying out decisions and estimating resources, and decreasing errors such as fixating on a single focus when there are several other demands as well.

Microworlds and virtual environments allow decisions that result in catastrophic consequences, yet only virtual catastrophic consequences (e.g., O'Neil & Perez, 2008). Future training programs, even using microworlds, could focus, for example, on strategy selection and application, and show how certain strategies and tactics can enhance success and other strategies and tactics can lead to decreased success (e.g., Gonzalez, 2005; Vollmeyer, Burns, & Holyoak, 1996), taking into consideration the unique demands of the simulated situations. This would lead to a strategy repertoire, allowing people to better perceive the circumstances to which a strategy best applies (see Schunn, McGregor, & Saner, 2005). Learning to expand one's planning and decision making strategies to be more proactive, to use more foresight, to focus on several things at the same time, and to keep from being impulsive, in a virtual environment will hopefully lead to the use of these learned strategies and tactics when confronted with real-life problem situations.

Conclusion

To conclude, the current study addressed the need to investigate strategies, tactics, and errors systematically in DDM using a relatively big sample size. Observing participants while working on the WINFIRE simulation was a reliable method which allowed for the assessment of strategies, tactics, and errors in their specific context. Successful participants showed more proactive planning, more active decisions, more multi-tasking, more flexibility in their strategy, and fewer decision-making errors. The decrease in errors and increase in successful planning strategies and tactics showed that participants' proficiency while solving problems increased. And in this assessment, it was gleaned that some strategies and tactics, such as multitasking and tactical planning that are intentional and directed, lead to greater success, more than number of responses or activity alone. The study also confirms that microworlds can be used to study DDM strategies in a non-Western culture, in a Filipino sample, where utilizing this can be robust for scientific endeavors.

Acknowledgements: This research was supported by a National Science Foundation Grant No. 0349997 to the first author and by a Humboldt Fellowship for Experienced Researchers also to the first author. We are thankful to those who helped us in data gathering and data analysis when we were in the Philippines: Lorraine Matulac, Oliver Pangan, Peter Tuason, Michelle Westerwelle, and to our colleagues in the Philippines for all their support and hard work. We also would like to thank Shannon McLeish for editing a previous version of this manuscript.

Declaration of conflicting interests: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

Author contributions: CDG designed the study. CDG and MTT conducted the study, collected the data, and coded the data together with the colleagues mentioned in the acknowledgements. LVO was primarily responsible for organizing the study and data collection in the Philippines. CDG conducted the data analysis. CDG and MTT wrote the manuscript.

Supplementary material: Supplementary material available online.

Handling editor: Andreas Fischer

Citation: Güss, C. D., Tuason, M. T., & Orduña, L. V. (2015). Strategies, tactics, and errors in dynamic decision making in an Asian sample. *Journal of Dynamic Decision Making, 1*, 3. doi:10.11588/jddm.2015.1.13131

Received: 01 April 2014
Accepted: 16 October 2015
Published: 10 November 2015

References

Anderson, J. R. (2005). *Cognitive psychology and its implications* (6th ed.). New York, NY: Worth Publishers.

Bacher, J., Wenzig, K., & Vogler, M. (2004). *SPSS Two step Cluster - A first evaluation. Social Science Open Access Repository.* Retrieved from: http://nbn-resolving.de/urn:nbn:de:0168-ssoar-327153

Baker, E. L., & Delacruz, G. C. (2008). A framework for the assessment of learning games. In H. F. O'Neil & R. S. Perez (Eds.), *Computer games and team and individual learning* (pp. 21-37). Oxford, UK: Elsevier.

Blackie, J., & Maharg, P. (1998). *The Delict Game.* Retrieved from: http://www.academia.edu/1960624/The_Delict_Game

Bonissone, P. P., Dutta, S., & Wood, N. C. (1994). Merging strategic and tactical planning in dynamic and uncertain environments. *IEEE Transactions on Systems, Man, & Cybernetics, 24,* 841-862. doi:10.1109/CAIA.1992.200016

Brehmer, B. (2004). Some reflections on microworld research. In S. G. Schiflett, L. R. Elliot, E. Salas, & M. D. Coovert (Eds.), (2004). *Scaled worlds: Development, validation, and applications* (pp. 22-36). Hants, England: Ashgate.

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior, 9,* 171-184. doi:10.1016/0747-5632(93)90005-D

Broadbent, D. E. (1977). Levels, hierarchies, and the locus of control. *Quarterly Journal of Experimental Psychology, 29,* 181–201. doi:10.1080/14640747708400596

Cañas, J. J., Antolí, A., Fajardo, I., & Salmerón, L. (2005). Cognitive inflexibility and the development and use of strategies for solving complex dynamic problems: Effects of different types of training. *Theoretical Issues in Ergonomics Science, 6,* 95-108. doi:10.1080/14639220512331311599

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, R. I., & Woods, D. D. (1994). Operating at the sharp end: The complexity of human error. In M. S. Bogner (Ed.), *Human error in medicine* (pp. 255-310). Hillsdale, NJ: Erlbaum.

Dörner, D. (1980). On the difficulty people have in dealing with complexity. *Simulation & Games, 11*, 87–106.

Dörner, D. (1996). *The logic of failure.* New York, NY: Holt.

Dörner, D. (1999). *Bauplan für eine Seele* [Blueprint for a soul]. Reinbek, Germany: Rowohlt.

Dörner, D., & Güss, C. D. (2013). PSI: A computational architecture of cognition, motivation, and emotion. *Review of General Psychology, 17,* 297-317. doi:10.1037/a0032947

Dörner, D., & Pfeifer, E. (1991). Strategisches Denken und Stress [Strategic thinking and stress]. *Zeitschrift für Psychologie, Supplement, 11,* 71-83.

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *The Journal of Problem Solving, 4,* 19-42. doi:10.7771/1932-6246.1118

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 288-299. doi:10.1037/0096-1523.1.3.288

Flach, J. M. (1999). Beyond error: The language of coordination and stability. In P. A. Hancock (Ed.), *Human performance and ergonomics* (pp. 109-128). San Diego, CA: Academic Press.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions.* New York, NY: Wiley.

Frensch, P., & Funke, J. (Eds.). (1995). *Complex problem solving: The European perspective.* Hillsdale, NJ: Erlbaum.

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Erlbaum.

Funke, J. (2003). *Problemlösendes Denken* [Problem solving thinking]. Stuttgart: Kohlhammer.

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing, 11,* 133-142. doi:10.1007/s10339-009-0345-0

Gerdes, J., Dörner, D. & Pfeiffer, E. (1993). *Interaktive Computersimulation "Winfire"* [The interactive computer simulation "Winfire"]. Otto-Friedrich-Universität Bamberg, Germany: Lehrstuhl Psychologie II.

Gonzalez, C. (2005). Decision support for real-time, dynamic decision-making tasks. *Organizational Behavior and Human Decision Processes, 96,* 142-154. doi:10.1016/j.obhdp.2004.11.002

Granlund, R. (2003). Monitoring experiences from command and control research with the C3Fire microworld. *Cognition, Technology, and Work, 5,* 183-190. doi:10.1007/s10111-003-0129-8

Greiff, S., & Funke, J. (2009). Measuring complex problem solving: The MicroDYN approach. In F. Scheuermann & J. Björnsson (Eds.), The transition to computer-based assessment. *New approaches to skills assessment and implications for large-scale testing* (pp. 157-163). Luxemburg: Office for Official Publications of the European Communities.

Güss, C. D. (2000). *Planen und Kultur?* [Planning and culture?]. Lengerich, Germany: Pabst.

Güss, C. D. (2011). Fire and ice: Testing a model on cultural values and complex problem solving. *Journal of Cross-Cultural Psychology, 42,* 1279-1298. doi:10.1177/0022022110383320

Güss, C. D., & Dörner, D. (2011). Cultural differences in dynamic decision-making strategies in a non-linear, time-delayed task. *Cognitive Systems Research, 12(3-4),* 365-376. doi:10.1016/j.cogsys.2010.12.003

Güss, C. D., Glencross, E., Tuason, M. T., Summerlin, L., & Richard, F. D. (2004). Task complexity and difficulty in two computer-simulated problems: Cross-cultural similarities and differences. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 511-516). Mahwah, NJ: Erlbaum.

Güss, C. D., Tuason, M. T., & Gerhard, C. (2010). Cross-national comparisons of complex problem-solving strategies in two microworlds. *Cognitive Science, 34,* 489–520. doi:10.1111/j.1551-6709.2009.01087.x

Hayes-Roth, B. & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science, 3,* 275-310. doi:10.1207/s15516709cog0304_1

Hunt, E., & Joslyn, S. (2000). A functional task analysis of time-pressured decision making. In J. M. Schraagen, S. F. Chipman, & V. Shalin (Eds.), *Cognitive task analysis* (pp. 119-132). Mahwah, NJ: Erlbaum.

Johansson, B. J. E., Trnka, J., Granlund, R., & Götmar, A. (2010). The effect of a geographical information system on performance and communication of a command and control organization. *International Journal of Human–Computer Interaction, 26,* 228-246. doi:10.1080/10447310903498981

Klein, G. (1999). Applied decision making. In P. A. Hancock (Ed.), *Human performance and ergonomics* (pp. 87-107). San Diego, CA: Academic Press.

Koop, G. J., & Johnson, J. G. (2011). Response dynamics: A new window on the decision process. *Judgment and Decision Making, 6*(8), 750–758. Retrieved from: http://journal.sjdm.org/11/m29/m29.html

Lipshitz, R. (2005). There is more to seeing than what meets the eyeball: The art and science of observation. In H. Montgomery, R. Lipshity, & B. Brehmer (Eds.), *How professionals make decisions* (pp. 365-378). Mahwah, NJ: Erlbaum.

Luchins, A. S., & Luchins, E. H. (1959). *Rigidity of behavior: A variational approach to the effect of Einstellung.* Eugene, OR: University of Oregon Books.

Montgomery, H., Lipshitz, R., & Brehmer, B. (2005). Introduction: From the first to the fifth volume of naturalistic decision-making research. In H. Montgomery, R. Lipshitz, & B. Brehmer (Eds.), *How professionals make decisions* (pp. 1-11). Mahwah, NJ: Erlbaum.

Motowidlo, S. J., Dunnette, M. D., & Carter, G.W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75,* 640-647. doi:10.1037/0021-9010.75.6.640

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Omodei, M. M., McLennan, J., Elliott, G. C., Wearing, A. J., & Clancy, J. M. (2005). "More is better?": A bias toward overuse of resources in naturalistic decision-making settings. In H. Montgomery, R. Lipshitz, & B. Brehmer (Eds.), *How professionals make decisions* (pp. 29-36). Mahwah, NJ: Erlbaum.

O'Neil, H. F., & Perez, R. S. (Eds.) (2008). *Computer games and team and individual learning.* Oxford, UK: Elsevier.

Putz-Osterloh, W., & Lemme, M. (1987). Knowledge and its intelligent application to problem solving. *German Journal of Psychology, 11,* 286-303.

Qudrat-Ullah (2008). Future directions on complex decision making: Using modeling and simulation decision support. In H. Qudrat-Ullah, J. M. Spector, & P. I. Davidsen (Eds.), *Complex decision making* (pp. 323-337). Cambridge, MA: Springer. doi:10.1007/978-3-540-73665-3_16

Quiroga, M. A., Martínez-Molina, A., Lozano, J. H., & Santacreu, J. (2011). Reflection-impulsivity assessed through performance differences in a computerized spatial task. *Journal of Individual Differences, 32,* 85–93. doi:10.1027/1614-0001/a000038

Ramnarayan, S., Strohschneider, S., & Schaub, H. (1997). Trappings of expertise and the pursuit of failure. *Simulation & Gaming, 28,* 28-43. doi:10.1177/1046878197281004

Rasmussen, J. (1990). Human error and the problem of causality in analysis of accidents. *Philosophical Transactions of the Royal Society of London, 12,* 449-460. doi:10.1098/rstb.1990.0088

Reason, J. (1990). *Human error.* New York, NY: Cambridge University.

Renner, K. H., & Beversdorf, D. Q. (2010). Effects of naturalistic stressors on cognitive flexibility and working memory task performance. *Neurocase, 16,* 293–300. doi:10.1080/13554790903463601

Rieskamp, J. (2008). The importance of learning when making inferences. *Judgment and Decision Making, 3,* 261-277. Retrieved from: http://journal.sjdm.org/bn6/bn6.html

Rigas, G., & Brehmer, B. (1999). Mental processes in intelligence tests and dynamic decision making tasks. In P. Juslin, & H. Montgomery (Eds.), *Judgment and decision making: Neo-Brunswikian and process-tracing approaches* (pp. 45-66). Mahwah, NJ: Erlbaum.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review, 115,* 101–130. doi:10.1037/0033-295X.115.1.101

Schaub, H. (2001). *Persönlichkeit und Problemlösen: Persönlichkeitsfaktoren als Parameter eines informationsverarbeitenden Systems* [Personality and problem solving. Personality characteristics as parameters of an information processing system]. Weinheim, Germany: Beltz.

Schmid, U., Ragni, M., Gonzalez, C., & Funke, J. (2011). The challenge of complexity for cognitive systems. *Cognitive Systems Research, 12,* 211-218. doi:10.1016/j.cogsys.2010.12.007

Schoppek, W. (1991). Game and reality. Reliability and validity of behavior-patterns in complex situations. *Sprache und Kognition, 10*, 15-27.

Schoppek, W., & Putz-Osterloh, W. (2003). Individuelle Unterschiede und die Bearbeitung komplexer Probleme [Individual differences and complex problem solving]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 24,* 163-173. doi:10.1024/0170-1789.24.3.163

Schunn, C. D., McGregor, M. U., & Saner, L. D. (2005). Expertise in ill-defined problem-solving domains as effective strategy use. *Memory & Cognition, 33,* 1377-1387. doi:10.3758/BF03193370

Sohn, M.-H., Douglass, S. A., Chen, M.-C., & Anderson, J. R. (2005). Characteristics of fluent skills in a complex, dynamic problem-solving task. *Human Factors, 47,* 742-752. doi:10.1518/001872005775570943

Spector, J. M. (2008). Expertise and dynamic tasks. In H. Qudrat-Ullah, J. M. Spector, & P. I. Davidsen (Eds.), *Complex decision making* (pp. 25-40). Cambridge, MA: Springer. doi:10.1007/978-3-540-73665-3_2

Strohschneider, S. (1991). Problemlösen und Intelligenz: Über die Effekte der Konkretisierung komplexer Probleme [Problem solving and intelligence: On the effects of concretizing complex problems]. *Diagnostica, 37,* 353-371.

Strohschneider, S., & Güss, D. (1999). The fate of the Moros: A cross-cultural exploration of strategies in complex and dynamic decision making. *International Journal of Psychology, 34,* 235-252. doi:10.1080/002075999399873

Thomas, M. J. W. (2009). *Integrating low-fidelity desktop scenarios into the high-fidelity simulation curriculum in medicine and aviation.* Retrieved from: http://www.unisanet.unisa.edu.au/staff/MatthewThomas/Paper/Thomas_DesktopScenarios.pdf

Toups, Z. O., Kerne, A., Hamilton, W. A., & Shahzad, N. (2011). Zero-fidelity simulation of fire emergency response: Improving team coordination learning. *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (CHI '11) (pp. 1959-1968). New York, NY: CHI. doi:10.1145/1978942.1979226

Vollmeyer R., Burns B. D., & Holyoak K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science, 20,* 75-100. doi:10.1207/s15516709cog2001_3

Wenke, D., Frensch, P. A., & Funke, J. (2005). Complex problem solving and intelligence: Empirical relation and causal direction. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence* (pp. 160–187). New York, NY: Cambridge University Press.

Wüstenberg, S., Greiff, S., Molnár, G., & Funke, J. (2014). Cross-national gender differences in complex problem solving and their determinants. *Learning and Individual Differences, 29,* 18-29. doi:10.1016/j.lindif.2013.10.006