Serious gaming for adaptive decision making of military personnel

Yelim Mun[1], Esther Oprins[2], Karel van den Bosch[2], Anja van der Hulst[2], Jan Maarten Schraagen[1&2]
[1]University of Twente, The Netherlands; [2]The Netherlands Organization for Applied Scientific Research TNO, The Netherlands

The importance of improving adaptive decision making for the military is ever increasing, particularly in operational environments that are unfamiliar, complex, and constantly changing. This paper presents the development and testing of a serious game for training military officers in adaptive decision making. Participants were to detect rule changes in the game world, and to adjust their decisions in accordance with these changes. In an explorative study, the effectiveness of the game was tested by using in-game and out-game measures. The findings on the in-game measure suggest that the game helps participants to detect rule changes and to adapt their decision making. Despite this effect, participants' cognitive flexibility did not increase based on the findings on the out-game measures. Discussions, future directions, and training implications for the Defense organization are described.

INTRODUCTION

Making effective decisions is challenging, especially for military personnel in operational environments that are new, changing, uncertain, and unexpected. Being able to quickly assess, and to adapt to such situations is a prerequisite for military personnel (Dandeker, 2006). Despite the importance of adaptive decision making, few studies have focused on training adaptability that is directly applicable in military training situations. This paper proposes a low-cost training application, a Serious Game (SG), as a means to train adaptive decision making of military personnel.

Adaptability and cognitive flexibility

Adaptability refers to the ability to adjust to changing, new, unexpected, and unpredictable situations (Pulakos et al., 2000). Pulakos et al. (2000) argued that individuals' adaptability is largely influenced by the nature of situations. They distinguished eight different situation-specific dimensions of adaptability. These are: (1) creative problem solving, (2) dealing with uncertain and changing circumstances, (3) learning new skills, knowledge and procedure, (4) interpersonal adaptability, (5) cultural adaptability, (6) physical adaptability, (7) handling emergency situations, and (8) coping with stress. Despite all eight dimensions of adaptability being relevant to successful military performance, this paper focuses on coping with uncertain and changing circumstances, because uncertainty management is critical for decision making in time-pressured, high risk, and complex environments (Lipshitz et al., 2001).

Adaptability is shaped by underlying dispositional properties. For example, by cognitive flexibility (CF): the ability to quickly reshape one's knowledge and to respond adaptively to changing, new, and uncertain situations (Spiro et al., 1988; Good, 2014). Cañas et al. (2006) found that when situations change, individuals with high CF quickly detect the changes and assess the new situation. Then they adapt their thinking or strategies suitable for the new situation. Finally, they perform or behave in an adaptive manner, thus maintaining effective performance. Instruments have been developed to measure an individual's CF, such as how fast individuals can learn and unlearn rules after sudden changes of the learned rules (Grant & Berg, 1948).

Training for adaptive decision making

A training, especially for military, in adaptive decision making requires components such as naturalistic and complex decision making, uncertainty, ambiguity, high stakes, high risks, and time pressure (Klein et al., 2003). Exposure to context-specific, non-routine situations requiring adaptive behavior is an important instructional strategy to encourage learning of CF (Spiro et al., 1988).

Serious Games (SG) may provide such a context for training. SGs are games for other purposes than just entertainment, with a strong focus on learning professional skills (Ritterfeld, Cody & Vorderer, 2009). SGs can provide a series of dynamic, uncertain, and new environments that are suitable for training adaptive decision making and in situated, specific contexts (Gee, 2005).

Existing SGs with the purpose of improving learners' adaptability contain specific didactical features that stimulate learners to perform in an adaptive fashion, such as diversity training (e.g., Brunstein & Gonzalez, 2010). One way to stimulate adaptive behavior in learners is to confront them with rule changes in their world. A rule change can be implemented in a contextually rich SG where learners have to make decisions. They then have to detect and appreciate the rule change, and have to adapt their decision making according to the modified rules (Lepine, Colquitt & Erez, 2000). The literature on training and transfer provides evidence for this notion. Although introducing a rule change often results in a performance decrease during training, it in the end generally pays off through better learning (Schmidt & Bjork, 1992). Initially learners become confused by the rule change, but after a period of practice, learners tend to apply the learned skills into a different, yet related task as the rule change may facilitate adaptation.

In line with the abovementioned literature, we designed and developed a PC-based, complex decision making SG that aims to train adaptive decision making of military personnel. In the game, players have to make a series of complex decisions, as the game requires constant information processing, situational assessment, and dealing with missing information. Two fictional scenarios with rich narratives were created. Through the game play, players naturally learn the underlying rules of the fictional world by making a series of decisions and receiving feedback. Then these rules change. Players have to detect the changes and adapt their decision making for successful play. We intentionally used fictional scenarios to make sure that familiarity was not a factor in the study. The scenarios are representative for military decision making, referring to uncertain, unexpected, and changing environments.

The present study

The purpose of this exploratory study is to test the intervention whether a SG-based training that confronts the participants with rule-changes stimulates adaptive decision making, and whether any effects of training extend to performance on CF tasks.

We expect that the group that plays the game with the rule change will demonstrate superior performance on out-game tasks, compared to the group that plays the game without the rule change. We also expect that their in-game performance will be lower than for the control group. This is because the rule change will confront them with problems, requiring the execution of adaptive strategies. This will decrease performance on the short term (i.e. the in-game measurements), to the benefit of learning (Schmidt & Bjork, 1992).

METHOD

Game design

The game consists of three main components (see Figure 1). These are the Learning Phase (LP), Consolidation Phase (CP), and Test Phase (TP). A briefing is given before and after the LP with information about the mission and the available resources. Two fictional scenarios with the context of a robot threat and nano weapon were created. Both scenarios contain 21 cases. A case is a snapshot situation in the progressing scenario where players have to decide how to act. Each case has four given actions (decision options). Players must choose two of the four given options. This helps players to discover the rules of the fictional world in an interactive manner and to feel control of the game. After each case, players receive feedback on the chosen actions. During the LP, players receive nine cases to discover three rules (i.e., abilities of the enemy robots). After the LP, players receive open questions to test whether they have mastered the rules. During the CP, players apply the mastered rules. Then, an event is introduced that causes a change in the rules of the fictional world. The event (i.e., solar storm) is communicated to players during the second briefing, but its resulting effect (i.e., changing functions and weaknesses of enemy robots) upon the rules is not. Players have to discover the changed rules during the TP by assessing the situation and processing the feedback. After completion of the TP, players answer open questions, and then the game ends.

To examine the effect of the rule change on adaptive decision making, we constructed two versions of S1 and S2 for this study. Both versions are identical until the end of the CP. During the TP, players who play version one will experience the game with changed rules, whereas players of version two follow the original rules. The open questions for both versions of the scenarios are almost identical.

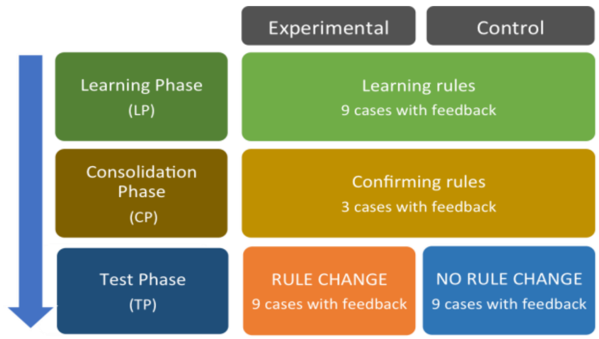

Figure 1. The structure of the SG.

Participants

A total of 33 (29 males, 4 females) Army officers, students at the Dutch Major's school, volunteered to participate in this study. Their age varied from 33 to 53 years (*M=41.2*, *sd=5.9*).

Experimental design and procedure

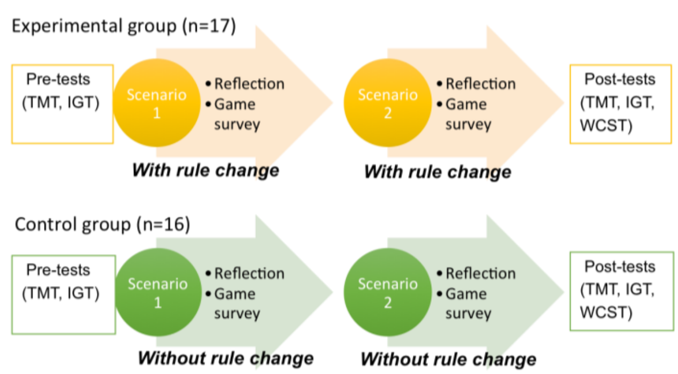

Figure 2. The design and procedure of the experiment.

Figure 2 shows the design and flow of this study. The experimental group received the SG training with the rule change (rules change in the TP), and the control group received the SG training without the rule change. The dependent measures were: scores on out-game measurements (three tasks measuring CF); and those of in-game measurements (game scores of the CP and TP in both scenarios). The scores on out-game tasks are used to test whether the SG training with rule change increased participants' cognitive flexibility. The scores on in-game tasks are used to test whether the learning objectives of the game, to adapt decision making if the rules change, have been achieved. Participants' LP game score was excluded from the in-game measures to avoid possible confounding variables such as intelligence. In addition, we examined how players experienced each scenario in terms of game difficulty, engagement, motivation, and concentration as these factors could be moderating variables that might influence learning and the effectiveness of the SG training.

At the Major's school, participants were randomly assigned to either the experimental or the control group, each to its own room with a facilitator present throughout the experiment. Participants were asked not to exchange information during breaks with peers from the other condition. Prior to the experiment, all participants filled out the informed consent form. A short introduction was given about the study, the procedure, and how to execute the CF tasks and the game. The whole session lasted five hours including breaks and waiting time. Due to the overlapping duty schedules and technical problems encountered during the experiment, the number of participants on the SG play and CF tasks differed.

Measures

Out-game measures. Three computerized CF tasks that contain implicit rule-change were used as out-game measures. They are the Trail Making Task (TMT), the Iowa Gambling Task (IGT), and the Wisconsin Card Sorting Task (WCST). The TMT (e.g., Lezak, 1995) consists of two parts. In part A, participants were asked to draw lines to sequentially connect numbers and letters that are serially distributed. This is repeated in part B but with a mixed distribution. Therefore, only part B completion time was used for the performance outcome in this study. Short time completion indicates high performance. In the IGT (e.g., Bechara et al., 1994), participants had to maximize their profit by choosing cards from four decks with varying risks and gains. The performance is measured as the number of chosen unprofitable cards deducted from the number of chosen profitable cards, and the remaining profit after 100 trials. In this study, we excluded the remaining profit as a performance measure due to a confounding variable, pure coincidence of selecting profitable cards. High values indicate high performance. The WCST (e.g., Grant & Berg, 1948) requires participants to sort cards according to changing rules of color, shape, and number. The number of the total trials, errors and perseverative errors were measured. Low numbers of the total trials, errors, and perseverative errors indicate high performance. From the three CF tasks, the way feedback is provided on changed rules is the most apparent in the WCST; therefore, it was only used as a posttest measure to avoid a priming effect on the game score.

In-game measures. Each phase of the SG contains three test cases that measure if participants made normatively correct decisions using the three learned rules. Each test case consists of two correct and two incorrect actions (scoring range between 0 to 2). The total game score of each phase is measured as the sum of correct actions (decision making performance), ranging from zero to six for LP, CP, and TP. These scores were used as in-game measures. For the experimental group, the game score in the TP indicates the adaptive decision making performance because it is the phase where the changed rules are applied. The in-game measures of S1 and S2 were analyzed separately as the difficulty and the contents of both scenarios are different.

Open questions. To test their knowledge of the rules, participants had to answer open questions for both scenarios. These were administered before the CP and after the TP. The first author evaluated the answers.

Game survey. To assess how players experienced the game, we asked all participants to fill in a survey upon the completion of each scenario. The survey consisted of five points-rating scale on difficulty, engagement, motivation, and concentration of the game play. For example, participants could rate the difficulty of the game play ranging from very easy (--) to very difficult (++).

RESULTS

Out-game measures

To investigate the effects of training, repeated measures of mixed ANOVA were conducted with condition (exp vs. con) as the between-subject factor, and scores on the TMT (pre vs. post), the IGT (pre vs. post), and the WCST (trials vs. errors vs. perseverative errors) as dependent variables. Table 1 shows scores on the three CF tasks. Contrary to our hypothesis that effects of conditions should be present of all CF tasks, significant differences were found only on the TMT performance ($F(1, 29) = 3.59$, $p = .04$), meaning the experimental group performed better than the control group on the post-TMT. No significant effects of conditions were found on the IGT and the WCST performance (both *p's* >.05).

Table 1. Descriptive statistics of scores on CF tasks.

Task	Measure	Exp (n=16) Mean (sd)	Con (n=16) Mean (sd)
TMT	Pre-Time B	58.87s (26.62)	47.96s (90.04)
	Post-Time B	45.95s (17.84)	52.46s (28.37)
IGT	Pre-Score	-23.00 (24.43)	-4.40 (48.46)
	Post-Score	10.00 (24.88)	15.10 (36.31)
WCST	Total trials	100.25 (5.20)	105.88 (4.49)
	Errors	27.50 (17.91)	34.81 (23.05)
	Perseverative errors	6.94 (2.08)	8.44 (4.03)

In-game measures

To investigate the effect of rule change on participants' decision making, we analyzed the data on the SG scores by conducting a mixed 2 (condition: exp vs. con) x 2 (phase: CP vs. TP) ANOVA as within-subject factors. Figure 3 shows the participants' SG scores for both scenarios. As predicted, the performance of the experimental group on S1 during the CP and the TP was significantly lower than that of the control group ($F(1, 24) = 3.26$, $p = .04$), suggesting that the rule change in the game stimulated adaptive decision making for the experimental group. No difference was found between the CP and TP for the control group during S1, meaning they continued to make decisions following the original rules. Analysis of S2 also showed the significant effect of the condition on the CP and the TP game scores, again supporting our hypothesis. The difference was significant for the experimental condition ($F(1, 27) = 42.46$, $p < .001$), meaning that the rule change during the TP required the experimental groups' decision making to be adaptive.

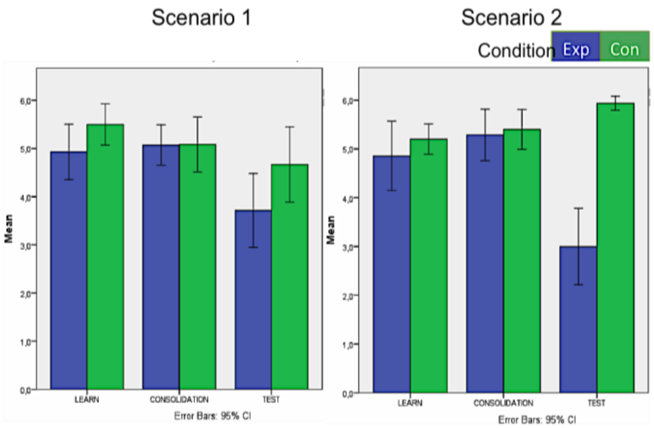

Figure 3. SG scores of the three phases, for each scenario.

Open questions

To investigate whether participants from the experimental group were able to detect the rule change during the TP, we analyzed their answers for both scenarios, after the rules were changed. For S1, 74 % of the participants were able to detect and answer correctly on the three changed rules. For S2, the average correct rate on three changed rule was 64 %.

Game survey

Next, we examined how the participants experienced the SG in terms of game difficulty, engagement, motivation, and concentration (See Figure 4). Participants' concentration and motivation were highly rated for both groups on both scenarios. Also, both group rated motivation and concentration higher in S1 than that of S2. The control group was more engaged in both S1 and S2 than the experimental group. The experienced difficulty varied per group and scenario. In general, S2 was viewed as more difficult than S1 for both groups.

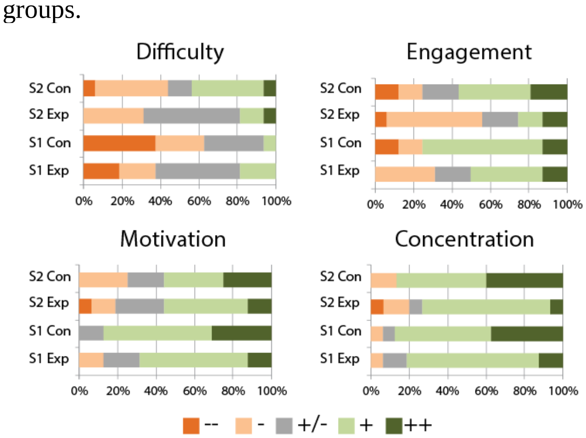

Figure 4. Participants' assessment on SG play.

DISCUSSION

The current study explored the effectiveness of the SG training for adaptive decision making using military participants. Regarding the performance on the out-game measures, we found no evidence that a SG-based training in a rule changing scenario increases cognitive flexibility, a cognitive part of adaptability (Good, 2014). This finding is in conflict with previous studies (Brunstein & Gonzalez, 2010; Good, 2014). Although statistically significant differences were found between conditions on the TMT, we nevertheless reject our hypothesis because the difference was caused by a decrease in performance of the control group from pre- to posttest, rather than caused by an increase of the experimental group. The absence of a main effect on out-game measures could be explained by the different nature of the CF tasks that were used in this study. The WCST and the IGT resemble the SG in terms of how the rule change is administered, whereas the TMT requires other skills such as motor skills (Lezak, 1995). From the debriefing log, participants in the experimental group expressed that they could find a relation between the game and the WCST but not with the TMT. Also, the short training time, and the lack of emphasis on the learning goal of the intervention during the training likely did not yield to increasing participants' cognitive flexibility, as they are underlined factors to facilitate learning and transfer (Baldwin & Ford, 1988). Furthermore, the low numbers of participants and hence low power as well as participants' lack of motivation and concentration to repeat the IGT and the TMT could explain the absence of an effect. To increase the effect of the game training with rule change on individuals' cognitive flexibility, further improvements could be made, such as on the game and on the outcome measures. Adding extra guidance during the game play, creating more scenarios to increase the duration of training, or embedding additional interventions into the game could strengthen the desired skills. Also, using CF tasks with a close relevance to the SG is likely to increase the effect following above-mentioned literature.

Regarding the in-game measures and open questions, we found that the experimental group scored significantly worse than the control group when the rules were changed, which supports our hypothesis. In particular, the contrast between their SG scores became more apparent in S2. This is probably due to scenario 2 being considered more difficult. Indeed, participants expressed that they found S2 more difficult than S1. This result aligns with previous literature in that performance during training will decrease when participants are exposed to new, uncertain, and unexpected circumstances (Brunstein & Gonzalez, 2010; Good, 2014). The overall high detection rate of the experimental group to the changed rules supports our idea that the decrease of their scores during the TP was caused by adaptation in their decision making; thus supporting our claim that the SG stimulates adaptive decision making. This aligns with the operationalization of individual CF described by Cañas et al. (2006), in that detecting a changed environment is the first step towards generating adaptive behavior.

With respect to the game survey, the majority of participants were highly motivated and concentrated in playing both scenarios. Both groups assessed motivation and concentration higher on S1 than on S2, possibly because the context of S1 is more familiar to both groups than that of S2. This notion is supported by the assessment on difficulty and from the debriefing log in that S2 was viewed as more difficult for both groups. Contrary to our underlying assumption that the experimental group will experience the game as more difficult than the control group due to the changing rules, the rating on difficulty varied, which could be due to the nature of self-assessment. Lastly, the control group was more engaged in the SG than the experimental group, especially in S2. It is possible that the unexpected and uncertain rule change could have caused frustration for the experimental group (Lepine et al., 2000), resulting in demotivation. Also, it could be that playing one scenario after another with confusion could have resulted in lack of engagement for the experimental group as participants complained during the debriefing session that their confusion was not resolved between S1 and S2 play.

The current study was an exploratory experiment in a real training environment, adding difficulties in exerting control on all aspects of the experiment. Therefore, it suffers from a number of limitations. Future studies need to increase statistical power and experimental control to assess the effectiveness of the SG. Also, we will focus on instructional methods that could increase individual CF and adaptability while maintaining uncertainty. Although we found that the SG allowed participants to detect the changed situation and adapt accordingly, this approach alone was not sufficient for learners to apply it in different tasks. To better examine the learning effect and to mitigate the different nature of CF tasks with the SG, a test scenario of the SG should be developed that can measure adaptive performance of both the experimental and the control group. After improving the game and further validating its effectiveness under more controlled laboratory conditions, we believe our results will have practical relevance for the training of adaptive decision making by military personnel and other professionals.

Acknowledgement

This study was funded by the Human and Organizational Adaptability program (V1520) at TNO. We thank the Dutch Major school for their participation in this study. We appreciate the support of TNO colleagues Ronald Jongen, Andrea Jetten, Pia de Boer, Marjoleine ’t Hart, and Renske Boswinkel in this study.

References

Baldwin, T., & Ford, J. (1988). Transfer of training: A review and directions for future research. *Personnel psychology*, *41*(1), 63-105.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1), 7-15.

Brunstein, A., & Gonzalez, C. (2010). Preparing for novelty with diverse training. *Applied Cognitive Psychology*, *25*(5), 682-691.

Cañas, J. J., Fajardo, I., & Salmeron, L. (2006). Cognitive flexibility. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (2nd ed., pp. 297-301). Boca Raton, FL: CRC Press.

Dandeker, C. (2006). Building flexible forces for the 21st century. In G. Caforio (Ed.), *Handbook of the Sociology of the Military* (pp. 405-416). New York, NY: Springer.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404-411.

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-Learning*, *2*(1), 5-16.

Good, D. J. (2014). Predicting real-time adaptive performance in a dynamic decision making context. *Journal of Management & Organization*, *20*(6), 715-732.

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems*, *18*(3), 81–85.

Lepine, J. A., Colquitt, J. A., & Erez, A. (2000). Adaptability to changing task contexts: Effects of general congitive ability, conscientiousness, and openness to experience. *Personnel Psychology*, *53*(3), 563-593.

Lezak, M. D. (1995). *Neuropsychological Assessment*. New York, NY: Oxford University Press.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Focus article: Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, *14*, 331-352.

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology*, *85*(4), 612-624.

Ritterfeld, U., Cody, M., & Vorderer, P. (Eds.). (2009). *Serious games: Mechanisms and effects*. New York, NY: Routledge.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains*. Champaign, IL: Center for the Study of Reading.