MULTI-MODAL STUDY OF THE EFFECT OF INFORMATION COMPLEXITY IN A CRISIS MANAGEMENT GAME

Paris Mavromoustakos Blom
Tilburg University
p.mavromoustakosblom@uvt.nl

Sander Bakkes
Utrecht University
s.c.j.bakkes@uu.nl

Pieter Spronck
Tilburg University
p.spronck@uvt.nl

KEYWORDS

Game-based training, information complexity, serious games, multi-modal player modelling, crisis management

ABSTRACT

In this paper, we study the effect of information complexity on player in-game behaviour and physiological responses during a dilemma-based crisis management game. We run a user study, where players attempt to solve a crisis scenario while their in-game and physiological activity is being monitored through game logs and wearable physiological sensors. Results show that information complexity has noticeable effects on players' decision making and physiological responses, while moderate correlation was found between specific in-game- and physiology-based behavioural features. This study is focused on exploring behavioural patterns correlated to various levels of information complexity. Our findings can be applied in future studies aiming at designing personalised crisis management training scenarios.

INTRODUCTION

The use of digital games as training tools has been steadily increasing in popularity over the last decades. Complex and highly stressful tasks such as crisis management, require a significant amount of resources for the proper training and education of expert staff. Digital games enable the simulation of such tasks, providing an alternative and cost-efficient solution.

Furthermore, the introduction of game-based training techniques has enabled multiple data collection channels during gameplay. Such modalities include in-game action logging (Bakkes et al., 2014), peripheral and wearable sensor data (Healey and Picard, 2005), and trainee- or trainer-generated feedback (Sabourin and Lester, 2013). Several multi-modal approaches have been presented in studies that involve player modelling (Pedersen et al., 2010), affect recognition (Kapoor and Picard, 2005) and crisis management training (Mackinnon et al., 2013). It has been shown that multi-modality increases the robustness of player behavioural monitoring (Holmgård et al., 2015).

We employ the Mayor's Game (van de Ven et al., 2014), which enables the simulation of text-based, dilemma solving crisis management scenarios. We have designed three variants of a crisis scenario, which involve low, medium and high information complexity. We run a user study, where participants attempt to "solve" a crisis scenario which has a pre-defined information complexity level. For the remainder of this paper, we define information complexity as the total amount of information that is made available to a player as in-game text, through non-player characters (NPCs). In the Mayor's Game, information is presented to players through discrete messages, each presenting a new perspective to an already described problem which needs to be solved.

The main motivation behind this study is the detection of behavioural patterns, with respect to specific information complexity levels. We expect information complexity to have an observable impact on (1) players' approach towards solving the crisis scenario, and (2) players' (perceived) stress levels during gameplay. To that end, we track player in-game actions to analyse their decision making processes, while we use physiological sensors and post-game questionnaires to retrieve objective and subjective measurements of player affect respectively. In the long term, the results of this study will be leveraged in order to implement personalised crisis management scenarios.

RELATED WORK

The detection, modelling, prediction and expression of human player characteristics which are manifested through cognitive, affective and behavioural patterns in games is often referred to as player modelling (Yannakakis et al., 2013). Multi-modal player modelling involves multiple sources, through which descriptive information about player behaviour is extracted. Methods of multi-modal player modelling have been applied in entertainment games (Pedersen et al., 2010), detection of player affect (Martínez and Yannakakis, 2011) and even diagnosis of PTSD symptoms (Holmgård et al., 2015). Yannakakis et. al. have conducted a thorough study on the methods and principles of player modelling (Yannakakis et al., 2013).

Any computing that relates to, arises from, or influences user emotions is called affective computing (Picard et al., 1995). Over the last decades, scientists have been developing new methods and devices to monitor and interpret humans' affective state. Such devices include wearable physiological sensors, which have been previously used for tasks such as user emotion recognition (Haag et al., 2004), entertainment modelling (Mandryk et al., 2006) and stress detection (Healey and Picard, 2005). Gaming and affective computing have also been combined in recent studies (Toups et al., 2006), where skin conductance and heart rate were used as descriptors of player affective state during gameplay. In particular, player experience analysis through affective signals has been conducted using shooter games (Nacke et al., 2011). In similar fashion, we leverage wearable physiological sensors in order to extract descriptive features of the affect of information complexity in players' physiological responses.

Regarding crisis management training, several key cognitive characteristics (also referred at as "soft skills" (Di Loreto et al., 2012)) of crisis managers have been identified, and require to be trained on in order to maintain experts' preparedness at a sufficient level (Stolk et al., 2001). Such soft skills have been investigated in previous studies and include teamwork & collaboration (Sagun et al., 2009), strategic planning (Turoff et al., 2005), leadership (Wallace and Suedfeld, 1988), stress resilience (Janka et al., 2015) and information complexity processing (Svensson et al., 1997). Crisis management training sessions have also been implemented in a large-scale real-life exercise (Helsloot, 2005), where team coordination played a major role in the effectiveness of multi-level decision making.

We believe each of the aforementioned soft skills need to be studied in isolation, in order to yield valid results. It has been shown that the outcomes of crises are likely to be linked to responders' personality traits (Wallace and Suedfeld, 1988). Recently, Steinrucke et al. (Steinrücke et al., 2019) studied the effect of time pressure as a stressor on players' analytical skills using the Mayor's Game. Similarly, this study identifies information complexity as a core characteristic of crises, and leverages the Mayor's Game to study player decision making under various levels of it.

A review on collaborative management games has been conducted by Di Loreto et. al. (Di Loreto et al., 2012). The authors define crisis management as "the anticipation of foreseeable events and minimisation of unexpected events during a crisis". Furthermore, they point out the importance of the use of predictive models within the context of crisis management games, in order to implement realistic and effective training scenarios. In the process of implementing such predictive models, a key step is to analyse player behaviour and extract descriptive features to interpret it. The current study focuses on the latter two topics, by collecting behavioural data through multiple modalities and analysing its correlation to training scenario-related variables.

Figure 1: A screenshot of the Mayor's Game. In the bottom half of the screen, players are provided with the scenario's texts, while in the top half, five advisors provide additional information when an information bubble appears on top of their avatars.

A key contribution to the field of crisis management training is Pandora (Bacon et al., 2013). It provides a complete training environment, which includes multi-modal player modelling through physiological sensors, trainee self-reports and expert observation of trainee performance. Furthermore, training scenarios are adaptable, based on trainees' real-time affective state assessment. In our study we aim to extract information on player affective state through physiological sensors. Similarly to Pandora, we aim to employ the extracted data in order to interpret players' in-game and physiological behaviour, and ultimately provide crisis management experts with personalised crisis management training scenarios.

METHODOLOGY

In order to study patterns in player decision making under various levels of information complexity in a crisis management game, we ran a user study using the Mayor's Game (van de Ven et al., 2014).

Mayor's Game

The Mayor's Game is a text-based, dilemma-solving digital game, originally designed to assess leadership skills of town mayors in the Netherlands. During a Mayor's Game scenario, players become the mayor of a fictional town which is undergoing a crisis. The crisis is initially described to the player by a short text, and unfolds through a sequence of dilemmas which have to be addressed in order for the scenario to end. Regardless of the players' answers to the dilemmas, each scenario has a pre-defined ending, meaning that every player will have to answer the same questions and reach the exact same outcome. Dilemmas are "yes or no" type of questions, and are designed in such a way where there is no right or wrong answer; each answer represents a different ap-

Table 1: Setup of our experiment's crisis scenario. For each of the eight dilemmas provided, additional information were either neutral (=), suggesting a "yes" answer (+), or suggesting a "no" answer (-). Advice was always given in the form of a two to one majority suggesting a "yes" or a "no" answer.

	Information ID					
Dilemma ID	1	2	3	4	5	Advice
1	=	=	=	-	+	+
2	=	=	=	+	-	-
3	=	=	=	+	-	+
4	=	=	=	+	-	+
5	=	=	=	-	+	-
6	=	=	=	+	-	+
7	=	=	=	+	-	-
8	=	=	=	-	+	-

proach towards solving the crisis.

During the game, players are also able to interact with a number of advisors (see Figure 1). These advisors represent various institutions such as the police and fire department, and provide additional information on the players' demand. They can also provide advice in the shape of a "thumbs up" or "thumbs down", which indicates what their suggested answer is for the current dilemma. Both additional information and advice are optional and require players to explicitly click on the advisors to request them. Moreover, depending on the crisis scenario's design, additional information and advice may not always be available.

Crisis scenario

The crisis scenario used in this experiment, discusses a train accident in the outskirts of a fictional town. A tanker train carrying hazardous toxic chemicals collided with a truck on a road crossing, which caused the derailing of a tanker car which is leaking toxic gasses. The scenario consists of eight dilemmas and the player is accompanied by five advisors, namely the communication & press advisor, the public health services advisor, the fire dept. advisor, the legal advisor and the police dept. advisor (see Figure 1).

The scenario starts with a short description of the crisis, followed by the appearance of the first dilemma. A total of eight dilemmas appear sequentially, in 15 second intervals. Each dilemma is accompanied by a short description, and five pieces of (optional) additional information. For each dilemma, all of the advisors are visible to the player, but whether or not they provide the additional information pieces depends on the experimental condition. Additional pieces of information are designed to arrive in 15 second intervals, and are indicated by an "information" text box above the advisors' avatars (see Figure 1). Incoming additional information does not interrupt gameplay. The scenario has a maximum playtime of 20 minutes, after which it automatically ends. Participants were instructed to answer each dilemma as soon as they believe they have read enough information to make a confident decision.

Table 2: First dilemma encountered in the current scenario. Experimental conditions 0, 1 and 2 provide information pieces 1, 1–3 and 1–5 respectively.

Dilemma question	Are you allowing civilians to assist with the evacuation?
Dilemma description	Because of the potential explosion hazard, the CoPI and ROT plan to evacuate the surrounding area within a radius of 300 meters around the burning truck.
Information 1	A daycare centre lies within the evacuation radius.
Information 2	It is expected that there will also be parents who want to pick up the children themselves.
Information 3	An emergency ordinance is being prepared. The police are sent to the nursery to assist with the evacuation, because every child care provider is able to attend to at most four children.
Information 4	Evacuating is our job. We cannot bear the responsibility that parents must take risks to help us with the evacuation.
Information 5	The parents might become emotional if we try to stop them. It is better to let them help with the evacuation. Many hands make light work.

Experimental conditions

Three experimental conditions were designed: Low, medium and high information complexity. These experimental conditions will be referred to as condition 0 (low), condition 1 (medium) and condition 2 (high).

In condition 0, players were provided with only one piece of additional information from the advisors. In similar fashion, in conditions 1 and 2 players were provided with three and five pieces of additional information respectively. Table 1 illustrates the distribution of additional information. For each dilemma, the first three pieces of additional information were designed to be neutral, meaning that the information did not bias the player towards a positive or negative answer. However, the fourth and fifth pieces of additional information would suggest either a negative or a positive answer. Regardless of experimental condition, for each dilemma, only three advisors would respond with advice in the shape of a "thumbs up" or "thumbs down". There was always a two to one majority towards a positive or negative answer to the specific dilemma (see last column of Table 1). An example dilemma is illustrated in Table 2.

Data collection

97 participants (29 male) were voluntarily recruited from Tilburg University's student participant pool, with an average age of 21.55 years ($sd = 3.37$ years). Due to sensor recording failure, 18 out of 97 participants were excluded from the dataset. This resulted in 81 valid participants (25 male), out of which 26 were allocated to condition 0, 25 allocated to condition 1, and 30 allocated to condition 2. Before playing the Mayor's Game crisis scenario, participants were given a presentation describing the goals of the study and the mechanics of the game.

In order to collect physiological sensor data, Shimmer3 GSR+ (Burns et al., 2010) sensors were used. These sensors are equipped with a Photoplethysmography (PPG) and Galvanic Skin Response (GSR) sensor to measure heart rate (HR) and skin conductance (SC) respectively. Sensor sampling rate was set at 20 Hz. For every participant, a two minute baseline measurement (resting state) of physiological activity was collected. Baseline data collection enables the comparison of physiological data between participants. Along with physiological data, in-game activity is logged by the game engine and provided in XML format. Lastly, post-game questionnaires were submitted by the participants, providing scoring of their experiences in the valence-arousal-dominance scale (Russell, 1980).

From the collected in-game and physiological data, several features were extracted. These features have been described in detail previously in (Mavromoustakos-Blom et al., 2018). The raw PPG signal has been processed using the HEARTPY (van Gent et al., 2019) python library, while the raw GSR signal has been processed using the BIOSPPY (Carreiras et al., 2015) python library.

RESULTS

In this section, we present the results acquired from experimentation. We divide results in three categories; in-game behaviour related, player affect related, and a correlation test between the two modalities.

In-game behaviour

Regarding in-game behaviour, we explore how information complexity affects participants' interactions with the Mayor's Game in three major aspects of the game:

1. Participant answers:

   Figure 2 illustrates the percentage of "yes" answers given by all participants per dilemma, in each experimental condition. Our first observation is that in four out of eight dilemmas, participants in different experimental conditions have shifted from a majority in "yes" answers to a majority in "no" answers and vice-versa. More specifically, in dilemmas 1,2,4 and 7 the majority of answers has shifted along with the complexity of provided information.

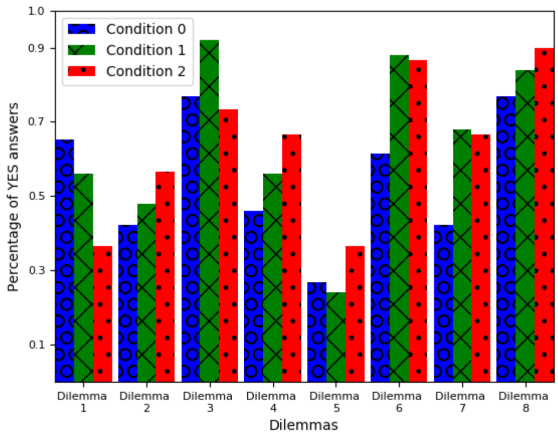

Figure 2: Percentage of participants that gave "yes" answers per dilemma, for each experimental condition.

   Furthermore, looking at Table 3, we observe the total number of participants (per dilemma) that requested advice from the advisors, in each experimental condition. In parentheses we indicate the number of participants that followed the advice (answered according to the advice majority). A Kruskal-Wallis statistical test showed no statistical significance in the number of participants that followed the given advice across experimental conditions ($s = 0.91, p = 0.63$). In all conditions, most participants tended to follow the majority of the advisors' suggestions.

2. Additional information read:

   In conditions 0, 1 and 2 participants had (optional) access to one, three and five pieces of additional information from the advisors respectively. We expected participants in different conditions would interact differently with the advisors' additional information.

   Indeed, looking at Figure 3, we observe that participants in condition 1 read 100% of the provided additional information by average, while in condition 2 they consulted less additional information than in the other two conditions. A Kruskal-Wallis significance test has shown a significant difference in the number of additional information read in condition 2. ($s = 7.77, p < 0.05$), . By average, participants took 56.9, 97.0 and 92.4 seconds to answer each dilemma in conditions 0, 1 and 2 respectively. This means, that even though participants had access to the fourth and fifth additional information in condition 2, most of them chose to proceed to the answer without consulting those.

Table 3: Total number of participants that requested advice (yes or no) per dilemma, for each experimental condition. In paretheses, the total number of participants that followed the given advice.

		Dilemma ID							
		1	2	3	4	5	6	7	8
Condition	0	13(11)	11(8)	14(8)	13(6)	14(10)	20(12)	16(9)	14(6)
	1	9(5)	12(10)	11(10)	12(9)	10(9)	18(17)	12(5)	11(3)
	2	9(4)	10(7)	9(6)	5(4)	11(9)	11(11)	13(7)	11(2)

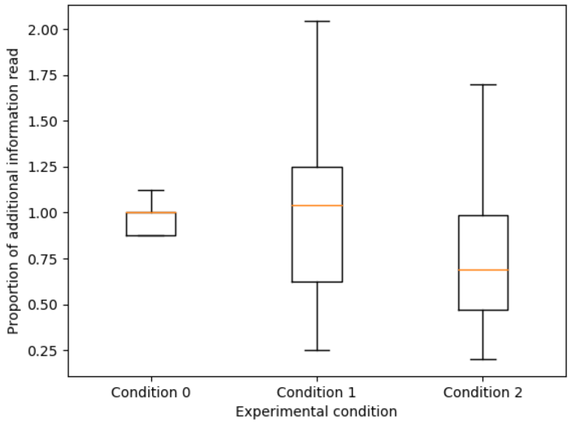

Figure 3: Percentage (out of total number available) of additional information read per dilemma, for each experimental condition. Values greater than 1 mean that additional information pieces were re-visited and re-read by participants.

3. Advice requested:

   Given that participants in different conditions had access to a different number of additional information pieces, we expected information complexity to affect the number of participants that requested advice from the advisors. Table 3 illustrates the total number of participants that requested advice per dilemma, in each experimental condition. We observe a tendency of participants in condition 0 to ask for advice more often, given the low information complexity. A Kruskal-Wallis significance test has shown a significant difference in number of times advice was requested across all experimental conditions ($s = 16.0, p < 0.001$).

Player affect

Player affect was analysed through two separate modalities, namely post-game questionnaires and physiological sensor data. From post-game questionnaire answers, we analyse participants' scores on three variables: valence, arousal, and dominance. Each variable was scored on a scale from 1 (low) to 9 (high). Looking at Table 4, we observe no statistical significance across experimental

Table 4: Player affect related features across conditions, described by mean and standard deviation, and tested for statistical significance through a Kruskal-Wallis test (significant difference found in condition 1).

Feature	Condition	μ	σ^2	s	p
valence	0	5.8	1.4	1.0	0.58
	1	6.0	1.2		
	2	5.6	1.4		
arousal	0	4.0	1.6	0.48	0.78
	1	3.8	1.4		
	2	3.9	2.0		
dominance	0	5.6	1.3	2.3	0.30
	1	5.9	1.5		
	2	5.1	1.8		
diffSC	0	2.7	2.3	11.4	< 0.01
	1	9.1	7.8		
	2	3.5	3.5		
shiftSC	0	1.2	2.4	16.3	< 0.001
	1	8.3	7.9		
	2	1.5	4.0		

conditions for these three variables. They were similarly scored by participants across the three experimental conditions.
Regarding physiological sensor features, we ran statistical tests for $avgHR$ & $avgSC$ (average HR & SC), $diffHR$ & $diffSC$ (subtraction of maximum minus minimum HR & SC), $shiftHR$ & $shiftSC$ (subtraction of last minus initial value of HR & SC) and $tsMaxHR$, $tsMinHR$, $tsMaxSC$ & $tsMinSC$ (timestamps of maximum/minimum HR & SC). $diffSC$ and $shiftSC$ were the only features that showed statistical significance across different experimental conditions, and are included in Table 4. While participants in conditions 0 (low information complexity) and 2 (high information complexity) showed relatively stable physiological responses (low standard deviation), participants in condition 1 (medium information complexity) showed significantly more variation in their physiological responses, as described by these two features.

Cross-modal correlation

We investigated correlation between features that belong to different modalities. Alongside looking for the effect that information complexity has on each modality separately, we would like to observe whether the effects

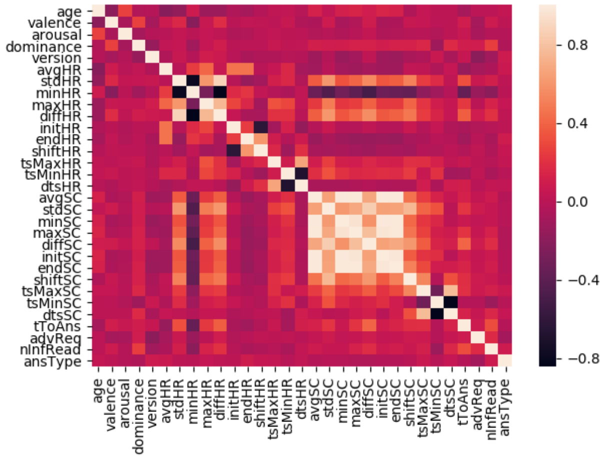

Figure 4: Correlation matrix of in-game, physiological sensor and questionnaire-related features. Each measurement is the average value across all dilemmas.

are shared across two or more modalities. We collected in-game and physiological sensor features during each dilemma answered by the participants along with their questionnaire scores, and created a correlation matrix of these features (see Figure 4).

Looking at Figure 4, we observe correlation between various physiological sensor-related features. Moreover, there is a moderate ($coeff > 0.3$ or < -0.3) linear correlation between the in-game feature $tToAns$ (time required to answer dilemma) with the physiological sensor features $minHR$ (minimum value of HR, $coeff = 0.36$), $diffHR$ (difference between minimum and maximum value of HR, $coeff = 0.37$), $diffSC$ (difference between minimum and maximum value of SC, $coeff = 0.40$). It is important to point out that time required to answer dilemma was measured as a percentage of total game time, to normalise measurements across participants.

DISCUSSION

In this section, we discuss the results obtained through experimentation. We identify how information complexity has affected both players' in-game behaviour and affective state, while indicating how these findings can be employed to supplement crisis management training through the Mayor's Game.

In-game behaviour

Table 3 and Figure 3 illustrate how differences in information complexity cause a shift in players' decision making processes. As Table 3 shows, the lack of additional information during gameplay, encouraged players to consult the advisors significantly more (condition 0) than in the other two conditions.

One may argue that this is the expected behaviour, however, Figure 3 shows that an increase in information complexity is not necessarily followed by an increase in the amount of information read by the participants. Participants in condition 2 read approximately 75% of the additional information provided, showing that in most cases, 75% of the information provided is enough for them to reach a final decision with confidence. In real-life crises, neglection of incoming information can lead to communication gaps, which may harm the efficiency of crisis management (Palttala et al., 2012). However, through a crisis simulation game, after identifying the amount of information deemed necessary for efficient decision making, the next step would be to indicate how information could be structured and communicated so that all involved parties are equally regarded. Towards that goal, we ultimately aim to design adaptive crisis management training scenarios, where the structure and source of incoming information will be based on the players' current information processing performance.

Moreover, Figure 2 illustrates how participants' average answer to each dilemma may shift across the three experimental conditions. Not only can information complexity affect the dominant answer type (dilemmas 1,2,4 and 7), but the average participant answer in condition 1 tends to follow the condition with dominant majority in each dilemma. This leads us to the conclusion that even if there is no right or wrong answer to a dilemma, various levels of information complexity have a significant effect on crisis responders' approach towards answering it. Such inconsistency in decision making can also be a major cause of sub-optimal crisis management response. From the above observations and in the context of this study, we conclude that condition 1 incorporates a level of information complexity that leads to confident decision making; participants in condition 0 consult their advisors to retrieve more information, while participants in condition 2 avoid consulting them.

Player affect

Regarding participants' affective state, no significant difference was found in post-game questionnaire scores regarding the valence-arousal-dominance model across different experimental conditions. This finding can be explained by the low-fidelity nature of the Mayor's Game. A text based game with a "static" non-playing character (NPC) setup does not seem to have a strong effect on player emotions, as would be expected in a real-life crisis.

However, we expect physiological sensor measurements to enable a more fine-grained analysis of player affect. In fact, Table 4 indicates that two skin conductance-related features, namely $diffSC$ and $shiftSC$, differed significantly across experimental conditions. More specifically, participants in condition 1 have significantly higher average measurement in both variables, meaning that that the difference between the minimum & maximum and initial & final SC measurement was significantly higher. From this observation, we may assume that the level of information complexity in condition 1 caused a distinguishable, yet ubiquitus physiological response from the players. While we cannot conclude whether this response derives from experienced stress, excitement, uncertainty or any other aspect of human cognition, we believe this result is worthy of further investigation. Identifying the source of this subtle effect, may be employed towards increasing the fidelity of virtual crisis management training scenarios. Moreover, the moderate correlation between $diffSC$ and the average time taken to answer dilemmas (see Figure 4), indicates that there is a relationship between players' playstyle and their physiological responses.

Condition 1

By assessing the aforementioned results and observations, we identify condition 1 as a salient condition. Participants in condition 1 show a difference in playstyle, which includes not only the patterns in which they process information before answering dilemmas, but also participants' answers themselves. In condition 1, participants consulted 100% of the additional information provided by the involved advisors. This pattern was not observed in condition 2, despite more information pieces being available to the players. In a real-life crisis scenario, uniform decision making and coordination of communication are needed (Palttala et al., 2012). To that end, we claim that participants in condition 1 were the highest performing by average, in terms of additional information processing. However, this does not mean that participants in other groups underperformed; an increase in information complexity (condition 2) may cause higher uncertainty, which the participants seem to avoid. Furthermore, physiological sensor data has shown that condition 1 causes a significant increase in skin conductance-related values, indicating that information complexity can have an effect on player physiology. Lastly, since a moderate linear correlation between in-game and physiological behaviour was observed, we believe that a multi-modal approach should be preferred in the context of this study.

Limitations

As this study's results show, information complexity has affected players' playstyle more than their physiological responses. In an attempt to extract player subjective scorings of affect through the valence-arousal-dominance module, we observe no significant differences across experimental conditions. The low fidelity nature of the Mayor's Game seems to have a mild impact on player affective state, regardless of experimental condition. That being said, we acknowledge that crisis management training scenarios through the Mayor's Game should mostly focus on player decision making and information processing instead of providing a realistic, immersive crisis management experience.
Lastly, with this study's participant pool consisting of university students, we recognise that results may vary when actual crisis responders are put under experimentation. We expect experienced experts to have developed specific decision making skills, probably decreasing the observed variety across experimental conditions. However, experts may relate to the scenario's texts more intimately, which could have noticeable effects on their physiological & affective responses.

Conclusion

In this paper, we presented a multi-modal study of player decision making under various levels of information complexity. Our goal was to determine whether, and to what extent information complexity can affect not only player playstyle, but also their physiological responses during gameplay.
Using the Mayor's Game, we ran an experiment in three separate conditions: low, medium and high information complexity. During gameplay, we collected game-generated logs to monitor player in-game behaviour, while at the same time measured players' physiological responses through wearable physiological sensors. Moreover, we requested players' objective affect scoring through post-game questionnaires, using the valence-arousal-dominance model. We ran statistical analyses on the collected dataset, to identify patterns of player decision making that correlate to the level of information complexity presented to them.
Results show that players did not perceive the three conditions differently in terms of experienced affect, as shown from the analysis of the questionnaires. However, we have identified that under the medium information complexity condition, participants' in-game and physiological behaviour significantly varies with respect to the other two conditions. Players in the medium information complexity condition tend to read 100% of the game's provided information pieces, whereas participants in the high information complexity condition tend to read less information, albeit more is available to them. Furthermore, participants in the medium information complexity condition show significantly higher variance in specific skin conductance-related features, indicating a higher impact in player's physiological responses, with respect to the other two conditions. Lastly, moderate correlation was found between heart rate- and skin conductance-related features and players' in-game actions.
Our long-term aim is to provide crisis management experts with an adaptive, personalised crisis management training environment. Our main focus in future studies is to drive crisis responders towards uniformly dis-

tributed information processing, while maintaining consistency in their decision making.

ACKNOWLEDGEMENT

This study is conducted within the Data2Game project, partially funded by the Netherlands Research Organisation (NWO).

REFERENCES

Bacon L.; MacKinnon L.; Cesta A.; and Cortellessa G., 2013. *Developing a smart environment for crisis management training. Journal of Ambient Intelligence and Humanized Computing*, 4, no. 5, 581–590.

Bakkes S.; Whiteson S.; Li G.; Vişniuc G.V.; Charitos E.; Heijne N.; and Swellengrebel A., 2014. *Challenge balancing for personalised game spaces.* In *2014 IEEE Games Media Entertainment*. IEEE, 1–8.

Burns A.; Greene B.R.; McGrath M.J.; O'Shea T.J.; Kuris B.; Ayer S.M.; Stroiescu F.; and Cionca V., 2010. *SHIMMER^{TM}– A wireless sensor platform for noninvasive biomedical research. IEEE Sensors Journal*, 10, no. 9, 1527–1534.

Carreiras C.; Alves A.P.; Lourenço A.; Canento F.; Silva H.; Fred A.; et al., 2015. *BioSPPy: Biosignal Processing in Python. Accessed on*, 3, no. 28, 2018.

Di Loreto I.; Mora S.; and Divitini M., 2012. *Collaborative serious games for crisis management: an overview.* In *Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2012 IEEE 21st International Workshop on*. IEEE, 352–357.

Haag A.; Goronzy S.; Schaich P.; and Williams J., 2004. *Emotion recognition using bio-sensors: First steps towards an automatic system.* In *Tutorial and research workshop on affective dialogue systems*. Springer, 36–48.

Healey J.A. and Picard R.W., 2005. *Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on intelligent transportation systems*, 6, no. 2, 156–166.

Helsloot I., 2005. *Bordering on reality: Findings on the bonfire crisis management simulation. Journal of Contingencies and Crisis Management*, 13, no. 4, 159–169.

Holmgård C.; Yannakakis G.N.; Martínez H.P.; Karstoft K.I.; and Andersen H.S., 2015. *Multimodal ptsd characterization via the startlemart game. Journal on Multimodal User Interfaces*, 9, no. 1, 3–15.

Janka A.; Adler C.; Fischer L.; Perakakis P.; Guerra P.; and Duschek S., 2015. *Stress in crisis managers: evidence from self-report and psychophysiological assessments. Journal of behavioral medicine*, 38, no. 6, 970–983.

Kapoor A. and Picard R.W., 2005. *Multimodal affect recognition in learning environments.* In *Proceedings of the 13th annual ACM international conference on Multimedia*. ACM, 677–682.

Mackinnon L.; Bacon L.; Cortellessa G.; and Cesta A., 2013. *Using emotional intelligence in training crisis managers: the Pandora approach. International Journal of Distance Education Technologies (IJDET)*, 11, no. 2, 66–95.

Mandryk R.L.; Inkpen K.M.; and Calvert T.W., 2006. *Using psychophysiological techniques to measure user experience with entertainment technologies. Behaviour & information technology*, 25, no. 2, 141–158.

Martínez H.P. and Yannakakis G.N., 2011. *Mining multimodal sequential patterns: a case study on affect detection.* In *Proceedings of the 13th international conference on multimodal interfaces*. ACM, 3–10.

Mavromoustakos-Blom P.; Bakkes S.; and Spronck P., 2018. *Personalized crisis management training on a tablet.* In *Proceedings of the 13th International Conference on the Foundations of Digital Games*. ACM, 33.

Nacke L.E.; Kalyn M.; Lough C.; and Mandryk R.L., 2011. *Biofeedback game design: using direct and indirect physiological control to enhance game interaction.* In *Proceedings of the SIGCHI conference on human factors in computing systems*. 103–112.

Palttala P.; Boano C.; Lund R.; and Vos M., 2012. *Communication gaps in disaster management: Perceptions by experts from governmental and non-governmental organizations. Journal of contingencies and crisis management*, 20, no. 1, 2–12.

Pedersen C.; Togelius J.; and Yannakakis G.N., 2010. *Modeling player experience for content creation. IEEE Transactions on Computational Intelligence and AI in Games*, 2, no. 1, 54–67.

Picard R.W. et al., 1995. *Affective computing.* Perceptual Computing Section, Media Laboratory, Massachusetts Institute of Technology.

Russell J.A., 1980. *A circumplex model of affect. Journal of personality and social psychology*, 39, no. 6, 1161.

Sabourin J.L. and Lester J.C., 2013. *Affect and engagement in Game-BasedLearning environments. IEEE transactions on affective computing*, 5, no. 1, 45–56.

Sagun A.; Bouchlaghem D.; and Anumba C.J., 2009. *A scenario-based study on information flow and collaboration patterns in disaster management. Disasters*, 33, no. 2, 214–238.

Steinrücke J.; Veldkamp B.P.; and De Jong T., 2019. *Determining the effect of stress on analytical skills performance in digital decision games towards an unobtrusive measure of experienced stress in gameplay scenarios. Computers in human behavior*, 99, 144–155.

Stolk D.; Alexandrian D.; Gros B.; and Paggio R., 2001. *Gaming and multimedia applications for environmental crisis management training. Computers in Human Behavior*, 17, no. 5-6, 627–642.

Svensson E.; Angelborg-Thanderez M.; Sjöberg L.; and Olsson S., 1997. *Information complexity-mental workload and performance in combat aircraft. Ergonomics*, 40, no. 3, 362–380.

Toups Z.O.; Graeber R.; Kerne A.; Tassinary L.; Berry S.; Overby K.; and Johnson M., 2006. *A design for using physiological signals to affect team game play. Foundations of Augmented Cognition*, 134–139.

Turoff M.; Chumer M.; Yao X.; Konopka J.; Van de Walle B.; and Carle B., 2005. *Crisis planning via scenario development gaming.* In *Proceedings of the 2nd International ISCRAM Conference, Brussels, Belgium.*

van de Ven J.G.M.; Stubbé H.; and Hrehovcsik M., 2014. *Gaming for Policy Makers: It's Serious!* In A. De Gloria (Ed.), *Games and Learning Alliance*. Springer International Publishing, Cham. ISBN 978-3-319-12157-4, 376–382.

van Gent P.; Farah H.; van Nes N.; and van Arem B., 2019. *HeartPy: A novel heart rate algorithm for the analysis of noisy signals. Transportation Research Part F: Traffic Psychology and Behaviour*, 66, 368–378.

Wallace M.D. and Suedfeld P., 1988. *Leadership performance in crisis: The longevity-complexity link. International Studies Quarterly*, 32, no. 4, 439–451.

Yannakakis G.N.; Spronck P.; Loiacono D.; and André E., 2013. *Player modeling.* In *Dagstuhl Follow-Ups*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, vol. 6.