Evaluation System for Game Playability
Using Emotion Sensor Based on AI

Shuh-Yeuan Deng[*] and Kuo-Kuang Fan

National Yunlin University of Science and Technology, Yunlin County, 640301, Taiwan

Game players almost always experience games through electronic devices, such as computers, cell phones, headphones, and microphones, which allow them to play and communicate. However, players' emotional changes during games cannot be sensed accurately in real time. Therefore, to evaluate whether players like a particular game, questionnaires are generally required afterwards to collect information on players' feelings, which do not provide a complete record or real-time information. Advances in AI have increased convenience in everyday life, and emotion sensors based on AI, which employ face recognition techniques and physiological sensors, have been introduced. Thus, through the interpretation of facial expressions, users' emotional reactions while they are using electronic devices can be recognized. Moreover, voice recognition can also assist in improving the accuracy of AI-based emotional recognition. In this study, we examine how an emotion sensor based on AI can enable a more immediate and complete assessment of game playability, serving as an important reference for game development.

1. Introduction

Various methods for facial expression recognition have been proposed by researchers. The facial expression dataset has been a benchmark commonly used by traditional methods in the past.[1–3] Unfortunately, traditional image processing techniques for face detection and expression recognition have the problems of low detection performance due to angle variation and limited recognition accuracy caused by monotonic recognition networks. Therefore, most of these techniques are only applicable to frontal and short-distance facial emotion recognition. However, these limitations in viewing angles and lighting were overcome by a breakthrough in deep neural networks. Subsequently, attempts have been made to address the unsatisfactory performance of image recognition through extensive recording and image training. For instance, Yang *et al.* have developed face detection models that can be applied in real-world scenes.[4] Ekman and Friesen conducted in-depth studies on facial expressions and classified human expressions into seven categories: happiness, sadness, fear, disgust, surprise, anger, and neutral.[5,6] After face detection is completed, facial expressions can then be recognized through

affective computing based on the captured facial images and trained facial expression recognition neural networks (e.g., ResNet-50, Inception v3, MobileNet).[7–9] These developments have enabled diverse applications of facial emotion recognition in the retail, medical, entertainment, and other fields.[10,11]

During the game development process, game designers refer to popular topics and trends in the market as well as previous experience and design concepts. They must undergo many iterations of scenario planning and simulation testing to identify whether the designed game concept will be successful. However, a game's playability must be evaluated through the real playing experience of different players. Therefore, in the process of designing a game, different testing methods must be applied to identify any issues in the game play design, mechanism, and game screen aesthetics. Problems that emerge closer to the completion of game development typically cost more to fix. Therefore, it is very important to conduct game playability evaluations that allow players to test the game under development and determine whether it is fun to play.[12]

1.1 Emotion sensor based on AI

Kort *et al.* proposed the learning companion theory, stating that students generate many emotions in the process of learning new things and that these emotions can be plotted on an emotional axis to represent students' emotional perceptions of learning and a vertical axis to represent students' motivation to learn.[9,10] This framework can be used to understand students' emotional reactions during learning to determine their learning challenges. We show such a plot in Fig. 1 and explain the relationship between learning process and transformation of emotions. Ashwin *et al.* and Ray and Chakrabarti proposed the use of physiological responses, such as voice, heartbeat, skin sensations, and blood pressure, to improve the accuracy of AI-based emotion recognition.[13,14] An emotion sensor based on AI applies electronic devices (i.e., smart phones, computers) to recognize player's emotions during game play, allowing researchers to
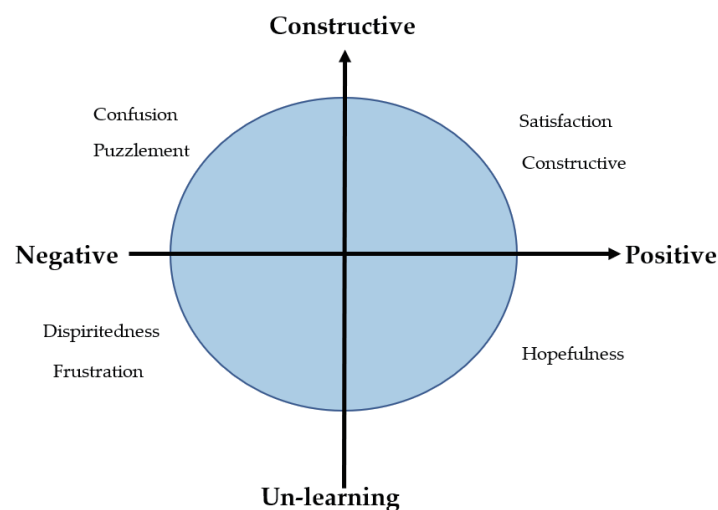

Fig. 1.   (Color online) Learning process and transformation of emotions.[5]

understand individuals' learning states in a real-time and continuous manner. This, in turn, can be very helpful to learners. Players involved in testing new games undergo a similar process to learners learning new concepts. If the technology of an emotion sensor based on AI can be applied to evaluate the playability of new games, designers can then effectively acquire real-time data and understand players' emotional changes and reactions throughout the process of experiencing games. Therefore, such data and information acquired via an emotion sensor based on AI can serve as an important reference for game designers.

1.2   Game playability

Fabricatore *et al.* have suggested that playability is influenced by parameters that operate and control game interactions; they also stated that improving a game's playability can make the game more appealing and interesting, encouraging players to spend more time playing, as shown in Fig. 2.[15] User testing is the benchmark for any playability evaluation since designers cannot fully predict user behavior. Therefore, in game design, expert reviews must be conducted via heuristic evaluation for playability (HEP) to identify playability issues and seek alternative solutions. HEP enables thinking from the user's perspective and is very helpful in problem solving.[16]

Clanton divided game playability into three components: game interface, game machines, and game play.[17,18] Each of these components can be defined as follows:
1. Game interface: This is the device through which players interact with a game (e.g., joystick, keyboard, mouse, etc.).
2. Game machines: Game machines combine the physical concepts in the game, the game animation, and the computer program, allowing players to adopt different roles and freely perform actions in the game (e.g., walking, jumping, and other actions).
3. Game play: This is the process of achieving goals in the game; players will encounter challenges in the process of playing the game, and they must face and try to overcome these challenges.
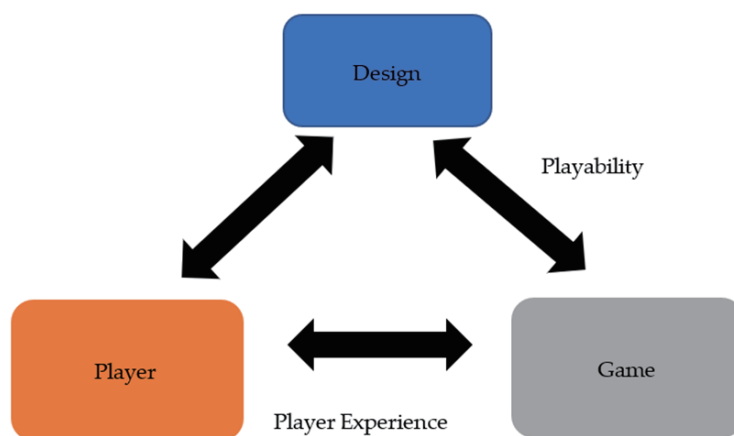

Fig. 2.   (Color online) Role of playability in the system.[16]

Whether players develop interest in a particular game depends on multiple factors, including the smoothness of the control interface, the ease of entering the game, the devices required to experience the game, the interactivity of the game content, and the game screen smoothness and accessibility. In this study, an emotion sensor based on AI is used to identify players' emotional reactions while experiencing a game to evaluate the game's playability. Players' emotional reactions are acquired continuously and systematically and then transformed into a continuous numerical record. In this study, artificial emotional recognition is presented on emotional axes, as shown in Fig. 3.

2.   Materials and Methods

The conceptual framework for testing in this study comprises four components: (1) the game player (that is, the game tester), (2) the sensing devices (cell phone and microphone), (3) the affective computing tool, and (4) the emotion recognition evaluation system. While testing the game, the game player is required to set up a cell phone and microphone as emotional input devices. The cell phone detects the player's face and captures emotional expressions while the microphone collects audio information from the player during the test. The player's facial expressions and audio information are then used to identify emotions through an affective computing tool, and emotional changes are transmitted to the evaluation system and recorded. The conceptual framework is shown in Fig. 4.

2.1   Graduation projects (tested games)

Final-year students of the Department of Animation and Games are required to produce games for their graduation projects. Students participating in this study were divided into 10 groups with 3–5 members per group. All games were designed to run on a computer, which excludes the control variables of different devices. Various types of games were developed by the students, including role-playing game (RPG), adventure game (AVG), and action game (ACT).
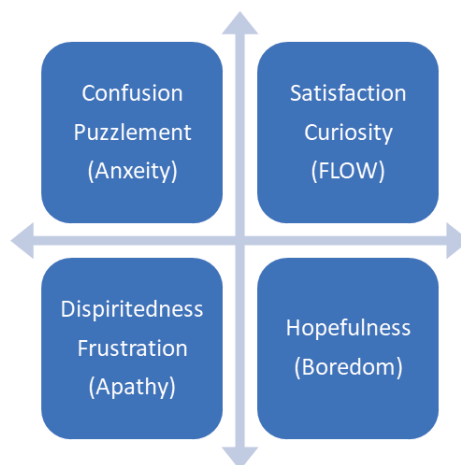

Fig. 3.   (Color online) AI-based recognition of emotions during game playing presented on emotional axes.[10,19,20]
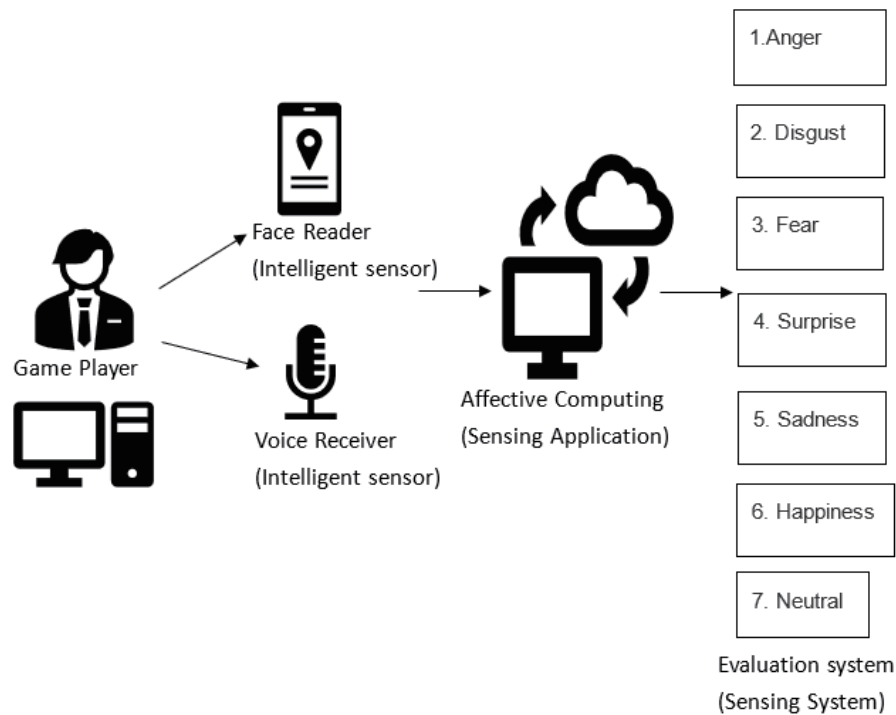
Fig. 4.   Conceptual framework.

2.2   Game testers (players)

The games developed for the above-mentioned graduation projects were used as test games, and 30 students from the Department of Animation and Games were recruited as game testers (players). The players were all male and between the ages of 19 and 22.

2.3   Test environment and equipment

The 10 games to be tested were installed in a Windows-based laptop computer. Moreover, a cell phone with an Android 4.4 operating system or above was set up in front of each tester to capture players' facial expressions. Finally, a microphone was set up to record testers' voice messages during the games.

The testers' facial expressions were captured as continuous data. Through the artificial emotion tool Affdex, seven types of emotions were identified. The system recorded each tester's emotional dimensions at each time point and presented the trend of emotions in a linear manner.

2.4   Test procedure

Each player was required to test all 10 games, devoting 10 min to each game. Within the same timeframe, each player tested 10 games in 2 h. The 10 games were preset in a numbering

sequence, and the order of games to be tested was determined by each player by drawing lots. The capture rate for emotion recognition was set as 10 frames/s. The basis for emotion recognition is shown in Table 1; these emotional dimensions were transmitted to the evaluation system. The testing flowchart is shown in Fig. 5.

Table 1
Emotional classification table.

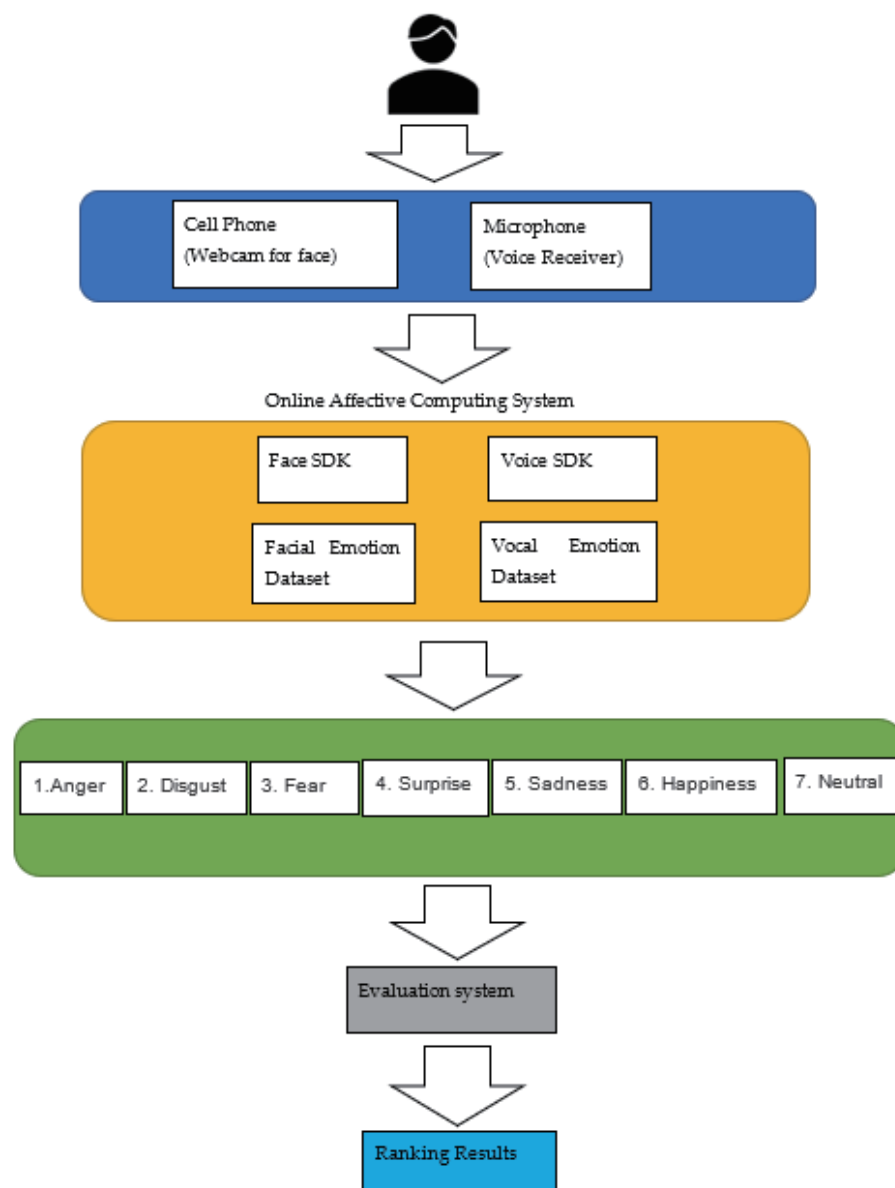| Emotional dimension	Emotions			
Positive	Surprise	Sadness	Happiness	Neutral
Negative	Anger	Disgust	Fear	—

Fig. 5.   (Color online) Testing flowchart.

3.   Results

The facial expressions of players captured during game testing included happy and neutral expressions, as shown in Fig. 6. The emotions recognized by the evaluation system during 3 min of testing are listed in Table 2. The evaluation system then converted the table into a line graph to present each player's emotional changes during game testing, which is presented as an emotional analysis in Fig. 7. However, the voice messages during game testing were not clear enough for recognition; thus, facial recognition was used as the study's main testing method.
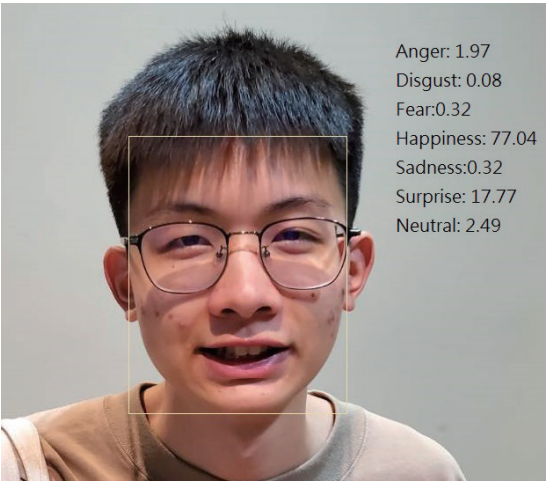

Anger: 1.97
Disgust: 0.08
Fear:0.32
Happiness: 77.04
Sadness:0.32
Surprise: 17.77
Neutral: 2.49

Fig. 6.   (Color online) Recognized emotions.

Table 2
Emotions recognized by the evaluation system over 3 min.

Time (s)	Valence	Emotion	Area
0	20	Happiness	210
10	22	Happiness	235
20	25	Happiness	260
30	27	Happiness	260
40	25	Surprise	310
50	37	Surprise	410
60	45	Neutral	615
70	78	Neutral	780
80	78	Neutral	725
90	67	Neutral	625
100	58	Neutral	630
110	68	Neutral	680
120	68	Neutral	685
130	69	Neutral	690
140	69	Neutral	680
150	67	Neutral	670
160	67	Neutral	615

Fig. 7.   (Color online) Emotional analysis graph.

After each game was tested, the evaluation system took 10 s to calculate the emotional dimension and generate an emotional analysis graph, as shown in Fig. 7. Meanwhile, the proportion of each player's positive emotions while testing each game was also calculated (presented in Table 2). The sum of the positive emotion areas was calculated using the area calculation method, and each player's ranking of the 10 games featured in this study is displayed in Table 3. In total, 30 game emotional analysis tables were derived and used to sum the results. Then, a playability evaluation summary table obtained through the artificial emotion recognition tool was used to score the games. The ranking of the games is shown in Table 4.

After completing the testing process, each player was given a voting sheet to select the top three games with the highest playability. The voting results in Table 5 were then compared with the ranking provided by the evaluation system to determine whether the results of the top three games matched. Although the ranking results were not completely consistent, a recognition rate of 70% was achieved, indicating that artificial emotion recognition tools can effectively help to evaluate the playability of games.

Game playability analysis is the main reference for game design. Players can use an emotion sensor based on AI during game testing, and then complete questionnaires after the test is completed. A comparison of these two playability evaluation methods is given in Table 6. The functions of these methods are complementary and together improve game playability evaluation.

Table 3
Player 1 ranking of games based on positive emotions.

Team number	Satisfaction (Positive area)	Rank
1	12860	7
2	13250	6
3	11260	10
4	13720	5
5	14580	4
6	14820	3
7	11530	9
8	16040	1
9	15020	2
10	12370	8

Table 4
Ranking results of evaluation system.

Rank	Team number (Positive area)	Satisfaction
1	8	520290
2	9	503670
3	6	471960
4	5	437640
5	4	432660
6	2	391050
7	1	390480
8	10	388110
9	7	331290
10	3	298590

Table 5
Voting results of players.

Rank	Team number	Number of votes
1	9	23
2	8	21
3	6	20
4, 5	5	7
	2	7
6	1	5
7	4	4
8	10	3
9, 10	7	0
	3	0

Table 6
Comparison of playability evaluation methods.

	Questionnaire survey	Emotion sensor
Advantages	1. Multi-aspect analysis with topic models to understand the thinking of players. 2. Able to analyze the player's reaction in different levels of the game.	1. Able to understand the emotional reaction of the test game in real time. 2. Able to continuously record the emotional state when the players test the game.
Disadvantages	1. Unable to understand the emotional reaction of the test game in real time. 2. The feeling of emotion cannot be accurately and quantitatively presented.	1. Too difficult to confirm the user's current emotional reaction. 2. Unable to understand the preference of the player accurately.

4.　Conclusions

　　The applications of emotion sensors based on AI are becoming increasingly diversified. We applied an emotion sensor based on AI to identify players' emotional perceptions throughout the process of testing games, thus developing a method that can be implemented in a relatively simple way and utilized in many fields in the future. In this study, games were not categorized because it would have been possible to identify players' preferences in terms of game types from the results. The emotional dimensions captured via an emotion sensor based on AI can be used as a reference for players' reactions during games. Identifying and harnessing these emotions would help game designers to develop more enjoyable and interesting games. The emotion sensor based on AI used in this study can also be introduced during the game testing stage to provide significant information and enable designers to modify games before they are completed and released.

References

1　M. J. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba: Proc. 3rd IEEE Int. Conf. Automatic Face and Gesture Recognition (IEEE, 1998) 14–16. https://dor.org/10.1109/AFGR.1998.670949
2　P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews: Proc. IEEE Workshop CVPR for Human Communicative Behavior Analysis (IEEE, 2010) 94. https://dor.org/10.1109/CVPRW.2010.5543262
3　R. El Kaliouby: https://www.technologyreview.com/2017/10/20/148315/we-need-computers-with-empathy/ (accessed October 2017).
4　S. Yang, P. Luo, C. C. Loy, and X. Tang: Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (IEEE, 2016) 5525–5533.
5　P. Ekman and W. V. Friesen: Facial Action Coding System: A Technique for the Measurement of Facial Movement (Consulting Psychologists Press, Palo Alto, CA, 1978).
6　P. Ekman, W. V. Friesen, and P. Ellsworth: Emotion in the Human Face (Cambridge University Press, Cambridge, UK, 1982) pp. 39–55.
7　D. Gukas, S. J. Leinster, and R. Walker: Med. Teach. **32** (2010) 5. https://doi.org/10.3109/01421590903398232
8　M. Sathik and S. G. Jonathan: SpringerPlus **2** (2013) 455. https://doi.org/10.1186/2193-1801-2-455
9　M. Yeasin, B. Bullot, and R.Sharma: IEEE Trans. Multimed. **8** (2006) 500. https://doi.org/10.1109/tmm.2006.870737
10　B. Kort, R. Reilly, and R. W. Picard: Proc. 2001 IEEE Int. Conf. Advanced Learning Technologies (IEEE, 2001) 42–46.
11　J. Qin, Q. Zheng, and H. Li: Edu. Technol. Soc. **17** (2014) 420. https://doi.org/10.0000/PMID99574683
12　J. L. G. Sánchez, F. L. G. Vela, F. M. Simarro, and N. P. Zea: Behav. Inf. Technol. **31** (2012) 1033. https://doi.org/10.1080/0144929x.2012.710648

13 T. S. Ashwin, J. Jose, G. Raghu, and G. R. M. Reddy: Proc. 2015 IEEE 7th Int. Conf. Technology for Education (IEEE, Warangal, 2015) 23–26. https://doi.org/10.1109/T4E.2015.21

14 A. Ray and A. Chakrabarti: Edu. Technol. Soc. **19** (2016) 112. https://www.semanticscholar.org/paper/b4e61b3da7a9211a199fddf16982490ca50b1ada

15 C. Fabricatore, M. Nussbaum, and R. Rosas: Hum. Comput. Interact. **17** (2002) 311. https://doi.org/10.1207/s15327051hci1704_1

16 H. Desurvire, M. Caplan, and J. A. Toth: Proc. 2004 Conf. Human Factors and Computing Systems—CHI '04 (ACM: New York, 2004) 1509–1512.

17 L. E. Nacke, A. Drachen, K. Kuikkaniemi, J. Niesenhaus, H. J. Korhonen, V. M. Van Den Hoogen, K. Poels, W. A. Ijsselsteijn, and Y. A. W. De Kort: Proc. 2009 DiGRA (DiGRA, London, 2009).

18 C. Clanton: Proc. 1998 ACM CHI 98 Conf. Summary on Human Factors in Computing Systems (ACM, 1998) 1–2.

19 M. Csikszentmihalyi: The Evolving Self: A Psychology for the Third Millennium (HarperCollins, New York, 1993).

20 M. Csikszentmihalyi and F. Massimini: New Ideas Psychol. **3** (1985) 115. https://doi.org/10.1016/0732-118X(85)90002-9

About the Authors

Shuh-Yeuan Deng received his B.S. degree from Soochow University, Taiwan, in 1999 and his M.S. degrees from Shu-Te University, Taiwan, in 2002 and 2004. From 2018 to 2020, he was an assistant professor at Shu-Te University, Taiwan. Since 2021, he has been an associate professor at Shu-Te University. His research interests are in digital game design, game script creation, and machine learning, design and test. (treefar@stu.edu.tw)

Kuo-Kuang Fan received his Ph.D. degree from Sun Yat-sen University, Taiwan. He has been a professor at Yunlin University of Science and Technology, Taiwan. His research interests are in design theory, media research, game and interactive design, design technology and computing, design humanities and art cross-cultural research, perceptual engineering, and mechanical learning. (fankk@yuntech.edu.tw)