J Multimodal User Interfaces (2015) 9:403–411
DOI 10.1007/s12193-015-0197-0

ORIGINAL PAPER

Investigating the effects of motion-based Kinect game system on user cognition

Akihito Nakai[1] · Aung Pyae[2,5] · Mika Luimula[2] · Satoshi Hongo[3] · Hannu Vuola[4] · Jouni Smed[5]

Received: 7 November 2014 / Accepted: 14 August 2015 / Published online: 28 August 2015

Abstract Kinect, which is useful as a multimodal input device, offers multiple features such as voice recognition, facial tracking, object recognition, and gesture recognition. Games with such human-centred, interactive interfaces create an environment that bridges between artificial and natural cognitive capabilities, enabling them to work together more efficiently. This study specifically examines a usability evaluation method that uses Kansei engineering knowledge to identify and remedy difficulties related to user interface development. As a practical application of the suggested evaluation method, we used a particular serious game, "The Glider," a prototype developed by Serious Games Finland. It exploits Kinect as its game controller. Serious games are usually aimed at promoting specific effects that can be achieved through playing the game. However, such effects can also present tradeoffs in terms of usability. Controlling motions in "The Glider" has been developed in close cooperation with medical doctors, physiotherapists, and patients in terms of well-being and light exercise. However, modulating the difficulty balance and usability of the game is also a very important aspect of game development. As described in this paper, our goal is to evaluate the control motion of the game from a usability perspective. Evaluation consists of state and appearance observations conducted with 12 Japanese subjects. Observed data were tagged into several categories to quantize their behaviours and speeches, and were analysed using both cognitive bias evaluation and statistical evaluation. Results show that difficulties for subjects in control motions of the game were curved out, which is a key to balancing usability and difficulty while maintaining the genuine purpose of the serious game.

Keywords Skeletal motion interface · Usability test · Kansei engineering · Serious game

✉ Akihito Nakai
nakai@ais.riec.tohoku.ac.jp

Aung Pyae
aung.pyae@turkuamk.fi

Mika Luimula
mika.luimula@turkuamk.fi

Satoshi Hongo
hongo@sendai-nct.ac.jp

Hannu Vuola
hannu.vuola@seriousgames.fi

Jouni Smed
jouni.smed@utu.fi

1 Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ward, Sendai, Japan

2 Turku University of Applied Science, Turku, Finland

3 Sendai National College of Technology, Natori, Japan

4 Serious Games Finland, Helsinki, Finland

5 University of Turku, Turku, Finland

1 Introduction

The Japanese word "Kansei" can be interpreted to mean emotion, feeling, receptivity, or sensibility. The field of Kansei Engineering, defined by Nagamachi [1], specifically focuses on consumers' psychological feelings towards engineering products. In essence, Kansei Engineering is a consumer-oriented technological research field that intends to support engineers in conforming to user needs via the quantification and evaluation of users' Kansei.

Recently, multimodal input devices have come to be used in various applications such as digital game controllers, motion tracking systems, and smart phones, drawing close attention from people in digital media industry, and digital game companies. Conventional input devices (e.g., keyboard, mouse, controller, and joystick) are still widely used. However, new interactive techniques are becoming ever more popular among users for their efficiency, ease, smoothness, innovativeness, novelty, and styles.

Especially, game controllers such as Microsoft Kinect, Nintendo Wii's multimodal devices such as Wii Remote, and PlayStation's EyeToy, are not only popular and trendy multimodal input devices among gamers and in the gaming market; they have also gained popularity and noticeable success in research and development fields. Those game controllers are being developed. In fact, many interactive systems exist today. They can be classified as A. cutting-edge methods developed to interact with information and B. newly developed interfaces applied to innovate interactions with and effects of systems.

As an example of A, groups of the wearable gestural interfaces researchers from MIT Media Lab such as Sixth Sense Interaction [2,3] are developing and exploring tangible interfaces as a new approach to enable users to interact with projected information. The interactive tele-olfaction system with a movie [4], a context-aware augmented reality interface [5], and The Kiss Communicator [6] are some new approaches.

In contrast, as examples of B, Leo and Tan [7] use infrared markers and OptiTrack motion tracking systems to detect player movements in the interactive floor projection system that is built for gait and balance training for children with cerebral palsy. Theng et al. [8] use Nintendo Wii as a socialization aid for elderly people to enhance their social communication with younger generation players such as family members and friends. Matsuguma uses a Nintendo Wii balance board [9] to motivate elderly people to exercise using standing up motions. These studies listed as B examples are all aimed at innovating the mode of achieving specific purposes via the cognitive effects of a multimodal interface on its users through the game experience. Such games with specific purposes are called serious games. This study introduces an investigation method based on Kansei engineering knowledge to a research and development phase of a serious game such as the examples listed for B above.

Development of serious games includes various purposes to achieve along with its themes and against such diversity of purposes. Multimodal input devices must perform to the greatest extent possible in terms of achieving their fundamental purposes in any case. Generally speaking, a game interface should have good usability. Such usability and complex purposes for serious games can present a tradeoff. Therefore, evaluating those balances is extremely important for maintaining both usability and effects for the fundamental purposes of facilitating serious games. The authors collected cognitive impressions related to usability from experiments that include both protocol analysis and behavioural observation. To quantize the observed data, the data were labelled into cognitive categories. Then those labels were analysed to reveal biases against usability. As a result, evaluation results clarified the cognitive biases in quantized forms. They are useful for orientation to improvement.

2 Relations between multimodal interface and Kansei engineering

2.1 Previous researches of multimodal interface

Among the multimodal input devices and interaction techniques, Microsoft's Kinect is well-regarded as a promising multimodal input device that can provide various interaction techniques, such that the user does not need to hold any controller. By virtue of its well-prepared SDKs, in the applications of Kinect as an input interface, interactions between users and systems are modifiable. Using such features of Kinect, many research topics are related to interface developments. Biswas and Basu [10] propose a method to recognize human gestures using a Kinect depth camera, which allows the efficient use of a depth camera to recognize multiple human gestures. For their study, they use eight human gestures: clap (clapping), call (hand gesture to call someone), greet (greeting with folded hands), wave (waving a hand), no (shaking the head sideways), yes (tilting the head up and down), clasp (hands clasped behind head), and rest (the chin resting on a hand). Thibaut et al. [11] build a Kinect-based system that enables any user to control the facial expressions of a digital avatar or character in real time. The authors advocate that Kinect-based realistic facial tracking facilitates a range of new applications such as digital gameplay, telepresence, and social interactions. Wu et al. [12] create a Kinect-based hand gesture recognition system that can address the problem of one-shot learning gesture recognition with a user-defined training and testing system. Such a system is useful as a remote control with which the user can allocate a specific function using a particular gesture that is defined by the users by performing it only once. Kinect is useful as a visual communication medium for online instant messaging users. Chao et al. [13] present a Kinect-based visual communication system that includes the following features: a sign language recognition module to make deaf-mute persons able to chat, an expression recognition module to enrich online chatting experiences simultaneously, a speech recognition module to free users' hands while chatting, and a cross-media multilingual visualized translation module to enable the users to catch the meanings of the conversation more easily. Their experiments related to this visual communication system demonstrate that it provides a powerful and efficient communication function and a good user experience. Regarding human detection, Xia et al. [14] develop a novel human detection method using depth information captured using Kinect for the Xbox 360. In this work, the authors propose a model-based approach that tracks humans using a 2D head contour model and a 3D head surface model and segmentation scheme to segment the human from the surroundings and to extract the whole contours of the figure based on the detection point.

Gamification [15] is known as a process of enhancing a service using gameful experiences to support a user's overall value creation. Because of the severity and difficulty of rehabilitative trainings, and through innovation of experiences, many research organizations all over the world are currently interested in studies related to motion-based games for application to the field of rehabilitation [16–19]. The development of a rehabilitative game with motion sensors consists of both content development by rehabilitation professionals to set up the rehabilitative effects, and the development of gamification by game developers to motivate players to undertake rehabilitation via game play. In development phases of such serious games, evaluations to balance the game between rehabilitation effects and gameful experiences are extremely important. Therefore, cognitive evaluations for control moves in the rehabilitation game are highly desirable.

2.2 Applying Kansei engineering to interface evaluation

Studies undertaken in the research field of Cognitive InfoCommunication (CogInfoCom) are undertaken to reveal the relations between the inner system of the human mind and information provided by the system. The word Kansei originates from the ancient Greek words *aesthesis* and *ethos*. Especially, in this paper, the word Kansei does not specifically signify task efficiency but rather a user's psychological state. Regarding this point, Kansei Engineering is expected be helpful to evaluate and to quantify a human's natural cognitive process in terms of their impressions and cognitive biases. Kansei Engineering is a study field aimed at development of engineered products that accommodate a user's needs and feelings via quantification and evaluation of a user's Kansei. Therefore, it engenders physical (e.g. ease, visibility), cognitive (e.g. memorability), and Kansei (e.g. desirability) attributes of a user's feeling while using the product [21] according to several techniques. For example, Kansei measurement is represented with semantic differential methods and objective evaluations. Objective evaluations are based on human physiological signals such as the elevation of the heart rate, body temperature, and perspiration. Kansei analysis technique is represented from statistical analysis [22]. The idea of quantification and distinctions related to elemental reactions and perceptions for observed data help us to find some relevance among factors of the system such as conceptual framework [23], which consists of Concepts, Messages, and Icons in CogInfoCom field.

Quantification by labels and distinctions given by the results of categorisation in this study is a good way to assess problems in the current development phase. We will obtain keys through analyses to construct hypotheses for orientation to improvement of development processes. As a case study of introducing Kansei engineering to such developments of the game with interactive interfaces, this report describes the investigation results over a serious game using evaluation methods based on Kansei engineering knowledge.

3 Evaluation of game prototype using Kansei engineering knowledge

The sample target we have used, "The Glider", is a Kinect-based serious game developed by a company called Serious Games Finland. The control moves in the game have been designed in close cooperation with medical doctors, physiotherapists, and patients at Carea–Kymenlaakso Social and Health Services in Finland, in terms of well-being, light exercise, and training for elderly people and patients recovering from a stroke. Main methods of usability evaluations are listed below.

1. Questionnaire
2. Behavioural observation
3. Speech observation
4. Test play
5. Analysis

Questionnaires were filled out after we received informed consent of participants. We had earlier informed the participants of the experiment purpose. Both speech and behavioural observations were used in the test play.

While pilot tests are also commonly used by developers and professional gamers to evaluate gaming software [24], our approach of "test play" involves deeper evaluations with respect to the psychology of software use. We perform further analyses based on the obtained results.

3.1 Preparation

First, the authors constructed a test play environment and conducted preparation play with three Finnish participants (university students) to assess their reactions and to ascertain how they actually play the game to make sure the environment is properly constructed for the evaluation. We conducted two preparation play sessions. In this preparation, a player sat at the player's position. Their play was recorded using video cameras (Fig. 1). Players were asked

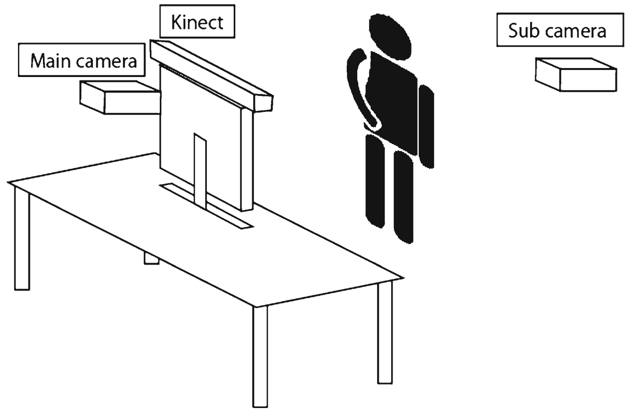

Fig. 1 Environment for preparation play

to express their current feelings during their play in English according to the think-aloud method. They played the first four stages of The Glider game. After the preparation play, if the player was asked to speak out their feeling in English, they felt difficulty in reporting their feelings, and rather murmured a few words during their play. This presents an important difficulty in using think-aloud methods because the analysis target is their speech: if the amount of speech is small, then results obtained through analysis will be few. The players appeared to be concentrating carefully on playing rather speaking about their feelings, which might explain why they did not speak a lot. Therefore, thinking aloud in the pilot test scheme was done with the player's mother language to simplify the task.

After the first preparation play, another Finnish student joined for the second preparation play before the main experiment. The second test play was conducted with the same environment set as that used the first time, but not in English. The student spoke about his feelings and impressions in the Finnish language during play. Those speech data were all translated into English after the experiment. Second preparation play (Fig. 2) revealed that the player was able to hide the playing screen with gestures from the sub camera. Therefore, recording a screenshot-movie was also introduced into the main experiment.

Fig. 2 Video observation of the second preparation play with two video cameras

3.2 Experiment

3.2.1 Questionnaire

Before speech and behaviour observations, we conducted a pre-test questionnaire. This questionnaire was administered to gather basic information on the subjects. After the observations, we administered a post-test questionnaire to gather the subjects' impressions.

3.2.2 Test play including player observations

Speech and behaviour observations were conducted. During this observation, the subjects' speech data were gathered based on a method called the think-aloud method. The obtained data were analysed after labelling work. In the think-aloud method, players are asked to state out loud what they are thinking, doing, and feeling. On this point, the experimenter must be careful to clarify the experiment purpose for the subjects, which is not to test the player's game play skills but to test the product. If the experimenter does not do so, then the possibility exists that subjects will falsely believe that their own playing skills are being scrutinized rather than the quality of the product. Users would then be unwilling to report their feelings towards the product whenever they failed in completing the required tasks. Additionally, the subjects reported their impressions and thoughts under the task. Analysis results depend on these protocol data. Protocol data are what a subject has spoken during the observation. They are recorded in a time sequence. Before conducting protocol analysis, we set up our experimental objectives and defined the task for the subjects. Subsequently, we identified hypotheses, and analysed them using statistical analysis after quantization with labels. Experimental objectives are evaluations related to several factors constructing the game, as already described in the introduction.

3.2.3 Labelling and categorization

Categorization is adopted after going through labelling work for observed data. Items of categories are used as labels for statistical analysis.

Additionally, we recruited test subjects and prepared a test environment with suitable software and hardware. For this experiment, we prepared Japanese and English localized versions from Finnish instructions.

– *Experimental environment* Conditions in this experiment are as presented below:
  Core i7 3630, GeForce GT 540 M, 16 GB RAM, Windows 7 Kinect and a 25-inch display (Recorded screen capture during play [25]).
  Two cameras (One is set in front of player; another is set behind the player as shown in Fig. 2).
  The Glider test version (The test version did not have BGM and sound effects, etc.).
– *Task* Ten minutes of playing the prototype of The Glider was the task. Within five out of ten minutes, all the players had completed five levels. Some players had completed the sixth level. Therefore, we decided to analyse data obtained for the first to sixth levels.
  Contents of each level consist of the following motions:
  Level 1: by rotating the player's body repeatedly from right to left and back, the glider flies faster (the leftmost two motions in Fig. 2).
  Levels 2 and 3: by moving the body forward or backward, the glider goes down or up together with the rotation of the body (the leftmost three motion in Fig. 2).
  Level 4: by moving the body left or right the glider turns left or right together with the rotation of the body (the leftmost two motions and the rightmost motion in Fig. 3).

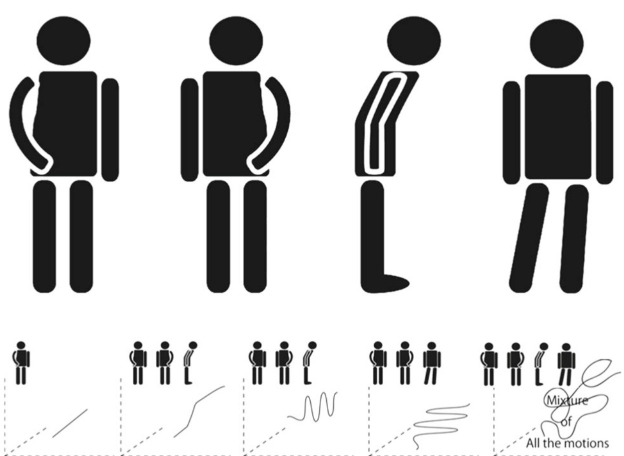

Fig. 3 Body motions used in The Glider are rotations, and movements *front–back* and *left–right* (*upper*). Movements required in each course (*lower*)

Fig. 4 Test subject participating a video observation in which behavior and speech were analysed

  Level 5: by combining these control methods, the glider can be controlled as a real glider.
  There are coins and circle-shaped gates in stages. All players were asked to gather those coins and gates to the greatest degree possible to promote them using many control motions during play.
– *Test subjects* Experiments were conducted with college students (Fig. 4).
  Twelve Japanese students (6 women, 6 men) The subject ages were 18–23 years. The 12 test subjects were students of Sendai National College of Technology. Their experiences of playing games varied from 1–15 years. Four students were not playing at all. The rest of the students were playing games from one up to seven hours per day. Seven students had some experience of using motion controllers in games.

4 Results

Here we present summaries of the experimentally obtained results, emphasizing the labelling work for protocol analysis.

4.1 Labelling work toward protocol analysis

We classified the observation results and emotive exclamations represented in three categories: *B*ehaviour (b), *E*motion (E) and *L*earning (L). *B*ehaviour includes the feelings of getting *T*ired (BT), *C*ontrolling the *V*ehicle (BCV) and *M*otor skills *F*ine/*C*oarse (BMF/BMC). *E*motion consists of *P*ainful feeling (EP), getting *B*ored (EB), getting *C*onfused (EC), *G*etting *F*rustrated (EGF), *M*otivation (EM), and *S*atisfaction (ES). The learning category (L) has no subcategories. To summarize, the features are labelled as follows:

– Behavior: BT, BCV, BM{F/C}.
– Emotion: EP, EB, EC, EGF, EM, ES.
– Learning: L.

Table 1 Example of a scripting file in the current experiment

Task	Time	Subject's speech	Labels	Observer's speech
6th level	4:05	Oh, it's hard	BMC, EM	
	4:10	Huh, moving sidewards is	BCV, EDF, BT	
	4:25	Do I have to gather coins?	EC	
	4:25			Yes, well, please stick to gathering as many coins and gates as you can
	4:25	Sense of modulating speed is not understandable. Well, it's vague how I accelerate	EC, EGF	

The subjects' speech is labelled based on their activities, atmosphere, and emotive-physiological responses with the labels explained above. Protocol data are examined and labelled in Japanese first, and are translated into English later. One scripting file is shown partially in Table 1. For example, the subject reported that "it's hard" while playing the sixth level, and the subject's emotion at this time is assumed to be a feeling of difficulty at control (BMC), but still motivated to play (EM). Those evaluations were inferred from a viewpoint based on the surrounding environment and situations of their activities, such as the playing state of the game monitor, controlling motions, and their outlooks of the emotive atmosphere recorded by two cameras, using the time tag of the protocol data.

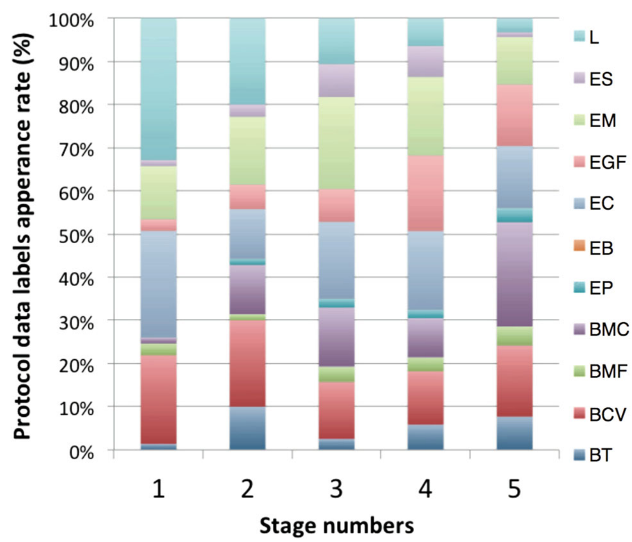

Fig. 5 Diagram describing the results of the analysed protocol data from the video observation

4.2 Cognitive bias evaluation

Results of the protocol analysis are shown in Fig. 5. The five bars show the average appearance rate of all the labels averaged over all the subjects in each stage. By the appearance rate, cognitive biases on some specific perspectives are analysed in this section. As an overview of the average appearance rate of all the labels, the game seemed not to overload the players physically because the test subjects were not getting so tired while playing because BT labels slightly appeared except at the second level. The labels BMF, EP, and EB almost did not appear through the experiments. Therefore, those labels do not provide us any effective information. As an assumption, the test subjects were not feeling motor skills exercises when we analyse their control ability (low BMF). However, according to appearance rates of other labels related to control (BMC, BCV), the subjects' good feelings related to good motor skills might have been included in ES/EM labels as described by "Yeah" or some other joyful exclamation. At this point, we should notice them clearly to separate those elements. Especially at the very beginning and at the end of the playing session, the subjects were losing control of the vehicle. Because the test subjects' fine motor skills were all quite weak at the time, and conversely, the coarse motor skills after the first level were quite strong, the game is apparently quite difficult to play even for young people.

We can make the following observations if the emotional categories are examined. The players were not getting bored during their play. Subjects were more or less confused at every level while playing. In addition, the subjects seemed to become increasingly frustrated. Simultaneously, their motivation was increasing during the first three levels; then it decreased quite dramatically in level 5. According to the learning category, test subjects were learning to play the game naturally mostly at the very beginning of the game. However, the players should learn while playing all the time. Now, after the first four levels, the learning factors seemed to decrease dramatically. The next classification was used to ascertain how comprehensible the game flow was. For example, when the players were motivated for a purpose, they will show us that the game is understandable or frustrating. It would be a sign of players' difficulties in understanding for the current game flow. The third category is ease: we classified labels also based on ease. For example, fine motor skill gave us information that game was easy to play or confusion showed

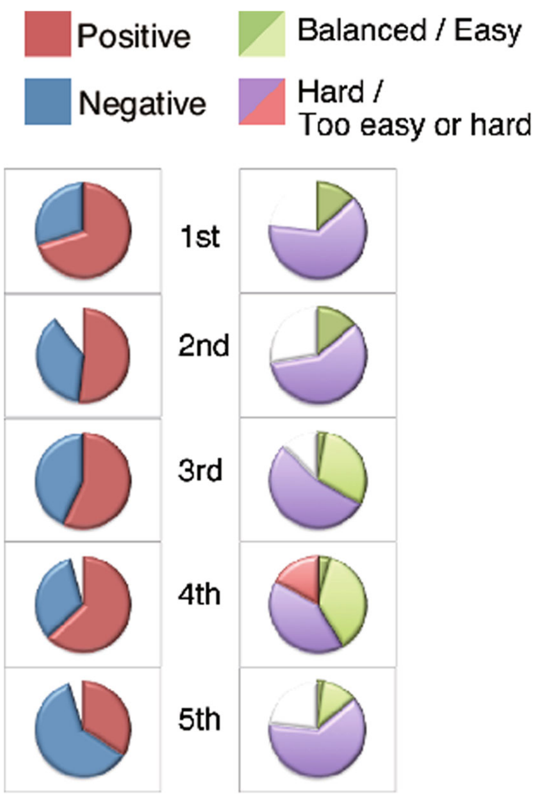

Fig. 6 Percentage based on classification in Table 2

Table 2 Classifying labels based on three points of view

	Points of view		
Labels	*Positive – Negative*	*Understandable*	*Hard – Easy to play*
BT	Negative	Mean	Mean
BCV	Positive	Good	Mean
BMF	Positive	Good	Easy
BMC	Negative	Bad	Hard
EP	Negative	Mean	Mean
EB	Negative	Bad	Too Easy or Too Hard
EC	Negative	Bad	Hard
EGF	Negative	Bad	Hard
EM	Positive	Good	Good game balance
ES	Positive	Good	Good game balance
L	Mean	Mean	Hard

us that the player had difficulties to play. Here is a short summary of the main findings. At the first level, the players seemed to control the vehicle well (BCV 21 %, BMC 1 %). That is to say, this level has an extremely simple structure to play. It does not frustrate the players (EGF 3 %). Merelyl by rotating the player's body repeatedly from right to left and back, the glider flies faster. Figure 6 is the result of classification (Table 2) shown in percentage; it also supports that point. At the same time, players were quite confused (EC 25 %) even though they had information related to the top left of the screen and the players were able to learn (L 33 %). Being confused at the very beginning indicates the necessity of providing more information about how to play before starting. At the next two levels, learning is decreasing (L 33 → 20 ∼ 11 %), which is reasonable because the amount of new things for the player decreases when the levels are gained. New things are occurring rarely after the first level, but each level has a new control motion. Therefore, coarse motor skills seem to appear, more or less (BMC 1 → 11 ∼ 14 %). At the same time, the players' motivation is increasing (EM 16–21 %). Overall, the game flow seems to build in these stages. According to their better understandings, decrement of learning labels led hard factors shift to balanced/easy in Fig. 6. However, the players felt tired in comparison with other levels (BT 10 %) in the second level. This fact implies that the motion to ascend and descend (Fig. 3) is a bit tiring even for young players. It should be modulated to be easier for elderly people. It is apparent that in the first three levels the players are feeling quite positive as shown in Fig. 5 (BCV 13–20 %, EM 12–21 %). It is actually increasing simultaneously when understandability is a bit decreasing because of new control motions. Especially, the third level has a good balance (EM 21 %, ES 8 %) in terms of understandability and difficulty (EC 18 %, EGF 8 %). The fourth level seems to keep the players motivated (EM 18 %). We can find that simultaneously the players' coarse motor skills are low (BMC 9 %). As Figs. 5 and 6 shows, this level has the best balance of understandability while players are practising turning left and right but does not include rising motions (EM 18 %, EGF 18 %). The fifth level, in turn, is apparently difficult to play with a low understandability. We can find these challenges also in the first diagram with high negative feelings.

4.3 Statistical evaluation

From a statistical viewpoint, chi-square tests and residual analysis are adapted to the result for eight frequently appearing labels in Table 3. The statistical views and statements below enrich the interpretability of the cognitive bias evalu-

Table 3 Label appearance frequency of 12 subjects and the result of residual analysis

	BT	BCV	BMC	EC	EGF	EM	ES	L
1st	1	15	1▽	18	2▽	9	1	24▲
2nd	7▲	14	8	8	4	11	2	14▲
3rd	5	26	27	35	15	42▲	15▲	21
4th	9	19	14	28	27▲	28	11	10▽
5th	7	15	22▲	13	13	10	1	3▽

▲ Significantly many $p < 0.05$
▽ Significantly less $p < 0.05$

ation above. At this point, the level of significance in the test was set as $p = 0.05$.

By using the residual analysis, we can know which labels remarkably appeared in each stage in the game.

For example, at the fourth level, the label EGF (Emotion, get frustrated) appeared significantly many and the label L (Learning) appeared significantly less. These emphasized labels suggest us that the players understand how to control the game, but they feel frustrations for the control movement that is newly installed at the fourth level therefore the difficulty increased compared to previous levels. Accordingly, to make some changes to the game toward lowering the appearance rate of EGF label in the fourth level in next version-up by adding an additional instruction level between the levels of three and four, or lowering motion detection threshold to make the control easier, etc., will prevent sudden increment of the play difficulty from previous levels and will reduce frustrations of players.

As described above, the analysis on the collected data in each level will provide orientations to improve the games. At this point, what game developers think improvement is depending on their development strategies. Further considerations to the obtained data follow; At the first level, a combination of learning (L) labels, emotion-get frustrated (EGF), and the behaviour-motor skill coarse (BMC) are emphasized. Results show that the first level is easy to play without stress. Only at the second level is the label behaviour-tired (BT) emphasized, which implies that the required motions in the second level have high priority compared with enjoyableness. In spite of the required motions, the third level resembles the second level. The emphasized emotion-motivation (EM) and emotion-satisfaction (ES) shows a better balance of enjoyableness. In the fourth level, new motions to turn the glider in lateral directions are demonstrated. Several test players described that in those motions, it was difficult to follow the instructions. Consequently, controllability of the motion or comprehensibility of the instruction in this level should be considered. In the fifth level, behaviour-motor skill coarse (BMC) is emphasized and learning (L) is significantly less. This result leads us to speculation that difficulty factors are stronger than enjoyableness.

5 Conclusion

As described in this paper, we have emphasized the usability evaluation of a serious game using evaluation methods based on the Kansei engineering concept. The 12 Japanese test subjects were asked to play the prototype of The Glider while reporting their thoughts and impressions. Two cameras recorded the proceedings. Subsequently we labelled the observed data and created categories. We conducted protocol analysis and confirmed the result both in cognitive bias evaluation and statistical analysis. Those analysis results are used to evaluate and construct a hypothesis with an orientation to improvement.

To be concrete, the results showed us that the first 3–4 levels were quite playable. The start of the game is extremely simple, with just one control motion per level. The players were learning exercises a lot while playing the first levels. A player should receive more information before starting the game. It implies a greater need for instruction in the game. Some motion control moves developed for rehabilitation naturally increased the complexity of the game control, which is apparent in challenges related to the subjects' states labelled as coarse motor skills (BMC) while they were playing. At the same time, when players were still motivated, the game flow looked promising: The game seemed to attract the players' attention to a great degree. The players were interested in playing and learning new things. Furthermore, understandability and difficulty were apparently in balance. After the first four levels, the players seemed to lose some motivation for the game. The fourth level is quite promising, with a high motivation level. These four levels iteratively remind the players how to control the glider in acceleration, during ascent or descent, and turning left or right. These results imply that more complex steering motions are needed. However, the next levels show that the players are not any more able to control left and right. The game flow is apparently in a balance between difficulty and usability from the beginning until the fourth level. Currently, research organizations are extremely active in this serious game development field. Ageing is a challenge not only for Japanese society, but also for Western societies such as in Finland. Accordingly, we believe that such rehabilitation concepts will be demanded continuously. In future studies, we intend to conduct similar usability testing for interaction techniques of different types.

Acknowledgments The authors thank all participants in the game development, doctors and physiotherapists in Carea, developers at Serious Games Finland, and students who helped us conduct experiments. Additionally, we would like to thank Mr. Fusao Sato, Director Hiroshi Yoshimura from the Sendai—Finland Well-being Center, and professors of the Sendai National College of Technology for their fruitful cooperation.

References

1. Nagamachi M (1995) Kansei engineering: a new ergonomic consumer-oriented technology for product development. Int J Ind Ergon 15(1):3–11
2. Pranav M, Pattie M, Liyan C (2009) WUW-wear Ur world: a wearable gestural interface. In: CHI EA 09 extended abstracts on human factors in computing systems, pp 4111–4116
3. http://www.ted.com/talks/pattie_maes_demos_the_sixth_sense. Accessed 13 June 2014

4. Choh N et al. (2008) Demonstration of interactive teleolfaction with movie. In: Proc. ACE 08 the 2008 international conference on advances in computer entertainment technology, pp 395–395
5. Allan O, Regina BA (2012) Creation and visualization of context aware augmented reality interfaces. In: Proc. AVI 12 the international conference on advanced visual interfaces, pp 324–327
6. Buchenau M, Fulton J (2000) Experience prototyping. In: Symposium on designing interactive systems 2000 (Brooklyn, NY). ACM Press, Neu York, pp 424–433
7. Leo KH, Tan BY (2010) User-tracking mobile floor projection virtual reality game system for paediatric gait and dynamic balance training. In: Proc. iCREATE10 the 4th international convention on rehabilitation engineering and assistive technology (article no. 25)
8. Theng YL, Chua PH, Pham TP (2012) Wii as entertainment and socialization aids for mental and social health of elderly. In: Proc. CHI12 extended abstracts. ACM, New York, pp 691–702
9. Matsuguma H, Fujioka S, Nakajima A, Kaneko K (2011) Standing growing blooming trees rehabilium-serious game for rehabilium. In: Asia digital art and design association
10. Biswas KK, Basu SK (2011) Gesture recognition using microsoft Kinect. In: Proc. 5th international conference on automation, robotics and applications (ICARA)
11. Thibaut W, Sofien B, Hao L, Mark P (2011) Kinect-based facial animation. In: Proc. SA 11 SIGGRAPH Asia 2011 emerging technologies (article no. 1)
12. Wu D, Zhu F, Shao L, Zhang H (2012) One shot learning gesture recognition with Kinect sensor. In: Proc. MM 12 the 20th ACM international conference on multimedia, pp 1303–1304
13. Chao S, Bing-Kun B, Changsheng X, Tao M (2012) Kinect-based visual communication system. In: Proc. the 4th international conference on internet multimedia computing and service, pp 55–59
14. Xia L, Chen CC, Aggarwal JK (2011) Human detection using depth information by Kinect. In: International workshop on human activity understanding from 3D data in conjunction with CVPR (HAU3D). Colorado Springs CO
15. Huotari K, Hamari J (2012) Defining gamification a service marketing perspective. In: Proc. the 16th international academic MindTrek conference, pp 17–22
16. Burke JW et al. (2010) Designing engaging, playable games for rehabilitation. In: International conference series on disability, virtual reality and associated technologies (ICDVRAT)
17. Chang YJ, Cheng SF, Huang JD (2011) A kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. Res Dev Disabil 32:2566–2570
18. Anderson F, Annett M, Bischof WF (2010) Lean on Wii: physical rehabilitation with virtual reality and wii peripherals. Ann Rev CyberTherapy Telemed 8:181–184
19. Pyae A, Yuen TB, Gossage M (2013) Understanding stroke patients needs for designing user-centered rehabilitative games. In: Proc. the 7th annual international conference on computer games multimedia and allied technologies, pp 151–156
20. Shiizuka H, Hashizume A (2011) The role of Kansei/affective engineering and its expected in aging society. In: Watada J et al (eds) Intelligent decision technologies, SIST10. Springer, Berlin, pp 329–339
21. KEIO Univ. Fukuda Lab, Ergonomics guide- the way to study Kansei scientificaly, Scientistsya, p 354 (ISBN 10-4860790367)
22. Kaiho H, Harada E, Protocol analysis what would you read from speech data, Sinyosya, p 249 (ISBN 10-4788504693)
23. Csapo A, Baranyi P (2012) A unified terminology for the structure and semantics of CogInfoCom channels. Acta Polytech Hungar 9(1):85105
24. Levy L, Novak J (2009) Game development essentials QA & Testing. Course Technology. ISBN 10-1435439473
25. http://www.bandicam.com/. Accessed 27 Aug 2015