

# Acting Evaluation As An Assessment Method Using Participant's Imagination

Yoshie Kiritani

*Chiba University, kiritani@faculty.chiba-u.jp*

**Abstract:** The present study examined a new evaluation method called as acting evaluation to make clear its features. Acting evaluation was a way of forming an ideal attitude of participants toward the task, which required them to imagine and to pretend the practical object person in the real evaluation. In the present study, they were schoolteachers who would be users of an instruction manual to ask a special class that a veterinary university went to elementary schools and conducted with real dogs. A prototype of the manual was prepared. Twenty-six undergraduate students were divided into 2 groups: the experimental group which adopted acting evaluation and the control group. Both group read the prototype, planned a class and answered questionnaires. As a result, 2 groups showed similar tendencies of preference for dogs and of impression about the special class and planned similar proper classes. However, the differences between 2 groups were seen in the required time for the task and its understanding. Acting evaluation could be a simple instruction method to make the participants concentrate on their task and be suitable for preliminary research. Difference between acting evaluation and other methods in design and in psychology using imagination was also briefly discussed.

**Key words:** *Evaluation, method, forming of attitude, communication design*

## 1. Introduction

It is difficult to recruit participants for the experiments dealing with more practical issues. For example, A psychological study [1] aimed to investigate effects of positive feedback by supervisor in workplaces but used undergraduate students, not employed persons, as participants. It was like a simulation and would be examined in more practical way. We can imagine how the researchers had had a difficulty to recruit 80 participants who were working people. Thus, the study would be easy, if it could use students of classes as participants or evaluators.

If the number of participants or evaluators is large or if they are a specific kind of people, studies will come up with a difficulty. At least, it would be useful to get a method to do preliminary research more easily. A previous study [2] proposed a new evaluation method for basic or preliminary researches. It was called as acting evaluation. Acting evaluation requires imaging of participants; in the previous study, they had to pretend to be schoolteachers and answered the questionnaire, although in reality the participants were graduate students. More specifically, they were asked to put themselves in the position of schoolteachers and to answer some questions. The study was about evaluation of a proposal, so that the understanding of task and the usability of proposal were estimated from the results obtained by acting evaluation. They were apparently legitimate outcomes, partially because the

participants were design-students who would be imaginative and have the power of empathy. However, the results were not compared with the results by another evaluation method in the previous study.

Purpose of the present study was to examine a new evaluation method called as acting evaluation to make clear its features. The previous study did not prepare a control group so that the present study compared the assessment by the acting evaluation with that by normal questionnaires. Effect to be expected, procedure relative to similar methods, and range of application of the acting evaluation would be discussed. What was evaluated was a manual that would be for schoolteachers to ask a veterinary university for a special class.

## **2. Experiment**

### **2.1 Method**

#### **2.1.1 Participants**

Firstly, 31 students of design major were divided into 2 groups, 16 for experimental group and 15 for control one. However, 5 students of them abandoned their task on the way and, finally, 26 students (14 males and 12 females, from 20 to 22 in age) participated in the experiment. Eight males and 7 females were in the experimental group and 6 males and 5 females were in the control group.

#### **2.1.2 Stimulus**

It was a prototype of a manual that the participants evaluated. The manual explained how to ask the Laboratory of Effective Animals for Human Health, Educational Research Center for Anthrozoology (ERCAZ) at Azabu University for a class using dogs [3]. Azabu University was one of veterinary universities in Japan. ERCAZ offers some classes using animals for children as animal-assisted education (AAE) [4]. The present study prepared a manual for a class in which dogs and trainers visited elementary schools. Eventually, schoolteachers would use the manual to ask ERCAZ for the visiting class. This manual could give schoolteachers an image of the visiting class and save ERCAZ's staff the time and effort of the initial step of communication.

The prototype was a manual modified from the version used in the previous study [2] which had revealed 4 problems, 1) failure of color coding, 2) that of layout, 3) inconspicuous of important information and 4) lack of information. The prototype used in the present study took action on these issues: 1) rejecting color-coded chapters but using colors to differentiate the contents of class, 2) using left justification with illustration in the right part, 3) changing presentation order of information and deleting wider information about AAE or dogs but specializing in ERCAZ's activities in the last chapter, and 4) increasing information about the class and adding illustration for supplemental information.

The prototype had a total of 6 chapters, 48 pages in total, containing "Chapter 1, Introduction", "Chapter 2, Preparation for the class", "Chapter 3, Contents of the class", "Chapter 4, About education support dogs", "Chapter 5, For preliminary discussion", and "Chapter 6, Information". Chapter 1 contained "What is the special class with dogs?", "What is this manual?" and "Procedure of the class". Schoolteachers as the future but proper users would grasp and order the visiting class by reading the first and the second chapters where an outline of the class and the procedure of order were explained.

Moreover, Chapter 5 contained an arrangement sheet for preliminary discussion in which schoolteachers would write their contact information, preferred day and time, class's goal as free description, required contents of the class as multiple choice, expectations in the class as multiple choice, other comments, and acceptance of shooting or filming.

### 2.1.3 Procedure

All of the participants read the prototype, perform the task of planning a class using the prototype and evaluated it according to the instruction. Concretely, they completed the arrangement sheet to plan a class while they read the prototype. After that they answered the questions in Table 1 to clarify how to plan the class and to evaluate the prototype. They also reported an impression about the ERCAZ's activity, preference for dogs and time required for the task and the answers.

The instruction for the experimental group was as follows. "You are a second grade teacher in an elementary school. The principal told you about a visiting class with dogs by a laboratory in Azabu University and required you to plan a class. The number of children in your class is 30. You have to plan a 80 minute class." After this instruction, the participants planned a class and answered the question in Table 2. This was acting evaluation in which the participants pretended the proper user of the prototype, schoolteachers. But, the participants were required to quit their role, when they answered the impression about the ERCAZ's activity, preference for dogs and time required for the task and the answers.

The participants in the control group was just instructed to plan a 80 minute class for 30 children of the second grade of elementary school according to the prototype. They were not required to pretend a teacher. All of questions they answered were the same as those for the experimental group.

The participants could do the experiment at their own pace and reported the time needed in Question 12.

Table 1. Questions that all of participants responded

No.	Questions	Kinds of question
1	What kind of lesson do you want to give?	About planning of the class
2	Do you have any worries in this lesson?	
3	What do you think is the role of schoolteacher in this class?	
4	Which part of the manual did you refer to write the goal in this class?	
5	Which part of the manual did you refer to write the strong expectation in this class?	
6	What is the most read part in the manual during the planning?	
7	What kind of information was insufficient for the planning?	About evaluation of the prototype (manual)
8	In which points did the manual was hard to use?	
9	What was the visual problem in the manual?	
10	What did you frankly think about ERCAZ 's activity, "a special class with dogs"?	
11	Do or don't you like dog?	
12	Please tell us how long it took for you to complete the task?	

## 2.2 Results

Difference in the preference for dogs (the answer to Q11) between the experimental group and the control group was not shown ( $\chi^2(3)=3.67$ , *n.s.*, *Phi*=.375). Difference in the impression about ERCAZ's activity (the answer to Q10) between two groups was not shown, neither ( $\chi^2(1)=.51$ , *n.s.*, *Phi*=.109). Thus, these 2 groups were equal in quality and did not propose a critical issue to perform the present study.

### 2.2.1 Results of the classes planned by 2 groups

The times required for the planning (the answer to Q12) significantly tended to be different in 2 groups ( $F_{1,25}=4.02$ ,  $.05 < p < .10$ ); the average time by the experimental group was 55.0 minutes (S.D. 14.8), while it was 72.3 minutes (S.D. 28.8) for the control group. Thus, the participants who did acting evaluation could finish the task earlier than those in the control group did.

There were no differences in the results between 2 groups about the set of goal ( $\chi^2(4)=1.92$ , *n.s.*, *Phi*=.122), about the required contents ( $\chi^2(4)=1.08$ , *n.s.*, *Phi*=.122), and about the expectations in the class ( $\chi^2(4)=.72$ , *n.s.*, *Phi*=.105). In both groups, over 40 % of the participants set a goal as learning, understanding or knowing. In both groups, over 40 % of the participants wanted the class about how to greet dogs. As for the expectations, trigger for the development of respect for life or that for consideration toward others reached the top 2 for both groups (36.1 % for respect for life and 38.9 % for consideration in the experimental group, 34.5 % for the former and 34.5 % for the latter in the control group).

Although there was no difference in the contents of classes planned by both groups, the time required for the task tended to be different.

### 2.2.2 Results of the questionnaire about planning of the class (from Q1 to Q6)

These results were about how the participants had planned the class. There were no differences in the results between 2 groups about the wanted class ( $\chi^2(2)=.035$ , *n.s.*, *Phi*=.024), about the worry ( $\chi^2(2)=1.48$ , *n.s.*, *Phi*=.215), about the role of schoolteacher ( $\chi^2(2)=3.04$ , *n.s.*, *Phi*=.227), about the reference pages to set the goal ( $\chi^2(5)=6.56$ , *n.s.*, *Phi*=.345), about the reference pages to write the expectation for class ( $\chi^2(6)=4.79$ , *n.s.*, *Phi*=.326), and about the most frequently read part ( $\chi^2(4)=5.45$ , *n.s.*, *Phi*=.373).

The participants in both groups wanted a class to interact with dogs or to learn the respect for life. It was the safety of children that the participants in both groups worried about. They considered that the role of schoolteacher was management of the classroom or deepening of the lesson. The most referred chapter to set the goal and to write the expectation for class was Chapter 3 that was about contents of the class. The most frequently read part was Chapter 3, too.

However, there were some differences between the experimental group and the control group, when the data were analyzed in detail. As for the role of schoolteachers, 28.6 % of the participants in the experimental group and 22.9 % of them in the control group commented the teaching. But breakdown of the data highlights that the participants in the control group were significantly more likely to consider the role as just a trigger than those in the experimental group ( $\chi^2(2)=6.56$ ,  $p < .005$ : Figure 1). As for the most frequently read part during the task, it was Chapter 3 as mentioned above, but the ratios of reference were different in 2 groups (Figure 2). In the experimental group ( $\chi^2(5)=35.19$ ,  $p < .01$ ), Chapter 3 was referred significantly more than Chapter 1 ( $p = .0032$ ),

Chapter 4 ( $p=.0008$ ), and Chapter 6 ( $p=.0032$ ); there were no significant differences between Chapter 3 and Chapter 2 and between Chapter 3 and Chapter 5. On the other hand, in the control group ( $\chi^2(5)=55.20, p<.01$ ), Chapter 3 was significantly the most referred past in all chapters: than Chapter 1 ( $p=.0002$ ), Chapter 2 ( $p=.0002$ ), Chapter 4 ( $p=.0002$ ), Chapter 5 ( $p=.0012$ ), and Chapter 6 ( $p=.0094$ ). Thus, the participants in the experimental group read not only Chapter 3 (“Contents of the class”) but also Chapter 2 (“Preparation for the class”) and Chapter 5 (“For preliminary discussion”), while those in the control group referred to just Chapter 3 to do the task.

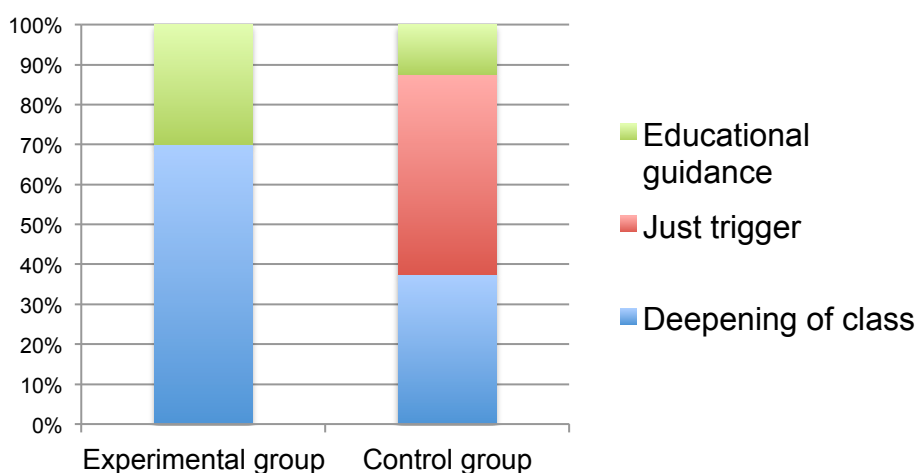


Figure.1 Answers to the role of schoolteacher

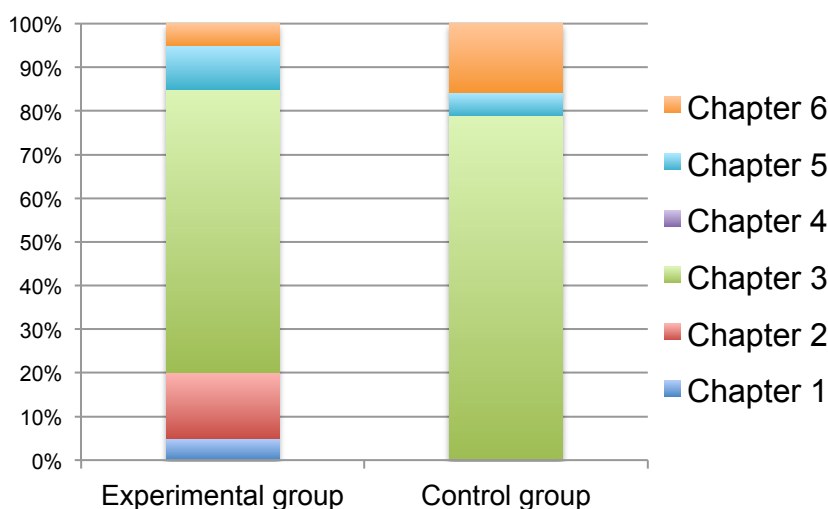


Figure 2. Answers to the most frequently read part

### 2.2.3 Results of evaluation of the prototype (from Q7 to Q9)

As for the missing information, there were 4 categories set by free description: “about the preparation of class”, “about relevant people”, “not missing information but structural problems”, and “nothing”. Although frequencies in these categories in 2 groups were not significantly different, the frequencies in each group were not equal (Figure 3). For the experimental group, they presented a significant tendency ( $\chi^2(3)=6.44, .05<p<.10$ ), but for the control group they were significantly different ( $\chi^2(3)=12.86, p<.01$ ). In other words, the participants in the

experimental group seemed to report “nothing” more frequently and those in the control group significantly reported that the information about the preparation or follow-up of class was insufficient.

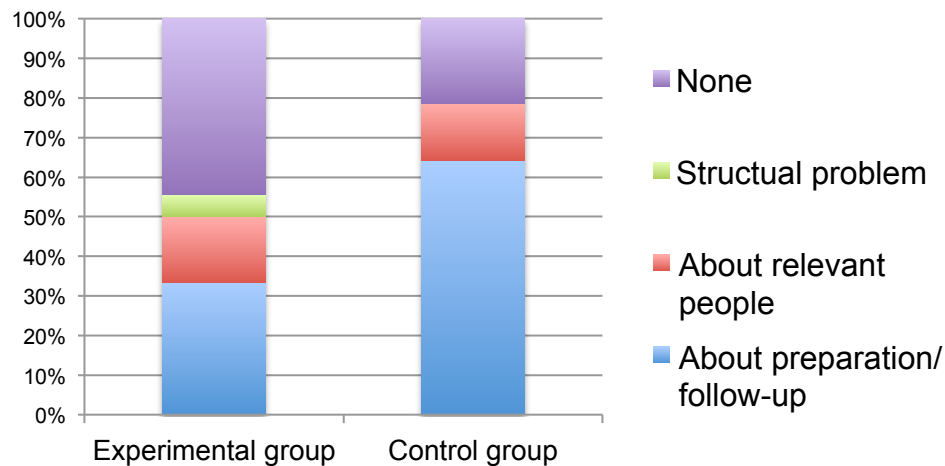


Figure 3. Answers to the missing information

As for the difficulty for using, 2 groups presented a significant tendency of the difference ( $\chi^2(3)=6.85$ ,  $.05 < p < .10$ ; Figure 4). A residual analysis revealed it in detail. In the experimental group, the participants significantly complained about the volume more than about the how-to-use ( $p < .05$ ). Six of them considered the prototype as redundant, but 4 of them felt the lack of information. On the other hand, in the control group, they complained about the how-to-use more than about the volume. Thus, the participants in the control group might not understand the contents of the prototype and the task that they had to do in the experiment.

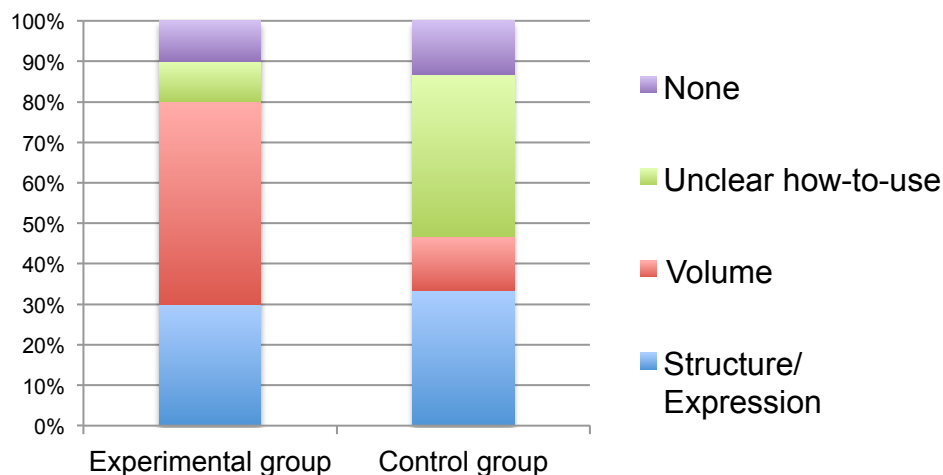


Figure 4. Answers to the difficulty to use

As for the problem about the visual expression, the frequencies showed a significant different tendency in 2 groups ( $\chi^2(3)=7.61$ ,  $.05 < p < .10$ ; Figure 5). A residual analysis revealed that the participants in the control group complained about the contents of the prototype more than those in the experimental group ( $p < .05$ ); no one pointed out the problem of contents as problematic points about the visual expression in the latter group. In this group, the difference of frequency was significant ( $\chi^2(3)=18.36$ ,  $p < .01$ ); the invisibility was significantly reported more than the contents ( $p = .0026$ ) and other problem ( $p = .0094$ ), and the no-problem was significantly reported more than

the contents ( $p=.0044$ ) and other problem ( $p=.0156$ ). The Q9 was not about the contents of the prototype, but about its appearance. Although the participants in the experimental group properly pointed out the problem or did not indicated any problem, those in the control group complained about the contents here again. It was suspected that they did not understand the question or the task, again.

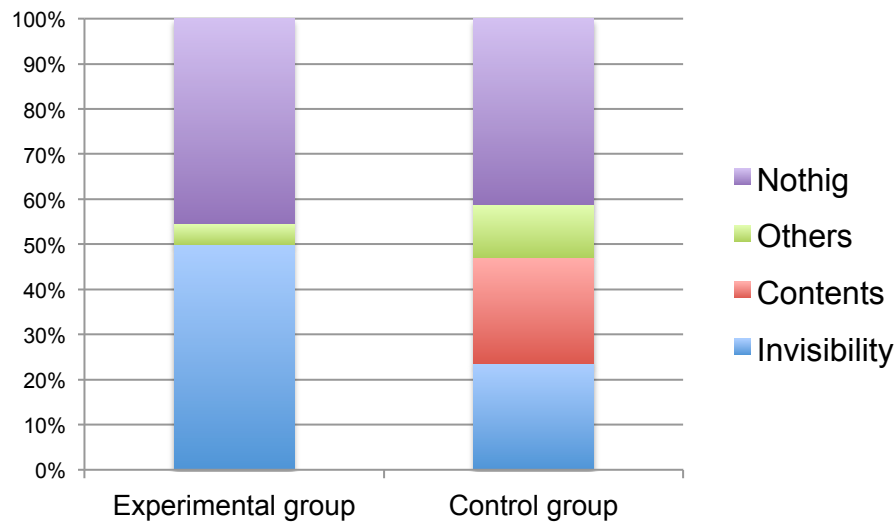


Figure 5. Answers to the problem about the visual expression

## 2.3 Discussion

In the experiment, both groups did the same task, but the participants in the experimental group had been instructed to pretend a schoolteacher in doing the task. Apparently, there was no difference in the results of tasks by 2 groups; both of them could plan a class using the prototype without problems. However, it took more time to complete the task in the control group than in the experimental group. Moreover, the questionnaire revealed a risk that the control group did not understand the task. Thus, the participants in the experimental group who did acting evaluation finished the task more speedy and more adequately. Acting evaluation could increase the participants' concentration, which can be an expected effect. It can be suitable for basic researches that have difficulties to recruit the real evaluators.

Now, we would like to get acting evaluation straight once and for all.

Firstly, since acting evaluation, in short, is usage of person's images, it may be reminiscent of personas. This is a strong tool in design [5, 6 as examples of review], which is often introduced as "persona marketing" in less academic but more practical journals [7,8, for example in Japan]. As more psychological techniques, psychodrama and role-playing should be mentioned in this context. However, acting evaluation is different from these methods. It can be summarized that personas are virtual users created by designers or marketers to understand their product or service. Usually, personas are made by more than one member of design or market team to share the images of users. Acting evaluation is an individualized task, which does not require the sharing of created images. Participants in acting evaluation can freely imagine the person instructed by researchers. The difference from some psychological techniques is also in the number of relevant player. Psychodrama is an improvisation and role-playing is originally one of technique of psychodrama [9], which are usually played by more than one. All of

techniques mentioned here need the ability to imagine or the creativity. But acting evaluation is performed individually.

Its procedure is very simple. Researchers just ask participants in the experiment or in the evaluation for imaginary acting performance. The participants image and do the task or answer questionnaires. The present study gave the participants an instruction to image a schoolteacher who taught the second grade of elementary school and his/her mission from the principle. The instruction did not define teacher's gender, age, teaching history and so on. A further study should examine the effect of amount of information for imagination.

Acting evaluation requires the participant's imagination, which may be more than confirmed by their life experience. Pretending working people may be difficult for freshmen who have no experience of working, for example. However, sympathism may help the imagination. There may be someone who is good at acting evaluation and others who is not. A further study will make clear the relation between the effectiveness of acting evaluation and the participants' personality or tendency of behavior.

### 3. Conclusion

The research question of the present study was rather simple. How are the results of acting evaluation different from those by ordinal procedure? We revealed that acting evaluation could increase the participants' concentration to their task.

However, acting evaluation was not a method of usability evaluation as a current hot topic [10, 11]. The present study proposed a procedure for questionnaire to get more reliable results, but did not revealed what kinds of question were effective. Acting evaluation prepares participants for the judgment. This method seems trivial, but makes them conscious of what they should do in the experiment or in the evaluation. Imagination, "if I were the person", can raise the participants' awareness of their role in the experiments.

Further research should reveal the effective instruction to make participants sympathize and imagine and the range of application.

### 6. Examples Citations

- [1] Yamaura, K., Horishita T., and Kanayama, M. (2013) *Positive feedback is not fully effective in all situations*, The Japanese Journal of Psychology, vol. 83, no. 6, pp 517-525.
- [2] Authors. (2012) *A test of acting evaluation to assess a proposal of communication support study*, Bulltin of Japanese Society for the Science of Design, vol. 59, no. 1, pp 31-38.
- [3] <http://ercaz.jp>.
- [4] Turner, D.C. (2009) *Anthrozoology – now past its juvenile years! Change by Anthrozoology in Japan*, vol. 1, pp 4-7.
- [5] Turner, P., and Turner, S., (2011) *Is stereotyping inevitable when designing with personas?* Design Studies, vol. 32, pp 30-44.
- [6] Massanari, A. L. (2010) *Designing for imaginary friends: information architecture, personas and the politics of user-centered design*, New Media & Society, vol. 12, no. 3, pp 401-416.



- [7] Authorless (2007) *Possibilities of persona marketing: making “a person” from diverse values*, Nikkei Information Strategy, vol. 16, no. 9, pp 46-48. (In Japanese)
- [8] Ueki, T. (2010) Economic new trends, “Persona” marketing: selling to just a person, Nikkei Business, vol. 1548, pp 86-89. (In Japanese)
- [9] Nakajima, y., Ando, K., Koyasu, M., Sakano, Y., Shigemasu, K., Tachibana, M., and Hakoda, Y. (1999) The Yuhikaku dictionary of psychology, Yuhikaku: Tokyo. (In Japanese)
- [10] Filippi, S., and Barattin, D. (2012) *Generation, adaptation, and tuning of usability evaluation multimethods*, International Journal of Human-Computer Interaction, vol. 28, pp 406-422.
- [11] Hornbæk, K., *Dogmas in the assessment of usability evaluation methods*, Behaviour & Information Technology, vol. 29, no. 1, pp 97-111.