ACTING EVALUATION AS AN ASSESSMENT METHOD USING PARTICIPANT’S IMAGINATION

Yoshie Kiritani

Chiba University Yayoi-cho 1-33, Inage-ku, Chiba 263-8522, Japan

Abstract: The present study examined a new evaluation method called as acting evaluation. It 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 the schoolteachers who would be users of an instruction manual. Twenty-six undergraduate students were divided into 2 groups: the experimental group which adopted acting evaluation and the control group. Both groups read prototypes of the manual, planned a class and answered questionnaires. As a result, 2 groups seemed to do the task similarly. 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.

Keywords: Evaluation method, forming of attitude, communication design

1. Introduction

This is a further study of the previous one that focused on communication support between specialist and non-specialist and proposed a new evaluation method of design proposals [1]. The present study is to systematically examine this new method, acting evaluation. Acting evaluation is used to prepare participants for the evaluation of design proposal. However the results of evaluation themselves are not concerned here. The present study was conducted to show that acting evaluation would be useful in the preliminary experiment that did not recruit the real user of design.

It is sometimes difficult to recruit participants for the experiment that deals with more practical or specific issues. For example, a psychological study [2] 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 a preliminary study. We can imagine how the researchers had a difficulty to recruit 80 participants who were working people. Thus, that study might use students as participants, because it would be a pilot study.

If the number of participants is large or if they are a specific kind of people, the studies will come up with a difficulty. If the researchers are released from the recruit of participants, they can do preliminary research more easily. Acting evaluation examined in the present study is one of solutions. This is a way of instruction whose experiment alternatively uses students as participants. Acting evaluation requires the participants their imagination; in the experiment, the participants pretend the real or eligible person for the study. The participants think about the position, the role or the work of the designated person, image them and act them. Take the previous study [1] for example.

The previous study was a case of communication support study and proposed a medium that helps the communication between university’s staff and schoolteacher. University’s staff were specialist on animal-assisted-education (AAE) [3]. The schoolteachers
were non-specialists in this case and clients of an activity of AAE. A booklet as medium that helps the schoolteachers to understand the activity was a design proposal and should have been evaluated. Students, not schoolteachers, participated in the evaluation and answered the questionnaire. Participants pretended schoolteacher according to acting evaluation; they were asked to put themselves in the position of schoolteachers and to answer the questions. The results showed that the proper images of the AAE activity were formed, that the preparation for the class and the role of schoolteachers were well understood, and that some design problems of the booklet were revealed. In the evaluation, the participants had no trouble acting schoolteachers. And the results seemed natural; for example, one of participants referred to the curriculum guidelines set by the Ministry of Education, Culture, Sports, Science and Technology.

But the question now arises: were these results due to acting evaluation? The previous study could reach to these results, if it did not use acting evaluation. The previous study did not compare the results with those by another evaluation method. Purpose of the present study was to examine this new evaluation method, 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 and procedure relative to similar methods would be discussed. What was evaluated was a booklet of AAE that the schoolteachers would use to ask a veterinary university for a special class. This evaluated object and its results themselves are not essential for the purpose of the present study. But the differences between the results by acting evaluation and those by the normal questionnaires will be examined.

If the present study shows the effect of acting evaluation, it will represent a route for more reliable preliminary experiments that use students as a substitute for the real participants. Although the eligible person should be used as participants, some constraints do not allow it at the first stage of study. Using students as participants is a realistic approach for preliminary experiments.

2. Experiment
2.1. Method
2.1.1. Participants
At the beginning, 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. Thus, 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. The participants in the experimental group do acting evaluation. All of students in both groups did not know the bellow-mentioned AAE activity before and got a book card of 1,000 yen as a reward.

2.1.2. Stimulus
It was a prototype [4] of a booklet that all of participants evaluated; Prototype in this case is a trial model to examine its design. The booklet explained how to make a request the Laboratory of Effective Animals for Human Health, Educational Research Center for Anthrozoology (ERCAZ) at Azabu University for a class using dogs [5]. Azabu University was one of veterinary universities in Japan. ERCAZ offered some classes using animals for children as AAE [3]. The present study prepared a booklet type of manual for the class in which dogs and trainers visited elementary schools. Eventually, schoolteachers would use this booklet to ask ERCAZ for the visiting class. This booklet could give the schoolteachers an image of the visiting class and save ERCAZ staff’s the time and effort of the initial step of communication.

The prototype was a booklet modified from the version used in the previous study [6]. 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”.

The 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 with ERCAZ. The arrangement sheet was a kind of order sheet in which the schoolteachers would write their contact information, preferred day and time, class’s goal, required contents of the class, expectations in the class, other comments, and acceptance of shooting or filming.

2.1.3. Procedure
All of the participants read the prototype, performed 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, they answered the questions in Table 1 to clarify how to plan the class and how 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.

As for making the arrangement sheet, the participants set the class goal as free description. They choose the class contents from the examples shown in the booklet, and expectations that ERCAZ had set in the arrangement sheet.

Questions from 1 to 3 were set to know how the participants understood this visiting class. Questions from 4 to 6 were set to know how they used the booklet. Questions from 7 to 9 were for the evaluation of booklet. Questions 10 and 11 were set to confirm whether the preference for dogs affected the results of above questions. Question 12 was to confirm the required time. Questions except questions 11 and 12 were free description; for the data collecting, key words were picked up, categorized and counted in each questions [7]. Chosen answers were counted and calculated. Question 11 was the choice between the four, “like”, “mostly like”, “mostly dislike” and “dislike”. Question 12 would give the quantitative data.

Table 1. Questions used in the present study

<table>
<thead>
<tr>
<th>No</th>
<th>Question</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What kind of lesson do you want to give?</td>
<td>About planning of class</td>
</tr>
<tr>
<td>2</td>
<td>Do you have any worries in this lesson?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>What do you think is the role of schoolteacher in this class?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Which part of the manual did you refer to write the goal in this class?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Which part of the manual did you refer to write the strong expectation in this class?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>What is the most read part in the manual during the planning?</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>What kind of information was insufficient for the planning?</td>
<td>About evaluation of the prototype</td>
</tr>
<tr>
<td>8</td>
<td>In which points did the manual was hard to use?</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>What was the visual problem in the manual?</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>What did you frankly think about ERCAZ’s activity, “a special class with dogs”?</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Do or don’t you like dog?</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Please tell us how long it took for you to complete the task?</td>
<td></td>
</tr>
</tbody>
</table>

The purpose of present study was to confirm the difference between the results by acting evaluation and those by the normal questionnaires. Acting evaluation was conducted by an experimental instruction. Thus, the instructions for both groups were important.

The instruction for the experimental group prepared the participants for the mind of schoolteachers and was, to be more precise, 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 1. 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 (Question 10), preference for dogs (Question 11) and time required for the task and the answers (Question 12).

On the other hand, 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.

All of the participants in both groups could do the experiment at their own pace and reported the time needed in Question 12.

2.2. Results

The present study considered that the experimental group and the control group were equal in quality. For reference, difference in the preference for dogs between 2 groups (Q11) was not significant, although the effect size [7,8] was medium ($\chi^2(3)$=3.67, $n.s., Phi$=.375). Difference in the impression about ERCAZ’s activity between two groups (Q10) was not significant and the effect size was small ($\chi^2(1)$=.51, $n.s., Phi$=.109).

2.2.1. Results of the classes planned by 2 groups (required time and completion of arrangement sheet)

The times required for the planning (Q12) tended to be significantly different in 2 groups and the effect size was large ($F_{1,25}$=4.02, .05<$p<.10$, $\eta^2$=.14); the average time by the experimental group was 55.0 minutes (95% CI [46.8, 63.2]), while it was 72.3 minutes (95% CI [53.0, 91.6] ) for the control group. Thus, there was a tendency that the participants who did acting evaluation finished the task earlier than those in the control group did. At least, there was a great difference among the participants in the control group about the required time, because the standard deviation of the experimental group was 14.26 and that of the control group was 27.42.

About the completion of arrangement sheet, there was
no difference between the experimental group and the control group. Concretely speaking, there were no differences in the results between 2 groups about the set of goal ($\chi^2(4)=1.92$, n.s., $\text{Phi}=0.122$), about the required contents ($\chi^2(4)=1.08$, n.s., $\text{Phi}=0.122$), and about the expectations in the class ($\chi^2(4)=0.72$, n.s., $\text{Phi}=0.105$). 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 34.9% 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 (Q1: $\chi^2(2)=0.035$, n.s., $\text{Phi}=0.024$), about the worry (Q2: $\chi^2(2)=1.48$, n.s., $\text{Phi}=0.215$), and about the role of schoolteacher (Q3: $\chi^2(2)=3.04$, n.s., $\text{Phi}=0.227$). 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.

There were not significant differences between 2 groups but the effect sizes were medium about the reference pages to set the goal (Q4: $\chi^2(5)=6.56$, n.s., $\text{Phi}=0.345$), about the reference pages to write the expectation for class (Q5: $\chi^2(6)=4.79$, n.s., $\text{Phi}=0.326$), and about the most frequently read part (Q6: $\chi^2(4)=5.45$, n.s., $\text{Phi}=0.373$). In short, 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, the details were different.

As for the most frequently read part during the task, the ratios of reference were different in 2 groups (Figure 1). In the experimental group ($\chi^2(5)=35.19$, $p<.01$), Chapter 3 was referred significantly more than Chapter 1 ($p=0.0032$), Chapter 4 ($p=0.0002$), and Chapter 6 ($p=0.0032$); 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 part in all chapters: than Chapter 1 ($p=0.0002$), Chapter 2 ($p=0.0002$), Chapter 4 ($p=0.0002$), Chapter 5 ($p=0.0012$), and Chapter 6 ($p=0.0094$).

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 most frequently read part

Another difference between the experimental group and the control group was revealed by a detailed analysis. As for the role of schoolteachers, the participants in the control group were significantly more likely to consider the role as just a trigger than those in the experimental group ($\chi^2(2)=6.46$, $p<.005$, $\text{Phi}=0.598$: Figure 2).

Figure 2. Answers to the role of schoolteachers
2.2.3. Results of evaluation of the prototype (from Q7 to Q9)

As for the missing information or lack of information (Q7), there were 4 categories set by free description: “about the preparation of class”, “about relevant people”, “no lack of information but presence of 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 (χ²(3)=6.44, .05<p<.10), but for the control group they were significantly different (χ²(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.

As for the difficulty for using (Q8), 2 groups presented a significant tendency of the difference and the effect size was medium (χ²(3)=6.85, .05<p<.10, Phi=.442: 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; the participants in the control group said that the booklet was hard to understand.

In the experiment, both group used the same booklet and did the same task, planning a class. But, the number of participants who complained about the how-to-use was significantly larger in the control group than in the experimental group. This result drives us to the question whether the participants in the control group might not understand what the booklet was and the task that they had to do in the experiment.

As for the problem about the visual expression (Q9), the frequencies showed a significant different tendency in 2 groups and the effect size was medium (χ²(3)=7.61, .05<p<.10, Phi=.441: 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 experimental group.

In the experimental group, the difference of frequency was significant (χ²(3)=18.36, p<.01); the invisibility was significantly reported more than the contents (p=0.0026) and other problem(p=0.0094), and the no-problem was significantly reported more than the contents (p=0.0044) and other problem (p=0.0156).

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 problems, 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, because Question 9 asked the apparent or visual problem; if the contents was insufficient, it should have be pointed out in Question 8.
2.3. Discussion

In the experiment, both groups did the same task, but the participants in the experimental group had been instructed to pretend 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, there was a tendency to take more time to complete the task for the control group than for the experimental group. Moreover, the questionnaire revealed a risk that the control group might not understand the task very well. 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; it can be a desirable effect. It can be suitable for basic researches that have difficulties to recruit the real evaluators.

Understanding of task is the most important thing for experimental participants. In the discussion of naïve participants who do not know the purpose of study, it is said that a good participant is not a person who has extraordinary talents nor who is particularly trained; it is essential that the participant really understand what (s)he is supposed to do in the experiment [9]. This is a discussion in the psychological research, but the carefulness or attentiveness for the task is an important and required capability of the experimental participants in any research domains. Thus, acting evaluation is a useful method to pull out the desirable capability of participants in questionnaires.

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 persona in design practice. Persona used in design practice is imaginary character that strictly describes a typical user who is envisioned by user survey [9]. Persona is a strong tool in design [10, 11, as examples of review]. It is often introduced as “persona marketing” in less academic but more practical journals [12, 13, for example in Japan]. As more psychological techniques, acting evaluation may remind psychodrama and role-playing. Each method utilized participant’s imagination or creativity.

It can be summarized that persona is virtual users created by designers or marketers to understand and to develop their product or service. Usually, persona is made by design or market team to share the images of users among the members. Psychodrama is an improvisation and role-playing is originally one of technique of psychodrama [15]; they are usually played by more than one. All of techniques mentioned here need the ability to imagine or the creativity; this point is common to acting evaluation. As for the difference of acting evaluation from persona, psychodrama and role-playing, acting evaluation is an individualized task that does not require the sharing of created images among the participants. In acting evaluation, they can freely imagine and act the person instructed by researchers.

The procedure of acting evaluation is very simple. Researchers just ask participants in the experiment or in the evaluation for imaginary acting performance. The participants image a role specified by researchers, adjust to the role 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 had his/her mission from the principle. The instruction did not define teacher’s gender, age, teaching history and so on. The concrete characters of teacher left to participants’ own choice. A further study should examine the effect of amount of information, its kind or its control for imagination.

There is another problem to be considered. Acting evaluation requires the participant’s imagination, which may be confirmed by their life experience. For example, pretending working people may be difficult for university’s freshmen who have no experience of working. Sympathism may help the imagination. There may be someone who is good at acting evaluation and others who are 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 [16, 17]. 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.

Imagination is an important ability for design practice. Not only above-mentioned personas but also a method of acting out [4, 10] depends on member’s imagination. Acting out is a kind of role-playing; members in the design project pretend the product, the system, the service or users according to the user’s scenario and examine the design proposal. In acting out, people play non-human object as well as human and realize points of view useful for the design. Acting out and acting evaluation in the present study depend on participants’ imagination that requires sympathy because these methods involve playing. Imitation behavior related to mirror neuron is one of topics of current cognitive neuroscience [18,19]; Imitation of other’s behavior helps to understand other’s experiences and one of ways to feel sympathy.

The present study has issues to consider. Firstly, the difference within group was not examined, because of the purpose of the comparison among groups. Secondarily, the participants in the control group were not checked for the power of empathy, which some of the participants in that group might use for the role-playing. Thirdly, to clarify the effect for concentration by acting evaluation, the results of the present study should be also compared with those by the participants sharpening concentration by another method.

4. Acknowledgement

This work was supported by JSPS KAKENHI No. 23611007 and originally reported in IASDR 2013. The author expresses her appreciation for helpful comments of a reviewer to improve the manuscripts.

5. References

3. Turner DC, Anthrozoology – now past its juvenile years! Change by Anthrozoology in Japan, 2009; 1, 4-7.
6. The previous study 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.
7. For instance, the answers for Question 1 firstly gave 15 key words: 1) learning or study of dogs, 2) communication with dogs, 3) softening the resistance against animals, 4) affection to dogs, 5) feeling with dogs’ heart, 6) sanctity of life, 7) consideration or understanding for others, 8) communication, 9) discussion or sharing, 10) serious study, 11) beyond the classroom lectures, 12) autonomous participation, 13) unforgettable activity, 14) amusing activity, and 15) emphasis of individual feelings and experiences. The key words from No. 1 to No. 5 made a category of “About dogs or animals”. The key words from No. 6 to No. 8 made a category of “About general communication”. Those from No. 9 to No. 15 made a category of “About style of the class”. A comment, “leaning of dogs’ feeling”, contained No. 1 key
word and was classified into the category of “About dogs or animals”. A comment, “a lesson in which children think over the life”, contained No. 6 key word and was classified into the category of “About general communication”. A comment, “a lesson in which children can freely discuss and share what they thought or felt”, contained No. 9 key word and was classified into the category of “About style of the class”.


10. Yamazaki K, Ueda Y, Go K, Takahashi K, Hayakawa S, and Yanagida K. Experience vision. Maruzen: Tokyo, 2012 (In Japanese); According to this book, a design approach using persona is called as goal-direct design, not as persona approach. In goal-direct design approach, user survey introduces a persona and the design should be satisfied with the persona’s goal.


