RESEARCH ON DESIGN EVALUATION INDICATORS DRAWN FROM THE GOOD DESIGN AWARD JURY MEMBERS′ COMMENTS

Haruka Sogabe[1], Morita Yoshitsugu[2] and Ishibashi Shinsuke[1]

[1] User Science Institute, Kyushu University
[2] Faculty of Design & User Science Institute, Kyushu University

ABSTRACT

In order to clarify the gap in design evaluation that is thought to exist between groups of people – designers, providers (including managers) and end users – we aimed to create and put into a practical format design evaluation indicators that would make it possible for people from all perspectives to make determinations. We wondered if the same object is being evaluated, whether there are gaps in design evaluations among people who are relatively well versed in creating objects, people such as sellers and marketers who are responsible for providing society with the ultimate fruit and the people who actually use them. Therefore, we created design evaluation indicators so that people with different perspectives can evaluate designs using the same criteria and clarified the gap in design evaluations by collecting data from each group of users through the design indicators. In addition, we think that discussion of the factors that lead to the gap between user groups based on the results of the survey can provide hints at new directions for creation.

Keywords: design evaluation, design award, craftsmanship, evaluation factor, evaluation indicator

1   RESEARCH BACKGROUND AND PURPOSE

The conditions under which users purchase items have dramatically changed. For Japan, the period between 1950 and 1970 was one of high economic growth after World War II and since this was a period of scarcity of products, developing new products and introducing them into the market made them extremely valuable and users actively purchased them. After that, between 1970 and 1990, a period when people's lives were richer, various high-added-value products that made people's lives more convenient were developed and introduced into the market so that people could live fuller lives. Then, from 1990 to the present, a period when people's lives are materialistically full, various market surveys are conducted when developing products. Even if manufacturers develop ingenious products, it can be extremely difficult to sell them. One possible reason for this is that there have been dramatic changes in both the times and the criteria users employ to evaluate products. In addition, the increase in the volume of information over the past couple of years has influenced the criteria that users employ to evaluate products. During a period with few products, there was little information since there were few media outlets and users evaluated the existence of the item. At the present time, there are many media outlets and users are able to actively obtain information; therefore, users, who provide their own interpretation of information from various perspectives, are establishing various criteria for evaluating items. In addition, compared to the period right after the war, it can be said that users are living materialistically fuller lives from birth, which has had a dramatic effect on the creation of criteria for evaluating objects. In an era of diversified user evaluation criteria, we decided to focus on the gap between people's criteria for evaluating items as one measure when creating products and to establish a creation support system that makes use of this gap. In this paper, we broadly categorize users and discuss this process with the goal of clarifying the evaluation gap that appears to exist between these users.

2   CREATING THE DESIGN EVALUATION INDICATORS

Users are a collective group of various individuals. Therefore, we decided to focus on items and to consider three broad categories of users. Considering the relationship to products, it is possible to broadly categorize users into three groups: designers including technicians who are in the position of actually creating the items; providers including sellers and marketers who provide the item to users and end users who actually purchase the items and use them (Figure 1). We attempted to clarify the gap in the criteria for evaluating items based on these three user groups. Therefore, we focused on language, which is mainly used to share information among people. The acquisition of language has made it possible for people to obtain accumulated cultural knowledge, including technology, worldviews and ceremonies from the past. If one thinks of design evaluation as perspective, the system of design awards,
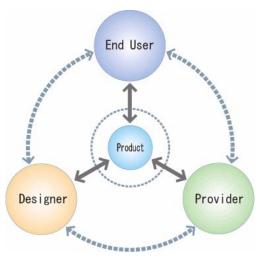

Figure 1. User concept

which exist in countries throughout the world, is a representative example of design evaluation that makes use of language. Design evaluation within the context of design awards differs from an approval system that evaluates physical qualities such as safety, durability and functionality, evaluates the value of products from a cultural and social perspective and is responsible for making people's lives both materialistically and spiritually rich by spreading this information to as many people as possible. Therefore, in order to determine the gap in criteria for evaluating items, we decided to create design evaluation indictors that make it possible for anyone to evaluate designs. There are many design awards throughout the world and for the majority of these awards, screening reviews of the winning items are made public through the award's records or official website. These screening reviews include the reason the item won the award and the superior points of the winning items, which is one set of criteria for determining the value of the item. Judges for design awards are mainly pundits such as designers, producers, university professors and curators. We think that the written reviews accurately represent the value factors behind design evaluations that general people are unable to express and decided to create design evaluation criteria based on these written reviews.

The screening reviews collected in order to create the design evaluation indicators are mainly ones from the Japan Good Design prize. The Good Design Prize, which was launched in 1957, is Japan's only comprehensive design evaluation and commendation system and is celebrating its 50th anniversary in 2006. Around 30,000 items have received the good design prize. We collected the screening reviews of entrants that were awarded the special prize or better over the past 10 years from the good design public records. While gaining an overall understanding of the context of the review, we closely analyzed the reviews and extracted, as short sentences, the reasons the entrant was awarded the prize. We took these short sentences as design evaluation sentences and extracted around 2,700 of them. In order to use the extracted 2,700 design sentences as design evaluation indicators, it was necessary to first organize the sentences based on content, which made use of the "evaluation factor list", a survey and analysis of European design awards undertaken as part of the department's activities in 2004. This "evaluation factor list" is a multifunctional list based on analysis of the "evaluation factors" in each sentence extracted from the screening reviews of five European design awards. The main sources of information were the summaries of winning designs for each award in 2004, as well as the public records and public websites, which contain the screening reviews. This "evaluation factor list" can also be called a list of "value factors that should be evaluated for designs" (Table 1). The evaluation factor list is composed of 8 main categories and 22 subcategories. In order to organize the actual design evaluation sentences, a database was created using the 22 subcategories. The extracted design evaluation sentences were matched against the evaluation factor list. However, it was impossible to organize the approximately 2,700 design evaluation sentences using only 22 subcategories, so we created new categories in the database when a new evaluation factor was discovered in the design evaluation sentences.

As a result of this work, an additional 100 items were added to the 22 subcategories for a total of 122 "value factors that should be evaluated for designs". In the process of this work, the point that should be kept in mind is that when creating the evaluation factor list the subcategories were classified using

keywords in order to provide multi-functionality and supplementary information was added to the content, but these keywords were not geared to design evaluations made from various cultural and social perspectives. Since the design evaluation evaluates items from various perspectives, there is the danger that the interpretations can differ depending on the user if one does not limit the content of the indicators to some degree and use concrete expressions. It became clear that when trying to create design evaluation indicators for use by various people, it is not desirable to have items with different interpretations for different users. Then, we determined that there is a contradiction in imparting multi-functionality to screening reviews of design evaluations made from a limited position but taken from various perspectives within a culture and society that has overlapping relationships. During the work to organize the additional items in the database, there was the danger of changing the content of the evaluation factor in the design evaluation sentences when converting them to keywords based on the decisions of individuals. We avoided the method of adding categories based on keywords, but extracted the item taken to be the evaluation factor within the design evaluation sentence as the sentence itself, which made it possible to avoid changing the meaning by making additions to the database. Then, the 22 subcategories represented by the keywords became the items for which the content was replaced with sentences (Table 1). This work organized the design evaluation sentences extracted based on decisions by individuals into the appropriate subcategory. After that, the design evaluation sentences were rechecked by various people, including us, using the KJ method in order to make sure they were organized into the appropriate subcategory. As a result of all this, it was possible to summarize the almost 2,700 design evaluation sentences into 185 short sentences, which become the model of the design evaluation indicators (referred to below as the indicator prototypes).

These indicator prototypes have been collected by the authors as "evaluation factors that should be used when evaluating designs" from the design prize screening reviews. Then, in order to verify whether it is possible to actually create such evaluation indicators, we conducted a verification evaluation test using the indicator prototypes and collected opinions regarding the indicators from a wide range of users, the evaluators. This verification test is given in Table 2. The main purpose of this verification test was to learn points from a wide range of evaluators regarding the content and method of using the indicator prototypes. The evaluation test was conducted over a six month period and information on items that were considered deficient during the evaluation were also obtained from the evaluators, which resulted in revisions to the content and use of the indictor prototypes. The results of the various evaluators during the evaluation verification test are summarized in Table 3.

Initially, revisions were made to the content of each indicator prototype in Table 3. In addition to revisions so that each indictor incorporated a single evaluation item, the parts of the item being evaluated were also clarified. Differences in the content of indicators that were similar were clarified and for items that could not be clarified, the indicators were combined. Next, it was determined that a system to extract indicators according to the item to be evaluated and the evaluation condition was necessary and using the basic requirements of Good Design (appearing in the Good Design Prize screening criteria) and the 7 good design rules (version 2) created by the Design Center of the Design Department at North Carolina University in 1997 as a reference, the groups of multiple indicators were classified into 7 factors (Figure 2). Since the part of the item to be evaluated was clarified for each indicator with revised content, the indicators were formed into a two-dimensional matrix with one dimension being one of the seven evaluation factors, such as safety and the other the evaluation part such as price or shape, which provided order to the groups of indicators (Table 4). This work resulted in around 400 design evaluation criteria being created from the 185 prototype design evaluation criteria (Table 5).
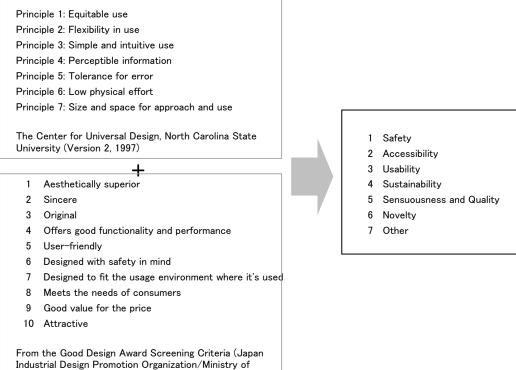
Table 1 Evaluation factor list

Main category		Sub-category		Details	Example evaluation
1	Newness	1.1	Innovation	The good or design is innovative (new in comparison to precedents in its industry or field, including existing goods and competitor products)	· design overturns stereotypes · unprecedented combination of form and function
		1.2	Originality	The good or design has originality	· unprecedented structure or shape
		1.3	Ideas	The good or design is the result of an excellent idea or insight	· invented a folding mechanism · understanding of target user is outstanding
2	Materials and functions	2.1	Performance and functionality	The performance and functionality of the good is superior	· includes innovative functions · achieves high level of durability and safety
		2.2	Reduced environmental load	Consideration given to reducing environmental loads throughout the good's life from production and distribution to use, disposal and recycling	· environmentally-friendly base materials, paints and other materials used · energy consumption during use is reduced
3	User perspectives	3.1	Usability	Use is easy, pleasant and comfortable	· frees people from pains caused by conventional goods · anyone can use without error
		3.2	Ergonomics	Applied ergonomic data and considerations	· contrivances made to prevent user physical stress · mechanism allow movement in coordination with user body movements
		3.3	Maintenance	Consideration given to make cleaning, upkeep and maintenance after start of use easy and economical	· changing light bulbs is easy · form and surface finish make cleaning easy
		3.4	Accessibility	Consideration given to make use easy for a greater number of users, particularly users who may have had difficulty using conventional products	· supports active behavior by children with disabilities · has appeal for both men and women
		3.5	Other user merits	Acquisition, use, action and lifestyle merits for users	· suitable for various family structures · provides a comfortable experience
4	Production processes	4.1	Materials	Use and selection of materials are suited to the manufacturing methods, applications and design features; and the materials are well understood and the design brings out their strengths	· revolutionary weight reduction through materials selection · makes advantageous use of natural materials in parts that users touch with their hands
		4.2	Production technology	Appropriate manufacturing methods, engineered structures, joining methods, surface finishing and other techniques chosen; design reflects good understanding of production	· makes use of traditional hand techniques · finish makes use of cast metal surface texture
		4.3	Cost	Cost reduction is considered in production and distribution	· unnecessary materials have been eliminated and the number of parts reduced, keeping down costs
5	Feeling and meaning	5.1	Aesthetics	Shape and appearance have a simple beauty that appeals to people's senses	· form is so close to perfect that it could be put in a museum · refined and stylish
		5.2	Meaning	Shape, appearance and other sensory aspects offer symbolism, affordance and other meaning	· the object gives the space where it is placed a special atmosphere · expresses symbolism and metaphor
		5.3	Communication	Shape, appearance and other sensory aspects convey messages and images	· iconic shape conveys its use directly · appearance is inviting
6	Significance of design work	6.1	Integration	Elements that tend to be self-contradictory are unified without contradiction through design	· functionality and aesthetics coexist · meets needs of both experts and ordinary users
		6.2	Challenge	Difficult problems are actively tackled and solutions sought through design	· realizes unique design while meeting strict requirements · tackles an issue that no one had been able to solve
		6.3	Forward-thinking	Suggests perspectives and values that link to the future of the item type or industry	· greatly surpasses the present standard level of the industry · gives birth to a new standard that links to the future
7	Social benefits	7.1	Degree of business contribution	Contributes concretely to the business of a client or others	· staff motivation was increased · visual identity established brand
		7.2	Commercial success	Design is valued by the market and contributes to sales	· design was a draw that led to product sales · design created an important product appeal
		7.3	Design popularization, awareness-raising and education	Design popularization has effects of social awareness-raising and education	· changed consciousness about design · spread awareness about design

Table 2 Verification evaluation test

Evaluation experiment	Good Design Presentation 2005	Fukuoka Prefecture Industrial Design Award	Minoshima area social experiment
Evaluators	5 division members	8 designers that live in Fukuoka Prefecture / 5 staff of award-winning enterprises / 100 attendees of an industrial design award symposium	400 ordinary people passing through the area
Objects of evaluation	Good Design Prize candidates	5 chairs that had won past awards	11 benches in public spaces
Question indices	185	180/180/80	71
Experiment details	As the first case study, objects were not evaluated, but each participant identified issues that might arise when the question indices were put into actual use in evaluation.	First, a workshop was held with the designers, and they identified problems with the indices, Based on this, the indices were revised and surveys of the attendees and staff were conducted.	Prototype benches were placed in a social experiment implementation area and people passing by and users were asked to evaluate the benches after actually sitting on them.
Results	Among the issues identified, the following were typical: (1) unclear what parts of objects were to be evaluated, (2) mixture of object indices that require expert knowledge with those that do not, (3) necessary indices vary according to the object, (4) importance of indices varies	Among the issues identified, the following were typical: (1) difficult to answer, (2) too many questions, (3) items difficult to understand	Among the issues identified, the following were typical: (1) words used in the indices were difficult, preventing response, (2) too many questions, (3) some indices are similar
		By creating frequency distribution and normal distribution tables for each item, we made clear that there are gaps between answer clusters. The averages for the manufacturers was high, and there was a tendency for their responses to be high values. The average scores for designers and end-users did not tend to be high, but rather varied according to the item. End-user responses tended to be dispersed rather than clustered.	The objects were 8 wood and 5 concrete benches. The highest ranked bench was a wooden bench, while the lowest was a wooden chair for one person. Furthermore, for "consideration has been given to Japanese style," an index that express differences in materials, the score for wooden ones was high, while those with plaster or tatami seats were even higher. This result verifies that the features of objects can influence the responses to different indices.

Evaluation experiment	Furniture investigation with people from 5 European countries and Japan	Investigation pf prototypes from a recycled tile development project	Public good investigation in Taiwan
Evaluators	Design professionals who live in Europe / Design professionals who live in Japan	Designers	Design students
Objects of evaluation	13 designer chairs	2 tile prototypes	Goods in the Living Mall and Taipei 101
Question indices	49	93	80
Experiment details	Investigations were conducted in France, Germany, the UK, Denmark, Italy and Japan. Evaluation indices were explained and feedback and suggestions were given.	At the Fukuoka Research Center for Recycling Systems, Tagawa Industrial Co., Ltd. is developing recycled tiles from sewerage sludge incineration ash. We used indices to investigate prototype tiles under development.	We selected evaluation indices for public spaces and public products, and visited 2 facilities to experience the spaces and evaluate them. Then, after collecting the evaluations, we interviewed the evaluators using Q Charts.
Results	Among the issues identified, the following were typical: (1) question contents are unclear, (2) the meanings of some questions are "Japanese" and difficult to understand, (3) English expressions are not good, (4) not clear what time is to be assumed, (5) some questions cannot be answered based on an evaluation with only photos.	Among the issues identified, the following were typical: (1) reading the indices again for each object is a nuisance, (2) too many questions (3) evaluation can only be done by estimation	Among the issues identified, the following were typical: (1) cannot answer without information about the objects, (2) indices are insufficient for evaluation of spaces
	Depending on the point of evaluation, gaps occurred between evaluations by people from different countries even for chairs that have been highly regarded around the world for many years. Evaluation of seating comfort made up over half of the overall evaluation, and the images resulting from these evaluations tended to influence the overall evaluation. Differences in evaluation results also occurred due to the amount of information possessed by the evaluator.	Since the fact that the objects were recycled tiles that include incineration ash from sewerage sludge as a raw material was presented, indices related to environmental issues received high scores. This investigation was an evaluation by designers in the process of development, and it allowed consideration of improvements based on both positively and negatively evaluated indices.	

Table3. Findings from the verification evaluation test

Points that were raised regarding the content of the indicator prototypes	Points that were raised regarding on ways to use the indicator prototypes
There were indicators that the evaluator did not understand how to handle the evaluation since it was not clear what part of the item was to be evaluated.	Since each indicator was created from the design evaluation sentences for award-winning items, some indicators were necessary for the object being evaluated while others were not.
Some indicators required specialist knowledge of the object being evaluated and some indicators did not	Evaluations are tiring since various sentences are used for the indicator.
There is an overlap in content.	There are differences in the importance of each indicator.
The content of some indicators was considered necessary and the content of some indicators was not.	There is a large number of question indicators for each evaluation.
The language was difficult to understand.	Some indicators could not ... actual information about the develop...

Principle 1: Equitable use

Principle 2: Flexibility in use

Principle 3: Simple and intuitive use

Principle 4: Perceptible information

Principle 5: Tolerance for error

Principle 6: Low physical effort

Principle 7: Size and space for approach and use

The Center for Universal Design, North Carolina State University (Version 2, 1997)

+

1　Aesthetically superior

2　Sincere

3　Original

4　Offers good functionality and performance

5　User–friendly

6　Designed with safety in mind

7　Designed to fit the usage environment where it's used

8　Meets the needs of consumers

9　Good value for the price

10　Attractive

From the Good Design Award Screening Criteria (Japan Industrial Design Promotion Organization/Ministry of Economy, Trade and Industry)

1　Safety

2　Accessibility

3　Usability

4　Sustainability

5　Sensuousness and Quality

6　Novelty

7　Other

Figure 2 Design evaluation factors

Table 4 Design evaluation indicator matrix

		Safety		Accessibility				Usability				Sustainability			Sensuousness and quality			Innovation		Other	
		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	
		Safety	Clear and full functionality and capabilities	Consideration for UD, elderly, and handicapped	Creates a sense of affinity	Provides user with information	Communication	Adaptability	Flexibility and selectability	Physical ease of use	Consideration of user's physical and spiritual needs	Economy	Durability	Contribution to society	Entails sensuousness and quality	Superiority of all factors	Embodiment of items such as theme and needs	Point that draws ones attention	Novelty, innovation	Process through which the product reaches the market	Subtotal
	Product Part	a	b	e	m	h	j	i	k	q	n	d	f	g	l	s	c	p	r	o	
A	Overall	11	4	3	13	10	4	11	10	12	9	7	10	14	16	2	3	7	14	11	171
B	Information	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2
C	Specifications and settings	0	0	0	0	0	0	5	3	0	0	0	1	0	0	0	0	0	0	0	9
D	Exterior appearance	8	0	2	6	10	0	5	7	1	5	0	3	0	8	2	0	4	4	0	65
E	Function	3	3	2	4	5	0	7	7	0	2	0	1	1	2	2	0	4	5	0	48
F	Development process	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	8	11
H	Concept	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2
I	Price	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	Technology	0	0	0	0	2	0	0	0	0	0	0	1	1	1	0	0	0	2	0	7
K	Structure	2	0	0	0	2	0	2	3	2	0	0	1	0	1	2	0	0	1	0	16
L	Capabilities	1	1	0	2	3	0	4	7	0	2	0	1	0	1	3	0	2	2	0	29
M	Material	2	0	2	3	3	0	4	5	0	1	0	2	1	1	2	0	2	5	0	33
N	Existence	1	1	0	2	0	0	0	0	1	1	0	0	0	1	0	0	0	1	0	8
O	Operation	0	0	0	0	0	0	0	0	4	3	0	0	0	0	1	0	1	0	0	9
P	Surroundings	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	4
Q	Variation	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	1	20
R	Image of Manufacturer	4	0	1	1	2	0	0	0	0	1	0	2	1	2	1	0	0	0	0	15
S	Each part	0	4	0	0	0	0	0	0	0	0	0	15	0	8	0	0	0	0	0	27
	Subtotal	32	13	10	31	41	4	39	61	22	24	7	36	18	41	18	3	20	36	20	476

Table 5 Created design evaluation indicators

Design evaluation indicators
Attention has been paid right down to the finest detail.
The item has an appropriate sense of luxury.
The item has a sense of superior sensuousness and structure.
The item gave the impression that its appearance when it is being stored was taken into consideration.
The item gave the impression of an integrated common form and expression.
The item has the appearance of innovation in expression.
The item has an appearance and sense of innovation in expression related to how the surfaces were processed.
The item has a sense of quality based on its appearance.
The item has a full range of functions.
Functions were appropriately selected for the item.
There is a sense of appropriately selected capability.
The material was appropriately selected.
There are no defects such as distortions in the wood.
There are no defects such as warps.
Since there are seams, the item may break.
There are no warps on parts such as the tops.
There are no warps on the sideboards.
There are no gaps between the body and the joints.
There are no gaps between the joints (details).
The item has been uniformly painted (polis...
...is fair

3 DESIGN EVALUATION SURVEY EMPLOYING THE DESIN EVALUATION INDICATORS

The following evaluation survey was conducted using the design evaluation indicators constructed through the above process. A wide range of items was evaluated, including designer chairs, a door-phone with a monitor, a TV, a washing machine, a soy-sauce dispenser and a robot. In this paper, the evaluation results for designer chairs are given as an example of the above-mentioned survey and the results discussed.

3.1 Verification process

The evaluated items were 10 randomly selected distinctive designer chairs and the evaluators included people working in the field of design regardless of type of design, people working at businesses such as furniture manufacturers who are involved in the marketing or selling of furniture and general consumers who would mainly use the items as furniture. The evaluations were conducted under the following conditions: the chairs were placed in two rows of 5 facing each other and information such as the product name, price, designer name, manufacturer and material was provided on a tag on the side of the each chair. We tried to conduct this survey under conditions that were close to those for daily life. The design evaluation indicators used in the actual evaluation were ones we considered to be important when evaluating chairs, including designer chairs and were chosen from 7 evaluation factor matrixes. We then analyzed each selected evaluation indicator and created a survey using 40 indicators determined to be appropriate and conducted the survey using these indicators. The evaluation was conducted by having the evaluator check one of 4 choices ranging from "think so" to "do not think so" that most closely represented their feeling. In addition, in the case that the evaluators could not make a determination, they would select the "not sure" item. Before the evaluation, the evaluators were told to answer the questionnaire after closely examining the chair, including sitting in and touching the target chair.

3.2 Results of and observations regarding the test

Survey data was obtained from 91 evaluators – 21 designers, 25 providers and 57 end users. For convenience, from the collected evaluation data, the average evaluation for each indicator was calculated with "I agree" being assigned a value of 4, "I agree somewhat" a value of 3, "I disagree somewhat" a value of 2 and "I disagree" a value 1 and a one-way factorial analysis of variance (multiple comparison employing Tukey HSD) was conducted. "I do not know" was statistically handled as a missing value in the analysis since a determination could not be made. Table6 provides the average value, standard deviation, F value and significance level and results of a multiple comparison for each answer group where the indicator provided a statistically significant result.

A survey was made for each object evaluated using 40 indicators (table 7). The results of the survey show that there were statistically significant gaps among designers, providers and end users for at least 1 indicator and at most 7 indicators for 9 of the 10 objects evaluated. For the chairs that were evaluated for this survey, when 40 perspectives were provided in order to evaluate the design, there were statistically significant gaps between designers, providers and end users for 9 of the 10 objects.

In addition, using multiple comparisons, we confirmed whether a gap existed between designers, providers and end users regarding the 9 chairs for which an initial evaluation gap was detected and there were either 1 or 2 evaluation gaps for each indicator, for a total of 29 gaps. Of these 29 gaps, the results of multiple comparisons revealed that there were 12 evaluation gaps between designers and providers, 13 gaps between designers and end user and 4 gaps between providers and end users. Of the 29 evaluation gaps, 25 of the gaps were between designers and either providers or end users, which shows that there are numerous gaps between the designers and the other groups. For the chairs that were evaluated, there were statistically significant gaps in the evaluation between designers, who are involved in creating the product and both providers and end users.

Therefore, it can be said that even if the same item is evaluated using the same indicators, there are gaps in the evaluation between designers, who are involved in creating the product and providers and end users who play a different role. It can be assumed that there is a tendency for designers to evaluate items differently from manufacturers and end users.

Table 7 40 indicators

The item has an appropriate sense of luxury.	There is a sense of originality that will not fade with time.
The item has visible ingenuity in display and expression.	The item is appropriately priced.
The item is of good quality.	There is a visible sense of ease.
The item has a high degree of completion	There is an overall sense of ease due to items such as material and feeling when sitting.
The item has an overall sense of design.	The item creates a feeling of trust based on the image of the manufacturer or seller.
The item has outstanding beauty	Considering the overall item, including material and feeling when one sits down, there is a sense of reliability.
The item was created carefully and precisely.	There is visible sense of stability.
The item uses appropriately materials.	When actually using and sitting in the chair, there is no rocking.
The item can be used for a long time.	There is a good balance between the product factors, such as appearance, function and price.
The item is easy to obtain and maintain.	When one sees the object, one wants to try it.
The item will remain interesting.	I find I have taken a liking to the item
The item is not influenced by the times or fashion.	One can feel attachment to the object.
The item is acceptable to people of all generations.	The item is appealing.
The item matches my lifestyle.	The item is desirable (stimulates desire).
When one sits in the chair, the seat is at an appropriate height.	The item creates a sense of satisfaction due to owning it.
It is possible to imagine using the chair in some part of one's life.	The item has a sense of comfort of use based on the image of the manufacturer or seller.
The item's appearance is innovation.	When I used the item, it gives users to comfort the mind and body.
The material used has never been used for chairs before.	It is possible to imagine physical properties, such as heaviness, lightness, sturdiness or flexibleness, from the appearance.
The item uses original materials	The item is simple.
The item will provide a sense of individuality to the space it is used in.	The item has presence.

Table 6 Analyzed design evaluation indicators (Item)

Evaluated product	Item	1 Designer	2 Provider	3 End-user	F	P
GO	The item has an appropriate sense of luxury. 1>2* 2>3**	2.75 (0.99)	3.78 (0.44)	2.74 (0.96)	$F_{(2,87)}=5.05$**	
GO	The item matches my lifestyle. 2<3*	1.83 (0.92)	1.11 （0.33）	1.98 (0.83)	$F_{(2,87)}=4.35$*	
Ero|s|	The item will remain interesting. 1>2*	2.41 (0.80)	1.56 (0.73)	2.20 (0.94)	$F_{(2,84)}=2.96$+	
Ero|s|	The item is not influenced by the times or fashion. 1>2*	2.64 (1.00)	1.67 (1.00)	2.11 (0,91)	$F_{(2,84)}=4.08$*	
Ero|s|	The item creates a feeling of trust based on the image of the manufacturer or seller. 1>3**	2.95 (0.69)	2.13 (0.99)	2.26 (0.88)	$F_{(2,68)}=5.28$**	
Ero|s|	The item creates a sense of satisfaction due to owning it. 1>3*	2.63 (1.10)	1.89 (1.05)	2. 02 (0.95)	$F_{(2,85)}=3.45$*	
Ero|s|	The item has a sense of comfort of use based on the image of the manufacturer or seller. 1>2**	2.75 (0.79)	1.75 (0.71)	2.24 (0.80)	$F_{(2,70)}=5.26$**	
OLIO	The item creates a feeling of trust based on the image of the manufacturer or seller. 1>3*	2.9 (0.79)	2.33 (1.00)	2.24 (0.98)	$F_{(2,71)}=3.46$*	
OLIO	The item is desirable (stimulates desire). 1>2*	2.42 (1.14)	1.44 (0.73)	1.88 (0.97)	$F_{(2,86)}=3.88$*	
OLIO	It is possible to imagine physical properties, such as heaviness, lightness, sturdiness or flexibleness, from the appearance. 1>3**	3.17 (0.65)	2.44 (1.01)	2.5 (0.95)	$F_{(2,85)}=4.98$**	
La Marie	The item is easy to obtain and maintain. 1<3*	2.46 (0.88)	3.11 (0.60)	3.04 (1.02)	$F_{(2,87)}=3.38$*	
La Marie	The item creates a feeling of trust based on the image of the manufacturer or seller. 1>2** 1>3*	2.63 (0.83)	1.50 (0.54)	2.02 (0.76)	$F_{(2,68)}=7.3$**	
La Marie	The item has a sense of comfort of use based on the image of the manufacturer or seller. 1>2*	2.30 (0.92)	1.50 (0.54)	1.89 (0.71)	$F_{(2,71)}=3.67$*	
Y Chair	The item has an appropriate feeling of luxury. 1>3**	3.33 (0.57)	3.00 (0.71)	2.74 (0.90)	$F_{(2,87)}=4.68$*	
Y Chair	The item is of good quality. 1>3*	3.58 (0.58)	3.00 (0.71)	3.17 (0.75)	$F_{(2,84)}=3.64$*	
Y Chair	The item can be used for a long time. 1>3*	3.41 (0.67)	2.67 (1.23)	2.79 (0.90)	$F_{(2,85)}=4.33$*	
Y Chair	The item will remain interesting. 1>2*	3.25 (0.74)	2.44 (1.13)	2.88 (0.83)	$F_{(2,87)}=3.39$*	
Y Chair	The item is appropriately priced. 1>3*	2.87 (0.82)	2.67 (0.87)	2.25 (0.94)	$F_{(2,85)}=4.13$*	
Y Chair	The item is appealing. 1>3*	3.61 (0.58)	2.89 (0.93)	3.13 (0.83)	$F_{(2,85)}=4.02$*	

	Item	Group 1	Group 2	Group 3	F
	The item is simple. 1<3*	2.87 (0.96)	2.67 (0.71)	3.28 (0.65)	$F_{(2,87)}=4.68$*
PUNTO	The item has an appropriate feeling of luxury. 1>2*	2.63 (0.92)	1.78 (0.67)	2.26 (0.81)	$F_{(2,87)}=3.65$*
SPAGHETTI	The item is of good quality. 1>3* 2>3*	2.74 (0.86)	3.00 (0.71)	2.23 (0.76)	$F_{(2,86)}=6.10$**
Meda Chair	The item is acceptable to people of all generations. 1>2* 2<3*	3.00 (0.66)	2.11 (0.93)	2.86 (0.79)	$F_{(2,87)}=4.53$*
	The item has a sense of comfort of use based on the image of the manufacturer or seller. 1>2*	3.25 (0.79)	2.44 (0.73)	2.81 (0.80)	$F_{(2,73)}=3.78$*
Multi Chair	The item is easy to obtain and maintain. 1<2*	1.54 (0.72)	2.33 (0.87)	1.91 (0.81)	$F_{(2,87)}=3.67$*

**$p<.01$ *$p<.05$ +$p<.06$

4 CONCLUSION

This survey has shown that there are design evaluation gaps among groups – designers, providers and end users. Therefore we think that it is possible to construct a design evaluation and diagnosis system that employs the evaluation gaps that exist between the three groups. The results of this survey reveal that there are various gaps in evaluation between designers and the other two groups – providers and end users. It is possible to think of various factors that would cause the gap in evaluations. There are likely differences among designers and others in their knowledge and experience regarding the evaluated object depending on whether they are involved in creating the object or not. Therefore, this difference in knowledge and experience could create a gap in the awareness of the product, which could be a factor behind this gap in evaluations. Therefore, there could be situations when this visible gap could be an issue or have significant meaning. It is important that designers have abundant knowledge and experience and this is also indispensable in order to create items that provide value. However, due to this abundant knowledge, designers might create items to flatter themselves. The point of this research is that designers realize that design evaluation gaps do exist. These design evaluation gaps that exist between groups – designers, providers and end users – will be useful when creating a strategy to develop products.

In this paper, we classified users into three broad groups based on their relationship to the item and created design evaluation indictors, which were the value criteria for items. Then, we conducted an evaluation test using these indicators. Among the three broadly grouped user groups, there were clear gaps in design evaluations. In the future, based on the data obtained from this test, we plan to examine the factors between the gaps among user groups based on workshop activities with the groups, centered on the designers. Ultimately we think that this research can lead to the creation of a design evaluation and diagnosis system that would be a useful tool for designers in order to provide value that people truly want and value that people potentially want.

REFERENCES

[1]   Japan Industrial Design Promotion Organization *Good Design Award 1996-1997 ~ 2004-2005*, 1996 ~ 2005, (Maruzen).

[2]   Sakuda Keiichi. *Kachi no Shakaigaku (Value Sociology)*, 1972 (Iwanami Shinsho).

[3]   Ray Jackendoff. *Kokoro no Patan – Kotoba no Ninchikagaku Nyumon (Patterns in the Mind: Language and Human Nature)*, 2005 (translated by Masanori Mizuko) (Iwanami Shinsho).

[4]   M.F.story,MS,IDSA,J.L.Mueller,MA,IDSA,M.Montoya-Weiss,PhD,*COMPLETION OF UNIVERSAL DESIGN PERFORMANCE MEASURES,* Proceedings of the RESNA2001 Annual Conference.

[5]   http://www.design.ncsu.edu/cud/pubs_p/pudperformproduct.htm

[6]   Research Institute of Human Engineering for Quality Life, *Wakushappu Ningenseikatsukogaku (Workshop – Human Engineering for Quality Life)*, 2004, (Maruzen).

[7]   Kyushu University User Science Institute Evaluation and Management Department, *2004 Report on the Results of the Evaluation and Management Department*, 2005.

[8]   Morita Yoshitsugu and Sogabe Haruka and Ishibashi Shinsuke, *Guddo Dezain To Univaseru Dezain No Kankei Kara Mita Dezain Hyoka/Shindan Shisutem No Arikata Ni Kansuru Kenkyu (Research on the Design Evaluation and Diagnosis System Based on the Relationship Between Good Design and Universal Design)*, IAUD '06, Vol. 2, Kyoto, October 2006

[9]   Atsushi Oshio,SPSS to Amos niyoru sinri tyousa data kaiseki inshi bunseki kyoubunsankouzou bunseki made, tokyotosho co.,Ltd,2004

[10]   Shiro Aoki, Discovery of New Evaluation Measure in Judgement of Good Design Award, 1999, Special Issue of Japanese Society for the Science of Design, Vol. 6, No. 4.

[11]   Katsuhiko Onai. Norihiko Mori. Kazuo Sugiyama, *Construction of Design Concept Words as a Fuzzy Structural Model,* 1989, Bulletin of Japanese Society for the Science of Design, No. 73.

[12]   Sakai Takashi, *Anketo Cyousa to Toukeikaiseki ga Wakaru hon (Guide to understanding questionnaire investigations and statistical analysis)*, 2003, Japan Management Association Management Center.

Contact: Haruka Sogabe
Kyushu University
User Science Institute
4-9-1 Shiobaru, Minami-ku
Fukuoka, 815-8540
Japan
Phone +81 92 553 9457
Fax +81 92 553 4674
e-mail sogabe@design.kyushu-u.ac.jp
URL http://www.usi.kyushu-u.ac.jp