This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2879944, IEEE Transactions on Affective Computing

Layered-modeling of affective and sensory experiences using structural equation modeling: Touch experiences of plastic surfaces as an example

Shogo Okamoto, *Member, IEEE,* Haruyo Kojima, Atsushi Yamagishi, Kyoichi Kato, Atsuko Tamada

Abstract—Developing a multilayered structure of adjective words that explains semantic relationships between human perceptual and affective responses to stimuli is instrumental in the design of affective aspects of products. However, the determination of multilayered structure is demanding and, thus far, it has been conducted by experienced developers in a trial-and-error manner. This study developed a method to systematically establish such structures through common tasks for sensory evaluation where products are rated along adjective labels. This method gradually expands the model from a simple two-layered to complex multilayered structures until it is accepted by structural equation modeling. The lower and higher layers of the initial two-layered model are composed of sensory and affective adjectives, respectively. The parts with weak fit indices of the higher layer are then remodeled, resulting in a multilayered affective structure. To validate the method, we built adjective structures based on responses to touching plastic plates. The method resulted in three- and four-layered structures that were quantitatively and semantically valid.

Index Terms—Structural equation modeling, sensory evaluation, tactile perception.

1 INTRODUCTION

It has become increasingly important in industrial design to develop products distinct from competitors, based on the affective values. Sensory, affective, and hedonic responses to products are commonly rated by sensory appraisals based on the use of adjectives. The reported rating value for each adjective is regarded as a variable, and it is popular to discuss the relationships between multiple variables by using multivariate analyses or machine learning techniques. In particular, hierarchical expressions of adjectives as shown in Fig. 1 are instrumental to understanding and for designing affective responses to products and have been applied to various types of products. Layered models were leveraged to seek relationships between affective and sensory responses and the design parameters of products from industrial products such as cars [1] and housing materials [2] to food products [3]. However, designing such structures with more than two adjective layers has been heuristically achieved by experienced developers. The present study develops a systematic method to produce statistically valid layered structures. We then apply the method to tactile feelings of plastic products as an example and discuss its semantic validity.

Fig. 1 is a hierarchical model that shows how the physical properties of products influence sensory and affective responses that are expressed with adjectives. Sensory responses, which we refer to as psychophysical responses, expressed by adjectives such as rough, warm, and soft are explained by linear combinations of the multiple physical properties of products that lie at the base of the hierarchy. Psychophysical responses discriminate between physical properties and are determined by the physical aspects of stimuli. Psychophysical responses then explain qualitative

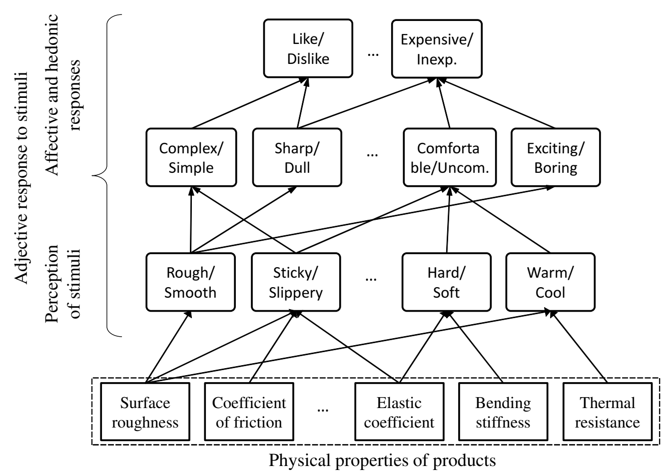

Fig. 1. Schematic figure of multilayered influence model of perceptual, affective, and hedonic responses to the tactile stimuli of products [4]. The physical properties of the products are located at the bottom of the figure although they are not considered in this study. Directional arcs indicate the direction of influences. A response using the word in a higher layer is explained by a linear combination of responses using the words in lower layers.

- *S. Okamoto and A. Tamada are with the Department of Mechanical Systems Engineering, Nagoya University, Japan.*
- *H. Kojima, A. Yamagishi, and K. Kato are with Kao Corporation.*

Manuscript received April 15, 2018; revised August xx, 201x. This study was in part supported by MEXT Kakenhi (15H05923). The authors have declared that no competing interests exist.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2879944, IEEE Transactions on Affective Computing

responses, such as the attributes of products in the middle layer. Finally, in the higher layer, affective or hedonic responses are linked with responses at lower layers. The classification of adjectives into psychophysical, affective, or hedonic layers is not deterministic and depends on scenarios including physical stimuli and adjectives used in subjective questionnaires. Such hierarchical models explain adjective responses better than two-layered models and enable us to graphically capture the relationships among human responses and physical properties. Furthermore, these models suggest how physical properties should be controlled for to improve or increase the ratings of certain adjective responses, serving to enhance the process of product design.

A problem of hierarchical modeling is that it is demanding to build hierarchical models of adjectives. For multivariate linear models, structural equation modeling (SEM) is commonly used to test and validate particular models. SEM validates the hypothesized linear structure by comparing between its covariance structure and the sample covariance matrix among adjective variables [5], [6]. Although SEM is a powerful statistical tool to test hypothetical models, it does not create models that explain the observed events. These hypothetical models have to be implemented by experienced developers on a trial-and-error basis. Thus far, there is no general method to design hypothetical hierarchical models of psychophysical, affective, and hedonic adjectives. Especially, in models with more than 2 layers, it is difficult to provide a determinate estimate of the number of layers and the loci of adjectives in the model, i.e. which adjective should be at which layer.

Most of the previous models in the literature are two-layered, with psychophysical and affective or hedonic adjectives respectively located at the bottom and higher layers [1], [2], [7], [8], [9], [10], [11]. Psychophysical words are commonly placed at the bottom. This is in part because with such a structure it is convenient to render the semantic chain that physical stimuli are perceived, and then affective responses are subsequently induced. However, there is no agreement or method to determine the position of adjectives when the model includes more than 2 layers.

Earlier studies on the multilayered-modeling of psychophysical and affective responses related to touch were based on their own rationale. Nagano et al. leveraged semantic influences among adjective words to determine their hierarchical structure, and then built three and four-layered models [4], [12], [13]. In their studies, subjects rated the degree of influence between two adjective words; for example, how much the roughness of products influenced their expensiveness. These degrees were then used to compute a graph structure to indicate influences among adjective words. Kidoma et al. used covariance selection to specify the structure in the middle to upper layers assuming that adjectives relating to tactile sensation were at the bottom [14]. Their method occasionally produces complex models with a large number of connections among adjectives because it started from inputting the most complex full model possible and gradually reducing its associations. Chen et al. proposed a three-layered model based on correlation coefficients among adjective ratings acquired through sensory evaluation [15]; however, the hierarchical modeling was not the primary focus of the study and thus was not elaborated on. Hashim et al. classified adjectives into three groups or layers *a priori* based on their meanings to complete a hierarchical model [16]. In the bottom layer, subjective responses to single physical quantities were located. In the middle layer, affective or emotional responses expressed by, for example, embracing and expensiveness, were used whereas the top layer included preference scores. These allocations of variables were decided by the researchers. Except for tactile or haptic sensations, Ueda et al. determined the hierarchy among adjectives based on their semantic abstractness [17] where words judged as more abstract in questionnaires were displaced at higher layers. Nishino et al. built a 3-layered model especially for coffee [3] where each layer was designed by the experts of coffee manufacturers. As suggested by these studies, the placement of adjectives was based on various approaches and there was no standard method. Some of them require questionnaire tasks which are unique to their own methods and not as common as sensory evaluation [4], [12], [13], [17]. Others still require experienced developers to semantically classify adjectives [3], [16]. In contrast, our approach approximates statistically valid models solely by using a standard sensory evaluation task.

In the present study, we build hierarchical models of psychophysical, affective, and hedonic responses caused by touching plastic materials as a specific example. We employed a general approach that does not depend on the class of materials or objects. Our method expands a simple model to a complex multilayered one with the initial model being the most simplified two-layered structure. During this process, the weaker parts of the model are improved. For each stage of the model development and expansion, a candidate model is validated by SEM. This cycle is repeated until statistically valid models are estimated. This arguably allows the estimation of the simplest statistically valid models compared to previous methods [14] that tended to estimate complex models with many associations among adjective variables.

We do not discuss the physical properties of stimuli, as although this is certainly significant information for product development, it is evident that they lie beneath the psychophysical layer as shown in Fig. 1. It is valuable to refer to earlier studies where physical properties and affective or perceptual responses were linked (e.g. [1], [18], [19]) to develop multilayered models of products.

2 Structural Equation Modeling

Structural equation modeling (SEM) [5], [6] is used to verify a linear model of continuous variables. The computation of SEM is based on a comparison between the covariance structure of a hypothesized model and the covariance matrix of the observed variables. If they are statistically close, then the hypothesized model is accepted. Otherwise, the hypothesized model is fixed and tested repeatedly. Hence, SEM is used to judge whether a hypothesized model that is scientifically grounded fits the observed data and to infer causality among multiple variables.

A linear connection between two variables indicates that a variation in one variable affects the variation in another variable. Such that the covariance structure of the model and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2879944, IEEE Transactions on Affective Computing

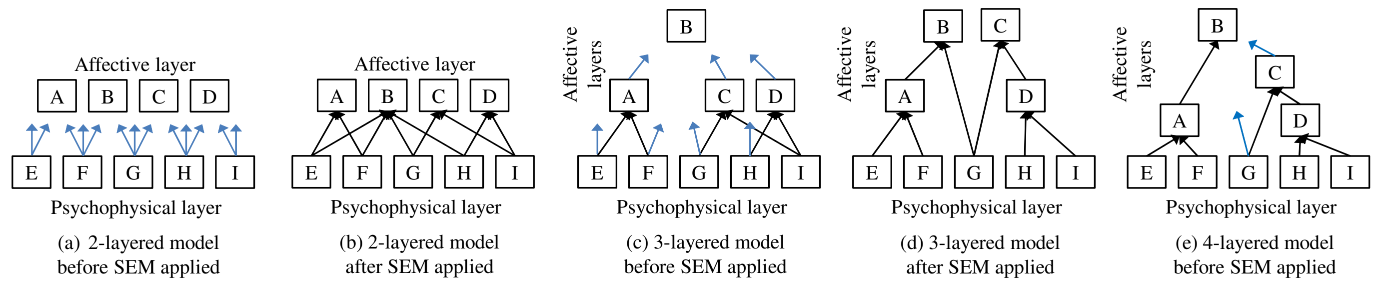

Fig. 2. Scheme of general flow. Models are developed from (a) to (e). Blue arcs are possible connections whereas black ones are specified by SEM. Nodes A–I are variables of adjective ratings.

sample covariance matrix become most similar, the magnitudes of linear connections in the model are computed. The large magnitude value indicates a strong connection between two variables. Each of the linear connections in a hypothesized model is statistically tested, and the entire model is also judged as to whether it conforms to the observed evaluation data. Unlike in multivariate regression analysis, multiple variables can be objective variables simultaneously. Furthermore, sequential or hierarchical relationships among variables can be modeled.

Models are validated from quantitative and semantic viewpoints. Quantitative indices are separated into gross and local ones. Gross indices evaluate the overall goodness of fit whereas local indices evaluate specific parts in a model.

As gross indices, goodness-of-fit index (GFI) and comparative fit index (CFI) are often used. Both indices range between 0–1 with 1 being a complete match between the observation and model. They indicate relative similarities between the covariance matrix of the observed data and covariance structure of the model, and the latter index takes into account the model's degrees of freedom [22], [23]. These values are preferred to be greater than .90–.95. The likelihood ratio test for these two matrices is also considered, and it is preferred that the observed and estimated covariance matrices are not significantly different ($p > .05$), although the model is not evaluated solely on this test [24]. These indices for gross assessment are criteria to evaluate whether the entire model is accepted.

The local index assesses how well specific variables are explained by the model, and R^2, which is a square value of the correlation coefficient between the observed and estimated value of a specific variable is often used. Hence, it ranges between 0–1 with a complete match being indicated by 1. Unlike gross indices, there is no standard value for R^2, and a greater value indicates a smaller prediction error. This index suggests which part of the model should be modified.

SEM may accept scientifically meaningless models, because it determines connections between variables primarily based on a mathematically optimal solution, with little consideration of each variable's context. Thus, to confirm the model soundness, the model also needs to be semantically validated. This is based on the consideration of whether the model can be interpreted without semantic contradiction.

3 General Flow for Constructing Multilayered Model

Herein, we introduce how the model is developed in general whereas specific examples using the plastic plates data are shown in Section 5 and Appendix A. Starting with a simple two-layered model, which is described in Section 3.1, we gradually develop it into complex models. The two methods used to develop the model are described in Sections 3.2 and 3.3, respectively. Fig. 2 shows the general flow for establishing the model. The model continuously repetitively develops until a valid statistical model is obtained that explains the observed sensory evaluations. Regarding the example of plastic plates, we reach the same final model by using either of two methods, although it remains to be investigated whether they will always necessarily estimate the same model.

3.1 Initial two-layered model

According to Guest et al. [9], psychophysical and affective adjectives are clearly segregated. In contrast, the layered structure among affective adjectives is not deterministic and thus depends on the stimuli and adjectives in the particular scenario. Hence, it is reasonable to start model development with a two-layered model as shown in Fig. 2(a) with psychophysical words placed at the lower layer and the remainder placed at the higher layer. Adjectives relating to psychophysical responses directly describe the physical properties of stimuli. To judge which adjectives are psychophysical, previous studies (e.g. [1], [9], [16], [20]) are informative. SEM is then applied to the two-layered model with possible connections assumed between the higher and lower layers. The connections that are judged as significant by SEM are employed in the model.

3.2 Reconstruction of upper layers: Step-wise method

In the previously described two-layered model, affective adjectives with high fitness indices are well explained by the psychophysical adjectives in the lower layer. In contrast, those with low fitness indices also need linkages with non-psychophysical adjectives. For this purpose, one of these adjectives is placed in a new higher layer. For example, we suppose that a two-layered model shown in Fig. 2(b) is acquired by the SEM analysis, and the fitness index of node B is low and thus should be considered for remodeling. Hence, this node is then located on a higher layer such that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2879944, IEEE Transactions on Affective Computing

it can be linked with other affective nodes A, C, and D as shown in Fig. 2(c). If this new model is found superior to the previous model by the SEM indices, then the new model is accepted. These processes are also applied to other variables with low indices.

As a result of the aforementioned processes, suppose that we have the model of Fig. 2(d). In the case that a node (node B) in the highest layer still exhibits low fitness index, this node is placed in a new top layer as shown in Fig. 2(e). Node B is then assumed to be explained by the variables in lower layers.

A specific example of the step-wise method using the plastic plates data is shown in Appendix A.

3.3 Reconstruction of upper layers: Method based on partial correlation matrix

In the former section, we introduced a method where variables with low fit indices are moved to higher layers one by one to improve the model. Here, we describe another method to estimate the structure of upper layers on the basis of partial correlation coefficients. The partial correlation coefficient corresponds to a correlation coefficient between two variables when the other variables are fixed and is used for judging whether two variables are directly influenced [21]. If the partial correlation coefficient is large, then a direct relationship is supposed between the two variables. We use the characteristics of hierarchical models for determining causality, that is, which variable is the cause and gives rise to another variable. In hierarchical models, adjectives in lower layers are the causes and explain resultant adjectives in higher layers. Lower layers are understood as not affected or determined by higher layers. Following these characteristics, some candidate models can be produced on the basis of partial correlation coefficients. The validity of the models is then tested by SEM. In an example described later in Section 5, we use this method to build a model to elaborate on the method.

4 Experiment: Sensory evaluation of plastic materials by touch

Participants performed sensory evaluation tasks where they touched each of ten types of plastic plates and rated their impressions by using adjectives as criteria. The experiment was approved by the ethical committee of the School of Engineering, Nagoya University (#15-12).

4.1 Material as stimuli

Ten types of plastic plates made of different materials were used for the experiment. They were cut into 3 cm by 7 cm pieces and comprised materials that are commonly used for daily products such as polyethylene, polypropylene, polyethylene terephthalate, and specially synthesized plastic[1]. Their surfaces were flat with the average roughness (R_a) values ranging 0.5–3 μm. Their thickness was $\sim$1 mm, and all the surfaces were unlubricated and dry. Nonetheless, some of them felt wet because of differences in friction and heat transfer properties. Also, because they were stored in a temperature-controlled room (28°C) for at least 1 hour before the task, their surface temperatures were considered the same. Plates were cleaned with water and ethanol before being tested by each participant.

4.2 Participants

Eleven university students (nine men and two women, age: 21 ± 1.8) participated in the task, after providing written informed consent. All participants were unaware of the objectives of the study and declared no noticeable deficit in tactile perception.

4.3 Tasks

Before the experiments, participants washed their hands with a piece of soap, dried them, and left them untouched by any objects for 20 min. The room temperature was controlled to 28°C.

Each participant explored the surface of each plastic plate by sliding his/her fingers over it and was restricted from lifting or bending the plate. Before and during the experiment, the participants did not see the plates; each plate was placed in a box with a curtain to ensure that the plate was not seen by the participants. Although there was no time limit, in most trials, the exploratory operation lasted at most 10 s. The participant then rated the plate using nine types of adjective dyads as criteria. The rating for each adjective dyad was performed on a 9-point Likert scale. Each of the adjectives in the dyad was located at either extreme of the scale with the center being labeled as neutral. Participants were told that the plastic materials would be used for a specific product[2]. They were prompted to conduct the rating task without contemplation. Before the main task, the participant experienced all the materials to familiarize themselves with the variation of their surface properties.

Each of the ten types of materials was presented to the participants in randomized order in a single set where two sets were performed for each participant. Hence, individual participants tested each stimulus twice. If the ratings for the first and second sets were inconsistent for a certain participant, then we excluded him/her from the statistics - this was ultimately unnecessary.

Criteria for ratings were nine types of adjective dyads as shown in Table 1. They were randomly presented in the native language of participants, i.e. Japanese. Each dyad included two words which were as semantically opposed as possible. We did not intend to exhaustively cover all the potential relevant adjectives. Instead, through a preliminary investigation where 30 dyads based on adjective lists of previous studies (e.g., [7], [9], [20]) were rated by a small participant group that did not join the main experiment, we selected those that related to the value of the supposed product. Dyads that exhibited high correlation coefficients were merged, such as *comfortable-uncomfortable* and *pleasant-unpleasant*. These two dyads were also highly correlated in another study [7]. Furthermore, dyads that did not substantially vary across the stimuli were excluded.

1. Because of a contractual agreement with the material provider, we cannot disclose the details of the materials.

2. It is a common item; however, the type of product is not disclosed because of a contractual obligation.

TABLE 1
Adjective dyads used in the questionnaire. Italic words in parentheses are Japanese words used in the experiment.

Soft-Hard	(*Yawarakai-Katai*)
Sticky-Slippery	(*Hikkakaru-Suberu*)
Rough-Smooth	(*Arai-Nameraka*)
Cold-Warm	(*Tsumetai-Atatakai*)
Wet-Dry	(*Shittori-Kawaita*)
Desirable-Undesirable	(*Suki-Kirai*)
Expensive-Inexpensive	(*Koukyuna-Yasui*)
Gentle-Harsh	(*Yasashii-Yasashikunai*)
Comfortable-Uncomfortable	(*Kokochiyoi-Fukaina*)

As part of preliminary data processing, the data from two trials against the same stimuli were averaged for each participant. Each participant's ratings for a specific adjective dyad were then normalized such that their mean and standard deviation were 0 and 1, respectively. We then computed a covariance matrix of the ratings of nine adjective dyads for the SEM analysis of Section 5.

5 Construction of Multilayered Model

We established multilayered models by using the methods described in Section 3. Although the two types of methods in Sections 3.2 and 3.3 reached the same conclusion, herein, a method based on a partial correlation matrix is elaborated. An example of the step-wise method described in Section 3.2 appears in Appendix A. For the SEM computation, we used the *SEM* (version 3.1) package for *R* (version 3.4.3).

5.1 Two-layered model: Establishment and validation

The two-layered model was composed of a lower layer of psychophysical words and an upper layer of affective and hedonic words. Among the adjectives used in the example, those directly related to the perception of physical properties were *sticky-slippery*, *rough-smooth*, *soft-hard*, and *cold-warm*. The *wet-dry* dyad indicates the physical status of materials; however, humans do not have perceptual mechanisms for sensing surface wetness and the sense of wetness is a byproduct of the integration of the senses of friction, coolness, and softness [25], [26], [27]. Hence, we placed *wet-dry* at the higher layer. The higher layer included adjectives relating to affective and hedonic aspects, which were *wet-dry*, *desirable-undesirable*, *expensive-inexpensive*, *gentle-harsh*, and *comfortable-uncomfortable*. We applied SEM on the two-layered model assuming that all adjectives in the lower layer could be correlated, and all adjectives in the lower layer could influence those in the upper layer.

Fig. 3 shows the two-layered model analyzed by SEM. Unidirectional arcs indicate the direction of influence and the accompanied values are their magnitude. Bidirectional arcs indicate correlation between the adjective nodes. All of the depicted arcs are statistically significant (z-test, $p < .05$). The R^2 values near the nodes in the upper layer indicate how well the nodes are explained by linking with the lower layer.

The two-layered model indicated that it well represented *comfortable* ($R^2 = .62$) and *wet* ($R^2 = .52$) whereas it did not represent *expensive* ($R^2 = .31$), and *gentle* ($R^2 = .31$) effectively. The gross assessment indices were $GFI = .75$, and $CFI = .75$, and the covariance matrix estimated by the model was not consistent with the observed one ($\chi^2 = 151.5$, $df = 15$, $p < .001$).

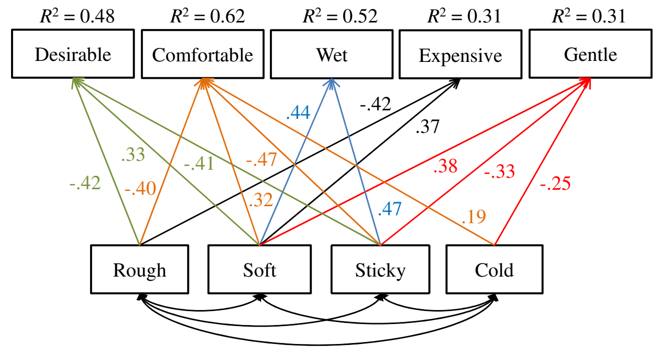

Fig. 3. Two-layered model of adjectives ($GFI = .75$, $CFI = .75$). Arcs and values are colored merely for visual clarity. Bidirectional arcs indicate that two connected nodes exhibit correlation. This model is significantly different from the observation ($p < .001$). For every adjective node, only the positive side is labeled. For example, a negative influence value between *rough* and *comfortable* indicates that rougher materials were felt as less comfortable.

TABLE 2
Matrix of partial correlation coefficients between *expensive*, *gentle*, *wet*, *comfortable*, and *desirable*

	Comfort	Expensive	Desirable	Gentle
Wet	.06	.29	.18	.06
Comfort.		.06	.52	.06
Expensive			.37	.07
Desirable				.32

5.2 Reconstruction of upper layer

To improve the goodness of fit for *desirable*, *expensive*, and *gentle*, we reconstructed the structure of the upper layer. For this purpose, we computed the partial correlation coefficients among the 5 adjective variables in the upper layer as in Table 2 and predicted their causality. When the partial correlation coefficient between two variables is large, a connection is suggested between them. Hence, hypothesized models are considered based on large coefficients. The partial correlation matrix merely provides the hypotheses of model, and their statistical validity is finally tested by SEM. Herein, we build models using coefficients greater than .20; however, if the models are judged invalid, then even smaller coefficients should be used to build models.

The partial correlation matrix suggests that *expensive*, which was not well explained in the two-layered model, would be linked with *wet* and *desirable*. Also, *gentle* would be linked with *desirable*. A high partial correlation coefficient between *desirable* and *comfortable* suggests their strong connection. The correlation coefficient of their ratings was .83, which was high considering the natural variation of participants' ratings. These adjectives might have been interpreted as very similar words as in [7].

Based on the aforementioned facts, 3 types of models shown in Fig. 4 could be viable candidates. For all these models, *expensive* and *gentle* of which the fitness indices were low in the two-layered model, were placed in the top layer and explained by *wet* and *desirable*. Model A includes

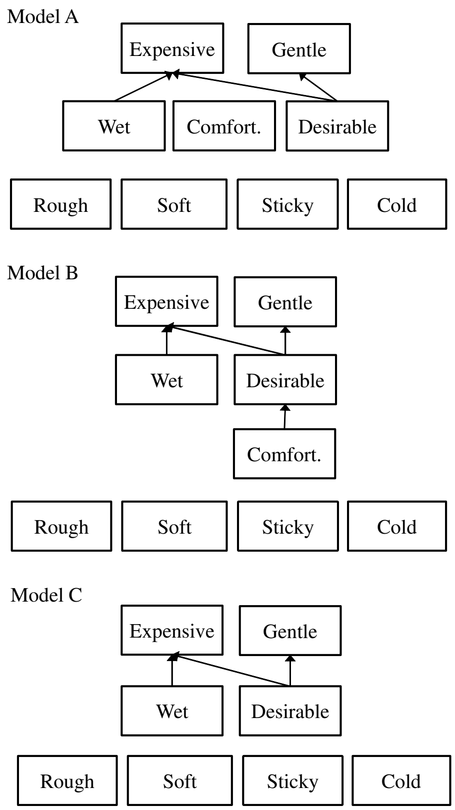

Fig. 4. Three types of model candidates. Models A and C are 3-layered structure whereas model B is a 4-layered structure. Model C does not include *comfortable* based on the suggestion that *comfortable* and *desirable* are semantically similar. Directional arcs are direct influences suggested by the partial correlation matrix of *expensive*, *gentle*, *comfortable*, *desirable*, and *wet*. Other arcs are specified through structural equation modeling.

wet, *comfortable*, and *desirable* in the middle layer. Model B is four-layered with *wet* and *desirable* in the middle layer whereas *comfortable* was located beneath *desirable*. Model C does not include *comfortable* based on the judgment that *desirable* and *comfortable* were practically the same. To indicate the aforementioned logics, the figures include only the directional arcs that are predicted by the partial correlation matrix. Nonetheless, adjectives in a certain layer can be explained by other adjectives belonging to the lower layers. Similar to the two-layered model, psychophysical adjectives in the bottom layer can be correlated with each other.

5.2.1 Fitness indices of model A

Fig. 5 shows model A of which parameters were determined by SEM. Its indices for gross assessment were $GFI = .88$, $CFI = .88$, and $p < .001$ ($\chi^2 = 84.4$, $df = 17$). These values are preferable to those for the two-layered model; however, the model does not statistically explain the variation of observed adjective variables. Nonetheless, the R^2 values for *expensive* and *gentle* improved by .22 and .15, respectively, compared with the two-layered model. In terms of these adjectives, three-layered models are more suitable. The R^2 value for *gentle* improved by being connected with an affective word (*desirable*) in the middle layer and a psychophysical word (*cold*) in the lower layer.

As suggested by the partial correlation matrix, *expensive* was significantly linked with *wet* and *desirable*, and *gentle*

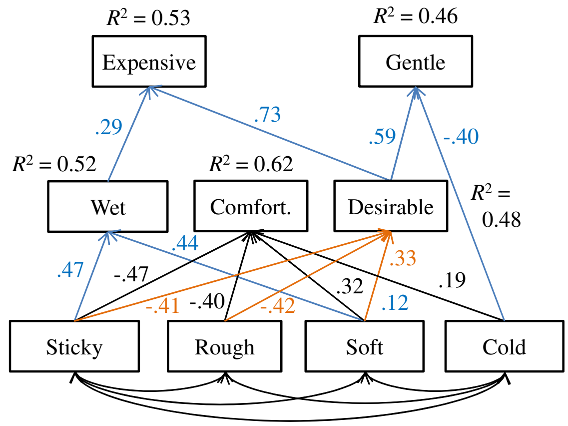

Fig. 5. Model A. A three-layered model with the *comfortable* node. $GFI = .88$, $CFI = .88$. Colors are used just for visual clarity. This model is significantly different from the observation ($p < .001$).

was significantly linked with *desirable*. SEM accepted these linkages as suggested by the partial correlation coefficient. In terms of *wet*, *comfortable*, and *desirable*, they were explained by the adjectives in the lower layer, which was the same as the two-layered model. Hence, their R^2 values were the same as those in the two-layered model.

5.2.2 Fitness indices of model B

Fig. 6 shows model B, which is a four-layered structure. The indices for gross assessment were $GFI = .95$, $CFI = .99$, and $p = .13$ ($\chi^2 = 24.7$, $df = 18$) which were preferable to those for the two-layered model and model A. Also, these indices are acceptably high and indicate that the model well explains the covariation of observed variables. Nonetheless, the number of participants in the present study was not large, and the power estimate [28] of the model was as low as 0.25 ($\alpha = 0.05$).

The R^2 values for *expensive* and *gentle* improved by .22 and .13, respectively, compared with those for the two-layered model. The R^2 value for *desirable* surged from .48 to .70, which is mainly because *desirable* was accompanied with *comfortable*. The *soft* node was directly linked with *desirable* and indirectly with *desirable* by way of *comfortable*.

In terms of *wet* and *comfortable*, their explanatory variables were the same as those in the two-layered model and model A.

5.2.3 Fitness indices of model C

Fig. 7 shows model C where *comfortable* was not involved. The indices for gross assessment were $GFI = .96$, $CFI = .99$, and $p = .17$ ($\chi^2 = 17.6$, $df = 13$), and in terms of goodness of fit, model C is better than the two-layered model and model A. Also, these indices are as good as those for model B. The power estimate of the model was 0.21 ($\alpha = 0.05$). The R^2 values for *expensive* and *gentle* improved by .22 and .15, respectively, compared with the two-layered model. Regarding *wet* and *desirable*, their explanatory variables are the same as those in the two-layered model and models A and B.

We could have continued to reconstruct the model to try to improve the goodness of fit. Nonetheless, according to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2879944, IEEE Transactions on Affective Computing

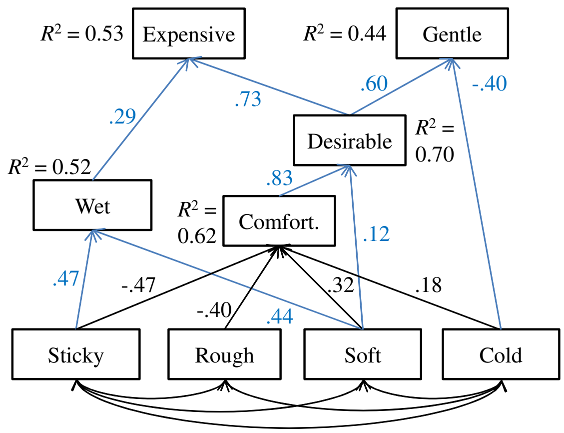

Fig. 6. Model B. A four-layered model. $GFI = .95$, $CFI = .99$. Colors are used just for visual clarity. This model statistically explains the observation ($p = .13$).

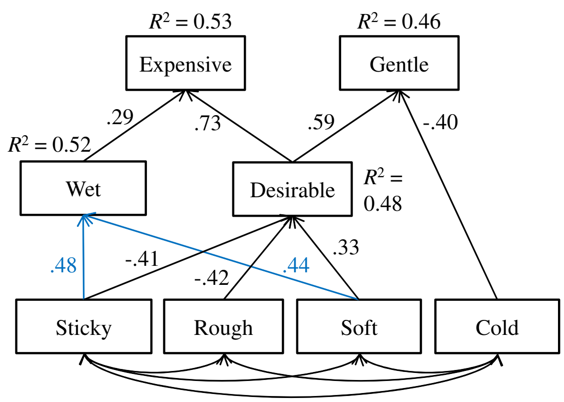

Fig. 7. Model C. A model without the *comfortable* node. $GFI = .96$, $CFI = .99$. Colors are used just for visual clarity. This model statistically explains the observed adjective values ($p = .17$).

Table 2, *expensive* and *gentle* are weakly related, and it was not rational to locate either *expensive* or *gentle* on a higher layer expecting an improvement in the model. Hence, we stopped the development of the model at this point.

5.3 Semantic validity of models B and C

Here, we semantically validate models B and C, both of which have fit indices that are acceptably high. Note that the models depend on the context of the experiments including the stimuli and set of adjectives, and the semantic validity is considered within the scenario of the present study.

The models suggested that *wet* is explained by a combination of *sticky* and *soft*, which is consistent with earlier studies. Guest et al. reported that the responses to *wet*, *damp*, *cold*, and *greasy* were correlated in a sensory task using cloths [9]. The connection between *wet*, *sticky*, and *cold* was also true in a study involving 47 types of materials [7]. Chen et al. reported that subjective reports of *wet* were correlated with physical compliance and friction in their study using 37 types of papers [15]. In a study by Tanaka and Sukigara, the sense of the wet of fabrics was accompanied with quantities related to heat flux and compressibility of fabrics [26]. Shibahara and Sato found that softer and colder cloths felt wetter [27]. Furthermore, a variation along the frictional dimension is accompanied with the variation along the wet dimension on the textural space [20]. The observed relationships between *wet*, *sticky* (frictional), and *soft* in our example are consistent with these earlier studies, although *wet* and *cold* were not statistically linked in our example.

In model B, the *comfortable* node in the middle layer suggested that smooth, slippery, soft, and cold materials were judged as comfortable. This means that in our experiments, comfort was similar to the weakness of the physical stimuli because the materials that were less stimulating to skin were comfortable. This also holds for diverse materials [7] where smooth, less sticky, and warm features led to comfort. These are understandable connections.

In model B, the ratings for *desirable* are mainly determined by those for *comfortable*. Their correlation coefficient was .83 and suggests that these adjectives were interpreted as possessing very similar meanings. Model B indicates that less stimulating materials led to higher scores in *desirable*. Hence, *desirable* might have meant physical comfort rather than preference, that is usually highly individual. Such similarities between *desirable* and *comfortable* may depend on the stimulus and adjective sets, and would not be a general conclusion, but likely a byproduct of the particular circumstances. Responses to *comfortable*, *desirable*, and other positive evaluative words such as *pleasant* and *relaxing* were highly correlated with each other in an experiment where various types of materials were touched [7].

Models B and C suggest that *expensive* was a combination of *desirable* and *wet*. Materials that were less stimulating and wet were judged expensive. This is consistent with our intuition considering that the specimens used in the experiment were hard plastic plates. It is intriguing that *sticky* influences the attribution of *expensive* in both positive and negative manners. Stickier materials felt wetter and positively lead to expensive ratings. In contrast, stickier materials are undesirable and negatively lead to expensive ratings. In this manner, hierarchical models can graphically indicate that the friction of materials influences the affective value of objects by way of multiple indirect effects. Although such relationships depend on contexts, regarding the tactile feelings of woods, Fujisaki et al. reached a similar result where ones that felt smooth, warm (less stimulating), and wet were judged expensive [2]. Furthermore, in a study using finely textured glasses, *wet* and *smooth* responses were strongly correlated with the subjective responses to *expensive* [29].

The rating for *gentle* was a combination of those for *desirable* and *cold*. More desirable (physically less stimulating) and warmer materials were appraised as gentler. As the surface temperatures of the material samples used in the experiment were the same, the perception of warmth relied on their thermal resistances or related properties that quantify how easily the heat flows from human skin to the sample [30], [31], [32]. As shown in model C (Fig. 7), interestingly, warmth of materials hardly affected *desirable* whereas it did affect *gentle*. The difference between two affective and indefinable words of *desirable* and *gentle* lay in the sense of warmth. Such results may not be general and may depend on the experimental conditions; however, in terms of *gentle*, our model is semantically sound.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2879944, IEEE Transactions on Affective Computing

Aforementioned interpretations do not include apparently semantic contradictions.

6 Discussion

As previously mentioned, a hierarchical model of adjectives relating to perceptual, affective, and hedonic aspects helps us understand human responses to physical stimuli. For example, *expensive* and *gentle* are strongly affected by *desirable*; materials that are less stimulating to the skin, that is, smooth, slippery, and soft, are judged as expensive and gentle. Furthermore, *gentle* is affected by *cold* the way that feeling warmer leads to and is associated with a gentler impression. Thus, gentleness felt through the material is improved by controlling its thermal properties. In contrast, *sticky* or frictional components negatively affect *gentle* by way of associations to *comfortable* and *desirable*, whereas it has a positive impact on *expensive*. In this way, the hierarchical model graphically describes how products' values described with adjectives mutually interfere, which facilitates product development.

The models built by our method were semantically and quantitatively valid except for the power value, which indicates that the same method is likely to be applied to other stimuli and adjective word sets. Nonetheless, there still exist inherent difficulties in building hierarchical models. Here, we discuss the possible problems and solutions.

The principal problem may be how we should select the candidate models. We proposed to narrow the possible models by gradually developing hierarchical models using determination coefficients (R^2) as local assessment indices and partial correlation matrix among adjective variables. However, when the number of adjectives used for the sensory evaluation task is large, the number of model candidates is also large. For such cases, it would be helpful to reduce the number of adjective variables by removing or grouping similar adjectives. Principal component analysis or related methods are effective for identifying similarities between the variables. Consequently, a reduced model with a smaller number of adjectives is established first. The reduced model determines the global structure of the model. It would be easier to add to the variables when the global structure is given.

We should take care when interpreting the capabilities of linear modeling. Human perceptual, affective, and hedonic responses are nonlinear [33]; however, their limited part or range can be approximated by a linear model. This means that the variety of stimuli used for experiments needs to be restrained. The stimulus set should include those that can be actually used for the product, and not include the haphazardly wide variety of stimuli, which would degrade the accuracy of linear models. When the fitness indices of SEM unacceptably low, we may exclude some stimuli that do not fit into linear models by outlier analyses.

An important role of adjective modeling is to understand their causal relationships. Observed variables should include both causal and resultant adjective variables. It is better to investigate responses using as many adjectives as possible; however, it is expensive. Otherwise, adjectives should be carefully selected by experts and product developers. For instance, in our example, it had been suggested by expert panels that the feel of premium denoted by *expensive* would be related to the sense of wetness. Hence, we included *wet-dry* and *expensive-inexpensive* in the sensory evaluation task. However, the local assessment index (R^2) for *expensive* was approximately .50, which is not very high. Other adjectives should be considered for inclusion to improve its fit index. SEM can deal with latent variables that occasionally complement variables unused in the experiment. However, the introduction of latent variables multiplies the number of candidate models that should be considered. Cases with latent variables are apparently more difficult than the one exemplified in the present study.

The individual difference in responses to stimuli is a concern for research involving numerous participants. When participants belong to more than two populations of different properties, we cannot expect that one model fits all participants. There exist individual differences in hierarchical models of adjectives [12]. Furthermore, the percept of physical properties (especially friction) also varies among participants [34], [35], [36]. For the cases where individualities are concerned, it would be beneficial to classify participants into groups based on the attributes of participants and develop a model for each group. Or, the classification of participants based on the similarities between individual covariance matrices of adjective ratings is also suitable for the SEM method [12].

Finally, it is not secure that our method always reaches the best model, although a definition of the best model does not exist for our problem class. The next step is to apply the method to other problems and clarify and resolve its limitations.

7 Conclusion

Hierarchical models of adjectives relating to psychophysical, affective, and hedonic responses are instrumental in comprehending and designing the affective effects of products' physical properties. However, there has not been a general method to build such models, and designing has been left to experienced developers. To solve this issue, we developed a general approach to establish the model by using the results of sensory evaluation where stimuli are rated on the basis of adjective criteria. In our approach, the model is gradually developed from simple to complex while each model is validated by SEM. We commenced with a two-layered model with psychophysical and affective or hedonic adjectives at lower and higher layers, respectively. Parts of the model with lower goodness-of-fit indices are reconstructed into new layers based on statistics comprising partial correlation coefficients.

We established a layered model of responses to plastic plates as an example stimulus set by using 9 types of adjective dyads. Three- and four-layered models that quantitatively explain the results of sensory evaluation tasks were then acquired. These models were also semantically valid. We cannot conclude the general validity of our approaches from one example of plastic plates; nonetheless, they should be pursued as promising methods.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2018.2879944, IEEE Transactions on Affective Computing

Appendix A
Model establishment based on the step-wise method

Here, using the data of plastic plates, the process to establish a model based on the step-wise method introduced in Section 3.2 is described. Similar to the method based on a partial correlation matrix, the process begins with the two-layered model of Fig. 3.

First, *expensive* node is targeted because its local fit index is lowest in the two-layered model. It is located at the higher position such that it can receive influences from the other affective nodes. SEM then approves significant connections from *desirable* and *wet* to *expensive* as shown in Fig. 8a). As a result, the local fit index of *expensive* increased to .53 from .31. The indices for the entire model are $GFI = .82$, $CFI = .83$, and $p < .001$ ($\chi^2 = 110.6$, *df* $= 16$) and are preferable than those for the two-layered model in Fig. 3.

Secondly, *gentle* node of which local fit index is lowest in the model of Fig. 8a) is to be improved. This node was displaced to the top layer from the middle one as shown in Fig. 8b). After being tested by SEM, *gentle* was found to be linked by *desirable* and *cold* while its direct connections with *soft* and *sticky* were removed. The R^2 value of *gentle* increased to .46 by .15. The gross indices are $GFI = .88$, $CFI = .88$, and $p < .001$ ($\chi^2 = 84.4$, *df* $= 17$) and better than those for the previous step; however, the entire model is not still statistically acceptable.

Furthermore, *gentle* node continues to be improved because its local fit index is still lowest in Fig. 8b). To improve the model, *gentle* is attempted to be located atop *expensive*, and the connection between these two nodes is tested. However, no significant link from *expensive* to *gentle* is approved by SEM, and a four-layered model with *gentle* and *expensive* being at the highest and second highest layers, respectively, is not accepted.

The node of *desirable* that has the second lowest local fit index in Fig. 8b) is then to be considered for remodeling. The node of *desirable* is placed between the second and third layers as shown in Fig. 8c) because *desirable* influences *expensive* and *gentle* in the third layer and these connections should be remained. The node of *desirable* is connected by *comfortable* and *soft*, and its R^2 value improved to be .70 from .48. The gross indices are $GFI = .95$, $CFI = .99$, and $p < .13$ ($\chi^2 = 24.7$, *df* $= 18$), and the model statistically explains the observed data. Hence, the model development is finished here. The resultant model in Fig. 8c) is of the same structure with the model acquired by the method based on the partial correlation matrix shown in Fig. 7.

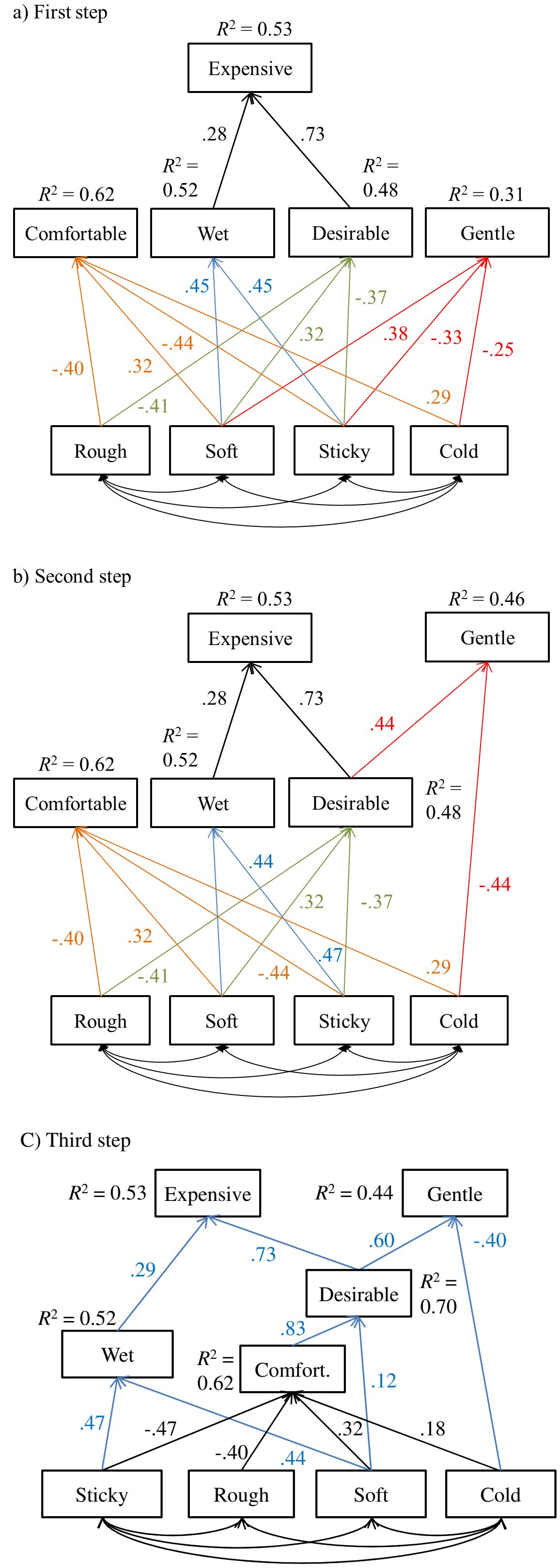

Fig. 8. Development of the model when the step-wise method is applied. a) First step. *expensive* node of the two-layered model is located at the higher layer. b) Second step. *gentle* node is moved to the highest layer. c) Third step. *desirable* node is placed in the second highest layer. This model is accepted by SEM, and same with the model of Fig. 7. Arrows and values are colored only for visual clarity.

References

[1] T. Matsuoka, H. Kanai, H. Tsuji, T. Shinya, and T. Nishimatsu, "Predicting texture image of covering fabric for car seat by physical properties," *Journal of Textile Engineering*, vol. 54, no. 3, pp. 63–74, 2008.

[2] W. Fujisaki, M. Tokita, and K. Kariya, "Perception of the material properties of wood based on vision, audition, and touch," *Vision Research*, vol. 109, pp. 185–200, 2015.

[3] T. Nishino, M. Nagamachi, and M. Sakawa, "Acquisition of kansei decision rules of coffee flavor using rough set method," *Kansei Engineering International*, vol. 5, no. 4, pp. 41–50, 2006.

[4] H. Nagano, S. Okamoto, and Y. Yamada, "Semantically layered structure of tactile textures," in *Haptics: Neuroscience, Devices, Modeling, and Applications*, ser. Lecture Notes in Computer Science, M. Auvray and C. Duriez, Eds. Springer, 2014, vol. 8618, pp. 3–9.

[5] P. M. Bentler, "Linear systems with multiple levels and types of latent variables," in *Systems under indirect observation: Causality, structure, prediction*, K. G. Joreskog and H. Wold, Eds., 1982, pp. 101–130.

[6] ——, "Structural modeling and psychometrika: An historical perspective on growth and achivements," *Psychometrika*, vol. 51, pp. 35–51, 1982.

[7] K. Drewing, C. Weyel, H. Celebi, and D. Kaya, "Feeling and feelings: Affective and perceptual dimensions of touched materials and their connection," *Proceedings of IEEE World Haptics Conference*, pp. 25–30, 2017.

[8] ——, "Systematic relations between affective and sensory material dimensions in touch," *IEEE Transactions on Haptics*, 2018.

[9] S. Guest, J. M. Dessirier, A. Mehrabyan, F. McGlone, G. Essick, G. Gescheider, A. Fontana, R. Xiong, R. Ackerley, and K. Blot, "The development and validation of sensory and emotional scales of touch perception," *Attention, Perception, & Psychophysics*, vol. 73, no. 2, pp. 531–550, 2011.

[10] R. Ackerley, K. Saar, F. McGlone, and H. Backlund Wasling, "Quantifying the sensory and emotional perception of touch: differences between glabrous and hairy skin," *Frontiers in Behavioral Neuroscience*, vol. 8, p. 34, 2014.

[11] S. Kawabata and M. Niwa, "Objective measurement of fabric hand," in *Modern Textile Characterization Methods*, M. Raheel, Ed. CRC Press, 1996, pp. 329–354.

[12] S. Okamoto, H. Nagano, K. Kidoma, and Y. Yamada, "Specification of individuality in causal relationships among texture-related attributes, emotions, and preferences," *International Journal of Affective Engineering*, vol. 15, no. 1, pp. 11–19, 2016.

[13] H. Nagano, S. Okamoto, and Y. Yamada, "Modeling semantically multilayered affective and psychophysical responses toward tactile textures," *IEEE Transactions on Haptics*, 2018.

[14] K. Kidoma, S. Okamoto, H. Nagano, and Y. Yamada, "Graphical modeling method of texture-related affective and perceptual responses," *International Journal of Affective Engineering*, vol. 16, no. 1, pp. 27–36, 2017.

[15] X. Chen, C. Barnes, T. Childs, B. Henson, and F. Shao, "Materials' tactile testing and characterisation for consumer products' affective packaging design," *Materials & Design*, vol. 30, no. 10, pp. 4299–4310, 2009.

[16] I. H. M. Hashim, S. Kumamoto, K. Takemura, T. Maeno, S. Okuda, and Y. Mori, "Tactile evaluation feedback system for multi-layered structure inspired by human tactile perception mechanism," *Sensors*, vol. 17, p. s17112601, 2017.

[17] K. Ueda, "A hierarchical structure for adjectives describing timbre," *The Journal of the Acoustical Society of America*, vol. 100, no. 4, p. 2751, 1996.

[18] S. Kawabata and M. Niwa, "Fabric performance in clothing and clothing manufacture," *Journal of the Textile Institute*, vol. 80, no. 1, pp. 19–50, 1989.

[19] G. Elkharraz, S. T. anf D. Akay, C. Eitzinger, and B. Henson, "Making tactile textures with predefined affective properties," *IEEE Transactions on Affective Computing*, vol. 5, no. 1, pp. 57–70, 2014.

[20] S. Okamoto, H. Nagano, and Y. Yamada, "Psychophysical dimensions of tactile perception of textures," *IEEE Transactions on Haptics*, vol. 6, no. 1, pp. 81–93, 2013.

[21] A. P. Dempster, "Covariance selection," *Biometrics*, vol. 28, no. 1, pp. 157–175, 1972.

[22] P. Bentler, "Comparative fit indexes in structural models," *Psychological Bulletin*, vol. 107, no. 2, pp. 238–246, 1990.

[23] D. Hooper, J. Coughlan, and M. Mullen, "Structural equation modelling: Guidelines for determining model fit," *Electronic Journal of Business Research Methods*, vol. 6, no. 1, pp. 53–60, 2008.

[24] R. P. McDonald and M. H. Ho, "Principles and practice in reporting structural equation analysis," *Psychological Method*, vol. 7, no. 1, pp. 64–82, 2002.

[25] M. J. Zigler, "An experimental study of the perception of clamminess," *The American Journal of Psychology*, vol. 34, no. 4, pp. 550–561, 1923.

[26] Y. Tanaka and S. Sukigara, "Evaluation of "shittori" characteristic for fabrics," *Journal of Textile Engineering*, vol. 54, no. 3, pp. 75–81, 2008.

[27] M. Shibahara and K. Sato, "Illusion of wet sensation by controlling temperature and softness of dry cloth," in *Haptics: Perception, Devices, Control, and Applications*, F. Bello, H. Kajimoto, and Y. Visell, Eds., 2016, no. 2, pp. 512–520.

[28] R. C. MacCallum, M. W. Browne, and H. M. Sugawara, "Power analysis and determination of sample size for covariance structure modeling," *Psychological Methods*, vol. 1, no. 2, pp. 130–149, 1996.

[29] C. J. Barnes, T. H. C. Childs, B. Henson, and C. H. Southee, "Surface finish and touch–a case study in a new human factors tribology," *Wear*, vol. 257, no. 7–8, pp. 740–750, 2004.

[30] S. Kawabata and Y. Akagi, "Relation between thermal feeling and thermal absorption property of clothing fabric," *Journal of Textile Machinery Society of Japan*, vol. 30, no. 1, pp. 13–22, 1977.

[31] M. J. Pac, M. Bueno, and M. Renner, "Warm-cool feeling relative to tribological properties of fabrics," *Textile Research Journal*, vol. 71, no. 9, pp. 806–812, 2001.

[32] H.-N. Ho and L. Jones, "Contribution of thermal cues to material discrimination and localization," *Perception & Psychophysics*, vol. 68, pp. 118–128, 2006.

[33] S. S. Stevens, *Psychophysics: Introduction to its perceptual, neural and social prospects*. John Wiley & Sons Inc., 1975.

[34] M. Hollins, S. Bensmaïa, K. Karlof, and F. Young, "Individual differencesinperceptualspacefortactiletextures: Evidencefrommultidimensional scaling," *Attention, Perception & Psychophysics*, vol. 62, no. 8, pp. 1534–1544, 2000.

[35] A. Klöcker, C. Arnould, M. Penta, and J.-L. Thonnard, "Rasch-built measure of pleasant touch through active fingertip explorations," *Frontiers in Neurorobotics*, vol. 6, p. 5, 2012.

[36] M. A. Masen, "A systems based experimental approach to tactile friction," *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 4, no. 8, pp. 1620–1626, 2011.

Shogo Okamoto received a Ph.D. in Information Sciences from Tohoku University in 2010. He has been with Nagoya University since 2010. Currently, he is an associate professor at the Department of Mechanical Systems Engineering, Nagoya University. His research interests include haptics, assistive robotics, and affective engineering.

Haruyo Kojima received a master degree in psychology, Nagoya University in 2003. She is a Research Scientist in Kao Corporation. Her research interests include haptics and affective engineering.

Atsushi Yamagishi received a Ph.D. in Dentistry from Aichi Gakuin University in 2013. He is a Research Scientist in Kao Corporation. His research interests include haptics, affective engineering and preventive dentistry.

Kyoichi Kato received a master degree in material science, Nagasaki University in 1990. He is a Research Scientist in Kao Corporation. His research interests include haptics and affective engineering.

Atsuko Tamada received Bachelor and Master degrees in Human Sciences from School of Human Sciences, Waseda University in 1995 and 1997, respectively. She is a research assistant at Department of Mechanical Systems Engineering, Nagoya University.