(03-020) - Methodological Proposal for Dimensionality Reduction in Semantic Space and Reflections on Sample Size in Kansei Engineering Applications

Molina Criado, Silvia [1]; Zamora Polo, Francisco [1]; Romero Lemos, Javier [2]; Luque Sendra, Amalia [1]

[1] Dpto. Ingeniería del Diseño, Escuela Politécnica Superior Universidad de Sevilla, Virgen de África 7, 41011 Sevilla, [2] Gabinete técnico en IAAP - Junta de Andalucía

Kansei Engineering is a user-oriented methodology for developing new products. The process begins with collecting kansei, which must be reduced to a manageable number of words. Once a sufficient number of kansei have been obtained, it is indispensable to reduce them to work with a manageable number of words. Affinity analysis has traditionally been used for this purpose. Firstly, repeated words are eliminated, and those considered less relevant are discarded. Secondly, words are grouped according to similar meanings, and preferably, one of each group is chosen to be representative. Affinity analysis may be subject to unintentional biases. Therefore, this paper proposes selecting the kansei to be used from the initial set using dimensionality reduction techniques, such as principal component analysis. Once the semantic and property space are defined, distributing the survey is necessary to obtain results. This paper reflects on the required sample size through analysis and statistics, applying Machine Learning techniques, such as the learning curve.

Keywords: Dimensionality reduction; sample sizing; product design; Kansei Engineering

Propuesta metodológica para reducción de dimensionalidad en espacio semántico y reflexiones sobre el tamaño muestral en aplicaciones de Ingeniería Kansei

La Ingeniería Kansei es una metodología para el desarrollo de nuevos productos orientada al usuario. Comienza con un proceso de recolección inicial de los kanseis. Una vez que se ha obtenido un número suficiente de kansei, se hace necesario reducirlos para poder trabajar con una cantidad manejable de palabras. Para ello, tradicionalmente se ha recurrido al análisis de afinidad. En este análisis se realiza una primera reducción, en la que se eliminan las palabras repetidas y se descartan aquellas que se considera que tienen menor relevancia. En una segunda reducción, se agrupan las palabras según tengan significados parecidos y, preferentemente, se elige una de ellas como la palabra representativa de cada grupo. El análisis de afinidad está sujeto a sesgos involuntarios. Por ello, en este trabajo se propone seleccionar los kanseis a utilizar del conjunto inicial de kanseis utilizando técnicas de reducción de dimensionalidad, tales como el análisis de componentes principales. Una vez definidos el espacio semántico y de propiedades es necesario distribuir la encuesta para obtener los resultados. En este trabajo se reflexiona sobre el tamaño muestral necesario, a través de análisis y estadísticos y mediante la aplicación de técnicas de Aprendizaje Automático, tales como la curva de aprendizaje.

Palabras clave: Reducción de dimensionalidad; tamaño muestral; diseño de productos; Ingeniería Kansei

Correspondencia: Silvia Molina Criado silviamolinacriado@gmail.com

Agradecimientos: Los autores agradecen la financiación obtenida por la Cátedra Goya – Antonio Unanue de Ingeniería en la Industria Agroalimentaria y la financiación del VII Plan Propio de la Universidad de Sevilla.

1. Introduction

In the 1970s, the market was governed by the "product-out" concept, where the manufacturers themselves made the decisions on the design of the products that were put on the market. However, the vastness of today's marketplace makes it inconceivable that such a perspective still exists. The concept of "market-in" refers to the phenomenon whereby consumer demand and preferences dictate the design, function and price of products brought to the market (Nagamachi, 2003). In this context, emotional design techniques, such as Kansei Engineering (KE) aid in designing products that have added value by establishing an emotional connection with the heart, in addition to connecting with the rational side of the brain (Val-Carreres Azofra & Aguayo González, 2012). Therefore, it is difficult for users to dispose of these items, encouraging more efficient management of natural resources.

KE begins by selecting the design domain, where a market research is conducted to identify the target market and gather a comprehensive collection of existing products, concepts, and design solutions that have yet to be conceived (Alves, 2018). Next, the Semantic Space and the Property Space are defined. For the former, it is common to collect a large number of kansei words (KWs), which are representations of the emotions or feelings both expressed by the product or elicited by the product to the consumer (Córdoba Roldán et al., 2010). When selecting properties for the property space, multiple sources are used to identify those with the greatest emotional impact. After defining the KWs and properties, psychological assessment instruments are used to draw up questionnaires. The most well-known instrument is the semantic differential (Osgood y Suci, 1969) and the Likert scale (Likert, 1932). The data collected from the surveys is analyzed during the synthesis phase, where the combination of properties that a particular KW elicits in the user is established (Val-Carreres Azofra & Aguayo González, 2012). The tools used in this phase are divided into manual methods, such as category identification, and statistical methods, including Regression Analysis, General Linear Model, and Quantification Theory Type 1 (QT1), as well as others like Genetic Algorithm, Fuzzy Set Theory, or Rough Set Theory (Prodintec, 2011). Finally, with all the information obtained above, the model is validated and built.

Currently, numerous authors suggest implementing Machine Learning (ML) in KE research. ML refers to the ability of systems to learn from specific training data related to a particular problem. This learning process enables the automation of analytical model construction and facilitates the resolution of associated tasks (Sharma & Chaudhary, 2023). Although implementing it may not be easy, this paper proposes a machine learning-based methodology for verifying results in the synthesis phase.

2. Objectives and methodology

This paper mainly focuses on two critical aspects: reducing the semantic space and selecting the number of respondents.

To define the semantic space, the first step is to gather a large number of KWs from various sources. Several papers have been reviewed to estimate the typical number of initial KWs collected in KE studies. Subsequently, the main methods to reduce this initial number of KWs have been examined. These points are presented in section 3.

Similarly, various application cases of KE have been studied to determine the criteria for the required number of respondents in the synthesis stage. Section 4 proposes a methodology for determining and verifying a solution to the issue.

3. Spanning the Semantic Space

3.1 Revision of the number of initial KW obtained in previous work.

Firstly, a literature search was conducted using the "Web of Science" database to find case studies that applied KE. Table 1 displays the results obtained, including the initial number of KWs collected and the final number of KWs obtained after applying reduction techniques.

Table 1 shows that there is no universal criterion for the recommended number of initial KWs to collect nor the final number, and the results vary significantly. However, it is typical to obtain enough words that including all of them in the survey distributed to users during the synthesis phase would result in an excessive and possibly redundant number of questions.

It is important to remember that the final number of survey questions in the synthesis phase depends not only on the number of selected KWs, but also on the number of product samples obtained in the property space. The number of samples will be influenced by the number of properties and their respective categories, which are the result of the designer's decision. Therefore, it may seem logical that the more profiles there are, the fewer KWs there would be, and vice versa. However, table 1 data confirms that this criterion is not followed.

Table 1: Number of kansei words selected in Kansei Engineering studies.

REFERENCE	PRODUCT	NUM. INITIAL KWs	NUM FINAL KWs	NUM SAMPLES
Zhang et al., 2021	Beverage bottle	46	5	65
X. Li et al., 2021	Hand drill and bicycle helmet	168	6	
Zhong et al., 2022	Outdoor leisure chairs	28	4	15
Y. Li & Zhu, 2020	Car profile design	258	20	10
Cai et al., 2023	In-flight service of a Chinese airline	89	17	12
Jia & Tung, 2021	Wrist wereables	120	40	8
Q. Zhang et al., 2022	Sedan (car)	60	4	24
J. Zhang & Mu, 2021	Suit	120	6	11

It is important to note that the quality of the obtained data is directly related to the number of decisions required of the respondent. This is because of fatigue. Additionally, a high number of questions can incentivize a failure to recruit a sufficient number of respondents. (Lokman & Kamaruddin, 2010).

Therefore, it is necessary to reduce the number of collected KWs initially. The first step would be to eliminate repeated words since there is no single source of KWs. Subsequently, quantitative, and qualitative methods are used.

3.2 Dimensionality reduction

In order to find a smaller set of variables that can represent all or most of the information from the original set, KWs are selected and grouped. Practically, all the works studied use a two-phase reduction scheme, using qualitative methods in the first phase and qualitative methods in the second. KE commonly uses the affinity analysis in the case of qualitative methods, and Factor Analysis (FA) and Principal Component Analysis (PCA) in the case of quantitative methods.

Qualitative methods: Affinity analysis (AA)

The AA, developed by Kawakita Jiro in the 1960s, is commonly referred to as the KJ method. It is one of the Seven Management Tools (Helmold, 2021). These tools were developed by the Union of Japanese Scientists and Engineers (JUSE) after World War II to promote innovation and improve the management of large projects. AA aims to “generate, organize, and consolidate information concerning a product, process, complex issue, or problem”(Helmold, 2021). According to Plain (2007), “the amount of data being organized using this method can be intimidating”, after conducting Affinity Analyses involving more than 600 concepts”.

However, numerous studies concur that this approach is highly susceptible to unintentional bias since it involves subjective decisions by the person(s) involved. Therefore, Lokman & Kamaruddin (2010) attempted to create a Kansei Affinity Cluster to facilitate the development of emotional connections between KWs. They accomplished this by enlisting the help of linguistic experts and successfully categorizing 820 KWs into 42 clusters. Some if this clusters are presented in table 2.

Table 2: Examples of clusters for Affinity Analysis. Source: Lokman & Kamaruddin, 2010

Cluster Name	Kansei Words
Elegant	Advance, Arousing, Artistic, Charming, Clear, Delicate, Deluxe, Dignified, Distinguished, Erotic, Exotic, Exclusive, Enchanting, Exceptional, Extraordinary, Famous, Fantastic, Fascinating, Fashionable, Flamboyant, Gallant, Genuine, Gorgeous, Gracious, Graceful, Impressive, Intellectual, Intelligent, Intentional, Magnificent, Mature, Novel, Outstanding, Passionate, Precious, Romantic, Sexy, Smart, Sophisticated, Splendid, Sporty, Stylish, Sweet, Symbolic, Tasteful, Thin, Trendy, Ultimate, Valuable, Versatile, Well-known
Sophisticated	Abstract, Branded, Distinctive, High Class, High Cost, High Impression, Elite, Expensive, Formal, High Style, Intellect, Limited, Luxury, Official, Premier, Professional, Special, Significant, Stunning, Top-class, Vogue

However, certain concepts, such as 'Advance', may be more similar to other concepts, such as 'Modern' or 'Sophisticated', when applied to certain case studies. It is clear that AA depends on context. Therefore, while this guide is undoubtedly useful, it should be taken as just that - a guide - and should not take precedence over the designer's criteria.

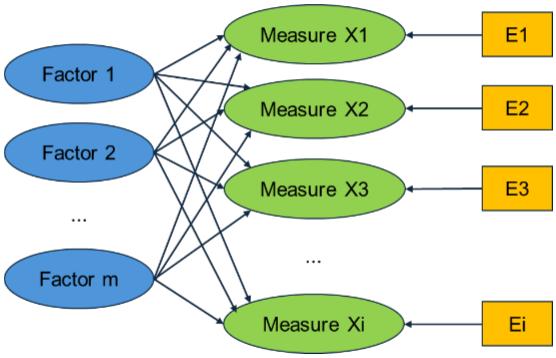
Nevertheless, the improvement that the previous method brings to AA is evident. However, it was developed at a time in history when tools based on data mining and artificial intelligence were not yet developed. With the advent of deep learning tools like Word2Vec, it is now possible to establish semantic and contextual relationships among groups of words. The application of these tools in KE can be highly beneficial. Although literature on this algorithm in the context of KE is still limited, some authors have used it to expand the property and semantic space, as well as to reduce the dimensionality of the semantic space by identifying similarity relations between words (Yang et al., 2024).

Word2Vec allows words to be converted into dense numeric vectors representing words in a multidimensional space, where the distance and direction between vectors reflects semantic and contextual relationships between words. Yang et al. (2024) established that words with a similarity higher than 0.6 will be grouped in the same cluster, and those with similarity lower than 0.4 will be kept in different clusters. Other related tools include Natural Language Processing (NLP)-based chatbots, such as ChatGPT. Although this tool could aid in generating clusters for AA, it is continuously undergoing training. Therefore, the obtained answers are subject to change, and as a result, long-term reliability of the results cannot be guaranteed.

Quantitative methods: Factor Analysis (FA)

The basic assumption of FA is represented in figure 1: for a set of observed variables there is a smaller set of underlying variables called “common factors” that can explain the interrelationships between those variables. A large part of each observed variable Xi depends on the common factors Fi, while the rest comes from an item-specific part and from noise (Ei). The contribution of each factor to the observed variables (matrix L-{lij}) is known as “Factor Loadings” (Equation 1).

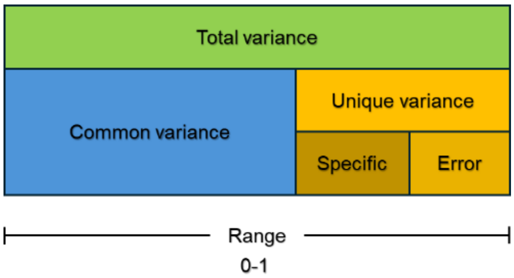
Figure 1: Factor Analysis data reduction



$$X_i = \sum_{m=1}^{M} l_{im} F_m + E_i$$ ñp (1)

In this model, the variance of each observed variable is divided into two components (figure 2): common variance (shared by a group of variables, with values close to 1 indicating that the extracted factors explain most of the observed value of a variable) and unique variance (which includes both variance that depends exclusively on the variable and measurement errors).

Figure 2: Total variance composition in Factor Analysis



Some confusion has been observed in many publications when factor analysis is compared with principal component analysis (PCA). As shown in Figure 3, PCA is a type of FA, as is common factor analysis (CFA). The difference between the two is that PCA considers the common variance and the total variance to be the same and tries to find the factors that maximize the variance by linear combinations. The variance contributed by each factor is obtained from the eigenvalues. The number of eigenvectors, if they exist, is equal to the number of variables, so the dimensionality reduction comes from selecting a subset of factors that explain a relevant part and not all of the variance. It is important that the factors are orthogonal. Figure 4 illustrates these ideas.

PCA is therefore a factoring technique. However, the confusion increases further when other techniques that are grouped under the term CFA come into play, since one of them, Principal Axis Factoring (PAF), uses PCA, albeit in a different way, to estimate the common variance. Other CFA techniques are Maximum Likelihood or Least Squares (LS), although this paper will not focus on them.

Figure 3: Methods for FA

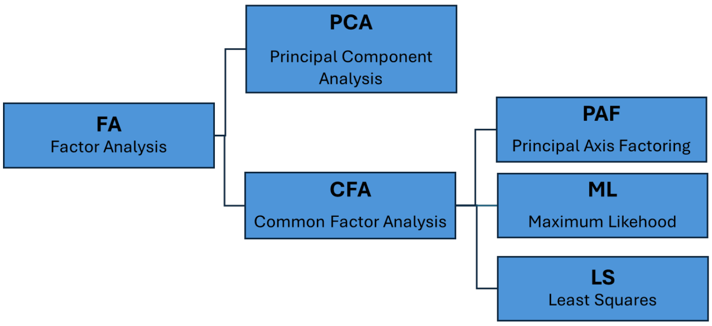
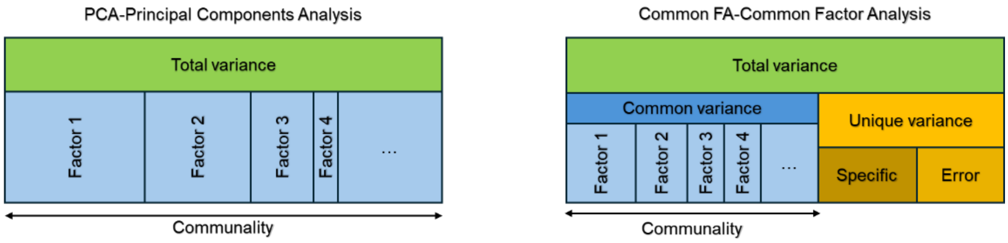


Figure 4: Differences in total variance composition between PCA and CFA



Most of the papers reviewed agree on two views that are contradictory in practice. On the one hand, many researchers claim that the factors obtained by PCA are generally more difficult to interpret. In the context of KE, this means that with PCA it would be more complex to find a single KW that describes or groups all those that have been associated with the same factor. However, many authors also point out that the results obtained using PCA and PAF, or another CFA technique, are similar. There is some consensus that the latter is more appropriate when surveys cannot directly measure the variables that motivate or condition responses, for example those related to the psychological profiles of the population being surveyed.

Methodology for the extraction of kansei factors or kansei reduction

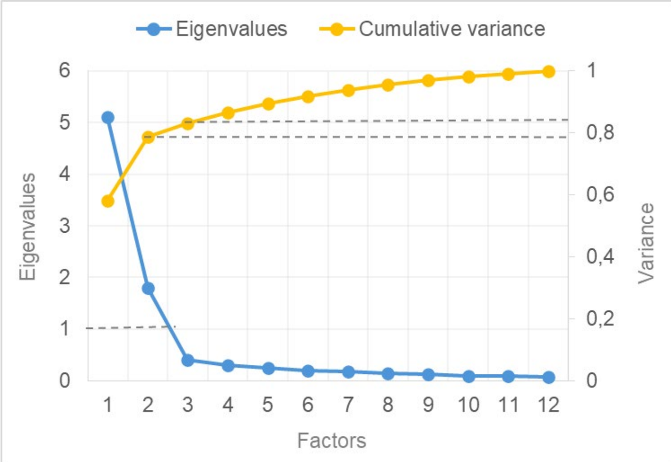
The procedure for extracting and interpreting the factors generally involves the following steps:

1. Sample size (n) determination. An *n* value between 5 and 20 times the number of observed variables (KWs in this case) is advised (Comrey & Lee, 1992). According to MacCallum et al (1999), the above rules and many others tend to overestimate the size of the survey because they do not take into account the nature of factor analysis. In this sense, some authors point out that with high correlations (> 0.6 or > 0.8) samples could be reduced very significantly, even 50 cases would suffice. As it is not possible to know the correlation a priori, this paper advises to carry out the analysis even with a small number of cases.
2. Data preparation. First, the variables should be measurable on an interval scale, a ratio scale, or at least a multi-level Likert scale, which is very useful for categorical or ordinary variables. Secondly, If the intervals or scales between different variables are very different, it is advisable to normalize them by scaling them to homogeneous ranges. This avoids one of them dominating the result.
   The two most common scaling methods are standardization and normalization. Standardization transforms each variable by subtracting its mean and dividing by its standard deviation, resulting in a variable with a mean of zero and a standard deviation of

one. Normalization transforms each variable by subtracting its minimum and dividing by its range, resulting in a variable with a minimum of zero and a maximum of one.
Ordinal or categorical variables (e.g. Likert scale) with a similar number of levels usually do not require scaling. On the contrary, continuous variables usually require scaling. Standardization does not change the structure of the distribution; normalization does, making it flatter and less skewed. The use of techniques such as PCA or Maximum Likelihood assumes that the variables are normally distributed, so this scaling technique would be the most appropriate.
With regard to the previous analyses, it should be noted that if the data is standardized, the subsequent analyses will have to work with the correlation matrix (which shows the relationship between the variables). In the case of standardized data, the covariance matrix (which shows how the variables vary with each other) will be used.

3. Data Suitability Test, to check whether the problem is suitable for factoring techniques. The most common methods are the following:
   - Based on the correlation matrix: Z-score. categorized the correlation loadings as 0.30 = minimal, 0.40 = important, and 0.50 = practically. Therefore, if there are values > 0.30 it is considered factorizable. This is the most reliable method, but also the most laborious.
   - Based on Barlett's test of sphericity. This is an indicator of the similarity of the correlation matrix to the identity matrix (which would indicate that the variables are not correlated). Very small values, less than 0.05, are considered sufficient for FA.
   - Based on the Kaiser-Meyer-Olkin (KMO) Measure (Kaiser, 1970b)(Kaiser, 1970a). This is an overall estimate of the proportion of variance in the variables that may be due to the underlying factors. Values > 0.5 are considered valid for FA.
4. Selection of factoring and extraction techniques. There is often little difference between the methods, and PCA is computationally cheaper. It is almost always advisable to use it, except perhaps in the case of suspected hidden causes, where CFA is likely to give a better result (Burton & Mazerolle, 2011). Common factor techniques are also more appropriate when the number of items per factor is small.
   According to (Costello & Osborne, 2005), once selected common factoring, Maximum Likelihood or PAF will give researcher the best results. The first one is better suitable for normally distributed data, and PAF for significantly non-normal distributions which is very common in variables with Likert-types scales.
   In any case, if in doubt, several types of analysis can be carried out and, if there are differences, the cause should be identified to choose the most appropriate solution.
5. Selection of the number of factors. This is a very important problem for which there is no formula. During the analysis process, it is necessary to decide based on different estimators or the nature of the variables being studied, always bearing in mind that a good solution is one that explains the most variance with the fewest number of factors.
   The most referenced and used methods are two that are also used in a complementary way: the Kaiser´s criterion and the Scree Plot. The Kaiser´s criterion suggests using all factors with an eigenvalue >1. The Scree Plot is a graphical representation of the eigenvalues. Figure 5 shows an example where the Scree Plot results, and the variance explained by including new eigenvalues are plotted. In this example, the Kaiser´s criterion would select 2 factors. In contrast, the Scree plot recommends the inclusion of the factor where an 'elbow' occurs, which in this example is the third factor. However, it is clear from the cumulative variance curve that this third factor does not significantly improve the explained variance.

Figure 5: An example of Scree Plot



The value chosen is indicative only; determining the final number of factors, like the rest of the analysis, is an iterative process of testing different values to arrive at a set of interpretable factors with clear meaning.

6. Interpretation and factor rotation. Factor loadings, also known as components matrix, describe the relationships between the factors and the observed variables. By evaluating the factor loadings of the selected factors is possible to understand the strength of the relationship between each variable and the factor. The final goal is to derive the observed variables corresponding to a specific factor.
   Comrey & Lee (1992) proposed the guideline in the Table 3a to qualify primary factor loadings. This is a good starting point to better understand many of the empirical criteria found in the scientific literature. Factor loadings table usually removes factors below 0.4 to simplify analysis (shown as blank spaces in Table 3b).
   In general, a factor structure is most interpretable if matches three conditions (Pedhazur & Schemelkin,1991):
   - Cond. 1: Each variable loads strongly on only one factor.
   - Cond. 2: Each factor shows 3 or more loadings. The greater the load, the greater the reliability.
   - Cond. 3: Most loadings are either high or low, and just few of them have intermediate values.

   Using the conditions above to evaluate the component table, the results are the followings:

   Variables such as 1, 2, 3, 4, 7, 11 and 15 clearly matches Cond. 1 because they load in just one factor with a “good”, “very good” or “excellent” qualification.

   All factors but factor number 4 match Cond. 2. The relevance of this factor is questionable.

   Some variables such as 10, 13 and 14 has intermediate and close qualifications. They do not match Cond. 3.

   The example above shows very common situations that make the interpretation of the factors very difficult. Frequently, initial factor loadings are difficult to be interpretated, because the most of variables load strongly on the first few factors. Rotation often helps to find a more interpretable factor structure.

Table 3: Guidelines proposed by Comrey & Lee (1992) and example table.

Factor Strength	
> 0.7	Excellent
> 0.63	Very gool
> 0.55	Good
> 0.45	Fair
> 0.32	Poor

Var	Components / Factor Loadings			
	1	2	3	4
1	.774			
2	.767			
3	.747			
4	.727			
5	.719		.527	
6	.515			
7	.872			
8	.610			.441
9	.580		-.420	
10	.542			
11		.630		
12	.410	.651		
13	-.534	.536		
14	.469	-.442	.531	
15				-.611

There are two basic types of factor rotations: orthogonal, where the new axes are also orthogonal, and oblique, where the new axes do not need to be orthogonal in order for the factors to be correlated.

Within the category of orthogonal rotations, there are three well-known algorithms. “Varimax” tries to make each variable representative in a single factor, thus reducing the number of variables within each factor. “Quartimax” tries to explain the variables from a small set of factors. Finally, “Equamax” seeks a compromise between the two. Varimax is by far the most used and mentioned. Its philosophy is best suited to Kansei reduction.

In oblique rotations, the new axes can theoretically take any position in factor space, but the degree of correlation allowed between factors is often limited because two highly correlated factors are better interpreted as just one. “Oblimin” is the most popular oblique rotation method. After performing the rotation, it is necessary to reinterpret the results to decide whether or not to keep the variables or to select other extraction algorithms (Krabbe, 2017).

The final condition of interpretability when the aim is to reduce the number of KWs is not whether or not there is a minimum set of factors that explain the survey scores, but to be able to group all the KWs that are strongly correlated with a factor around a single KW which describes the whole.

4. Determination of the number of respondents in the synthesis phase

As mentioned in the introduction, the choice of survey size in Kansei engineering studies is an important issue that can affect the quality of the analyses. In this paper, we propose to investigate this problem using statistical techniques and a methodology based on ML.

Although numerous authors, such as Almagro (2011), pointed out the lack of studies on the selection of the number of respondents some time ago, little has changed since then. After analyzing the papers searched in section 3.1 of this document, and taking into account the compilation carried out by De las Heras et al. (2023), it was concluded that there is not a criterion for the selection of the number of respondents, as there were results of all kinds. Moreover, none of the publication gave a justification for this choice. In the context of research, it is recommended to always include a final estimate of the confidence level of the results that justify, among other aspects, a right choice of the number of respondents.

In research, it is necessary to be methodologically rigorous and not leave any variable to chance. This is usually done by taking a sample from the target population. "The target population is the population that has the characteristic to be studied and to which the results found in the sample can be generalized" (Canales et al., n.d.). The sample must therefore be representative of the population, i.e. it must follow the distribution of the population from which it is drawn in terms of the variable(s) being studied, such as sex, age, etc. The selection of a sample also depends on many factors, such as the resources available, the heterogeneity of the variables, the type of analysis to be carried out (Canales et al., n.d.).

This paper proposes the need to decide the appropriate sample size in two complementary ways: first, using statistical techniques, then validating the sample size with a model and evaluating the result using learning curves.

4.1 Determination of the number of respondents using statistical techniques.

Various techniques exist for determining the sample size, and the choice of method depends primarily on the initial data. In statistical research, several formulas have been proposed to determine the number of participants in a survey, such as Slovin, Cochran, Murray, and Larry (table 4).

In general, the choice of the method depends on the number and nature of the variables: Sloving is suitable for small and known populations, while Cochran's formula is more appropriate for large and unknown populations.

Table 4: Some available formulas used in statistical research

SLOVIN	COCHRAN	MURRAY & LARRY
$n = \frac{N}{1 + N * e^2}$	$n = \frac{n_0}{1 + \frac{n_0}{N}}$ where $n_0 = \frac{z^2 * p * (1 - p)}{e^2}$	$n = \frac{N * Z^2 * \sigma^2}{e^2 * (N - 1) + Z^2 * \sigma^2}$

N (Population size), e (Margin of error), σ (Standard deviation), Z^2 (Z Score, from desired risk), p (estimated proportion of the population which has the attribute in question).

It is important to note that these parameters are interrelated. For the same Z^2, a larger *n* results with a smaller *e*, and vice versa. Additionally, the *e* is influenced by the respondents' choices. If the majority choose a certain option, the *p* for that choice is high, resulting in a low *e*, regardless of the *n*. However, since this information is not available when determining the *n* the maximum possible variability is usually considered. This means that all survey options are assumed to have the same probability (p has a value of .5).

4.2 Determination of the number of respondents by machine learning techniques.

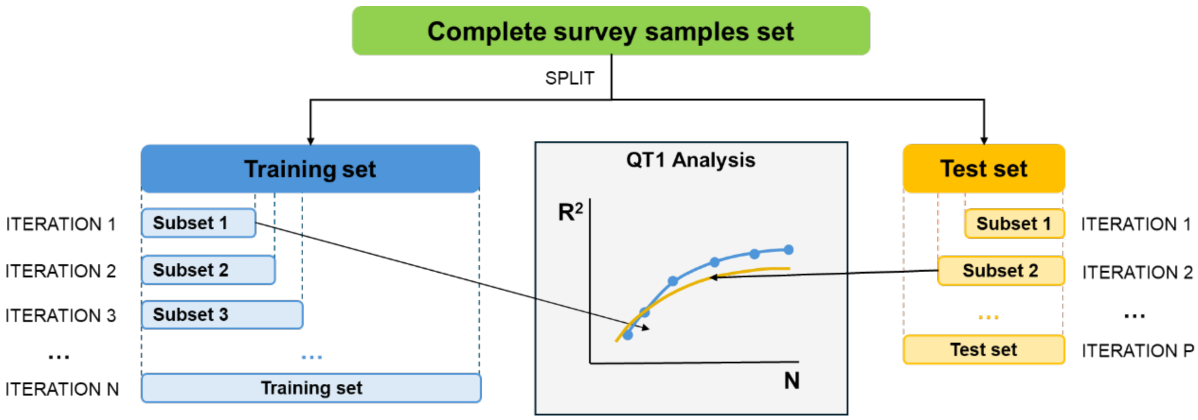
To be able to relate population profiles, KWs and the sample size of the surveys, this paper proposes a procedure based on ML techniques with which to extract relationships between variables which, by their nature, are evidently very complex to model.

Figure 6 illustrates the procedure, which begins with a survey. The survey size will be estimated using statistical procedures, overestimating by a percentage the most pessimistic valuation obtained with classic formulas.

Following classical ML procedures, the initial set of surveys will be divided in two (i.e. 70% for training and 30% for validation).

Starting with a small subset of the Training Set, randomly selected, successive iterations are carried out including an increasing number of participants, in other words, increasing the number of surveys. In principle, we add new data to the previous subset, emulating the behavior of a lifelong learning system. It is expected that the knowledge and results obtained do not have a large dispersion, especially in the initial stages where the small subset size may increase this undesirable effect. At each iteration, multiple linear regression analysis QT1 (Quantification Theory type I) is run for each subset to obtain the relationship between the KWs and the design parameters. Statistical software programs such as SPSS are commonly used for this purpose. One of the data obtained is the Multiple Correlation Coefficient (MCC), which relates the explained variance to the unexplained variance. The MCC parameter is used as a measure of the quality or reliability of this regression, taking a value between 0 and 1. The closer the solution is to 1, the better the regression.

Figure 6: Proposed method for the verification of sample size in synthesis phase´s questionnaires



Plotting the evolution of the MCC parameter over successive iterations enables the creation of a learning curve that correlates it with experience, represented by the number of samples included in training QT1. This curve allows for the evaluation of the model's learning performance. The same procedure is repeated with the test subset to plot the learning validation curve, which reports the model's ability to generalize.

The shape and dynamics of the learning and validation curves make it possible to diagnose the learning of the model and make it easier to modify it to improve its performance. There are three types of evolution associated with underfitting, overfitting, or good fit. Ideally, the validation and training curves should be similar.

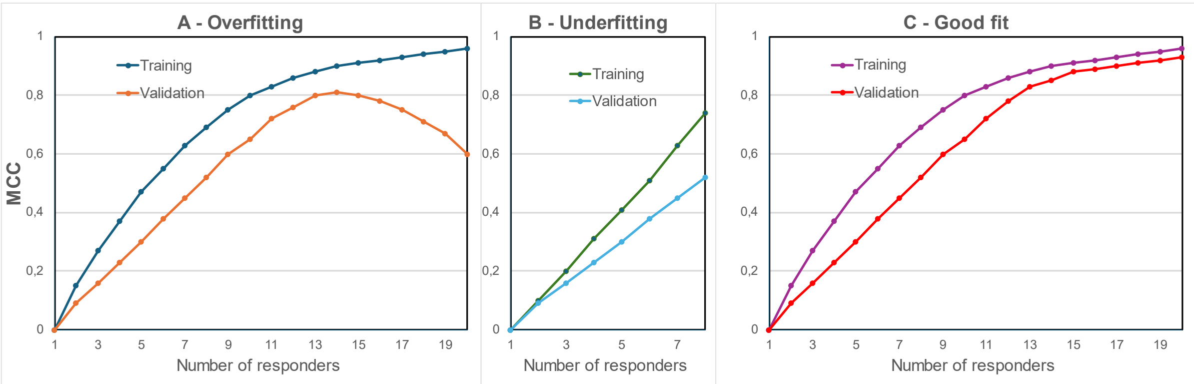
Figure 7a illustrates an example of overfitting. This phenomenon arises when the training curve does not stop increasing and the validation curve increases until an inflection point where it changes to an upward trajectory. This occurs because the model has been trained for too long or with too much data, causing it to perform too well for that dataset, to the point that it has 'memorized' it. Consequently, the algorithm will provide solutions with very low error for training data, but very high error for new data. To address this issue, consider adding new data sets, simplifying the model, or limiting the training time.

Figure 7b illustrates underfitting, which occurs when both curves keep rising. This suggests that the model is too simple and/or insufficient relevant features have been considered. To address this issue, the model's complexity or the number of input variables should be increased, the training time should be extended, or noise should be eliminated from the data.

Finally, figure 7c illustrates the case of Good Fit. A good fit is achieved when both the training and validation curves increase until they reach a point of stability with a decreasing separation between them. Although a zero error is ideal, it is not achievable in reality. If training is prolonged for too long, overfitting occurs, and the error begins to increase.

Finally, a good learning curve should demonstrate the model's ability to generalize well. This means that it should perform well not only with the training data but also with new and unseen data. If the learning curve shows a significant difference between the model's performance on the training data and the test data, it is a sign of a problem.

Figure 7: Learning curve. Types of evolution.



5. Conclusions

The literature review conducted in this paper reveals that there are still areas within the field of Kansei Engineering that require further research. Firstly, there is no established pattern for determining the appropriate number of Kansei words to use during the synthesis phase. Moreover, although in many works it is common to combine qualitative and quantitative techniques for dimensionality reduction of the semantic space, there is still a lot of confusion surrounding the latter. Consequently, this paper elucidates the principal concepts, furnishing a systematic methodology to facilitate the reduction of the number of kansei words to those that are meaningful.

It is also surprising how little research has been done on the number of respondents needed during the synthesis phase, given the importance of the correct choice of this value for the reliability of the final results. For this reason, an incremental method of analysis based on learning curves has been proposed. The first advantage of this proposal is the reduction of the risk of underestimating or overestimating the number of respondents in the accuracy of the solution. Furthermore, it enables the optimization of resources in terms of time and money, both in the preparation of the surveys and the collection and analysis of their results. However, the limitations of this proposal lie in the iterative nature of the process and the lack of a priori criteria for establishing the initial number of respondents.

A series of studies will be initiated to test the effectiveness of this proposal in different case studies (given the changing nature of each case) and to compare the results obtained in this way with those obtained by other methods.

6. References

Almagro, L. M. (2011). *Doctoral Thesis Statistical Methods in Kansei Engineering Studies Presented by Lluís Marco Almagro Xavier Tort-Martorell Llabrés*.

Alves, J. (2018). Kansei Engineering methodology to support the development of a water bottle packaging. https://api.semanticscholar.org/CorpusID:212720395

Burton, L. J., & Mazerolle, S. M. (2011). Survey instrument validity part I: Principles of survey instrument development and validation in athletic training education research. *Athletic Training Education Journal*, *6*(1), 27–35.

Cai, M., Wu, M., Luo, X., Wang, Q., Zhang, Z., & Ji, Z. (2023). Integrated Framework of Kansei Engineering and Kano Model Applied to Service Design. In *International Journal of Human-Computer Interaction* (Vol. 39, Issue 5, pp. 1096–1110). https://doi.org/10.1080/10447318.2022.2102301

Canales, F. H. De, Alvarado, E. L. De, & Pineda, E. B. (n.d.). *Manual para el desarrollo de personal de salud*.

Córdoba Roldán, A., Aguayo González, F., & Lama Ruíz, J. R. (2010). Ingeniería kansei: Diseño estético de productos. In *Dyna (Spain)* (Vol. 85, Issue 6, pp. 489–503).

de las Heras, A., Zamora-Polo, F., Ferramosca, A., & Luque-Sendra, A. (2023). Emotional design engineering for packaging of olive oil using Machine Learning techniques. *Available at SSRN*.

Helmold, M. (2021). *Successful Management Strategies and Tools*. Springer.

Jia, L.-M., & Tung, F.-W. (2021). A study on consumers' visual image evaluation of wrist wearables. *Entropy*, *23*(9), 1118.

Kaiser, H. F. (1970a). A Second generation little jiffy. *Psychometrika*, *35*(4), 401–415. https://link.springer.com/article/10.1007/BF02291817

Li, X., Su, J., Zhang, Z., & Bai, R. (2021). Product innovation concept generation based on deep learning and Kansei engineering. In *Journal of Engineering Design* (Vol. 32, Issue 10, pp. 559–589). https://doi.org/10.1080/09544828.2021.1928023

Li, Y., & Zhu, L. (2020). Extracting knowledge for product form design by using multiobjective optimisation and rough sets. *Journal of Advanced Mechanical Design, Systems and Manufacturing*, *14*(1), 1–16. https://doi.org/10.1299/jamdsm.2020jamdsm0009

Lokman, A. M., & Kamaruddin, K. A. (2010). Kansei affinity cluster for affective product design. *Proceedings - 2010 International Conference on User Science and Engineering, i-USEr 2010*, 38–43. https://doi.org/10.1109/IUSER.2010.5716719

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

Nagamachi, M. (2003). Kansei engineering: A new consumer-oriented technology for product development. *Occupational Ergonomics: Design and Management of Work Systems*, *15*, 25-1-25–14. https://doi.org/10.1201/9780203010457

Prodintec, F. (2011). Diseño afectivo e ingeniería Kansei. *Guía Metodológica, Asturias*.

Sharma, S., & Chaudhary, P. (2023). Machine learning and deep learning. *Quantum Computing and Artificial Intelligence: Training Machine and Deep Learning Algorithms on Quantum Computers*, 71–84. https://doi.org/10.1515/9783110791402-004

Val-Carreres Azofra, J., & Aguayo González, F. (2012). Diseño de un producto lumínico mediante ingeniería kansei para el entorno de Metropol Parasol. *Sevilla Técnica*, *40*, 44–53.

Yang, Y., Huang, J., Wu, G., Ma, Y., & Guo, Q. (2024). Applying kansei engineering and text mining to design fresh food logistics services. *International Journal of Logistics Research and Applications*, 1–25. https://doi.org/10.1080/13675567.2024.2316640

Zhang, J., & Mu, Y. (2021). Clothing Design Methods Based on Kansei Engineering: Example of Suit Design. *Lecture Notes in Networks and Systems*, *260*, 1111–1117. https://doi.org/10.1007/978-3-030-80829-7_135

Zhang, Q., Liu, Z., Zhang, X., Mu, C., & Lv, S. (2022). Target Mining and Recognition of Product Form Innovation Design Based on Image Word Similarity Model. *Advances in Mathematical Physics*, *2022*. https://doi.org/10.1155/2022/3796734

Zhang, S., Wang, S., Zhou, A., Liu, S., & Su, J. (2021). Cognitive Matching of Design Subjects in Product Form Evolutionary Design. *Computational Intelligence and Neuroscience*, *2021*. https://doi.org/10.1155/2021/8456736

Zhong, D., Fan, J., Yang, G., Tian, B., & Zhang, Y. (2022). Knowledge management of product design: A requirements-oriented knowledge management framework based on Kansei engineering and knowledge map. *Advanced Engineering Informatics*, *52*(January), 101541. https://doi.org/10.1016/j.aei.2022.101541

Communication aligned with the Sustainable Development Goals