An empirical approach to optimal experimental design selection and data analysis for the synthesis phase of Kansei Engineering

Antonio Lanzotti, Pietro Tarantino, Giovanna Matrone
University of Naples "Federico II"
Department of Aerospace Engineering
antonio.lanzotti@unina.it pietro.tarantino@unina.it giovanna.matrone@unina.it

Abstract

Purpose

Kansei Engineering (KE) is a methodology able to incorporate, systematically and concretely, people's emotions into product design solution, above all in concept design phase. This work aims at testing the appropriateness and the robustness of statistical methods which are at this moment new in KE applications. In particular, in order to reduce the length and the reliability of the evaluation session in Virtual Reality environment, optimal experimental designs and methods of analysis are suggested.

Methodology

Statistical methods are used in each phase of KE. In this work, we focus our attention on the choice of experimental design for the synthesis phase and on the analysis methods for the model building. In particular, we compare the KE results by using classical fractionated designs, with the efficiency of saturated designs and supersaturated designs. Two methods of analysis are tested: categorical regression analysis (CATReg) and Ordinal Logistic regression (OLR). Moreover, a comparison between the results of ranking and rating procedures are discussed for the saturated design.

Findings

The comparison among the suggested statistical methods is performed through a study on railway seats design in a virtual reality environment. The results of the analysis support the use of Fractional Factorial Design instead of saturated and supersaturated design. The two methods of data analysis give the same results. No evident differences emerge from the comparison of rating and ranking procedures.

Value of paper

This paper propose optimal experimental design selection strategies to reduce the number of product concepts to design, test and evaluate, and data collection analysis strategies in order to improve the appropriateness and the robustness of model building phase at the end of the synthesis phase. If applied faster and more reliable, a KE approach can overcome the distrust of industrial designers toward the methods belong to the emotional design area.

Key words : *Kansei Engineering, Saturated Design, Nested Experimental Design, Ordinal Logistic Regression, Categorical Regression.*

Category : Research paper

1. Introduction

Kansei Engineering (KE) has seen a growing diffusion in the last years among product designers. It is a methodology able to incorporate, systematically and concretely, people's emotions into product design solutions, above all in concept design phase. Part of this increasing interest was due to European and Japanese researchers, that are still working for quantitatively aiding the methodological flow of KE. In this direction, statistical methods can provide a valid support for designers to help them in every critical phase of product development. Some of these methods are already applied in KE studies. However three important considerations should be underlined:

1) the current use of statistical methods is "sparse" in the methodological flow;
2) the main used techniques are the traditional one proposed by KE inventors;
3) the research trend on KE topic reveals a major focus on design methods more than on statistical ones.

These considerations are partly confirmed by the results of a simple survey conducted on the papers presented at the first European Conference on Affective Design and Kansei Engineering, hosted by the 10th QMOD conference in Helsingborg (Sweden). Among the 34 presented papers (25 of which classified as research paper), 10 did not make use on any statistical methods (29%), while the remaining used the methods showed in Table 1. Moreover, only 4 papers made use of experimental design versus the 30 (88%) that did not arrange any design for concept construction and evaluation.

Table 1: Frequency distribution of statistical methods used for the paper presented at the First European Conference on Affective Design and Kansei Engineering (2007).

	Statistical methods	**Frequency of use**
PCA	Principal Component Analysis	9
QT1	Quantification Theory Type I	6
RSA	Rough Sets Analysis	5
DOE	Experimental Designs	4
PLS	Partial Least Squares	2
FA	Factor Analysis	2
ANOVA	Analysis of variance	2
NPT	Non parametric test	2
LRM	Local Regression Models	1
OLR	Ordinal Logistic Regression	1
CA	Correlation Analysis	1
RD	Robust Design	1
ANOM	Analysis of mean	1

In a previous work (Lanzotti and Tarantino, 2007) the authors underlined the importance of using innovative statistical methods in every phase of KE process, above all in the most critical activities, i.e. the choice of the experimental design in the synthesis phase and the choice of the model of analysis for the collected data. A particular attention was given to Supersaturated Design, a technique for the construction of a reduced number of concept, and the Ordinal Logistic Regression for the analysis of data collected by a Likert scale.

In this work, the authors emprically compare the level of agreement of *p*-efficient Design with that of classical fractional factorial design and a nested combination of the previous ones. Moreover, Ordinal Logistic Regression (OLR) is compared with Categorical Regression analysis (CATReg). Both methods pursue the same scope, i.e. to find the relationship among predictors variables and response variables when these are categorical in nature, but with a different approach. Moreover, a comparison between the results of ranking and rating procedures is discussed for the *p*-efficient Saturated Design.

The paper is organized as follow. Section 2 describes the main properties of *p*-efficient Design and the principles of Categorical Regression analysis. Section 3 presents the empirical approach for optimal experimental design and data analysis selection in the synthesis phase of KE. Section 4 describes the results of an application of this approach in a case study on railway seats design. The last section is dedicated to discussions, conclusions and the outline of future research.

2. Innovative statistical methods for Kansei Engineering

The diffusion of KE among researchers and industrial designers depends, for a great extent, on the adoption of quantitative methods able to concretely support the decision process above all in the concept design phase. This adoption can be facilitated if:

1) the proposed methods allow a simplification of the experimental effort with the minimum possible loss of information;
2) the proposed methods can be integrated with the other design activities such as tests in Virtual Reality (VR) environment or evaluation sessions involving users;
3) the proposed methods can be easily implemented through the available statistical packages;
4) the results are easy to be interpreted and discussed.

The central role of statistical methods is particular evident in the central parts of the KE procedure, i.e. the synthesis and the model building phases (Barone *et al.*, 2008). For these phases, two statistical methods are here presented: *p*-efficient Design for concept configuration and Categorical Regression analysis for results analysis.

2.1 A class of Saturated Design: *p*-efficient Design

When an experimentation is carried out for testing the impact of k design factors on a response variable (Kansei word), the minimal number n of product concept required to estimate all main effects is equal to $n = k + 1$. In such a case the experimental design is called *saturated*, in the sense we don't have more degrees of freedom to perform other estimations. Saturated Designs, together with the more bound Supersaturated Designs, are useful when it is impossible or inconvenient to prepare several product concept, both from an economical and experimental perspective. This class of design are often used in technological screening situations, where many potential relevant factors are investigated but reasonably only a part of them are active (Box *et al.* affirm that the percentage of active factors is about 25%). The same situation is encountered in product design field, where at the beginning of the project (concept design phase), many design factors should be considered but only a small portion of them will be further developed. When $k > n-1$ the use of Supersaturated Design is obligatory (Lin, 1993a), instead when $k = n-1$ Saturated Design should be applied. Among the proposed construction method for Saturated Design, *p*-efficient Designs have the appealing property of projectivity, i.e. for a subset of design factors it is possible to arrange a design with at least the near equal occurrence and the near orthogonality properties (Lin, 1993b). Moreover, these design are more efficient than D-optimal design for the estimation from the sub-model.

2.2 Categorical Regression analysis

Since the relationship between the response (respondent's agreement on a Likert scale) and the design factors is not linear, the regression model has to take into account this nonlinearity. Two approach can be used in such a case: Generalized Linear Models (McCullagh and Nelder, 1989) and Regression with transformation (Kruskal, 1965). In the first approach, a non linear function (link function) is used for linearizing the non linear relationship among response and predictors. Ordinal Logistic Regression belongs to this class of models. In the second approach, the relationship between the response and the predictors is linearized through separate nonlinear transformation of the variable (both non-monotonic or monotonic trasformation). In particular, by using optimal scaling it is possible to quantify categorical variables (nominal or ordinal) and at the same time optimize the relationship between response and predictors. Categorical Regression belongs to this class of models (Meulman, 2003). Even if, many preliminary decisions have to be considered before to perform this model (as the properties of the original variable to be preserved with the transformation), the results are more similar to those of linear regression and thus easier to be interpreted in comparison with Ordinal Logistic Regression. For example, the squared multiple regression coefficient R^2 and the regression coefficients assume the same form that in the case of linear regression analysis, while in Ordinal Logistic Regression Log-likelihood ratio test and odds ratio are used respectively. Moreover, Categorical Regression is nowadays implemented in statistical software as SPSS® and R (isoreg function). The case study will clarify how to interpret the result of CATReg.

3. Empirical approach for experimental design and data analysis selection

In some experimental situations, the reduction of the number of product concepts to model and, subsequently, to evaluate in a virtual environment, can determine the success of the evaluation session in terms of respondents' involvement and then reliability of results. However, this reduction is always paid in terms of loss of information and predictive ability of the built model. The trade off between the experimental effort and the results validity needs to be supported by methodological test able to indicate the most suitable experimental design and data analysis. The proposed approach is here described for the case of five design factors. However, its extension to a more general experimental situation is straightforward. As a general consideration, the more the number of runs in an experimental design the more the ability of the design to detect active factors. By following this reasoning a *p*-efficient Design with six run (a Saturated Design in this case) is compared with a 2^{5-2}_{III} fractional factorial design. These design have common runs (combination of factors level), so they can be nested in a 12-run experimental design (Table 2). This design constitutes the benchmark for the evaluation of individual design.

Product concepts, built according to the indication of the 12-run experimental design are then evaluated by respondents in a virtual environment on a five-point Likert scale. Ordinal Logistic Regression and Categorical Regression are both applied to the three designs in order to compare the results of these analysis in all the experimental situations. Moreover, a ranking procedure is performed for a 6-runs *p*-efficient Design. By comparing the results of ranking and ratings by using Categorical Regression with that obtained from 12-runs experimental design, it is possible to have an indication of which scale respondents prefer for evaluating product concepts.

4. A case study on railway seats design

The empirical approach for the choice of optimal experimental design and data analysis method is exploited in a study on new seats design for a regional train. It was developed at the University of Naples "Federico II" by involving undergraduates students of Faculty of Mechanical Engineering, attending the course of "Industrial technical design".

Table 2. Experimental design used for the empirical choice of design and technique of analysis.

Run	A	B	C	D	E	$y_{rat.}$	rank
1	-1	1	-1	-1	-1	y_1	r_1
2	-1	-1	-1	1	1	y_2	r_2
3	1	1	-1	1	1	y_3	r_3
4	1	-1	1	1	-1	y_4	r_4
5	-1	1	1	-1	1	y_5	r_5
6	1	-1	1	-1	-1	y_6	r_6

+

Run	A	B	C	D	E	y
1	-1	-1	-1	1	1	y_2
2	1	-1	1	1	-1	y_4
3	-1	-1	1	1	-1	y_7
4	1	1	-1	1	-1	y_8
5	1	-1	-1	-1	-1	y_9
6	1	1	1	1	1	y_{10}
7	-1	1	1	-1	-1	y_{11}
8	-1	1	-1	-1	1	y_{12}

=

Run	A	B	C	D	E	y
1	-1	1	-1	-1	-1	y_1
2	-1	-1	-1	1	1	y_2
3	1	1	-1	1	1	y_3
4	1	-1	1	1	-1	y_4
5	-1	1	1	-1	1	y_5
6	1	-1	1	-1	-1	y_6
7	-1	-1	1	1	-1	y_7
8	1	1	-1	1	-1	y_8
9	1	-1	-1	-1	-1	y_9
10	1	1	1	1	1	y_{10}
11	-1	1	1	-1	-1	y_{11}
12	-1	1	-1	-1	1	y_{12}

4.1 Design factors selection

The study of seat design for regional transport was initiated with a previous work (Di Gironimo *et al.*, 2008). There, the authors focused on the deep separation between the style and the engineering activities. Moreover, it was underlined the difficulty of identifying the user needs through the only adoption of Kano model. Starting from the results of that work, five factors are here selected in order to improve the users' needs identification phase through the adoption of a KE procedure. For each factor two different design solutions were characterized, assumed as levels for the experimental design (Table 3). These five factors span the space of characteristics for the KE analysis.

Table 3. Experimental design tested for the best choice of design and technique of analysis.

Design Factors		Levels -1	+1
A.	Style	Yesterday	Today
B.	Direct Light	Posterior	Lateral
C.	Folding table	No	Yes
D.	Armrest	Fixed	Mobile
E.	Footrest	No	Yes

According to the experimental design of table 1, twelve virtual concepts were generated in a 3D CAD environment complied with standards. Figure 1 shows an example of seat concept.

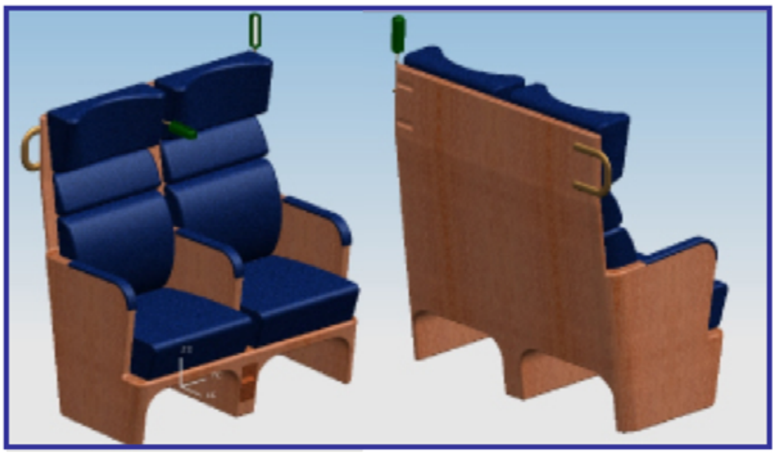

Figure 1: Two sides of concept 1.

4.2 Evaluation session

To identify the correlation between the physical properties of seat design and the Kansei word *"Comfortable"*, an evaluation session was performed at the Faculty of Engineering. Twenty students were asked to express their opinion about each concept, randomly showed on a pc desktop. The respondents were nearly 22 years old. Moreover, the 70% of them regularly uses regional trains while the 50% almost every day.
For the rating analysis each concept was displayed in a dynamic way, with the possibility of a virtual navigation around the seat. The students could state their opinion on a five-point Likert scale. For the ranking analysis the concepts were all simultaneously displayed (Fig. 2) in order to give to respondents the possibility of classifying them.

Figure 2: The concepts showed for the ranking analysis.

4.3 Analysis of experimental results

The collected data were analyzed by MINITAB® and SPSS®. These software are both able to perform Ordinal Logistic Regression and produce almost the same report. However, they have different additional options of analysis. In particular, MINITAB® produces summary measures of association between response variables and predicted probabilities, while SPSS® performs the important test of parallel lines (for verifying the hypothesis of same slope coefficients across response categories) and pseudo R-square. All the Ordinal Logistic Regression models fitted well data (Pearson and Deviance Goodness of fit tests > 0.05) and were significant (Log-likelihood ratio test with p-values less than 0.05). An example of Ordinal Logistic Regression output can be found in (Barone *et al.*, 2007). Categorical Regression was instead performed by SPSS® (Optimal Scaling). The results of CATReg for the 12-runs design are summarized in Table 4. The selected scaling level for response variable was numerical whereas design factors were leaved as nominal. The multicollinearity among predictors was not a concern (high values into tolerance column). The *F*-test for design factors indicates if omission of the corresponding factor from the model, with all other factors present, significantly worsen the predictive capabilities of the model. In this case, design factors A, C, D and E are important. The relative importance of the design factors are also calculated by the Pratt's measure (Importance column), with the four significant factors accounting for the 99.8% of the importance. Moreover, by squaring the value in the part correlation column, it is possible to measure the proportion of variance explained by factor relative to the total variance of response. Even if, the regression of design factors over response variable is statistical significant (*p*-value of *F*-test less than 0.05), the multiple regression coefficient is quite low. However, this value is strictly connected to the selected scaling level of variables. In general, more restrictive transformation (properties of variables persevered during transformation) results in less fit. In summary, the results of CATReg model applied to the data from the 12-runs experimental design are quite significant. The same analysis was executed with the other three design and also for the ranking procedure. The active factors detected with the whole procedure are summarized in Table 5.

Table 4. Summary of CATReg output with the 12-runs design for Kansei word *Comfortable*.

	Standardized coefficents					Correlations				
	Beta	**s_{Beta}**	**Df**	**F**	**p-value**	**Zero-Order**	**Partial**	**Part**	**Importance**	**Tolerance**
A	-0.213	0.053	1	16.2	0.000	-0.252	-0.254	-0.195	0.119	0.834
B	0.011	0.054	1	0.038	0.846	0.123	0.013	0.009	0.003	0.796
C	0.450	0.049	1	82.93	0.000	0.369	0.512	0.441	0.368	0.956
D	0.128	0.054	1	5.52	0.020	0.134	0.152	0.114	0.038	0.793
E	0.470	0.056	1	71.65	0.000	0.455	0.484	0.409	0.473	0.759

	SS	**MS**	**Df**	**F**	**p-value**	**Multiple *R***	R^2	R^2_{adj}
Regression	108.6	21.7	5	38.67	0.000	0.673	0.452	0.441
Residual	131.4	0.562	234					
Total	240.0		239					

Table 5. Active factors resulting from the analysis of OLR and CATReg with the studied design

	Rating		Ranking
Design	**OLR**	**CATReg**	**CATReg**
6-runs *p*-efficient	B-C-E	C-E	A-B-E
2^{5-2}_{III}	A-C-D-E	A-C-D-E	
12 run nested design	A-C-D-E	A-C-D-E	

Ordinal Logistic Regression and Categorical Regression produce similar results in all experimental situations. The only exception is for the 6-runs design. However, factor B was nearly significant with OLR and was not nearly significant with CATReg. With the other designs the two methods produce identical results also for the strength of factors' significativity. Fractional factorial design works better than 6-run p-efficient Design if compared with the 12-runs design. The addition of only two runs allows experimenter to detect the same significant factors than with the 12-runs design. No evident solution seems to emerge from the comparison between rating and ranking procedure. In both model two significant factors emerge, i.e. C and E in rating procedure and A and E in ranking procedure. However, in ranking model also factor B emerges as significant. This is consonant with the results of 6-runs *p*-efficient Design but different from the indication of 8-runs and 12-runs designs.

4.4 Confirmatory study

From the previous analysis the classical Fractional Factorial Design turned out to be better than *p*-efficient Design for detecting the active factors. To confirm this results a new simplified experimental phase was carried out. In particular, the same factors of the previous analysis (Table 3) were used except for the factor A "*Style*", fixed initially at its low level ("*Yesterday*") and then at its high level ("*Today*"). In this phase a 2_{IV}^{4-1} Fractional Factorial Design (FFD) was compared with a Supersaturated Design (SSD) nested in it and generated according to (Lin, 1993a) (Table 6).

Table 6. The classical Fractionated Factorial Design (left) and a Supersaturated Design (right).

Run	**D**	**B**	**E**	**C**	**y**
1	-1	-1	-1	1	y_1
2	-1	- 1	1	-1	y_2
3	-1	1	-1	-1	y_3
4	-1	1	1	1	y_4
5	1	-1	-1	-1	y_5
6	1	-1	1	1	y_6
7	1	1	-1	1	y_7
8	1	1	1	-1	y_8

Run	**D**		**B**	**C**			**E**				**y**
1	-1	1	-1	1	1	1	-1	-1	-1	1	y_1
2	1	-1	-1	1	-1	1	1	1	-1	-1	y_6
3	-1	1	-1	-1	1	-1	1	1	1	-1	y_2
4	-1	-1	1	-1	-1	1	-1	1	1	1	y_3
5	1	1	1	-1	-1	-1	1	-1	-1	1	y_8
6	1	-1	1	1	1	-1	-1	-1	1	-1	y_7

The generated 8 product concepts were shown in an immersive Virtual Reality environment at the Virtual Reality laboratory of the Competence Center for the Qualification of Transportation Systems founded by the Campania Region. Fifteen students of the Faculty of Industrial Design at the Second University of Naples were asked to give their opinion about each product concept on a ten-point Likert scale. The collected data were analyzed through Pareto Anova. The results (Table 7) underline once again the different results obtainable by classical Fractionated Factorial Design and Supersaturated Design. In particular, in both cases of style, Supersaturated design produced discordant results both in terms of active factors and strength of importance. These results, even if in a particular case, confirm the inadequacy of Supersaturated Design in experimental context with non-metric scale and highlight the Fractionated Factorial Design as the best design for detecting the active factors.

Table 7. The results from the Pareto Anova.

Style	***Design***	***Pareto Anova***			
Yesterday	FFD	C	E	D	B
		63%	20%	14%	3%
	SSD	C	D	B	D
		43%	34%	14%	9%
Today	FFD	C	E	B	D
		50%	35%	14%	1%
	SSD	D	C	E	B
		64%	26%	9%	1%

5. Conclusion and Discussion

This paper proposes an empirical experimental design selection strategies to reduce the number of product concepts to design, test and evaluate, and data collection analysis strategies in order to improve the appropriateness and the robustness of model building phase at the end of the synthesis phase. In this strategy two design with a similar number of runs are nested in a 12-runs experimental design. According to the run of this design twelve product concepts were built and evaluated into an immersive Virtual Reality environment on a five-point Likert scale. A ranking procedure was also performed for the 6-runs design. The results of Ordinal Logistic Regression and Categorical Regression are concordant and indicate that classical Fractional Factorial Design works better than saturated Design in terms of ability to detect active factors. This result was confirmed by a simplified experimental session in which Fractional Factorial Design was compared with Supersaturated Design. All conditions being equal, Categorical Regression presents an output similar to that of linear regression and easier to interpret if compared with that of Ordinal Logistic Regression. Moreover, since *p*-efficient Design are applied in technological field, the poor results can be due to the use of non-metric response variable. The comparison among ranking and rating procedure for the 6-runs design does not solve the dilemma about which methods to use in respondents evaluation session. However, the poor results of this test can be due to the correlation pattern of the 6-runs design, heavily biasing the estimation algorithm in CATReg. Since the choice of performing a rating procedure rather than a ranking one is critical, further researches need to be carried out in this context.
If applied faster and more reliable, a KE approach can overcome the distrust of industrial designers toward this methodology belong to the emotional design area. Researches for the choice of optimal experimental design and the most suitable methods of analysis address this goal.

References

Barone, S., Lombardo, A., Tarantino, P., (2008), “Analysis of user needs for re-design a Wheelchair”, in Erto, P., *Statistical Design for Innovation*, Springer, Italy.

Barone, S., Lombardo, A., Tarantino, P., (2007), “A Weighted Logistic Regression for Conjoint Analysis and Kansei Engineering”, *Quality and Reliability Engineering International*, 23, 689-706.

Box, G.E.P., Hunter J.S., Hunter W.G., (2005), *Statistics for Experimenters*, John Wiley and Sons, New Jersey.

Di Gironimo, G., Lanzotti, A., Matrone, G., Patalano, S., Renno, F., (2008) “Modellazione di assiemi ed esperimenti in Realtà Virtuale per migliorare la qualità di nuovi sedili ferroviari”. *Proceedings of 20th INGEGRAF Conference, Valencia, Spain, June 04-06*.

Kruskal, J. B., (1965), “Analysis of factorial experiments by estimating monotone transformations of the data”, *Journal of the Royal Statistical Society Series B,* 27, 251-263.

Lanzotti, A., Tarantino, P. (2007), “Kansei Engineering Approach for Identifying Total Quality Elements: A Case Study on Train Interior Design”, *Proceedings of 10th QMOD Conference, Helsingborg, June 18-20*.

Lin, D.J.K., (1993a), “ A new class of Supersaturated Design”, *Technometrics*, 35, 28-31.

Lin, D.J.K. (1993b), “Another look at first-order Saturated Designs: The p-efficient Designs”, *Technometrics*, 35, 284-292.

McCullagh, P., Nelder, J.A., (1989), *Generalized Linear Models*, Chapman and Hall, London.

Meulman, J. J., (2003), “Prediction and classification in nonlinear data analysis: Something old, something new, something borrowed, something blue”, *Psychometrika*, 4, 493-517.