UNIVERSITY OF NAPLES "FEDERICO II"

POLYTECHNIC SCHOOL AND OF BASIC SCIENCES

DEPARTMENT OF INDUSTRIAL ENGINEERING

DOCTORAL PROGRAM IN AEROSPACE, NAVAL AND QUALITY ENGINEERING - TOTAL QUALITY MANAGEMENT

A Participative Approach for Objective and Subjective Evaluation of Seat Discomfort and User Interface Usability

DOMENICO MARIA DEL GIUDICE

XXVII DOCTORAL CYCLE

THESIS FOR DEGREE OF DOCTOR OF PHILOSOPHY

DOCTORAL PROGRAM IN AEROSPACE, NAVAL AND QUALITY ENGINEERING / TOTAL QUALITY MANAGEMENT

ING-IND/15 – DESIGN AND METHODS FOR INDUSTRIAL ENGINEERING
SECS-S/02 – STATISTICS FOR EXPERIMENTAL AND TECHNOLOGICAL RESEARCH

A Participative Approach for Objective and Subjective Evaluation of Seat Discomfort and User Interface Usability

DOMENICO MARIA DEL GIUDICE

Department of Industrial Engineering,

University of Naples "Federico II",

P.le V. Tecchio n. 80,

80125 Napoli (Italia)

domenicomaria.delgiudice@unina.it

Tutor:

Prof. Eng. Antonio Lanzotti

Candidate:

Eng. Domenico Maria Del Giudice

Co-Tutor:

Prof.ssa Amalia Vanacore

Coordinator of doctoral program:

Prof. Eng. Luigi de Luca

a mia figlia e mia moglie,

a mio padre, mia madre e mia sorella,

ai miei nonni

Respirare l'aria,
parlare, passeggiare,
essere questo incredibile Dio che io sono!
Oh meraviglia delle cose, anche delle più piccole particelle!

Oh spiritualità delle cose!

Io canto il sole all'alba e nel meriggio, o come ora nel tramonto:
tremo, commosso, della saggezza e della bellezza della terra
e di tutte le cose che crescono sulla terra.
E credo che una foglia d'erba non sia meno di un giorno di lavoro delle stelle
e dico che la Natura è eterna, la gloria è eterna.
Lodo con voce inebriata
perché non vedo un'imperfezione nell'universo,
non vedo una causa o un risultato che, alla fine, sia male.
E alla domanda che ricorre:
Che cosa c'è di buono in tutto questo?

La risposta è:
che tu sei qui,
che esiste la vita,
che tu sei vivo,
che il potente spettacolo continua,
e tu puoi contribuire con un tuo verso!

Walt Whitman (1819-1892)
Tratto liberamente da *Leaves of Grass*

Ringraziamenti

È mio desiderio, in conclusione del percorso di dottorato, esprimere la mia riconoscenza e gratitudine a tutte quelle persone che in modi diversi hanno permesso e incoraggiato questo lavoro.

Ringrazio prima di tutti, con profonda stima, il prof. Antonio Lanzotti, Tutor, e la prof.ssa Amalia Vanacore, Co-Tutor: senza il loro supporto e la loro guida sapiente questo lavoro non esisterebbe.

Ringrazio il prof. Luigi de Luca, Coordinatore del corso di dottorato, per l'efficiente organizzazione e il costante aggiornamento su corsi, seminari e eventi.

Ringrazio con rispetto il prof. Pasquale Erto per i preziosi insegnamenti di vita.

Ringrazio il prof. Stanislao Patalano, il prof. Salvatore Gerbino, il prof. Giuseppe Di Gironimo, il prof. Massimo Martorelli e il prof. Michele Staiano per la disponibilità, le opportunità e i consigli che mi hanno dato.

Un ringraziamento particolare va ai colleghi e ai nuovi amici per tutte le esperienze e i momenti condivisi.

Ringrazio infine le persone a me più care: mia figlia per questi primi dieci mesi di felicità, mia moglie per essermi stata accanto sempre incondizionatamente e per avermi sostenuto nelle scelte, mio padre, mia madre e mia sorella per esserci sempre e comunque, e i miei nonni per il loro verso libero. A tutti loro dedico questo lavoro.

Abstract

The main aim of the work is the objective and subjective evaluation of two aspects of the interaction user-product, the seating discomfort and the user interface usability, relevant to industrial design, by using innovative methodologies for generating interpretative and predictive models that allowed the development of analysis strategies useful to improve the satisfaction of use of the types of industrial products considered.

On the first aspect investigated, research in the field of medicine and epidemiology has shown that, over the past decades, the incidence of work-related musculoskeletal disorders (WMSDs) has considerably increased due to sedentary modern lifestyle, closely related to prolonged period of sitting. The importance of good office seating design in improving human wellness, greatly motivates the interest of specialized literature in topics related to the investigation of the biomechanical aspects of sitting and their effect on perceived discomfort. Typically discomfort assessment is realized on the basis of subjective evaluations and/or postural analysis by the interface pressures. In such context, the experimental sessions and the related data analysis were aimed to investigating on three critical aspects of seat discomfort assessment: 1) the relationship between subjective and objective measures of seat discomfort; 2) gender-based differences in seat interface pressure distribution; 3) discriminant effectiveness of indexes based on seat interface pressure.

On the second aspect investigated, it's helpful to recognize that, today, design team can speed up the process of managing information related to design process by adopting digital pattern tools. These tools, as Knowledge Based Engineering (KBE) systems, can assist engineers in capture and re-use the multidisciplinary knowledge in an integrated way, in order to reduce time and cost of designing, to automate repetitive tasks and to support activities in conceptual design. The KBE analyzed in this study is a new digital pattern tool that supports the designers of automotive gearboxes. In such context, the research has focused on the evaluation of interface usability that represents a critical point in the development of a KBE system to demonstrate an effective reduction in the time and cost of designing and increased satisfaction in its use.

The methods used for the two aspects studied are both theoretical and experimental and can be summarized in four main steps:

1) development of participative protocols and execution of experimental sessions with collecting of objective measures related to the interaction user-product and subjective measures related to user perceptions;
2) organization, classification and synthesis of experimental data collected by using techniques of descriptive statistics;
3) definition of interpretative and predictive models of phenomena investigated including by developing synthetic indexes, by using techniques of multivariate and multicriteria analysis;
4) statistical validation of these models and indexes.

The main results achieved concern the assessment of user-product interaction for different types of industrial products where such evaluation is essential. The outcomes are originals because they allowed to find the factors that had most influence on case studies and to develop synthetic indexes useful for identify some critical issues related use. Statistical data analysis provided new information relating to phenomena examined. Furthermore, the proposed data analysis strategies can be easily adapted to other experimental contexts, involving different target populations, and could have important effects in the industrial field, because they allow the reduction of design time (with obvious consequences on cost) and improvement of products in terms of end-user satisfaction.

Key Words: usability assessment, comfort and discomfort assessment, participatory design, objective and subjective evaluation

Contents

1. Preface 19

1.1 Organization of the thesis 19

1.2 Appended papers 19

1.3 Paper A: Seat design improvement via comfort indexes based on interface pressure data 19

1.4 Paper B: Getting insight into Seating Discomfort via a comprehensive statistical data analysis and new diagnostic tools 20

1.5 Paper C: GUI usability improvement for a new digital pattern tool to assist gearbox design 21

1.6 Paper D: On the usability assessment of a new digital pattern tool graphical user interface 21

1.7 Paper E: Computational Procedure for Location Sensor Network Monitoring Volcanic Ash 22

1.8 Paper F: On the Influence of Scanning Parameters on the Laser Scanner based 3D Inspection Process 23

1.9 Paper G: On the Geometric Accuracy of RepRap Open-Source 3D Printer 23

2. Seating Discomfort Assessment via comprehensive Statistical Analysis and New Diagnostic Tools 25

2.1 A brief literature review 25

2.1.1 *On comfort and discomfort division by Helander and Zhang (1997)* 25

2.1.2 *New knowledge in the field of comfort and discomfort* 26

2.1.3 *The model by de Looze (2003)* 27

2.1.4 *The model by Moes (2005)* 28

2.1.5 *Pros and cons of the models* 29

2.1.6 *A new comfort model by Vink* 29

2.1.7 *Critical remarks and future challenges* ... 30

2.2 Introduction ... 31

2.3 Materials and Methods ... 33

2.3.1 *Seat Conditions* ... 33

2.3.2 *Subjective and objective measures of seat discomfort* ... 33

2.3.3 *Participants* ... 34

2.3.4 *Experimental protocol* ... 35

2.3.5 *Data collection and processing* ... 35

2.3.6 *Data analysis* ... 37

2.4 Results and Discussion ... 41

2.4.1 *Explaining the relationship between subjective evaluations and objective measurements of seat discomfort via multivariate data analysis* ... 41

2.4.2 *Effects of anthropometric variability and seat conditions on interface pressures* ... 49

2.4.3 *Discriminant effectiveness of indexes based on seat-interface pressure in predicting discomfort* ... 51

2.5 Conclusions ... 53

2.6 References ... 54

3. A New Approach for GUI Usability Assessment ... 57

3.1 User Interface Problems and the "Magic number 5" ... 57

3.2 Introducing the case study ... 58

3.3 The Methodological approach ... 60

3.3.1 *User Profile Definition and GUI characteristics* ... 61

3.3.2 *The AHP model for GUI usability* ... 63

3.3.3 *The Usability Index definition* ... 64

3.4 Experimental phase ... 67

3.4.1 *Overview of experiment* ... 67

3.4.2 *Experimental protocol* ... 68

3.4.3 *Data collection and processing* 68

3.5 Results and discussion 70

3.6 Conclusions 77

3.7 References 78

Appended Papers 81

Paper A 82

Paper B 90

Paper C 103

Paper D 111

Paper E 127

Paper F 129

Paper G 149

List of Tables

Table 2.1: Factors, influencing comfort or discomfort during sitting according to Zhang (1996). 26

Table 2.2: The tested seat conditions. 33

Table 2.3: Subjective and objective measures. 34

Table 2.4: Anthropometric characteristics of participants. 34

Table 2.5: Contact area and pressure variables. 38

Table 2.6: Results for association analysis on the subjective evaluation scales. 41

Table 2.7: Five principal components after varimax rotation (underlined values are >0.4 and maximal across factors in absolute value). 42

Table 2.8: Standard coefficients for regression models relating PCA factors to DD. 43

Table 2.9: Model Selection and Validation for DD (10 components cross-validated and 2 selected). 44

Table 2.10: ANOVA for DD on the mixed sample (PLSR case)... 44

Table 2.11: Normalized weights of standard coefficients of the stratified regression models connected to the PLS components (only significant coefficients given)............................... 46

Table 2.12: ANOVA for DD on filtered data (PLSR case)... 48

Table 2.13: Model Selection and Validation for DD on filtered data (10 components cross-validated and 5 selected). 48

Table 2.14: Maps of the MPCP for the different sub-samples stratified by gender and seat condition. 49

Table 2.15: Results of non-parametric Wilcoxon-Mann-Whitney test. 52

Table 2.16: Pressure maps related to maximum values of WPCL and PCP. 53

Table 3.1: Normalisation techniques adopted for UFs. 65

Table 3.2: The sources of UFs. 69

Table 3.3: Normalized measures of UFs. 70

Table 3.4: Weights and values of UFs.. 70

Table 3.5: Weights and values of UDs.. 71

Table 3.6: Root causes and corrective actions of the critical issues identified...................... 75

Table 3.7: Normalized measures of UFs in the validation test. 76

Table 3.8: Weights and values of UFs in the validation test.. 76

Table 3.9: Weights and values of UDs in the validation test.. 76

List of Figures

Figure 2.1: How to the new knowledge has been linked to the model of de Looze (2003).28

Figure 2.2: The process of discomfort experience by Moes.. 28

Figure 2.3: The model of Vink (2012) heavily inspired by the models of Moes and de Looze. 30

Figure 2.4: Division of pressure mat for four local body parts (left, number of sensors in parentheses) and exemplar pressure distribution (right, a higher peak pressure on left buttock). 36

Figure 2.5: Partial Least Squares Model Selection Plot (response is DD for mixed sample) 44

Figure 2.6: PLS Standard Coefficients Plot – 2 components. 45

Figure 2.7: PLS Loading Plot. 46

Figure 2.8: Comparison of the PLSR models 48

Figure 2.9: Quantile-Quantile plot for the unloaded weight. 50

Figure 2.10: Mean effect plots assuming WPCL (a) and PCP (b) as response. 51

Figure 3.1: The logic flow chart of methodological approach. 60

Figure 3.2: GUI for the gearbox CAD modelling software tool. 62

Figure 3.3: GUI Usability hierarchical decomposition. 63

Figure 3.4: Number of Operations related to each sub-activity 72

Figure 3.5: Average Additional Time to complete Task 1 and Task 2. 72

Figure 3.6: Cumulative frequency of number of errors related to each task and average values 73

Figure 3.7: Stacked bar chart showing different levels of success based on task completion. 73

Figure 3.8: Average subjective ratings split by statement 74

Figure 3.9: The new release of GUI. 76

List of abbreviations

DoE: Design of Experiment

RD: Robust Design

OR: Operations Research

PCA: Principal Component Analysis

PCR: Principal Component Regression

PLS: Partial Least Squares

PLSR: Partial Least Squares Regression

AHP: Analytic Hierarchy Process

MCDA: Multiple-Criteria Decision Analysis

ANOVA: ANalysis Of VAriance

VDU: Visual Display Unit

WMSD: Work-related MusculoSkeletal Disorder

MVC: Maximum Voluntary Contraction

PCP: Peak of Contact Pressure

WPCL: Weighted Pressure Comfort Loss

MPCP: Maximum Peak Contact Pressure

KBE: Knowledge Based Engineering

DP: Digital Pattern

GUI: Graphical User Interface

UI: Usability Index

UD: Usability Dimension

UF: Usability Function

NE: Number of Errors

TC: Task Completion

NO: Number of Operations

T: Time

PSR: Post Session Ratings

1. Preface

1.1 Organization of the thesis

This work is divided into two basic lines of research. In section 2 an insight into seating discomfort via a comprehensive statistical data analysis and new diagnostic tools is presented. In Section 3 a new methodological approach is discussed, in order to improve the usability of a new digital pattern tool graphical user interface.

1.2 Appended papers

The two research lines have been the starting point for the development of several papers, appended at the end of the thesis.
In particular, Paper A and C are the results of the first chronologically contributions to research and have allowed the successive discussions in the user experience context (Paper B and D); whereas the paper E, F, and G are not directly on the research lines mentioned, but have enhanced the experience developing of experimental protocols and using statistical and optimization methods as Robust Design (RD), Design of Experiment (DoE) and Operations Research (OR). These tools are useful in product innovation.
A short summary of each paper is following given.

1.3 Paper A: Seat design improvement via comfort indexes based on interface pressure data

Lanzotti A., Vanacore A., Del Giudice D.M., Proceedings of Joint Conference on Mechanical, Design Engineering & Advanced Manufacturing 2014, Paper n. 72, Toulouse (France), June 18th–20th 2014, 7 pp. It's waiting to be published on Research in Interactive Design Vol. 4 book by Springer Verlag (ISBN available soon).

Literature on seat comfort recognizes that seat interface pressures are the objective comfort measures that most clearly relate to users' comfort perceptions about sitting experience. In this paper, the above relationship is quantitatively investigated by performing simple but effective explorative analyses on seat comfort data collected during experimental sessions involving 22 volunteers who tested 4 office chairs (differing in terms of cushion stiffness). Statistical data analyses show that subjective sitting comfort/discomfort ratings are significantly related to several combinations of pressure variables. The joint analysis of synthetic indexes based on seat interface pressures reveals to be a useful tool for comparative seat comfort assessment. Besides valuable suggestions for the definition of an effective strategy for seat comfort assessment, the results of data analyses provide useful information to support the product design phase. In fact, the sitting experience results to be significantly improved by: (1) a balancing of pressures between the bilateral buttocks; and (2) a balancing of contact areas between buttocks and thighs.

1.4 Paper B: Getting insight into Seating Discomfort via a comprehensive statistical data analysis and new diagnostic tools

Lanzotti A., Vanacore A., Del Giudice D.M., on the 24th of March in 2015 it was submitted to Applied Ergonomics, ISSN 18729126 and 00036870 (Q1 nel 2013), 12 pp.

This paper provides new insights in the evaluation of seating discomfort with respect to three major concerns: 1) the relationship between subjective and objective measures of seat discomfort; 2) the gender-based differences in the distribution of seat-interface pressure; 3) the discriminant effectiveness of indexes based on seat-interface pressure. Seating discomfort data (both subjective and objective measures) were collected performing a designed experiment involving 22 volunteers who tested 4 office chairs (differing in terms of cushion stiffness). Statistical data analyses showed that subjective sitting discomfort ratings were significantly related to several combinations of pressure variables. This result, together with the evidence of gender-based differences in the distribution of seat-interface pressure, pushes forward a better exploitation of all information available in a pressure map. For this purpose, two novel methods for both graphical (Maximum Peak Contact Pressure - MPCP map) and analytical (Weighted Pressure Comfort Loss - WPCL index) analysis of seat-interface pressure data are discussed. Their joint use can provide useful information to support the product design phase being effective for comparative seat discomfort assessment. Though the paper focus is on the comparative assessment of *office* seating discomfort across a *gender* stratified population of *healthy* users, the proposed data analysis strategy can be easily adapted to other experimental seating contexts involving different target populations.

1.5 Paper C: GUI usability improvement for a new digital pattern tool to assist gearbox design

Patalano S., Del Giudice D.M., Gerbino S., Lanzotti A., Vitolo F., Proceedings of Joint Conference on Mechanical, Design Engineering & Advanced Manufacturing 2014, Paper n. 71, Toulouse (France), June 18th–20th 2014, 7 pp. It's waiting to be published on Research in Interactive Design Vol. 4 book by Springer Verlag (ISBN available soon). One of the best presented papers proposed to be published in the International Journal of Interactive Design and Manufacture (indexed in Scopus, Q1 in 2013; ISSN 1952513 and 19552505).

Design team can speed up the process of managing information related to gearbox design process by adopting digital pattern tools. These tools, as KBE systems, can assist engineers in re-using previous knowledge in order to improve time-consuming task as retrieval and selection of previous architectures and to modify and virtually test a new gearbox design. A critical point in the development of a KBE system is the interface usability to demonstrate effective reduction of development time and satisfaction in its use. In this paper, the authors face the problem of usability improvement of the Graphical User Interface (GUI) of the KBE system previously proposed. An approach based on Analytic Hierarchy Process (AHP) and Multiple-Criteria Decision Analysis (MCDA) has been used. A participatory test has been performed for evaluating the Usability Index (UI) of the GUI. Taking into account the data analysis some changes have been carried out and a new GUI release has been validated with new experimentations.

1.6 Paper D: On the usability assessment of a new digital pattern tool graphical user interface

Patalano S., Del Giudice D.M., Gerbino S., Lanzotti A., Vitolo F., to be submitted to International Journal of Interactive Design and Manufacture, ISSN 1952513 e 19552505 (Q1 nel 2013), 15 pp.

Design team spend up to 30% of their time to searching data and this percentage rises up to 50% if you take into account the time spent to their validating. To speed up these design processes, the use of a knowledge-based engineering (KBE) system is recommended. A critical point in the development of a KBE system is the interface usability for demonstrating an effective reduction of development time and satisfaction in its use. This work tackles the usability improvement of the KBE system's Graphical User Interface (GUI) which assists designing of automotive manual transverse gearboxes, through participatory tests. It's worthwhile to note that this work is an extended version of paper B. In particular, the paper is focused on a new validation experiment for highlighting how important it's use an iterative design when tests on usability assessment are performed. Results have demonstrated an significant improvement.

1.7 Paper E: Computational Procedure for Location Sensor Network Monitoring Volcanic Ash

Malmo F., Del Giudice D.M., Sterle C., Proceedings of XIII International Symposium On Locational DEcision 2014 (ISOLDE), Naples/Capri (Italy), June 16th– 20th 2014, ISBN 9788898273072, abstract. ISOLDE is a triennial Symposium in conjunction with the XXI Meeting of EURO Working Group on Locational Analysis (EWGLA).

Global air traffic is significantly affected by the volcanic ash especially when unfavourable weather conditions occur. About 500 active volcanoes are in the world and the plume thrown up by the eruptions provoked several crisis. Therefore, managing the problem of volcanic ash is a new important challenge for civil aviation, which if neglected can cause significant damage to aircrafts and large economic loss. In order to define no-flight levels, to re-route scheduled flights and to give warning massages to planes already on flight, we propose to use a permanent monitoring system. The aim is to place the sensors of monitoring system optimizing an objective function which is a linear combination of cost and performance, guaranteeing the required safety level. In this paper we tackle this problem by the usage of covering optimization models. The proposed model has been optimally solved by the usage of Xpress optimization software and tested on real test cases using the northern Italy.

1.8 Paper F: On the Influence of Scanning Parameters on the Laser Scanner based 3D Inspection Process

Gerbino S., Del Giudice D.M., Staiano G., Martorelli M., Lanzotti A, on the 21th of January in 2015 it was submitted to International Journal of Advanced Manufacturing Technology, ISSN 14333015 and 02683768 (Q1, IF 1.78 in 2013), 19 pp.

The quality of 3D scanned data is influenced by many factors both related to internal elements to the acquisition device, such as scanner resolution and accuracy, and external to it, such as proper selection of scanning parameters, ambient illumination and characteristics of the object surface being scanned (e.g. surface colour, glossiness, roughness, shape). Today it is of great industrial interest to study and correctly setting the scanning parameters that allow to improve the quality of the 3D acquisitions so to increase the massive usage of these systems in the product inspection activities. In this paper the effects of some scanning parameters and the ambient illumination were analysed by using a commercial triangulation 3D laser scanner. The test geometry chosen was a commercial sheet metal part more complex than the ones commonly used in laboratory and documented in literature. The outcomes of tests confirmed some suggestions documented in literature but also pointed out that the most influencing factor is the relative orientation of the object with respect to the scanner, as well as, its position of the measurement device within the field of view.

1.9 Paper G: On the Geometric Accuracy of RepRap Open-Source 3D Printer

Lanzotti A., Del Giudice D.M., Staiano G., Martorelli M., On the Geometric Accuracy of RepRap Open-Source 3D Printer, on 15th of February in 2015 it was submitted to Journal of Mechanical Design – Transactions of the American Society of Mechanical Engineers (ASME), ISSN 10500472 (Q1, IF 1.17 in 2013), 13 pp.

In the field of Additive Manufacturing (AM) processes, there is a significant lack of scientific data on the performance of open-source 3D printers in relation to process parameter values. The purpose of this paper is to assess the impact of the main process parameters on the accuracy of a set of typical geometrical features, as obtained with an open-source 3D printer, the RepRap Prusa-Mendel I2. A benchmarking part was set up, composed of elementary shapes, representing a series of different features. By means of a DoE approach, we were able to assess the effect of two process parameters - layer thickness and flow rate – on five geometrical features: cube, sphere, cylinder, cone and angled surfaces. A high resolution Laser Scanner was used to evaluate the variation between real features and nominal geometry. On the basis of the experimental results, it was possible to analyze and discuss the main effects of the process parameters on each feature. These results can help RepRap users in the correct selection of the process parameters with the aim of improving the quality of prototypes.

2. Seating Discomfort Assessment via comprehensive Statistical Analysis and New Diagnostic Tools

2.1 A brief literature review

People use products related to comfort everyday, like clothes, tools, electric appliances, computers and their workstations at the office and home, as well as, seats at the office and in airplanes, trains, bus and cars. You think, for instance, to how many hours students spend sitting from primary school to university, as well as, the office workers in their working lives. So, if you watch at the trends such as *"attention to well-being"*, *"attention to health"*, *"graying of the workforce (and population)"* and *"environmental awareness"*, you realize that comfort and discomfort are closely related to these issues as well. It's clear that the knowledge on comfort and discomfort are critical, but at the present this knowledge is still at the early stage.

Recognized by specialized literature the definitions of comfort and discomfort are: *"comfort is seen as pleasant state or relaxed feeling of a human being in reaction to its environment"* and *"discomfort is seen as an unpleasant state of the human body in reaction to its physical environment"* (Vink, 2012).

The theories of comfort and discomfort have been investigated by Helander and Zhang (1997), de Looze *et al.* (2003), Kuijt-Evers *et al.* (2004), Moes (2005) and Vink (2012). These authors have provided models and frameworks that convince experts. In particular, Helander and Zhang have provided a division between comfort and discomfort scales, de Looze *et al.* have added the physical dimension to the discomfort definition, Moes has theorized a simple and linear model of discomfort process and Vink has proposed a new synthesis model based on those previous.

2.1.1 On comfort and discomfort division by Helander and Zhang (1997)

Helander and Zhang (1997) distinguished comfort and discomfort. According to their theory, the absence of discomfort does not automatically result in comfort. Comfort will be felt when more is experienced than expected. Based on questionnaires (Zhang *et al.*, 1996; Helander and Zhang, 1997) discomfort is related to physical characteristics of the environment, whereas comfort is related to luxury, relaxation or well-being (Table 2.1).

This division is confirmed by the fact that comfort scales did not seem useful for high physical load (>65% MVC).

Table 2.1: Factors, influencing comfort or discomfort during sitting according to Zhang (1996).

Discomfort related factors	Comfort related factors
Fatigue	Luxury
Pain	Safe
Posture	Refreshment
Stiffness	Well-being
Heavy legs	Relaxation

2.1.2 *New knowledge in the field of comfort and discomfort*

In recent years, authors as De Korte *et al.* (2012), Vink *et al.* (2012), Groenesteijn *et al.* (2012), Ellegast *et al.* (2012), Franz *et al.* (2012), Kong *et al.* (2012), Kamp (2012), Noro *et al.* (2012), Kee and Lee (2012) and Zenk *et al.* (2012) have added a new specific knowledge in the field above. As regard,

- *Sensory input:*

  De Korte *et al.* (2012) focusing on comfortable VDU or computer work, have found that the use of different sensory channels can influence the comfort experience. So, you need to be aware of this fact. Furthermore, Vink *et al.* (2012) investigating on airplane passengers' comfort have highlighted that psychosocial factors like personal attention influence comfort.
- *Activities influence comfort:*

  Ellegast *et al.* (2012) and Groenesteijn *et al.* (2012) compared five office chairs during the execution of office tasks both in the office and in the laboratory. They proved how important it is search for the correct context and specific activity when experiments on comfort or discomfort are performed.
- *Different body regions:*

  Franz *et al.* (2012) tested various foam characteristics to define the most comfortable headrest. They described that the head needed different foam firmness than the neck. In addition, Kong *et al.* (2012) found that comfort in the palm in of the hand has been more related to the force levels than at the fingers. So, these results demonstrate that the product design is more complex, because the material characteristics need to be different for various locations having contact with the human body.

- *Contour:*

  Measuring the emotional reaction to a tactile experience in seating, the study by Kamp (2012) proved that contour and sporty or luxurious feel and appreciation influences comfort. Furthermore, Noro *et al.* (2012) found effects of contour affecting comfort for long-term static sitting. In particular, they have demonstrated, through a new surgeons seat inspired by that of the Zen priests, the importance of following the form of the human body in product design for comfort.

- *Physical loading:*

  In research or evaluation of products in development, the use of comfort and discomfort scales are useful to estimate the physical loading (Kee and Lee, 2012), especially above 65% MVC discomfort scales are more useful (Kong *et al.*, 2012). In general, long testing periods may be useful when rating comfort or discomfort for lower forces (Kyung and Nussbaum, 2008; Zenk *et al.*, 2012).

2.1.3 The model by de Looze (2003)

According to the division on comfort and discomfort, de Looze *et al.* (2003) proposed a model which highlights the relationship between the product physical features and comfort and/or discomfort experience. Figure 2.1 (on the next page) shows the above model and how new specific knowledge (see section 2.1.2) are connected to it.

The left side of this theoretical model concerns discomfort. The physical processes that underlie discomfort incorporate model parameters on the etiology of WMSDs (Winkel and Westgaard, 1992; Armstrong *et al.*, 1993), which consider the exposure, the dose, the response and the capacity. The exposure refers to the external factors producing a disturbance of the internal state (dose) of a person. The degree to which external exposure leads to an internal dose and the response depends on the physical capacity of the person.

The right side of the model regards the comfort only. The influential factors are described on human, product and context levels. At the human level, the individual expectations and other individual feelings or emotions are supposed that influence comfort. At the product level, the product aesthetic design as well as its physical features may affect the feelings of comfort. At the context level, the psychosocial factors together with the physical features play a role on comfort.

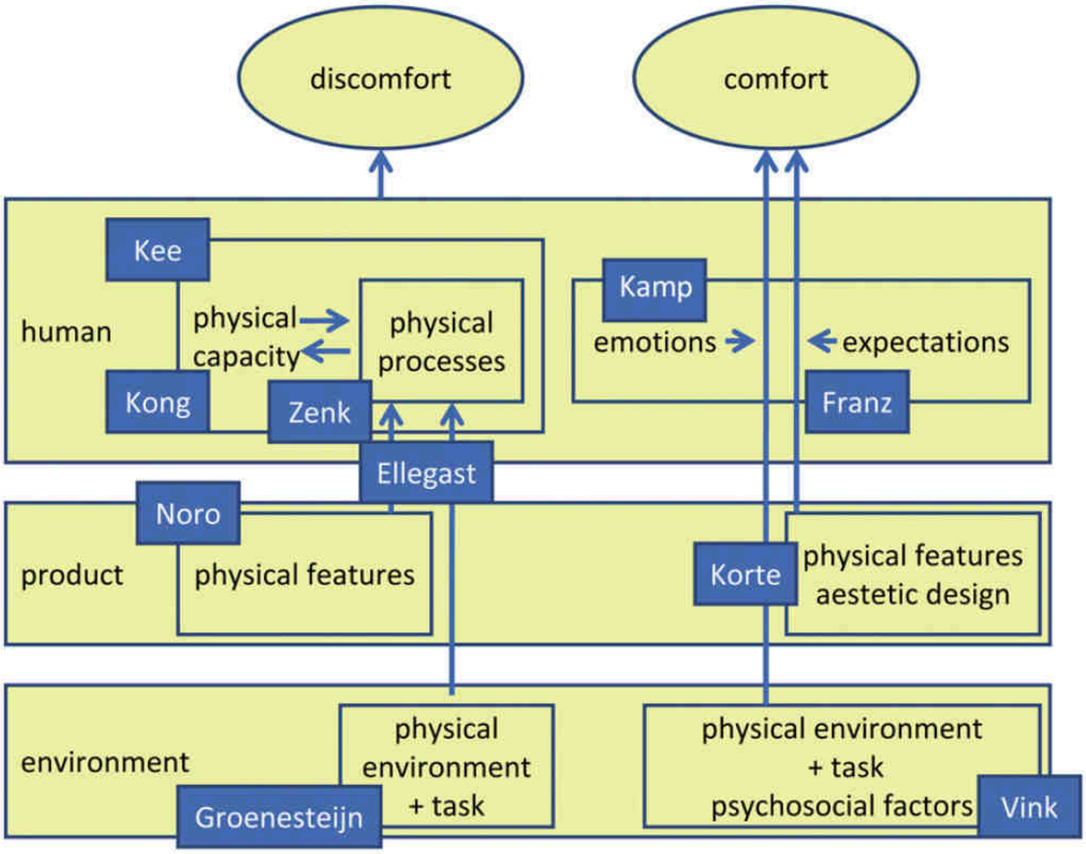


Figure 2.1: How to the new knowledge has been linked to the model of de Looze (2003).

2.1.4 The model by Moes (2005)

Also, the model of Moes (2005) could be used for explaining the process of discomfort experience. According to this model, there are five phases in the process before discomfort is experienced (I – interaction, E – effect in the internal body, P – perceived effects, A – appreciation of the effects and D – discomfort; Figure 2.2).

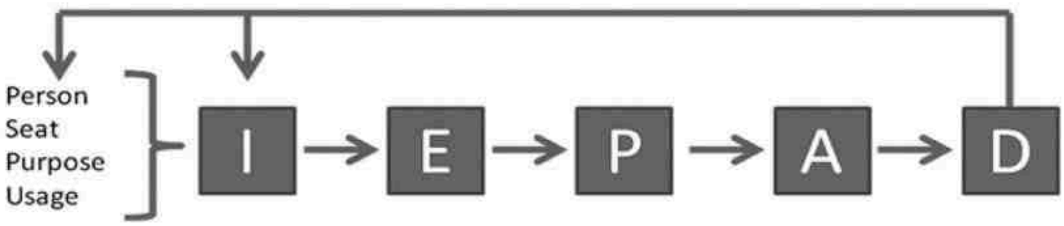


Figure 2.2: The process of discomfort experience by Moes.

This process is dependent on the person, the seat, the purpose and why the seat is used. The interaction (I) arises when a person uses a seat with a specific purpose. If we consider the seat-interface pressure, there is an interaction that results in internal body effects (E), such as tissue deformation or the compression of nerves and blood vessels. These effects can be perceived (P) and interpreted, such as pain. The following phase is the appreciation (A) of the perception. So, if these factors are not appreciated, it can lead to feelings of discomfort (D).

2.1.5 *Pros and cons of the models*

The model by de Looze reflects the prevailing concept of two distinct scales, one for discomfort and one for comfort (not just lack of discomfort), as shown by Kong *et al.* (2012). Often *"more comfort than expected"* is reflected in a comfort experience, which is a valuable result of the de Looze model.

The model of Moes (2005) is simple and linear and explains the process more clearly as the step between interaction and internal effects and weighting the internal to check whether it is appreciated are explicitly shown (Franz *et al.*, 2012; Kamp, 2012).

The advantages of the model of de Looze *et al.* are that:

- the environment is explicitly shown (as in Noro *et al.*, 2012; Ellegast *et al.*, 2012);
- the connection to expectations can be made, which is important in the mental process of deciding whether or not a product is comfortable (as in Vink *et al.*, 2012);
- the *"comfort"* can be an outcome.

Both models point out the probability of a relationship between discomfort and musculoskeletal complaints, but Hamberg-van Reenen *et al.* (2008) confirms that discomfort may influence the chance of having musculoskeletal disorders in the long-term.

2.1.6 *A new comfort model by Vink*

Vink presented a new comfort model (see Figure 2.3 on the next page) inspired by the model of Moes (2005) and de Looze (2003). The interaction (I) between a product and a person starts in an environment where the person is doing a specific task. This may cause in internal human body effects (H), such as changes in the human sensors, postural changes, tactile sensations, muscular activation and blood flow changes. The human body effects (H) as well as the expectations (E) influence the perceived effects (P). These are interpreted as comfortable (C) or you feel nothing (N) or it can lead to feelings of discomfort (D). Discomfort could lead in musculoskeletal complaints (M). The expectations (E) are often linked to comfort (C) as shown by the circle around E-C. If discomfort is too high or the comfort not good enough there is a feedback loop to the person who could do something like shifting in the seat, adapt the product or to change the task/usage. Also, the author supposes that both comfort and discomfort could be simultaneous experiences that occur not in one form.

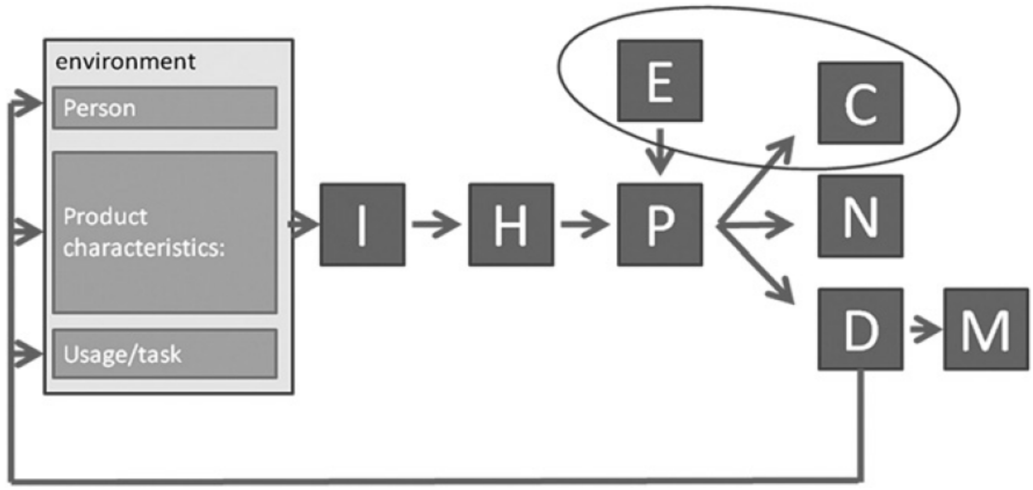


Figure 2.3: The model of Vink (2012) heavily inspired by the models of Moes and de Looze.

2.1.7 *Critical remarks and future challenges*

Every year an incredible amount of products are designed and put on the market, but these are rarely tested and iteratively refined/redesigned for comfort. Thanks to the work of Vink (2012) based on the previous models (see Moes, 2005; de Looze *et al.*, 2003) new scientific knowledge became available and a further step towards the conceptualization of comfort and discomfort has been taken.

It's clear that you should define first the tasks and the characteristics of users performing such tasks and only then should you proceed with product design and required tests. But, with the knowledge that the outcomes of tests should be fed back into the iterative design process. Additionally, data on the internal human body effects are essential to understand the process towards experiencing comfort or discomfort.

As environmental and sustainability issues become more important, you need to design products that consume less energy. "*For this reason it is important to know what the minimum requirements are for user feelings of comfort and what makes a product comfortable*" (Vink, 2012).

2.2 Introduction

Research in the field of medicine and epidemiology has shown that, over the past decades, the incidence of work-related musculoskeletal disorders (WMSDs) has considerably increased (Harkness *et al.*, 2005; Rubin, 2007) due to sedentary modern lifestyle characterized by prolonged period of time spent in a seated position (Ehrlich, 2003; Dul and Hilderbrandt, 1987; Annetts *et al.*, 2012). More than 60% of people experience at least one episode of lower back pain (LBP) at work, in almost 45% of cases the first attack of LBP happens while working, with an incidence in the office workers of at least one episode backache every 3 years (Lengsfeld *et al.*, 2000; Rezaee *et al.*, 2011).

Sitting on an ergonomic chair with a correct posture is undoubtedly one of the most useful remedy in preventing WMSDs, (Nelson and Silverstein, 1998; Herbert *et al.*, 2001; Loisel *et al.*, 2001). The importance of good office seating design in improving human wellness greatly motivates the interest of specialized literature in topics related to the investigation of the biomechanical aspects of sitting and their effect on perceived (dis)comfort.

Typically, discomfort is associated to *"an unpleasant state of the human body in reaction to its physical environment"* (Vink, 2012) and its assessment is realized on the basis of subjective evaluations and/or postural analysis. Subjective evaluations are collected by surveying potential seat users who are asked to express their feelings of discomfort with the seat and/or compare, in terms of perceived discomfort, similar seats.

Postural analysis is realized by measuring one or more objective parameters, several of which are listed in (Andreoni *et al.*, 2002):

- the pattern of muscle activation measured through electromyography (EMG) (Lueder, 1986; van Dieën *et al.*, 2001);
- the stress acting on the spine measured through pressure transducer and radio waves (Lueder, 1986; Zenk *et al.*, 2012);
- the postural angles obtained using contact or non-contact (like photogrammetric) techniques in real experiments (Dreyfuss, 2002) or using virtual manikins in virtual experiments (Lanzotti, 2008; Barone and Lanzotti, 2009);
- the seat-interface pressure measured through capacitative or resistive mats (Kyung and Nussbaum, 2008).

Many researchers have tried to deepen the relationship between such measurements (Zhang *et al.*, 1996). Among all objective parameters, pressure distribution results the objective measure with the clearest correlation with subjective evaluation (de Looze *et al.*, 2003; Hamberg-van Reenen *et al.*, 2008; Kyung and Nussbaum, 2008; Stinson and Crawford, 2009; Noro *et al.*, 2012). In particular, in several studies on seating design (Kamijo *et al.*, 1982; Reed and Grant, 1993; Park *et al.*, 2000; Fujimaki and Mitsuya, 2002; Franz *et al.*, 2012), the effects on seat (dis)comfort due to specific product features (*e.g.* cushion shape and materials) have been qualitatively verified by correlating the information obtained from pressure maps with users' (dis)comfort perceptions. In their recent review on the effectiveness of pressure measurements in the assessment of office chair comfort/discomfort, Zemp *et al.* (2015) highlight that investigations on the pressure-comfort/discomfort relationship are mainly based on seats other than the office one (*e.g.* car seats, wheelchairs, tractor seat and surgery seat); they call for further investigations in order to definitively answer whether pressure measurements are suitable for assessing the comfort/discomfort experienced while sitting in office investigating empirical chairs.

Independently from the specific investigation context, a further concern in studies on sitting comfort/discomfort assessment is that pressure measurements are not fully exploited being pressure distribution mostly described by the maximum (peak) pressure and/or the average pressure. Hitherto, little effort has been made to properly synthesize all the information provided by a pressure map and to highlight the usefulness of seat-interface pressure measures for specific purposes defined by designers (*e.g.* design for a specific user or design for a generic user).

In this work the main results of an experiment aimed at deepening knowledge on office seat discomfort are described. In particular the experiment and the related explorative data analysis were aimed at investigating three critical aspects of seat discomfort assessment: 1) the relationship between subjective and objective measures of seat discomfort; 2) the gender-based differences in the distribution of seat-interface pressure; 3) the discriminant effectiveness of indexes based on seat-interface pressure.

The dependency of subjective discomfort ratings from contact area and pressure variables was explored via (a) Principal Component Regression (PCR) and (b) Partial Least Squares Regression (PLSR); gender-based differences in seat-interface pressure distribution were investigated by analysing the sampling distributions of the unloaded weight for male and female users and building new pressure maps of the Maximum Peak Contact Pressure

(MPCP); finally, the discriminant effectiveness in predicting seat discomfort was evaluated for two indexes based on seat-interface pressure: the Peak of Contact Pressure (PCP) and the Weighted Pressure Comfort Loss (WPCL; Lanzotti *et al.*, 2011).

2.3 Materials and Methods

Data were obtained from an experiment performed at the Department of Industrial Engineering, University of Naples Federico II, in a suitable room cleared of furnishings and according to a well-defined experimental protocol. The whole experiment consisted of 88 experimental sessions during which 22 volunteers tested four ergonomic office chairs performing a task of reading a text on a Visual Display Unit (VDU).

2.3.1 Seat Conditions

The four office seats have a typical architecture of market product (*i.e.* a five-pointed base, a backrest and two armrests), but differ for the stiffness of the seat pan foam. The seats were named with fantasy names (Oslo Chair, OC, Madrid Chair, MC, Chicago Chair, CC, and Toronto Chair, TC) so as to avoid any conditioning of the brand name or the model name on the evaluation (Table 2.2). The codes 0, 1, 2, 3 used to distinguish different Seat Conditions refer to increasing levels of cushion stiffness with extremes low (*i.e.* soft cushion) and high (*i.e.* rigid cushion).

Table 2.2: The tested seat conditions.

Office Seat	OC	MC	CC	TC
Seat Condition (Stiffness)	0 (Low)	1	2	3 (High)

2.3.2 Subjective and objective measures of seat discomfort

During each experimental session subjective measures of discomfort perception as well as seat-interface pressures were collected.

In order to collect users' evaluations about seat discomfort, three different scales were used: 1) the Discomfort Rating (DR) based on a 10-points ordinal scale with extremes 1 (no discomfort) and 10 (maximum discomfort); 2) the Discomfort Degree (DD) based on a 4-level scale of agreement with the statement *"I feel uncomfortable"*; 3) the Chair Ranking

(CR) based on ordinal ascending ranks assigned to chairs consistently with the level of perceived discomfort.

Objective measures were obtained from pressure measured at the seat-interface; these measures consisted of both overall and local pressures (Table 2.3).

Table 2.3: Subjective and objective measures.

Type	Name	Area to measure
Objective	PCP, Peak Contact Pressure (N/cm^2) CP, Contact Pressure (N/cm^2) CA, Contact Area (cm^2) UW, Unloaded Weight (kg) WPCL, Weighted Pressure Comfort Loss index	• Left/right thighs (TL/TR) • Left/right buttocks (BL/BR) • Sum of 4 local body part pressures
Subjective	DR, Discomfort Rating	Whole body
	DD, Discomfort Degree	Whole body
	CR, Chairs Ranking	Whole body

2.3.3 Participants

Twenty-two volunteers, including 8 Females (F) and 14 Males (M), participated in four short-term experimental sessions. Participants were recruited from a university student population. This population was deemed to be relevant to this study as university students tend to spend a large amount of time performing seated work. All participants were free of low back pain for 12 months prior to the testing period. Before experiment began, participants gave informed consent and their personal details (*viz.* gender, age and main occupation) as well as anthropometric data (*viz.* stature and weight) were collected and reported in Table 2.4.

Table 2.4: Anthropometric characteristics of participants.

Gender	Age	N	Anthropometric variable	Mean	SD	Min	Max
F	20-31	8	Stature (cm)	164	8	153	178
			Weight (kg)	67.2	13.3	52.8	96.1
M	20-31	14	Stature (cm)	182	8	170	198
			Weight (kg)	79.4	9.3	64.4	93.0

2.3.4 Experimental protocol

Participants tested the four office chairs in a random order, in order to prevent the disturbance due to the testing sequence. For each testing session, a pressure mat was put on the seat cushion and secured with masking tape to facilitate seat adjustments. Each participant was instructed to sit carefully to minimize wrinkles on the pressure mat. Besides, in order to avoid that discomfort assessments could be affected by the visual impact with the tested chair, the participant was introduced into the room blindfolded and made to sit. Subsequently, the participant was asked to take off the blindfold and adjust the chair in such a way that the legs were in rest conditions and the feet were comfortably on the floor so as to form an angle between the thigh and the leg equal to 90°. Few minutes (≤5) were devoted to initial seat and posture adjustments, then the test session started. In each session, participants performed the task of reading a text on VDU for 20 minutes. At the end of the testing session, the participant was blindfolded again and taken back out of the room.

The specific task of reading a text on VDU was chosen in order to minimize differences in postures among the participants due to the peculiarities in performing more complex task. Indeed, previous studies have shown that the type of computer workstation task performed has an effect on postural responses while sitting (van Dieën *et al.*, 2001; Gregory *et al.*, 2006; Dunk and Callaghan, 2005; Moes, 2005; Ellegast *et al.*, 2012; Groenesteijn *et al.*, 2012). The choice of short-term experimental session is recommended when using pressure mats in order to prevent the well-known effects of creep and/or hysteresis (Fay and Brienza, 2000). Moreover, long-term sessions are generally suggested when investigating sitting discomfort due to fatigue resulting from sources other than chair design (Helander and Zhang, 1997; Kyung and Nussbaum, 2008).

2.3.5 Data collection and processing

The collection of subjective ratings was organised in such a way as to minimize confusion: the forms for the collection of the Discomfort Rating (DR) and the Discomfort Degree (DD) were administered to each participant immediately after each testing session; instead, the Chair Ranking (CR) was collected only after the participant had tested all four office chairs.

Pressure data were divided into four groups (Figure 2.4) and were collected continuously during the reading text on VDU, using a Novel Gmbh (Munich, Germany) pressure mat (S2027 Pliance™).

The above pressure mat comprises 256 (16x16) thin (<1.2 mm) capacitive sensors that could easily conform to the contour of the seat and measure pressures typically in a range from 0.2 N/cm^2 up to 6 N/cm^2. Thanks to its flexible structure the mat is a minimally invasive instrument, which does not interfere with user perception of seat discomfort. The mat has an active area of 392 mm x 392 mm, and sensor pitch is 24.5 mm (0.167 sensor/cm^2).

Pressures were recorded at 50 Hz. This sampling rate was considered sufficient given the frequency of postural changes and resultant pressure changes (Kyung and Nussbaum, 2008).

Contact area and contact pressure were calculated by including only data from sensors that were pressed at least once and average values were determined for the last 15 minutes of each session. Earlier data (5 min) were excluded since they were transient due to settling into the chair (Reed *et al.*, 1999).

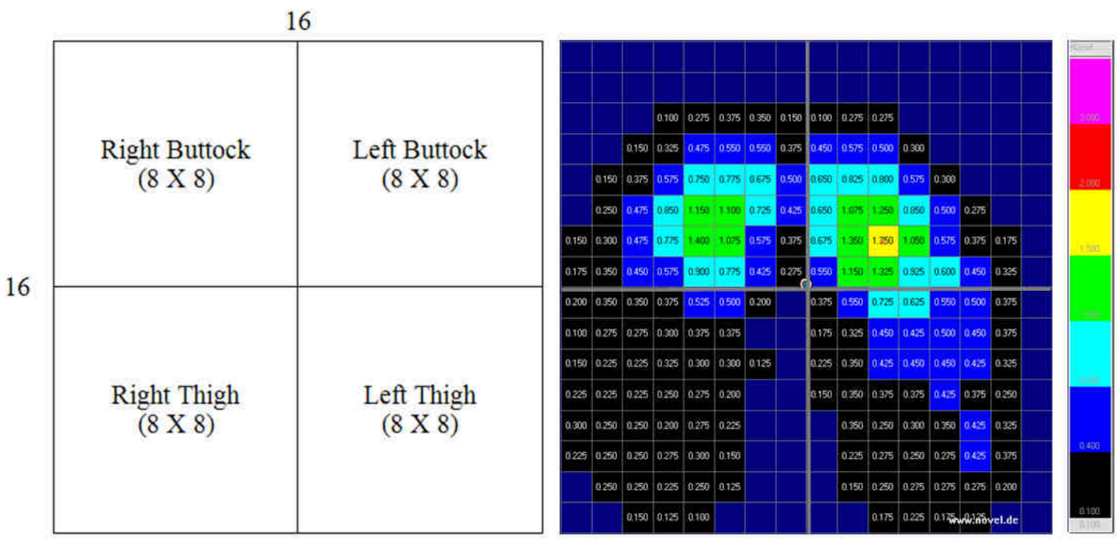


Figure 2.4: Division of pressure mat for four local body parts (left, number of sensors in parentheses) and exemplar pressure distribution (right, a higher peak pressure on left buttock).

2.3.6 Data analysis

Data analysis aimed at answering the following three research questions:

- *Does a relationship between subjective evaluations and objective measurements of seat discomfort exist?*
- *How do anthropometric variability and differences in seat conditions affect contact pressures?*
- *Are indexes based on seat-interface pressure effective in predicting discomfort?*

The first question was investigated by adopting two different multivariate approaches for the statistical analysis of collected data: the (a) PCR and the (b) PLSR.

The (a) PCR data analysis procedure developed into three steps: at the first step the association among the three adopted evaluation scales was evaluated via the Goodman and Kruskall's index in order to test the consistency of the subjective data and select the best proxy for perceived discomfort; at the second step pressure and contact variables were analysed via Principal Component Analysis (PCA) in order to reduce the number of explanatory variables; at third step a multiple regression of perceived discomfort on the PCA factors (obtained at step 2) was performed. The number of PCA factors was determined by two criteria, the size of the eigenvalue (>1) and the cumulative percentage (≈90%) of variance accounted for.

The (b) PLSR data analysis procedure provides a dimension reduction strategy in a single step. Such procedure tries to find the multidimensional direction in the X space, set of predictor variables, that explains the maximum variance direction in the Y space, one or a set of response variables. For this procedure has been used the same best response setting for PCR data analysis. The optimal number of components was determined by the cross-validation procedure with 'Leave-one-out' technique.

Following the data analysis strategy proposed by Kyung and Nussbaum (2008), all the collected contact area and pressure data were divided into four groups corresponding to four local body parts (*i.e.* right/left buttock and right/left thigh see Figure 2.4) and a total of 27 explanatory variables were derived (Table 2.5 on the next page) to be used at step 2 of the data analysis procedure. The 1-9 variables were related to average contact areas and ratios; the 10-18 variables described average contact pressures and ratios; and the 19-27 variables indicated average peak contact pressures and ratios. The overall pressure variables (caSUM, cpSUM, and pcpSUM) were only used to derive the ratio variables but they were not further analysed. Statistical results were considered 'significant' or 'marginal' when $p \leq 0.05$ and $0.05 < p \leq 0.10$, respectively.

Table 2.5: Contact area and pressure variables.

Variable	Description	Unit of measure
caTL (caTR)	Average contact area Thigh Left (Right)	cm^2
caBL (caBR)	Average contact area Buttock Left (Right)	cm^2
caSUM (caTL+caTR+caBL+caBR)	Sum of average contact areas	cm^2
caTL/caSUM, caTR/caSUM, caBL/caSUM, caBR/caSUM	Relative Average contact areas	
cpTL (cpTR)	Average contact pressure Thigh Left (Right)	N/cm^2
cpBL (cpBR)	Average contact pressure Buttock Left (Right)	N/cm^2
cpSUM (cpTL+cpTR+cpBL+cpBR)	Sum of average contact pressures	N/cm^2
cpTL/cpSUM, cpTR/cpSUM, cpBL/cpSUM, cpBR/cpSUM	Relative average contact pressures	
pcpTL (pcpTR)	Average peak contact pressure Thigh Left (Right)	N/cm^2
pcpBL (pcpBR)	Average peak contact pressure Buttock Left (Right)	N/cm^2
pcpSUM (pcpTL+pcpTR+pcpBL+pcpBR)	Sum of peak contact pressures	N/cm^2
pcpTL/pcpSUM, pcpTR/pcpSUM, pcpBL/pcpSUM, pcpBR/pcpSUM	Relative peak contact pressures	

The second question was investigated by building new pressure maps of the Maximum Peak Contact Pressure (MPCP) and by analyzing the sampling distributions of the unloaded weight for male and female users. Pressure data were stratified by gender and seat condition so as to obtain 8 (*i.e.* 2x4) strata. For each stratus a MPCP map was built (Table 2.14) so that each map cell represents the greatest value among all (peak) contact pressures sampled from a particular sensor for a given stratus.

Finally, the third question was evaluated by analysing the discriminant effectiveness of two specific indexes based on seat-interface pressure: the Peak of Contact Pressure (PCP) and the Weighted Pressure Comfort Loss (WPCL). The PCP index is the overall maximum pressure value registered on the mat (de Looze *et al.*, 2003; Dunk and Callaghan, 2005; Hamberg-van Reenen *et al.*, 2008).

The Weighted Pressure Comfort Loss (WPCL; Lanzotti *et al.*, 2011) is a discomfort index formulated under the assumptions that an ideal distribution of seat-interface pressures exists and that every deviation from it causes an increase in user's seat discomfort. Under the reasonable assumption that small deviation are not relevant and that larger deviations become increasingly important (*i.e.* the larger the deviation, the larger the increase in user's seat discomfort) the comfort loss is assumed to be a quadratic function of the deviation from the ideal pressure value. The existence of an ideal seat pressure distribution is accepted in the specialized literature (Kyung and Nussbaum, 2008; Fujimaki and Mitsuya, 2002) and it is generally believed that the ideal pattern of pressure distribution is obtained by uniformly distributing the body weight over the seating surface.

Coherently with this assumption, for each user, the ideal pressure (*i.e.* target pressure x_{0j}) can be defined as the mean pressure over the whole contact area (or any partition of it).

Let n_j be the number of activated cells in the pressure map for the *j-th* user and x_{ij} the pressure value measured by the *i-th* cell when the *j-th* user is seated, the target pressure is given by:

$$x_{0j} = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j} \quad \text{(eq. 2.1)}$$

For each user and for each cell of the pressure map it is possible to evaluate the deviation of the observed pressure value, x_{ij}, from the target pressure, x_{0j} and thus identify the associated Pressure Comfort Loss (PCL) based on a (Nominal the Best) quadratic loss function.

For the *j-th* user the pressure comfort loss at the *i-th* activated cell of the contact surface can be defined as:

$$PCL_{ij}(x_{ij}) = k_{ij}\left(\frac{x_{ij} - x_{0j}}{x_{0j}}\right)^2 \quad \text{(eq. 2.2)}$$

where k_{ij} is a proportionality coefficient that for each cell measures the loss corresponding to the maximum accepted deviation from the target pressure. In particular, let k_{ij} be the maximum accepted relative deviation from ideal pressure at the *i-th* cell activated by the *j-th* user and let C_0 be the comfort loss due to uneven pressure, the proportionality coefficient k_{ij} can be defined as follows:

$$k_{ij} = \frac{C_0}{\Delta_{ij}^2} \quad \text{(eq. 2.3)}$$

For the sake of simplicity and without loss of generality, hereafter C_0 can be assumed unitary and Δ_{ij} can be assumed constant over all activated cells and over all the users belonging to the same target population Ω (*e.g.* female or male users).

Assuming the hypothesis of additivity of comfort loss function, the PCL index for the *j-th* user belonging to Ω is given by:

$$PCL_{\Omega j}(\vec{x}) = \Delta_{\Omega}^{-2} \sum_{i=1}^{n_j} \left(\frac{x_{ij} - x_{0j}}{x_{0j}} \right)^2 \qquad \text{(eq. 2.4)}$$

where $\vec{x}$ is a vector of dimension n_j with generic element x_{ij} and the proportionality coefficient Δ_{Ω} can be calculated by averaging the maximum relative deviations from ideal pressure over all pressure maps rated at the lowest level on the scale for perceived discomfort (*i.e.* no discomfort).

Starting from eq. 2.4, the Weighted Pressure Comfort Loss for target population Ω can be defined as:

$$WPCL_{\Omega} = \sum_j W_{\Omega j} \cdot PCL_{\Omega j} \qquad \text{(eq. 2.5)}$$

where $W_{\Omega j}$ is a weight that allows to account for the degree of anthropometrical representativeness of the *j-th* user inside the target population. The weights, $W_{\Omega j}$, via the discrete approximation of a continuous random variable (*e.g.* the stature of potential users) taken as representative of the population anthropometrical variability (for further details see Lanzotti and Vanacore, 2007)

Moreover, the overall Weighted Pressure Comfort Loss for a mixture of sub-populations can be obtained as follows:

$$WPCL = \sum_k \theta_k \cdot WPCL_{\Omega k}; \qquad \sum_k \theta_k = 1 \qquad \text{(eq. 2.6)}$$

being θ_k is the mixture parameter accounting for the representativeness of the (sub)population Ω_k inside the overall population.

Thus when dealing with an overall composed by female and male users, the WPCL index is obtained by summing up the gender specific WPCL indexes taking into account their mixture weight:

$$WPCL = \theta_F \cdot WPCL_F + (1 - \theta_F) \cdot WPCL_M; \qquad 0 \leq \theta_F \leq 1 \qquad \text{(eq. 2.7)}$$

2.4 Results and Discussion

2.4.1 Explaining the relationship between subjective evaluations and objective measurements of seat discomfort via multivariate data analysis

The Goodman and Kruskall's index was calculated for all possible combinations of binary association among the adopted subjective evaluation scales. Results (Table 2.6) show a substantial consistency of the tested evaluation scales. In fact, the minimum value for Goodman and Kruskall's index in Table 2.6 is 0.653 revealing a medium-high level of association between the CR and DR scales. Since responses given on the DD scale were highly associated with both CR and DR scales (0.984 and 0.860, respectively), this scale was selected as a good proxy of perceived discomfort and set as a robust response function for explorative data analysis via PCR.

Table 2.6: Results for association analysis on the subjective evaluation scales.

Rating (DR)	0.984	0.653
	Degree (DD)	0.860
		Ranking (CR)

(a) <u>PCR Analysis</u>

From the PCA on the set of 27 variables listed and described in Table 2.5 (page 34), resulted five principal components with an eigenvalue >1 accounted for 86.9% of the total variance (Table 2.7 on the next page). After varimax rotation, principal components appeared to have a more general interpretation; indeed, for each of them a predominant subset of (2-4) pressure variables was found. Since these subsets of variables were mutually exclusive and distinguishable (*e.g.* in terms of body part) the components were termed accordingly to them. It is worthwhile to note that Factor 4 shows coefficients with opposite signs for thigh average contact area ratio and buttock average contact area ratio (*i.e.*, caTL/caSUM vs. caBL/caSUM), providing some evidence of negative association in terms of contact area ratio between thigh and buttock.

Fitted DD regression models were significant ($p \leq 0.01$) for the whole sample of users (*i.e.* group of mixed users) as well as the two sub-samples obtained by stratifying by gender (*i.e.* group of male users and group of female users); however the five factors account for the DD of female users somewhat better than for the DD of male users and mixed users (R^2 equals to 52.9%, 23.7% and 25.7%, respectively).

Table 2.7: Five principal components after varimax rotation (underlined values are >0.4 and maximal across factors in absolute value).

		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
n.	Variable	Left buttock (pressure)	Buttock (area)	Right buttock (pressure)	Left buttock vs. thigh (area)	Right thigh (pressure)
1	caTL	0.024	-0.104	0.025	<u>-0.422</u>	0.028
2	caTR	-0.003	-0.247	-0.011	<u>-0.501</u>	0.014
3	caBL	0.101	<u>-0.477</u>	-0.007	0.025	0.021
4	caBR	-0.052	<u>-0.568</u>	0.050	-0.111	0.043
5	caTL/caSUM	0.035	0.257	0.029	-0.288	-0.014
6	caTR/caSUM	-0.020	0.070	-0.040	<u>-0.445</u>	-0.024
7	caBL/caSUM	0.057	-0.082	-0.030	<u>0.443</u>	0.019
8	caBR/caSUM	-0.092	-0.286	0.044	0.271	0.018
9	cpTL	0.001	0.092	0.067	0.042	0.380
10	cpTR	-0.004	-0.108	-0.031	-0.007	<u>0.497</u>
11	cpBL	<u>-0.440</u>	0.070	-0.151	-0.036	0.106
12	cpBR	-0.094	0.024	<u>-0.453</u>	-0.022	0.166
13	cpTL/cpSUM	0.161	0.113	0.260	0.014	0.108
14	cpTR/cpSUM	0.157	-0.128	0.151	-0.032	0.261
15	cpBL/cpSUM	-0.381	0.058	0.121	-0.007	-0.241
16	cpBR/cpSUM	0.084	-0.032	<u>-0.413</u>	0.020	-0.078
17	pcpTL	-0.126	0.253	0.037	0.010	0.280
18	pcpTR	-0.109	0.022	-0.121	0.005	<u>0.499</u>
19	pcpBL	<u>-0.478</u>	-0.019	-0.057	0.001	0.095
20	pcpBR	-0.025	0.087	<u>-0.419</u>	-0.003	0.033
21	pcpTL/pcpSUM	0.115	0.256	0.246	-0.024	0.042
22	pcpTR/pcpSUM	0.178	-0.052	0.121	-0.020	0.249
23	pcpBL/pcpSUM	<u>-0.459</u>	-0.135	0.184	0.025	-0.076
24	pcpBR/pcpSUM	0.235	-0.021	<u>-0.423</u>	0.006	-0.118
	Eigenvalue	11.036	4.027	2.574	1.887	1.330
	Cum percent	46.0	62.8	73.5	81.3	86.9

As coefficients in Table 2.8 show, increasing Factor 2 (significant for the whole sample and for the sub-sample of male users) and decreasing Factor 1 (significant for the whole sample and for the sub-sample of female users) and Factor 5 (marginal for the whole sample and for the sub-sample of female users) would be effective at decreasing DD. In particular, the coefficients for Factor 2 (-0.348 and -0.360 for the whole sample and for the sub-sample of male users, respectively) indicated that increasing contact areas at the buttocks (specifically, caBL and caBR) would be the most effective method for decreasing DD in particular in the group of male users. Similarly, the coefficients for Factor 1 (0.174 and 0.350 for the whole sample and for the sub-sample of female users, respectively) suggest that decreasing average (peak) contact pressures and ratios relevant to the left buttock (specifically, cpBL, pcpBL e pcpBL/pcpSUM) would be the second most effective way of decreasing subjective perception of discomfort especially in the group of female users. Finally, the coefficients of Factor 5 (-0.206 and -0.901,for the whole sample and for the sub-sample of female users, respectively) provides one more suggestion for seat design improvement consisting in decreasing contact pressure and peak at the right thigh (specifically, cpTR and pcpTR), this action will be particularly effective on the group of female users.

Table 2.8: Standard coefficients for regression models relating PCA factors to DD.

Term	Mixed		Males		Females	
	Coeff	p.value	Coeff	p.value	Coeff	p.value
Intercept	2.602	0.000	2.648	0.000	2.449	0.000
Factor 1 - Left buttock (pressure)	<u>0.174</u>	0.024	0.132	0.209	<u>0.350</u>	0.004
Factor 2 - Buttock (area)	<u>-0.348</u>	0.001	<u>-0.360</u>	0.003	-0.276	0.166
Factor 3 - Right buttock (pressure)	0.013	0.836	0.018	0.810	0.157	0.247
Factor 4 - Left buttock vs. thigh (area)	-0.037	0.575	-0.076	0.333	-0.054	0.784
Factor 5 - Right thigh (pressure)	<u>-0.206</u>	0.064	-0.139	0.273	<u>-0.901</u>	0.024

(b) <u>PLSR Analysis</u>

From the PLSR analysis performed on the set of 27 variables listed and described in Table 2.5, two optimal components were selected (as the vertical line indicates in Figure 2.5 on the next page). The amount of variance in the predictors explained by the model selected is 55% (Table 2.9 on the next page). The Analysis of Variance (ANOVA) for DD shows that the fitted regression model was significant ($p \leq 0.01$, Table 2.10 on next page).

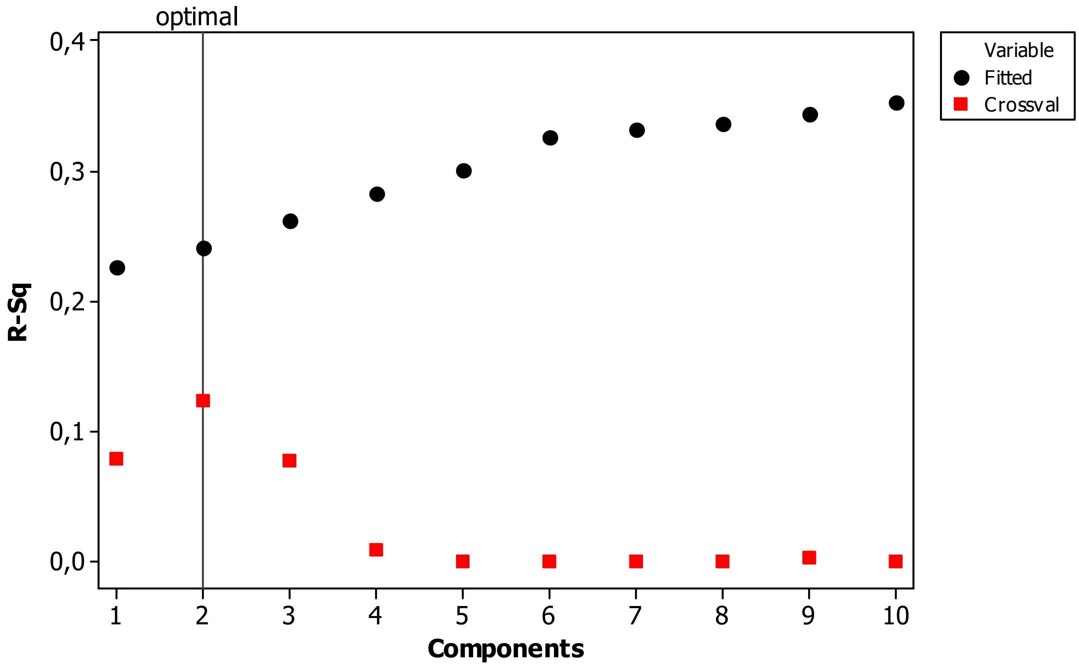


Figure 2.5: Partial Least Squares Model Selection Plot (response is DD for mixed sample)

Table 2.9: Model Selection and Validation for DD (10 components cross-validated and 2 selected).

Components	X-Var	Error SS	R-Sq	PRESS	R-Sq (pred)
1	0.167	53.428	0.226	63.656	0.079
2	0.550	52.362	0.242	60.495	0.124
3		50.925	0.263	63.754	0.077
4		49.554	0.283	68.482	0.009
5		48.270	0.301	71.729	0.000
6		46.500	0.327	75.004	0.000
7		46.105	0.332	74.514	0.000
8		45.844	0.336	70.725	0.000
9		45.287	0.344	68.930	0.002
10		44.691	0.353	70.870	0.000

Table 2.10: ANOVA for DD on the mixed sample (PLSR case).

Source	DF	SS	MS	F	p.value
Regression	2	16.718	8.359	13.57	0.00
Residual Error	85	52.362	0.616		
Total	87	69.079			

The standardized coefficient plot (Figure 2.6) depicts the sign and the magnitude of the relationship between predictors and response. In particular, the coefficients show that increasing contact areas at the buttocks (specifically, caBL and caBR) would be the most effective method for improving DD. Similarly, the coefficients suggest that decreasing average (peak) contact pressures and ratios relevant to the left buttock (specifically, cpTL, cpBL, pcpTL, pcpBLandpcpBL/pcpSUM) would be the second most effective way of improving the subjective ratings. Finally, the coefficients suggest a third strategy, the decrease of the peak pressure at the right thigh (specifically, pcpTR).

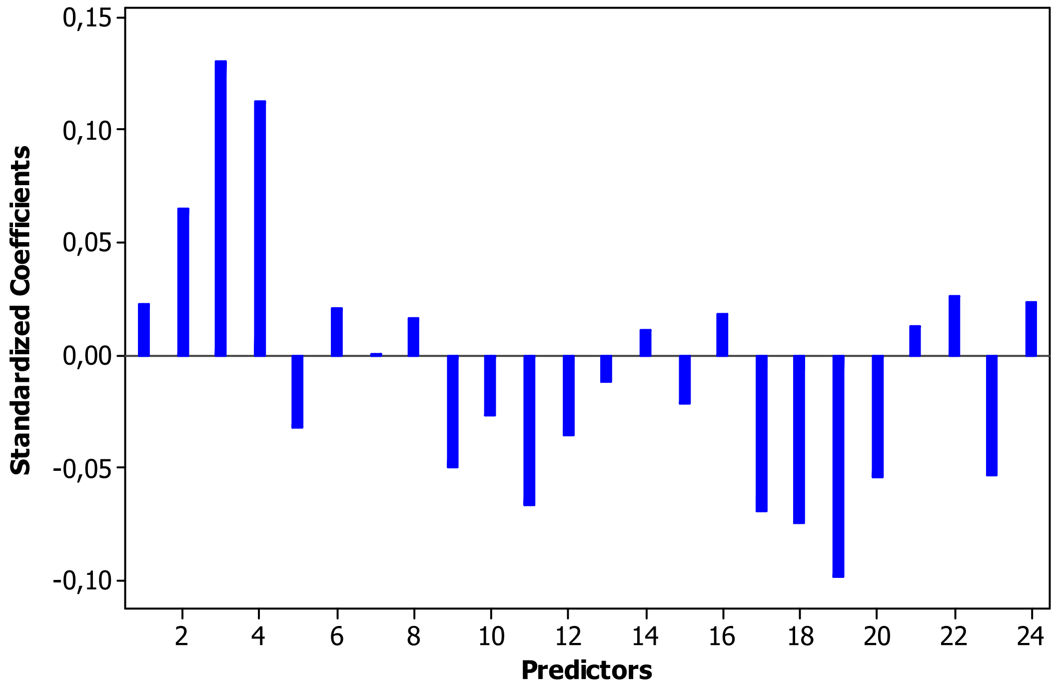


Figure 2.6: PLS Standard Coefficients Plot – 2 components.

Figure 2.7 (on the next page) shows the correlation between the loadings of each predictor on the first and second components comparing the importance of these to the model. A subset of variables was found in each optimal component that predominantly determined the respective component level, as evidenced by largest standardized coefficients and the biggest impact on DD. It's worthwhile to note that for component 2 the variables linked to thighs are positively related to DD, while the variables linked to buttock are negatively related.

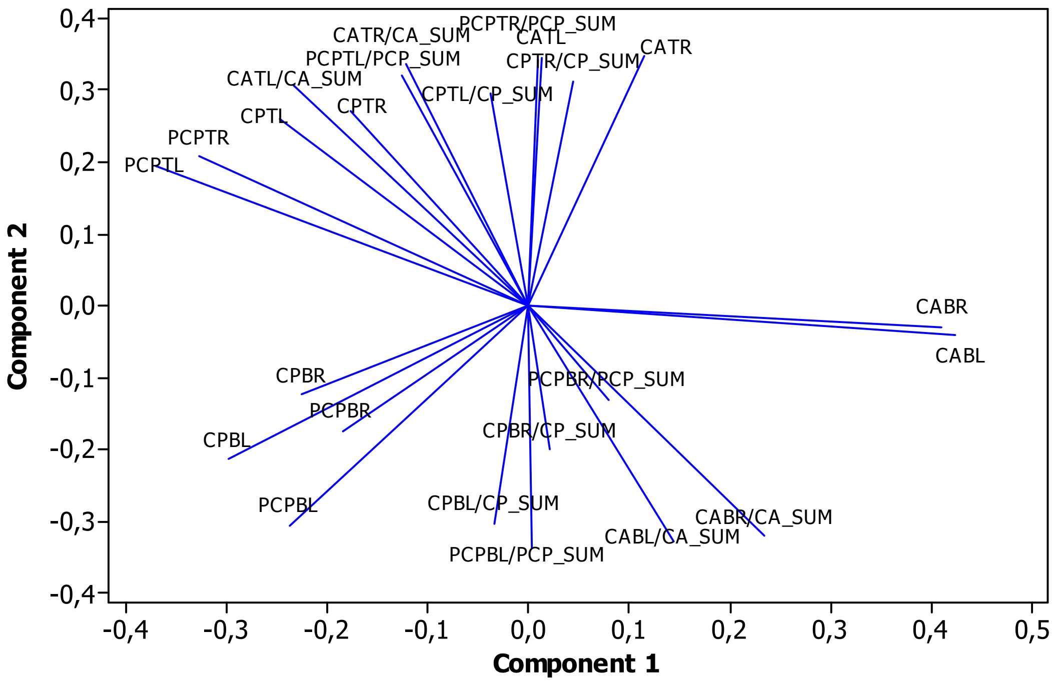


Figure 2.7: PLS Loading Plot.

By repeating the regression analysis both for the sub-samples of Males and Females, as done with PCR, it has come at the same conclusions just carried out for the mixed sample (Table 2.11). However, it seems not worth differentiate the improvement strategy between sub-sample of Males and Females.

Regression analysis showed that the components selected could explain better the DD for the Females sub-sample (R^2=52,64; X-Var=60,72%) than for the Males (R^2=25,12%; X-Var=32,82%). All three fitted regression models for DD were significant ($p \leq 0.01$).

Table 2.11: Normalized weights of standard coefficients of the stratified regression models connected to the PLS components (only significant coefficients given).

Variable	Mixed	Males	Females
caTL			
caTR	6,2%	7,9%	
caBL	12,3%	13,2%	7,6%
caBR	10,7%	8,7%	7,8%
caTL/caSUM			-10,1%
caTR/caSUM			
caBL/caSUM			
caBR/caSUM			6,2%
cpTL			
cpTR			

Variable	Mixed	Males	Females
cpBL	-6,3%	-5,4%	
cpBR			
cpTL/cpSUM			
cpTR/cpSUM			
cpBL/cpSUM			
cpBR/cpSUM			
pcpTL	-6,6%	-6,9%	-6,5%
pcpTR	-7,1%	-5,8%	-7,1%
pcpBL	-9,3%	-8,1%	-8,4%
pcpBR	-5,1%		
pcpTL/pcpSUM			5,7%
pcpTR/pcpSUM			
pcpBL/pcpSUM	-5,0%		-5,7%
pcpBR/pcpSUM			

(c) Filtering data and results comparing

Both multivariate approaches presented, in subsections (a) and (b), have identified similar improvement strategies for DD response and have found significant regression models, however these models do not obtain a high goodness of fit. This consideration has suggested that the data could be affected by a noise higher than expected, although a strict experimental protocol was used. For this reason it was decided to filter the data by applying two criteria validation: 1) the consistency of the subjective evaluation in relation to the Seat Conditions; 2) the Symmetry Index (SI) based on the body weight to detect a correct posture of the users. As regard the first criteria, the subjective evaluations clearly inconsistent were discarded. As regard the second criteria, data that reached a reliable score on the SI scale were selected. The SI index is expressed by the following formula:

$$SI = \frac{2 \cdot (W_{Loaded} - W_{unloaded})}{(W_{Loaded} + W_{unloaded})} \cdot 100 \qquad \text{(eq. 2.8)}$$

Filtered data represent about the 25% of the initial database. Subsequently on these, the approaches showed in subsections (a) and (b) were again performed.

Rerunning the PCR the model was not significant due to the reduced filtered sample size. Instead, rerunning the PLSR the model was significant (Table 2.12 on the next page) and its performance indexes are higher than those of the model on unfiltered data (Table 2.13 on the next page). In particular, the goodness of fit is about 79% and the amount of variance in the predictors explained by the model selected is about 84%.

Table 2.12: ANOVA for DD on filtered data (PLSR case).

Source	DF	SS	MS	F	p.value
Regression	5	11.181	2.236	10.37	0.00
Residual Error	14	3.019	0.216		
Total	19	14.200			

Table 2.13: Model Selection and Validation for DD on filtered data (10 components cross-validated and 5 selected).

Components	X-Var	Error SS	R-Sq	PRESS	R-Sq (pred)
1	0.462	9.976	0.297	14.719	0.000
2	0.646	7.306	0.486	16.268	0.000
3	0.747	5.187	0.635	18.254	0.000
4	0.784	3.272	0.770	16.312	0.000
5	0.844	3.019	0.787	13.126	0.076
6		2.791	0.803	16.369	0.000
7		2.329	0.836	23.837	0.000
8		1.926	0.864	29.093	0.000
9		1.552	0.891	31.275	0.000
10		1.422	0.900	31.718	0.000

Figure 2.8 shows the comparison between performance indexes of the PLSR pre-filtering model and of the PLSR post-filtering model. It was not possible to repeat separately the analysis on the subsample of Males and Females due to the reduced filtered sample size. The PLSR procedure has proved to be a more robust approach than the PCR both for the sample size and for the interpretative and explanatory ability.

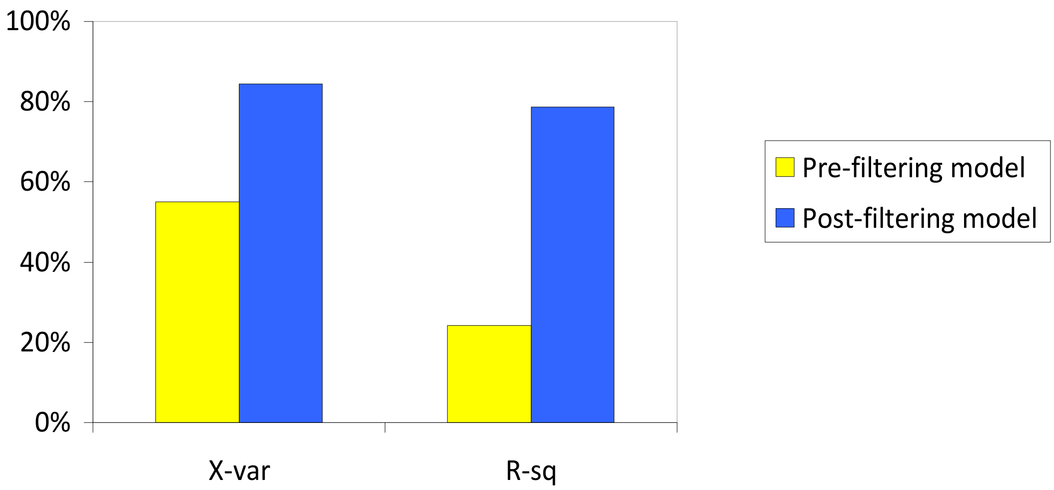


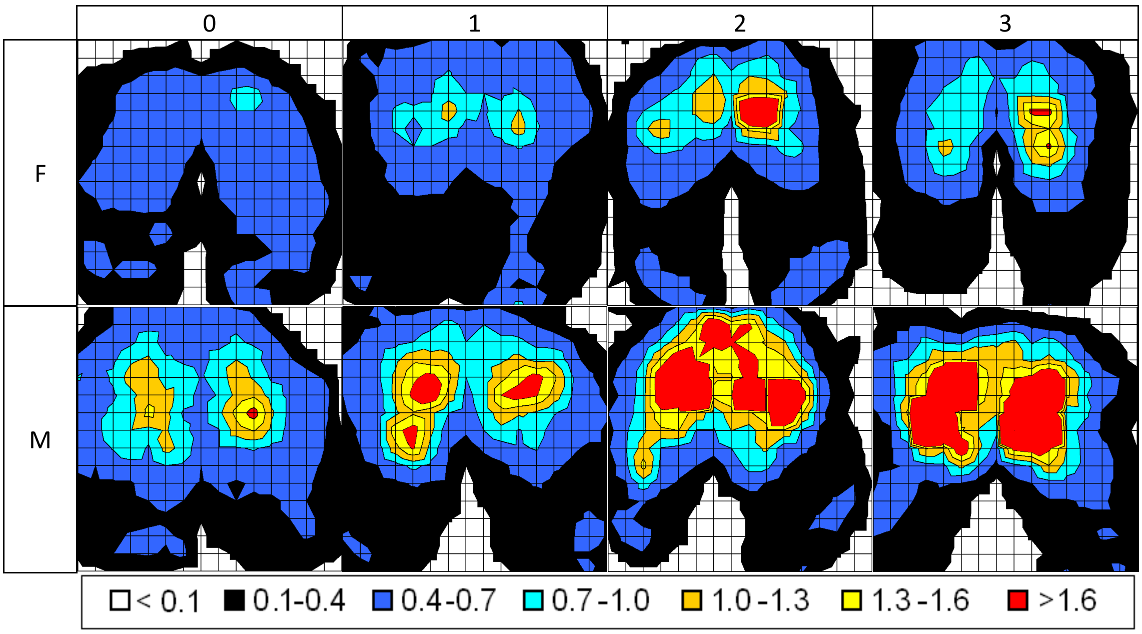
Figure 2.8: Comparison of the PLSR models

2.4.2 *Effects of anthropometric variability and seat conditions on interface pressures*

The analysis of MPCP maps shows that the contact pressure distribution of males is different from the contact pressure distribution of females. Table 2.14, reporting The MPCP maps can be arranged, can be read both by rows and by columns. In particular, the comparison by rows provides information on the effects of the anthropometric variability on seat-interface pressures. In fact, the female maps (first row of Table 2.14) show PCP values lower than the corresponding male maps (second row of Table 2.14). On the contrary, the comparison by columns provides information on the effects of the Seat Conditions. It's worthwhile to note that moving from the first column (low Stiffness) to the fourth one (high Stiffness), PCP values gradually increase. Thus, it can be said that the first two Seat Conditions show pressure levels lower than the last two Seat Conditions for both males and females.

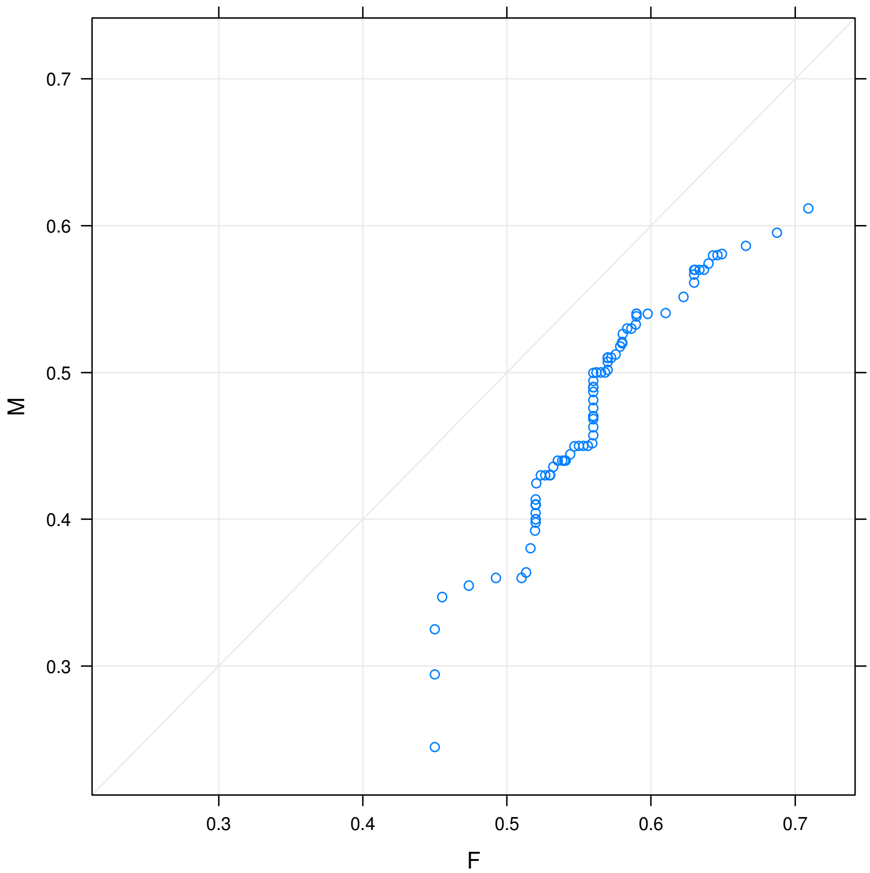
Briefly, the MPCP maps in Table 2.14 point out that: 1) pressure levels and contact areas vary between males and females and 2) it would seem that the males are more sensitive to changes in the seat condition and this could mean an amplification of discomfort effects in the long period.

Table 2.14: Maps of the MPCP for the different sub-samples stratified by gender and seat condition.



In addition the Quantile-Quantile plot in Figure 2.9 shows that, independently of the seat condition, the female users significantly differ from male users in terms of unloaded weight. In particular, quantiles of the unloaded weight are higher for female users than for male users. This result means that the location value of the unloaded weight is higher for female users than for male users; however, the non-linearity in the Quantile-Quantile plot implies that the difference between the two samples is not explained simply by a shift in location.

Figure 2.9: Quantile-Quantile plot for the unloaded weight.



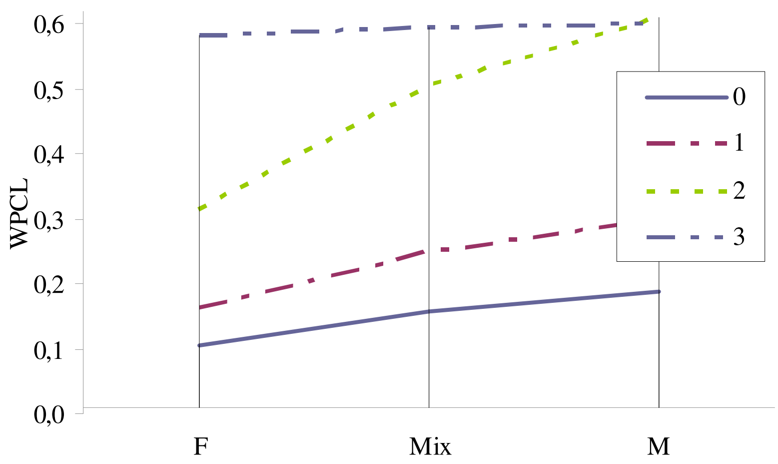
The obtained experimental results are consistent with the findings of previous studies (Dunk and Callaghan, 2005; Gregory *et al.*, 2006; Beach *et al.*, 2008) investigating the influence of personal characteristics (*i.e.*, gender and flexibility) on postures adopted when performing seated computer work. These findings generally evidence that males and females react differently to seated exposures; in particular, the study of Dunk and Callaghan (2005) suggests that men tend to slouch against the back rest while females perch closer to the front of the seat pan. The above gender-based differences have been related to inter-individual variations in hip, hamstring, and low-back flexibility.

2.4.3 Discriminant effectiveness of indexes based on seat-interface pressure in predicting discomfort

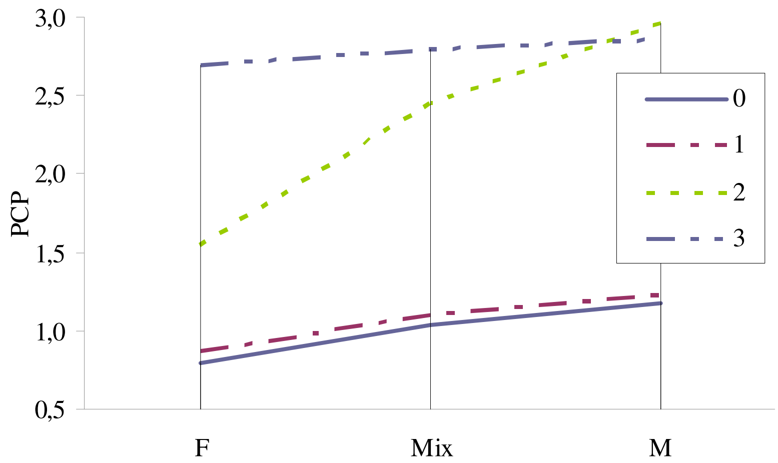
Mean values of PCP and WPCL were compared for the four tested chairs in order to verify the consistency of discriminant information provided by these indexes. The mean values of both indexes were calculated over the whole sample of users as well as over the two sub-samples of female users and male users. The results are shown in Figure 2.10(a) and Figure 2.10(b) for WPCL and PCP, respectively.

Figure 2.10: Mean effect plots assuming WPCL (a) and PCP (b) as response.

(a)



(b)



The two diagrams in Figure 2.10 show substantial coherence of the results provided by PCP and WPCL against increasing levels of Stiffness (S): the Seat Condition characterized by Low Stiffness (coded as 0) was the best one in terms of both WPCL and PCP, whereas the worst results were obtained for the Seat Condition characterized by High Stiffness (coded as 3).

For the Seat Condition 3, the mean effect plots of WPCL and PCP are constant against changes in the composition of the reference (sub-)sample. This result could be explained by a saturation effect due to the rigid cushion which produces very high pressure values which are comparable to those ones obtained in previous studies on a hard flat surface (Brienza and Karg, 1998; Ragan *et al.*, 2002). On the other hand, a similar effect is shown for the Seat Condition 0: the mean effect plots of both indexes, though not constant, show a lower slope compared to Seat Conditions 1 and 2 characterized by intermediate levels of Stiffness. This result could be explained as there were a point of diminishing returns beyond which decreasing the cushion stiffness is not really effective in further reducing the seat–interface pressure. A similar effect was found in a previous study with regard to cushion thickness (Ragan *et al.*, 2002).

Though the mean effects plots of WPCL and PCP show similar patterns, it's worthwhile to highlight that, in the WPCL diagram the differences between the mean effect plots of Seat Condition 0 and 1 are clearer than in the PCP diagram.

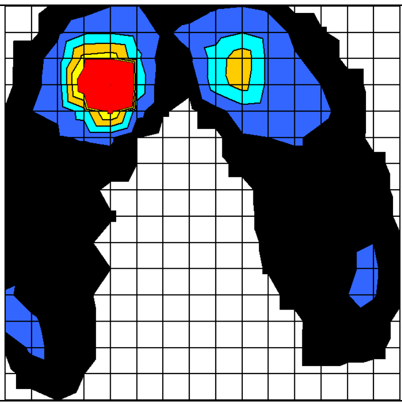
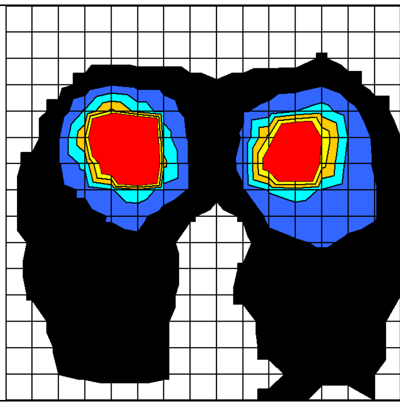
In order to verify if the two indexes significantly differ in discriminating among the four Seat Conditions, the Wilcoxon-Mann-Whitney test was performed. So for both indexes, three binary comparisons of Seat Conditions (0 vs. 1, 1 vs. 2 and 2 vs. 3) were carried out for the whole sample of users as well as the two sub-samples of Females and Males. Results (in Table 2.15) show that, independently from the composition of the reference (sub-)sample, the WPCL is able to discriminate between the Seat Conditions 0 and 1, and between 1 and 2; whereas, the PCP only distinguishes between the Seat Conditions 1 and 2.

Table 2.15: Results of non-parametric Wilcoxon-Mann-Whitney test.

	Mixed			Males			Females		
	0 vs. 1	1 vs. 2	2 vs. 3	0 vs. 1	1 vs. 2	2 vs. 3	0 vs. 1	1 vs. 2	2 vs. 3
PCP		X			X			X	
WPCL	X	X		X	X		X	X	

Finally, it's worthwhile to point out that, though their overall results are consistent, PCP and WPCL do not provide the same information. Indeed, for the sub-sample of male users, the maximum values of these indexes refer to different pressure maps and so identify different users (Table 2.16 on the next page).

Table 2.16: Pressure maps related to maximum values of WPCL and PCP.

	PCP: P.18 – Seat Cond. 2	WPCL: P.10 – Seat Cond. 3
M		

2.5 Conclusions

This study provides satisfactory answers to some relevant issues related to the assessment of sitting discomfort due to office chairs.

Subjective discomfort evaluations resulted significantly correlated to several combinations of pressure variables, derived in terms of average contact area and average (peak) contact pressure. Consequently, these variables can be used across anthropometric variability for the assessment of static sitting discomfort in short-experimental sessions.

In particular, the perceived discomfort appears to be mainly due to the lack of pressure balance between the bilateral buttocks and the lack of balance in contact areas between buttocks and thighs. Thus, asymmetries in pressure distributions and in contact areas should be considered undesirable as they lead to increasing unpleasant state of human body.

The experimental results confirm the hypotheses that due to fundamental biomechanical differences in their sitting behaviours, males and females are exposed to different loading patterns and experience different discomfort pathways.

The adopted statistical approach can effectively support the designer in diagnosing seat discomfort (via the MPCP maps) and testing (via the WPCL index) alternative design solutions (*e.g.* in terms of both shape and materials) for specific purposes (*e.g.* design for a specific user or design for a generic user).

Though the paper focus is on the comparative assessment of office seating discomfort across a gender stratified population of healthy users, the proposed data analysis strategy can be easily adapted to other experimental seating contexts involving different target populations.

2.6 References

Andreoni, G., Santambrogio, G.C., Rabuffetti, M., Pedotti, A., 2002. Method for the analysis of posture and interface pressure of car drivers. Applied Ergonomics, 33 (6), 511–522.

Annetts, S, Coales, P., Colville, R., *et al.*, 2012. A pilot investigation into the effects of different office chairs on spinal angles. European Spine Journal, 21 (Suppl 2), 165–170.

Armstrong, T.J., Buckle, P., Fine, L.J., Hagberg, M., Jonsson, B., Kilbom, A., Kuorinka, I., A.A., Silverstein, B.A., Sjogaard, G., Viikari-Juntira, E.R.A., 1993. A conceptual model for work-related neck and upper limb disorders. Scandinavian Journal of Work, Environment and Health, 19, 73–84.

Barone, S., Lanzotti, A., 2009. Robust Ergonomic Virtual Design. In: Statistics for Innovation: statistical design of "continuous" product innovation. Springer, 61–83.

Beach, T.M., McDonald, K.A., Coke, S.K., Callaghan, J.P., 2008. Gender responses to automobile and office sitting - influence of hip, hamstring, and low-back flexibility on seated postures. The Ergonomics Open Journal, 1, 1–9. Retrieved January 29, 2014, from: http://benthamopen.com/toergj/articles

Brienza, D.M., Karg, P.E., 1998. Seat cushion optimization: a comparison of interface pressure and tissue stiffness characteristics for spinal cord injured and elderly patients. Archives of Physical Medicine and Rehabilitation, 79, 388–394.

De Korte, E.M., Huysmans, M.A., De Jong, A.M., Van de Vene, J.G.M., Ruijsendaal, M., 2012. Effects of four types of non-obtrusive feedback on computer behaviour, task performance and comfort. Applied ergonomics, 43(2), 344–353.

de Looze, M.P., Kuijt-Evers, L.F.M., Van Dieën, J.H., 2003. Sitting comfort and discomfort and the relationships with objective measures. Ergonomics, 46(10), 985–997.

Dreyfuss, H., 2002. The measure of man and woman. Wiley.

Dul, J., Hilderbrandt, V.H., 1987. Ergonomic guidelines for the prevention of low back pain at the workplace. Ergonomics, 30(2), 419–429.

Dunk, N.M., Callaghan, J.P., 2005. Gender-based differences in postural responses to seated exposures. Clinical Biomechanics, 20, 1101–1110.

Ehrlich, G. E., 2003. Low Back Pain. Bulletin of the World Health Organization, 81(9), 671–676.

Ellegast, R.P., Keller, K., Groenesteijn, L., Krause, F., Berger, H., Vink, P., 2012. Comparison of four specific dynamic office chairs with a conventional office chair: impact upon muscle activation, physical activity and posture. Applied ergonomics, 43(2), 296–307.

Fay, B.T., Brienza, D.A., 2000. What is interface pressure? Paper presented at the 22nd Annual EMBS International Conference, Chicago, IL.

Franz, M.M., Durta, A., Zenk, R., Desmet, P.M.A., 2012. Comfort effects of a new car head-rest with neck support Applied ergonomics, 43(2), 336–343.

Fujimaki, G., Mitsuya, R., 2002. Study of the seated posture for VDT work. Displays, 23 (1-2), 17–24.

Gregory, D.E., Dunk, N.M., Callaghan, J.P., 2006. Stability ball versus office chair: Comparison of muscle activation and lumbar spine posture during prolonged sitting. Human Factors, 48(1), 142–153

Groenesteijn, L., Ellegast, R.P., Keller, K., Krause, F., Berger, H., de Looze, M.P., 2012. Office task effects on comfort and body dynamics in five dynamic office chairs. Applied ergonomics, 43(2), 320–328.

Hamberg-van Reenen, H.H., Van der Beek, A.J., Blatter, B.M., Van der Grinten, M.P., Van Mechelen, W., Bongers, P.M., 2008. Does musculoskeletal discomfort at work predict future musculoskeletal pain? Ergonomics, 51, 637–648.

Harkness, E.F., Macfarlane, G.J., Silman, A.J., McBeth J., 2005. Is musculoskeletal pain more common now than 40 years ago?: Two population-based cross-sectional studies. Rheumatology, 44 (7), 890–895.

Helander, M.G., Zhang, L., 1997. Field studies of comfort and discomfort in sitting. Ergonomics, 40(9), 895–915.

Herbert, R., Dropkin, J., Warren, N., Sivin, D., Doucette, J., Kellogg, L., Bardin, J., Kass, D., Zoloth, S., 2001. Impact of a joint labor-management ergonomics program on upper extremity musculoskeletal symptoms among garment workers. Applied Ergonomics, 32(5), 453–60.

Kamijo, K., Tsujimara, H., Obara, H., Katsumatu, M., 1982. Evaluation of seating comfort. SAE Technical Paper Series 820761.Society of Automotive Engineers, Troy, MI.

Kamp, I., The influence of car seat design on its character experience, 2012. Applied Ergonomics, 43(2), 329–335.

Kee, D., Lee, I., Relationships between subjective and objective measures in assessing postural stresses, 2012. Applied Ergonomics, 43(2).

Kong, Y.U., Kim, D.M., Lee, K.S., Jung, M.C., 2012. Comparison of comfort, discomfort, and continuum ratings of force levels and hand regions during gripping exertions, Applied Ergonomics, 43(2), 283–289.

Kuijt-Evers, L.F.M., Groenesteijn, L., de Looze, M.P., Vink, P., 2004. Identifying factors of comfort in using hand tools. Applied Ergonomics, 35, 453–458.

Kyung, G., Nussbaum, M.A., 2008. Driver sitting comfort and discomfort (part II): relationships with and prediction from interface pressure. International Journal of Industrial Ergonomics, 38(5-6), 526–538.

Lanzotti, A., 2008. Robust design of car packaging in virtual environment. International Journal on Interactive Design and Manufacturing (IJIDeM), 2(1), 39–46.

Lanzotti, A., Vanacore, A., 2007. An efficient and easy discretizing method for the treatment of noise factors in robust design. Asian Journal on Quality, 8(3), 188–197.

Lanzotti, A., Vanacore, A., Trotta, M., 2011. Validation of a new index for seat comfort assessment based on objective and subjective measurements. Proceedings of the IMProVe International conference on Innovative Methods in Product Design – ADM-INGEGRAF, June 15th – 17th, Venice, Italy.

Lengsfeld, M., Frank, A., van Deursen, D.L., Griss, P., 2000. Lumbar spine curvature during office chair sitting. Medical Engineering and Physics, 22 (9), 665–669.

Loisel, P., Gosselin, L., Durand, P., Lemaire, J., Poitras, S., Abenhaim, L., 2001. Implementation of a participatory ergonomics program in the rehabilitation of workers suffering from sub acute back pain. Applied Ergonomics, 32 (1), 53–60.

Lueder, R.K., 1986. Seat comfort: A review of the Construct in the Office Environment. Human Factors, 25(6), 701–711.

Moes, N.C.C.M., 2005. Analysis of sitting discomfort, a review. In: Bust, P.D., McCabe, P.T. (Eds.), Contemporary Ergonomics 2005. Taylor & Francis, London, 200–204.

Nelson, N.A., Silverstein, B.A., 1998. Workplace changes associated with a reduction in musculoskeletal symptoms in office workers. Human Factors, 40 (2), 337–50.

Noro, K., Naruse, T., Lueder, R., Nao, N., Kozawa, M., 2012. Application of Zen sitting principles to microscopic surgery seating. Applied Ergonomics, 43(2), 308–319.

Park, M.Y., Kim, J.Y., Shin, J.H., 2000. Ergonomic design and evaluation of a new VDT workstation chair with keyboard-mouse support. International Journal of Industrial Ergonomics, 26(5), 537–548

Ragan, R., Kernozek, T.W., Bidar, M., Matheson, J.W., 2002. Seat-interface pressures on various thicknesses of foam wheelchair cushions: a finite modeling approach. Archives of Physical Medicine and Rehabilitation, 83, 872–875

Reed, M., Grant, C., 1993. Development of a measurement protocol and analysis techniques for assessment of body pressure distributions on office chairs. TR. Ann Arbor, MI: Univ. of Michigan, Center for Ergonomics.

Reed, M.P., Manary, M.A., Schneider, L.W., 1999. Methods for measuring and representing automobile occupant posture. SAE Technical Paper Series 1999-01-0959.

Rezaee, M., Jafari, N.J., Ghasemi, M., 2011. Low Back Pain and Related Factors among Iranian Office Workers. International Journal of Occupational Hygiene, 3 (1), 23–28.

Rubin, D., 2007. Epidemiology and risk factors for spine pain. Neurological Clinics, 25(2), 353–371.

Stinson, M., Crawford, S., 2009. Optimal Positioning: Wheelchair Seating Comfort and Pressure Mapping. In International Handbook of Occupational Therapy Interventions, Springer, 83–90.

van Dieën, J.H., de Looze, M.P., Hermans, V., 2001. Effects of dynamic office chairs on trunk kinematics, trunk extensor EMG and spinal shrinkage. Ergonomics, 44, 739.750.

Vink, P., 2012. Editorial: Comfort and discomfort studies demonstrate the need for a new model. Applied Ergonomics, 43(2), 271–276.

Vink, P., Bazley, C., Kamp, I., Blok, M., 2012. Possibilities to improve the aircraft interior comfort experience. Applied Ergonomics, 43(2), 354–359.

Winkel, J., Westgaard, R., 1992. Occupational and individual risk factors for shoulder neck complaints: Part II – The scientific basis (literature review) for the guide. International Journal of Industrial Ergonomics, 10, 85–104.

Zemp, R., Taylor, W.R., Lorenzetti, S., 2015. Are pressure measurements effective in the assessment of office chair comfort/discomfort? A review. Applied Ergonomics, 48, 273–282.

Zenk, R., Franz, M., Bubb, H., Vink, P., 2012. Spine loading in automotive seating. Applied Ergonomics, 43(2), 290–295.

Zhang, L., Helander, M., Drury, C., 1996. Identifying factors of comfort and discomfort. Human Factors, 38, 377–389.

3. A New Approach for GUI Usability Assessment

3.1 User Interface Problems and the "Magic number 5"

A well-thought-out study have to consider questions how to select participants, how to order tasks, what participants perform what tasks, and how many participants you need to get a reasonably reliable feedback. Only if you design in this direction you can save time and effort, and answer research questions that your study arises clearly (Tullis and Albert, 2008).

One of the most debated issues in specialized literature is related to the choice of the right number of users to be involved in the usability tests. Early some researchers suggested that about five to six users could detect the most of the problems in a usability test (Al-Awar *et al.*, 1981). Lewis (1982) published the first study describing how the binomial distribution can be used to model the sample size required to reveal usability problems. It's based on the probability of discovering a problem with probability p for a given set of tasks and user population given a sample size n.

During the '90 years, the use of GUIs spreads rapidly and the need for more precision in sample size estimates generates some studies which proposing the binomial model (Virzi, 1990; Wright and Monk, 1991; Virzi, 1992; Nielsen and Landauer, 1993; Lewis, 1993; Lewis, 1994).

However, the controversy arose in 2000 when Nielsen published *"Why you only need to test with 5 users"*. Ever since many strong opinions about the "magic number 5" were stated (Caulton, 2001; Spool and Schroeder, 2001; Perfetti and Landesman, 2001; Turner *et al.* 2002; Wixon, 2003; Lewis, 2001; Hertzum and Jacobsen, 2003; Woolrych and Cockton, 2001; Bevan *et al.* 2003; Turner *et al.* 2006; Lewis, 2006; Lindgaard and Chattratichart, 2007; Schmettow, 2008). The magic number 5 is derived from the number of users required to detect 85% of the problems in an interface, assuming that the probability that a user would have to tackle a problem is about 31%.

Sauro in 2010 seems to have settled the problem as it has been shown convincingly that does not exist a specific number of users that will always be the right number but testing with 5 users may be all you need to find out the problems in an interface. The discussion has not focused on the use of the binomial formula (or Poisson equivalent) but on the value of the model parameter. Such parameter represents the average frequency with which problems really occur. Given that the problems do not affect users evenly, it is not easy to know how frequently they occur. As a general rule you could use a probability of 31% (or more) for early design whereas 10% (or lower) for applications in use that have many users. Nielsen recommends to test not more than 5 users at a time. This does not mean that the 5 users are in total, in fact, you could test up to 20 users, *i.e.* 4 or 5 set of 5 users. For this reason the best approach when you plan an usability study is an iterative design and test strategy.

3.2 Introducing the case study

Large and small companies develop products through structured work teams supported by software toolsets aimed to keep up their design (Sharmin *et al.*, 2009). These tools are generally complex and require skilled users dealing with design, test and check activities. The main issue is that these users are geographically dispersed and interdepartmental (Stenzel and Pourroy, 2008) besides the design and manufacturing process often aren't concurrent but they turn in the loop. This induce a data flow loop which move through some division in the world. Systems, procedures and software to capture and manage design and manufacturing data are necessary to ride out these issues. Some authors (Elgueder *et al.*, 2010; El Hani *et al.*, 2012) propose software tools to concurrent manage design and manufacturing process data. In such context, a "Digital Pattern" (DP) platform is recommended. A DP platform is a set of geometric and numeric data structures, as well as of preconfigured and parametric models, which can be adapted to specific contexts. Therefore, a DP acts as a knowledge-based engineering (KBE) system aimed to improve quality and reduce times and costs for product development through a massive use of knowledge and tools integration inside company. Such aims are accomplished through the fast and the best re-use of company knowledge, *i.e.* through technical and technological predefined solutions that quickly marry new projects, allowing fast performance evaluations and immediate checks. Such solutions should also be able to give feedbacks on design and production costs. In papers by Lanzotti *et al.* (2013) and Patalano *et al.* (2013) a DP system, developed to assist gearbox design, is described.

When developing a KBE system, the evaluation of interface usability, to demonstrate the effective reduction of development time and satisfaction in use, is a critical point. Indeed, usability can be defined as *the extent to which specific users, in a specific context of* *utilisation, can use a product to their satisfaction, in order to effectively and efficiently achieve specific goals* (Madhavan and Alagarsamy, 2013). Sohaib and Khan (2010) claim that agile projects needs to adopt aspects of *usability engineering* by incorporating user scenarios and including usability specialists in the team. During the designing of mechanical parts, designers need to verify the correctness of the hypotheses in use, especially in relation with multi-objective tasks (Patalano *et al.*, 2013). Furthermore, the use of a KBE system is strategic for designers if we consider that they spend up to 30% of their time to searching data (Sandeberg, 2003) and this percentage rises up to 50% when we take into account the time spent to validating the data (Bourke, 2013).

Following the approach used by Di Gironimo *et al.* (2013), this work proposes the usability improvement of the KBE system's Graphical User Interface (GUI) through a participatory testing. This GUI assists designing automotive manual transverse gearboxes. An approach based on Analytic Hierarchy Process (AHP) and Multiple-Criteria Decision Analysis (MCDA) is used. A participatory test was performed for evaluating the Usability Index (UI) of the GUI. The AHP approach implies the decomposition of the problem into several levels (Saaty, 2008). In the present work, three dimensions of UI are considered: Effectiveness, Efficiency and Satisfaction. The MCDA methodology implies that all the measures corresponding to the factors of the problem, being of different nature and magnitude, are first normalized and weighed, and then could be combined into one overall value through a bottom-up approach (Sauro and Kindlund, 2005a; Kim and Han, 2008).

For the experimental phase, a set of selected end-users has to complete two specific tasks: 1) to design an automotive manual transverse gearbox; 2) to modify an existing gearbox model re-using previous knowledge. The goals of both Task 1 and Task 2 are clearly defined. Then, measures of subjective ratings and objective metrics are collected. So, the UI is calculated and the effects of the usability dimensions are analysed.

Taking into account the experimental data analysis, some frequent critical issues are identified. Before making any changes to the GUI, a questionnaire is administrated to the same users of the previous experiment to confirm the validity of new conceptual features proposed.

In this perspective, a new release of the GUI is developed and the validation test is again performed for a new assessment of the GUI, according to a continuous improvement loop.

3.3 The Methodological approach

The traditional design process tends to favour the functional aspects of a product at the expense of the cognitive-emotional ones, not considering that a product can even have only an emotional function (Norman, 2004). The usability assessment must take into account the analysis of both objective and subjective aspects that are closer to the emotional sphere of the individual. In this respect, the participation of the end-user into design process is crucial according to a User-Centred Design approach. Using such assumptions and starting from (Di Gironimo *et al.*, 2013), the approach adopted to achieve the purpose of this study requires the involvement of potential users during all phases of the usability evaluation process (Nielsen, 1993). Figure 3.1 shows the logical flow chart of proposed methodological approach.

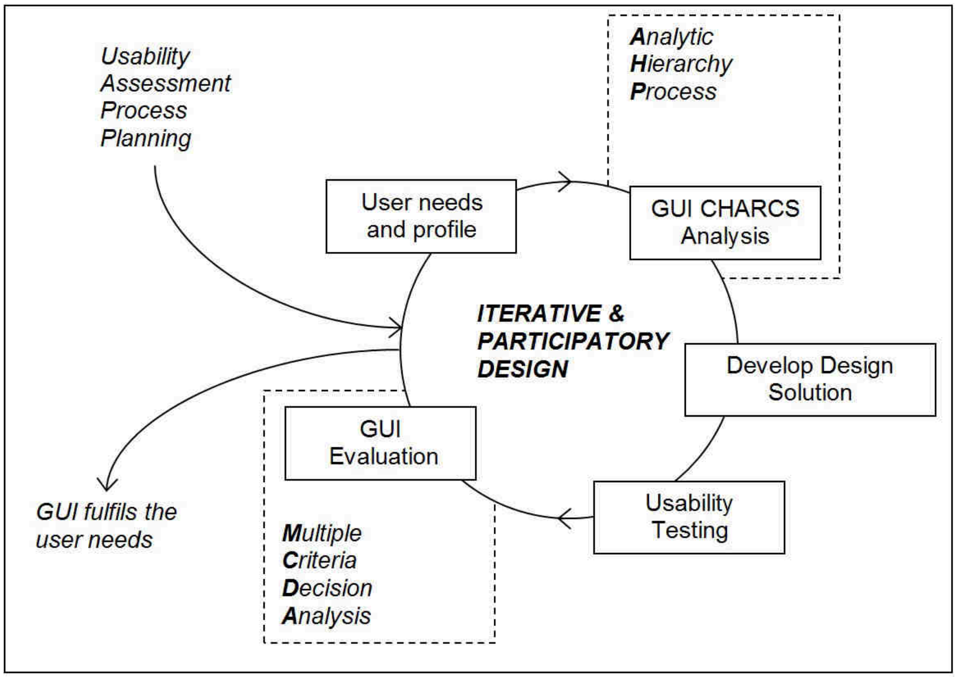


Figure 3.1: The logic flow chart of methodological approach.

In summary, both the user profile and the GUI characteristics are identified. Given these requirements and the context of use, the develop design solutions are implemented and the usability testing are planned. Two specific tasks, devoted to translate the usability characteristics factors into measurable usability functions, are properly defined. In order to reduce the noise related to the user's skill, a training phase is conducted for all users. Then, the experimental data are collected and the GUI evaluation is settled.

As stated before, the GUI usability assessment is carried out by using Saaty's AHP (Saaty, 2008) and MCDA methodology (Figueira *et al.*, 2005). The analysis can be summarised as follows:

- decomposition of the problem into several hierarchical levels and factors;
- scoring of the factors related to each identified level by means of pairwise comparison.

In particular, MCDA methodology allows combining the values of the individual usability functions into a single usability index (UI) by means of a bottom-up approach (Sauro and Kindlund, 2005a; Kim and Han, 2008).

Starting from results of experimental data, some changes are proposed. Finally, the validation experiment is performed to verify the goodness of these changes.

3.3.1 User Profile Definition and GUI characteristics

A KBE system can assist engineers in re-using previous knowledge in order to improve time-consuming tasks, as retrieval and selection of previous architectures, and to modify and virtually test a new product design. A critical point in the development of a KBE system is the interface usability to demonstrate effective reduction of development time and satisfaction in its use. Specifically, the present work deals with a KBE system previously proposed and providing to assist within the design of automotive manual transverse gearboxes. Then, the GUI for the KBE system is released (Figure 3.2).

GUI interaction depends primarily on the kind of user and the context of use. All characteristics, which identify specific needs, desires and interests and even behaviours, contexts of use and personal preferences (Ghosh and Dekhil, 2009), define a specific user profile. The MatLAB®-based GUI is accomplished for junior designers belonging to automotive industry. Their minimum skills is defined as follows: 1) good knowledge (at least theoretical) of a gearbox, 2) good expertise with the use of graphical user interfaces, and more generally, with the use of specialised software. Hence, the GUI should easily satisfy user needs with no cognitive overload.

The GUI is designed to perform two major tasks: 1) to design gearboxes rapidly, reducing the risks of using incomplete information when making product development decisions; 2) to assist the designer in redesigning the gearboxes previously developed (for more details see Lanzotti *et al.*, 2013; Patalano *et al.*, 2013).

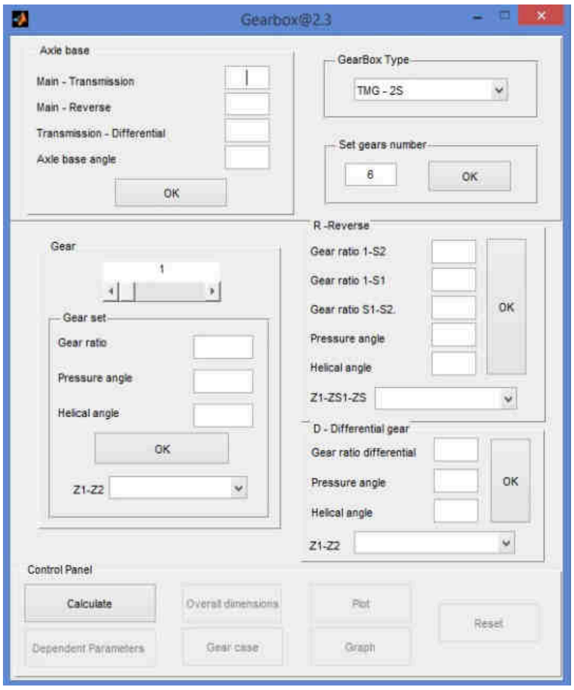


Figure 3.2: GUI for the gearbox CAD modelling software tool.

The window is divided in three main fields (Figure 3.2): the upper field where the gearbox is pre-configured; the middle field where the gearbox is configured and the gears are characterised; the lower field where the post-processing and evaluating commands are located.

The designers could set: type of gearbox (selecting it from a list box); number of gears (up to six); layout parameters (axle bases and angle between them). Besides, designers could set three characteristic parameters for each gear: gear ratio, pressure angle and helical angle (all in viable range). Teeth numbers of the gear are automatically generated by means of an algorithm that pulls out ten set of teeth numbers that meet the three parameters of the gear.

Specific panels to set reverse and differential gear are developed. A new gearbox can be defined by setting these parameters, but a previous *i.e.* existing design can also be edited.

The computational structure is guided by a directed graph (digraph). The nodes are associated to parameters (dependent and independent), while directed edges represent the mathematical relationships among parameters. The “Graph” button displays the digraph in a new window where the designer is able to interact with the graph: for example removing relations and generating an isolated node, in order to set a constant value during the calculation step.

As regards the post-processing, the “Dependent Parameters” command displays all computed parameters useful to determine the geometry correctness (modules, pitch diameters, addendum (or outside) diameters, root diameters, base diameters, etc).

The “Plot” command and “Gear Case” command display the mesh representation of the gears and the gear case, respectively. The automatically generated models can be exported both as *txt* and *stl* files, so ensuring the generation of the corresponding 3D solid models in any CAD environment.

The bounding box of generated meshes can be shown and this option helps the designer to interactively set the gearbox parameters as to fit the whole gearbox within a desired volume.

3.3.2 The AHP model for GUI usability

Figure 3.3 shows the decomposition of the usability according to the AHP approach. At first level we set the GUI usability (U) that is decomposed according to (ISO 9241-11-1998; Hornbaek, 2006) in Usability Dimensions (UDs), whit in the second level. The UDs are defined as follows:

- *Effectiveness*, the level of accuracy and completeness with which users achieve a specified goal;
- *Efficiency*, the level of effectiveness achieved to the expenditure of resources;
- *Satisfaction*, the condition of freedom from discomfort and positive attitude towards the use of the product.

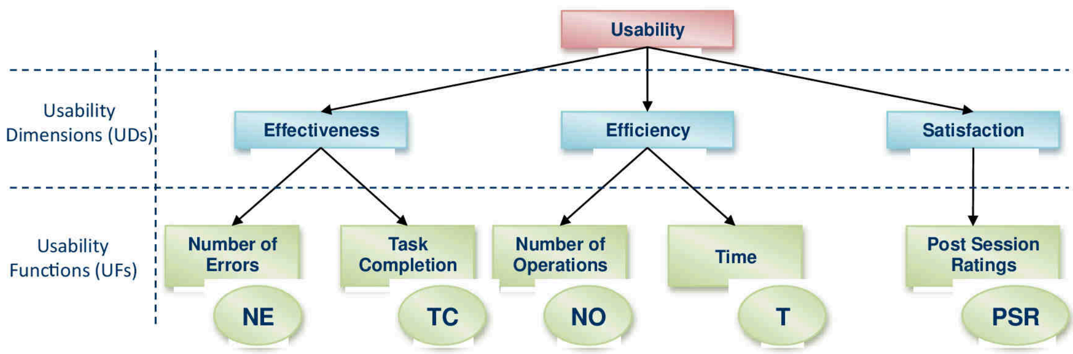


Figure 3.3: GUI Usability hierarchical decomposition.

In turn, the UDs are broken down, at the third level, in Usability Functions (UFs) that are strictly related to the experimental tasks. These UFs, in fact, are accurately determined during the experimental phase and they tackle critical aspects for GUI usability assessment.

According to the hierarchical decomposition above described, the analysis of GUI characteristics provides the following UFs:

- *Number of Errors* (NE), measure of Effectiveness, is the number of error messages reported by the GUI during the task execution;
- *Task Completion* (TC), measure of Effectiveness, is the level of completion and accuracy in achieving the goals of the task;
- *Number of Operations* (NO), measure of efficiency, is defined as the number of operations used to complete a task in terms of mouse clicks and keystrokes;
- *Time* (T), measure of efficiency, is the effective time to perform a task or sub-activities;
- *Post Session Ratings* (PSR), measure of Satisfaction, is a score, which expresses the feeling of users about the GUI use.

3.3.3 The Usability Index definition

Starting from the assumption that the factors of the hierarchy, for each level, are preferentially independent to each other, a simple linear additive evaluation model could be applied. By means of Multi-Criteria Decision Analysis (MCDA) all the measures corresponding to the factors could be combined into one overall value (Sauro and Kindlund, 2005a; Kim and Han, 2008). In particular, the measure of each factor is multiplied by a weight based on a specific criterion, and then the weighted scores are summed up. The calculation of the index starts from the UFs, by using experimental data. Being data of different nature and magnitude, a preliminary normalisation is required, in order to ensure the comparison between them.

The normalisation techniques, adopted for the specific UFs, are briefly described in the following:

- *0-Max* normalisation performs a linear transformation of the original data. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the equation (eq. 3.1):

$$e'_{ij} = \frac{e_{ij}}{\max_i} \qquad \text{(eq. 3.1)}$$

- Min/e_{ij} normalisation performs a linear transformation of the original data that reverses the direction of preferences. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the equation (eq. 3.2):

$$e'_{ij} = \frac{\min_i}{e_{ij}} \qquad \text{(eq. 3.2)}$$

- e_{target}/e_{ij} normalisation performs a linear transformation of the original data that reverses the direction of preferences and requires a target value, lower than the minimal value. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the formula (eq. 3.3):

$$e'_{ij} = \frac{e_{target}}{e_{ij}} \qquad \text{(eq. 3.3)}$$

The above techniques adopted for each UF are reported in Table 3.1.

Table 3.1: Normalisation techniques adopted for UFs.

Normalisation technique	Usability Functions
$0 - \max$	PSR
$\frac{\min_i}{e_{ij}}$	NE, TC
$\frac{e_{target}}{e_{ij}}$	NO, T

The outcomes of the normalisation procedure are the usability measures (um_i) that range from 0 to 1. Then, for each subgroup of usability measures, the Usability Dimension Index (UDI) is defined (eq. 3.4):

$$UDI_i = \sum_{i=1}^{n} w_i \cdot um_i \qquad \text{(eq. 3.4)}$$

where w_i is the weight of each usability measure, that could be different, based on the level of priority of usability measures in the specific application. The three usability dimension indexes are: 1) the Effectiveness index; 2) the Efficiency index; 3) the Satisfaction index.

The weighted sum of these three indexes provides the overall results for the UI (eq. 3.5):

$$UI = \sum_{i=1}^{n} w_i \cdot UDI_i \qquad \text{(eq. 3.5)}$$

In details, the AHP is applied in order to evaluate the relevance of the factors in the hierarchy, taking into account the analysis of GUI interaction. Starting from the hierarchy structure of the model, the matrix of weights is defined. Such matrix is accomplished for each level of the hierarchy and for each group (elements in the lower level hinge on the same element in the upper level), by placing the elements of the group both on matrix rows and columns. Hence, all the elements of the same group are compared in pairs. The generic matrix element a_{ij} is the result of the pairwise comparison between the attribute of the row *i-th* and the column *j-th*, with respect to a certain task, using the Saaty scale *i.e.* a 9-points scale anchored at the end with the terms *"Equivalent alternatives"* and *"The chosen alternative is absolutely better than the other one"*. Thus, the main diagonal of the matrix consists of unit elements only, while the values of the other cells are always positive, according to the reciprocity property (eq. 3.6):

$$a_{ij} = \frac{1}{a_{ij}} \qquad \text{(eq. 3.6)}$$

Once the pairs comparison matrix is defined, the weight of each element is assumed as (eq. 3.7):

$$w_i = \frac{\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}}; \quad i, j = [1, n] \qquad \text{(eq. 3.7)}$$

In eq. 3.7, *n* is the dimension of the metrics related to the element at issue. In particular, the allocation of weights is done with a bottom-up logic, from the lowest level of the hierarchy (UFs) to the highest (Usability).

3.4 Experimental phase

3.4.1 Overview of experiment

Based on of the requirements identified in section 3.3.1, participants are 12 newly graduated engineers (*i.e.* mechanical, electrical and management) aged between 27-32 years, attending a specialised course in Computer-Aided Design within a project named *Digital Pattern Product Development*.

The experiments are performed at the Fraunhofer JL IDEAS-COGITO laboratory, Department of Industrial Engineering of the University of Naples (IT) in a suitable room, with no furniture and equipped with a Visual Display Unit (VDU). Preliminarily, a GUI *Tutorial* is defined to present the graphical interface, to explain the procedures for data entry and to discuss about the functions of the interface. An example is also illustrated.

An experimental session is performed. In such session, each user has to complete two specific tasks: 1) to design an automotive manual transverse gearbox; 2) to modify an existing gearbox model re-using previous knowledge.

The goals of the Task 1 are the follows:

- to design a new gearbox according to the specifications assigned (i.e. the parameters of six gears, of differential and of reverse);
- to plot the gearbox designed;
- to assess the overall dimensions;
- to save the model;
- to export the model.

Whereas, the goals of the Task 2 are the follows:

- to modify the gearbox designed in the first task, according to new instructions (*i.e.* it was asked to change the some parameters of the gears and of the layout controller);
- to plot the gearbox modified;
- to assess the overall dimensions;
- to save the model;
- to export the model.

Therefore, the measures of subjective ratings (TC and PSR) and objective metrics (NO, T and NE) are obtained.

3.4.2 Experimental protocol

Several days before the test session, a preparatory meeting with the participation of users involved in the tests is accomplished. The purpose of the incoming experimentation and the functionality of GUI are presented. A detailed description of the *Tutorial* is given.

On the day fixed for the tests and before starting, users are informed, once again, that the aim of the experiment is to evaluate the GUI usability, and not the user's ability to quickly perform a set of assigned tasks. In this way, we try to minimize the "stress" that, generally, may affect the outcome of a proof. The inspectors show the procedures of the experimental session, with particular attention to the rules of test performing. Then, they provide further details about the *Tutorial* and they administrate the short questionnaire for the personal details and for the informed agreement to users. The questionnaire is filled and returned before the start of the test. Finally, an ID code to each user is assigned.

Users test the GUI individually and in random order to avoid noise factors. During the test, the inspectors record many details: the start and the ending time of the tasks, any notes on the bringing of the test, the number and kind of assistance provided. Specifically, if the user explicitly required the assistance, then the inspector invites him to consult the *Tutorial* (classifying this as a *level 1* assistance). Otherwise, if the user was not able to continue the test, the inspector removes all doubts (classifying this as a *level 2* assistance). The time limit for each test is set at 30 minutes after which the user is asked to suspend the operations (Sauro and Kindlund, 2005b). After completing both tasks the questionnaire is administered to each user for detecting the PSR measurement.

The procedures described above is also applied to the validation test.

3.4.3 Data collection and processing

All the UFs measures are collected. Table 3.2 (on the next page) summarises the sources related to UFs. In particular, an open source software is used to record all user's activities carried out during the experimental phase. Such tool is used to collect the NE, TC, NO and T metrics. The Effectiveness metrics are described in the following.

Table 3.2: The sources of UFs.

UDs	UFs	Source
Effectiveness	Number of errors (NE)	Video test
	Task Completion (TC)	Panel of experts
Efficiency	Number of operations (NO)	Video test
	Time (T)	
Satisfaction	Post-Session Ratings (PSR)	Questionnaire

Effectiveness metrics:

- The *Number of Errors* (NE) are derived from the video by counting each time the GUI reports an error message.
- The *Task Completion* (TC) are measured using a rating given by a panel of experts who are asked to assess the completeness of the goals reached for all the activities performed in the test, by using the following six-point scale: (1) Complete success without assistance, (2) Complete success with assistance, (3) Partial success without assistance, (4) Partial success with assistance, (5) Failure: the user does not understand that the task is not complete, (6) Failure: the user does not complete the task despite the assistance. The references to determine the level of completion in task execution is decided beforehand (Tullis and Albert, 2008).

Efficiency metrics:

- The *Number of Operations* (NO) are derived from video by counting, from time to time, the operations that are performed to complete the task.
- The *Time* (T) is measured by the inspector as the difference between the ending and the beginning time of the session. This measure is subsequently validated by a comparison with the clock of VDU shown in video recordings.

Satisfaction metrics:

- the *Post-Session Ratings* (PSR) are gained from the specific questionnaire that users filled out at the end of the session test (*i.e.* both Task 1 and Task 2). In particular, they are asked to express their agreement related to ten statements, all set in a positive sense, by using a seven-point scale, whose ends were the positions: *"strongly agree"* and *"strongly disagree"*.

In the calculation phase, the total value of each UF is obtained as the sum of the measures/ratings respectively noted to perform both the Task 1 and Task 2, for the same user. This operation is repeated for all users. For the calculation of the UI, the average values (arithmetic mean) of all aforementioned UFs are used.

The procedures described above to collect and process data are also applied to the validation test.

The experiment in numbers: 1 laboratory was used; 1 usability team consisting of 6 engineers, 1 panel of experts and 12 end-users were involved; 2 usability testing sessions were performed; over 4 hours of video footage and ca. 200 questions were examined.

3.5 Results and discussion

The following data are obtained. Table 3.3 shows the normalized measures of UFs for each user. According to the above UI definition (section 3.2.3), the average values of UFs are obtained using the collected measurements. The weights (w) of UFs are obtained submitting Saaty's questionnaire to a panel of experts and using the Equation 3.7, as summarised in Table 3.4.

Table 3.3: Normalized measures of UFs.

	Effectiveness		Efficiency		Satisfaction
ID user	Number Errors (NE)	Task Completion (TC)	Number of Operations (NO)	Time (T)	Post Session Ratings (PSR)
S.1	0.50	0.25	0.84	0.43	0.56
S.2	0.20	0.33	0.80	0.57	0.95
S.3	0.10	0.33	0.47	0.36	1.00
S.4	0.17	0.33	0.75	0.37	0.87
S.5	0.25	0.40	0.77	0.49	0.82
S.6	0.25	0.40	0.70	0.46	0.93
S.7	0.11	0.40	0.53	0.45	0.46
S.8	0.08	0.33	0.49	0.24	0.46
S.9	0.20	0.40	0.63	0.32	0.90
S.10	0.10	0.50	0.78	0.43	0.72
S.11	0.33	1.00	0.85	0.58	0.92
S.12	1.00	0.40	0.64	0.32	0.84

Table 3.4: Weights and values of UFs.

	NE	TC	NO	T	PSR
w_i	0.74	0.26	0.25	0.75	1.00
um_i	0.27	0.42	0.69	0.42	0.79

Likewise, the weights and values of UDs are obtained (Table 3.5 on next page). So, it is calculated the usability index: the value obtained is equal to 0.42, but this is not acceptable.

Table 3.5: Weights and values of UDs.

	Effectiveness	Efficiency	Satisfaction
w_i	0.59	0.31	0.10
UDI_i	0.31	0.49	0.79

The Table 3.5 shows that the Satisfaction index ($UDI_{satisfaction}$) is the highest (79%). However, this value could be smaller as the degree of Satisfaction of the users could be influenced by the achievement of the goal (*i.e.* the effective gearbox modeling) rather than the difficulties they overcome in using interface. In this case, the Satisfaction is the usability dimension that has a "reduced" effect (10%, see Table 3.5) on the calculation of the global index. Hence, results in Table 3.5 suggest that the primary strategy for improving the usability of the GUI is to increase in Effectiveness values by acting mainly on the usability functions *Number of Errors* (NE). Whereas achieving a higher Efficiency value, by leveraging on the usability functions *Time*, may be the second strategy to improve the GUI usability.

Furthermore, both Task 1 and Task 2 are divided into the following critical sub-activities: the choice of the gearbox architecture, the choice of the number of gears, the setting of the dependent parameters, the setting of the wheel parameters, the setting of the reverse gear, the setting of the parameters of *differential*, the overall dimensions, the procedure for file exporting. In this perspective, the measurement of Efficiency is analysed.

Figure 3.4 (on the next page) shows the radar chart that highlights how the normalized average value, related to the number of operations due for each sub-task, departs from the normalized optimal value (equal to 1). For example the value 3, related to one of the axes, means that, on average, the number of operations necessary to accomplish that specific subtask is three times higher than the ideal value. Figure 3.4 points out that the more critical sub-tasks, involved in the Task 1, are (in descending order): the setting of the parameters of the differential (6.10), the setting of the reverse gear (5.29), the file exporting (4.42), the setting of the gear parameters (2.43). Further results for Task 2 are: the setting of the reverse gear (4.17) and setting of the differential parameters (2.35).

Time is another critical UF. It's worthwhile to note that if we consider only the users who complete the tasks with success and without assistance, the average *Time* recorded is almost double than the predetermined optimum value. More generally, the average additional time to complete the Task 1 is much greater than the one related to the Task 2 as well as the variability of the measures (Figure 3.5 on the next page). This may indicate a good level of learnability of the GUI.

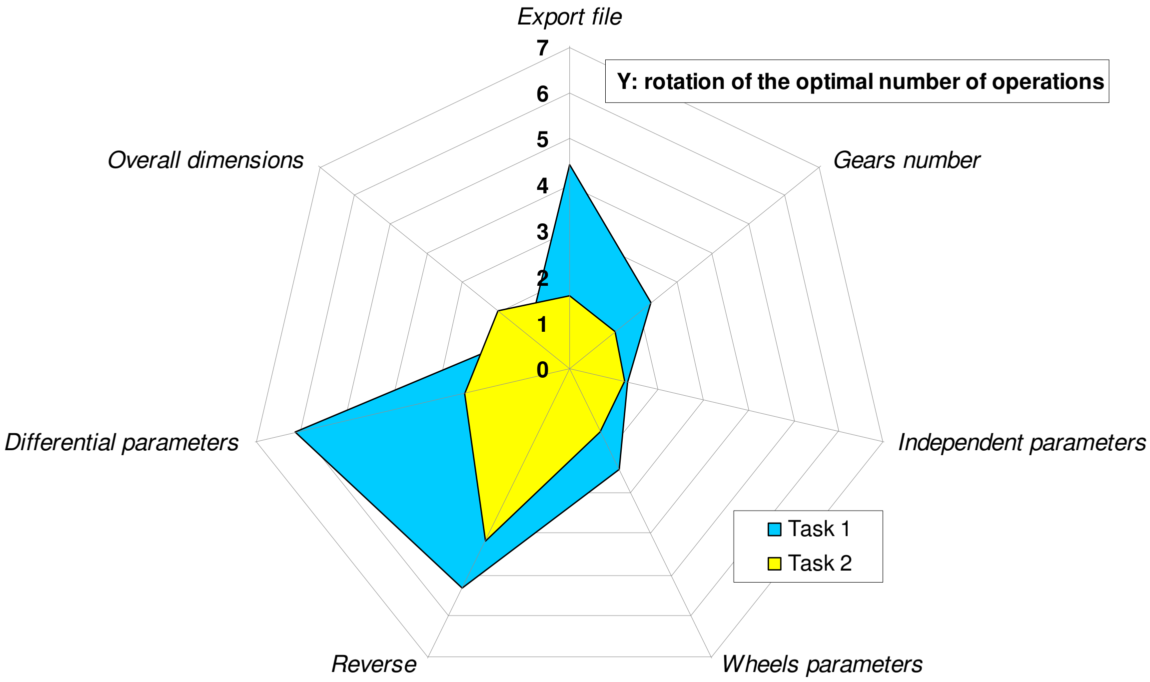


Figure 3.4: Number of Operations related to each sub-activity.

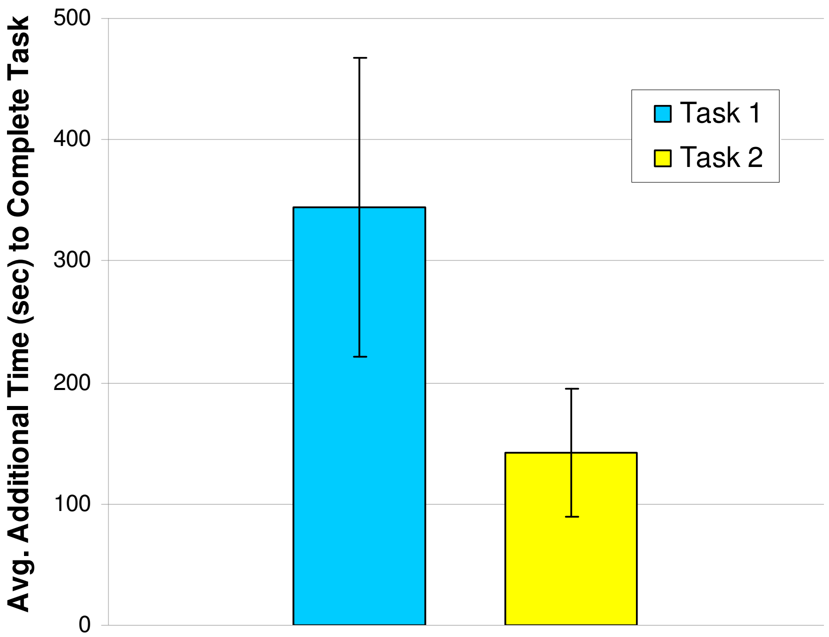


Figure 3.5: Average Additional Time to complete Task 1 and Task 2.

Tackling the measures of Effectiveness, it's worthwhile to note that there are no significant differences between the number of errors related to the Task 1 and Task 2, but the average values are not negligible (Figure 3.6 on the next page).

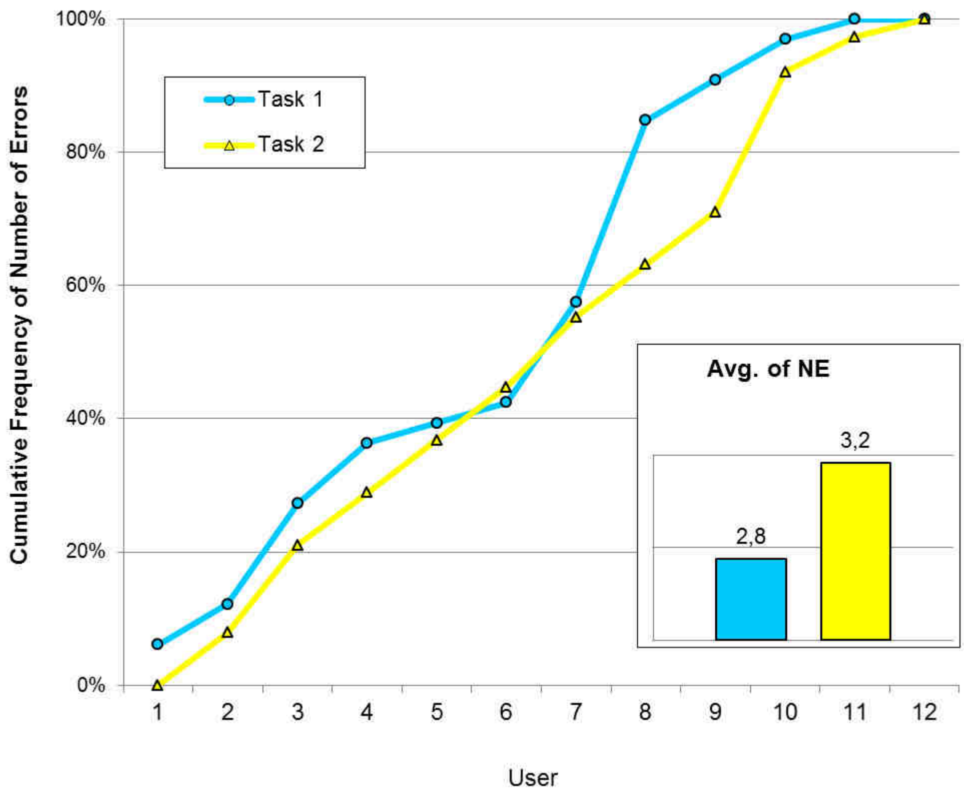


Figure 3.6: Cumulative frequency of number of errors related to each task and average values.

The values of *Task Completion*, grouped using the level of success, are depicted in Figure 3.7. In particular, most users carry out the Task 1 in a complete success. Otherwise, in the accomplishment of Task 2, only 1 user completely achieves the goals (complete success), while 10 users get a partial success. There is also 1 user who fails.

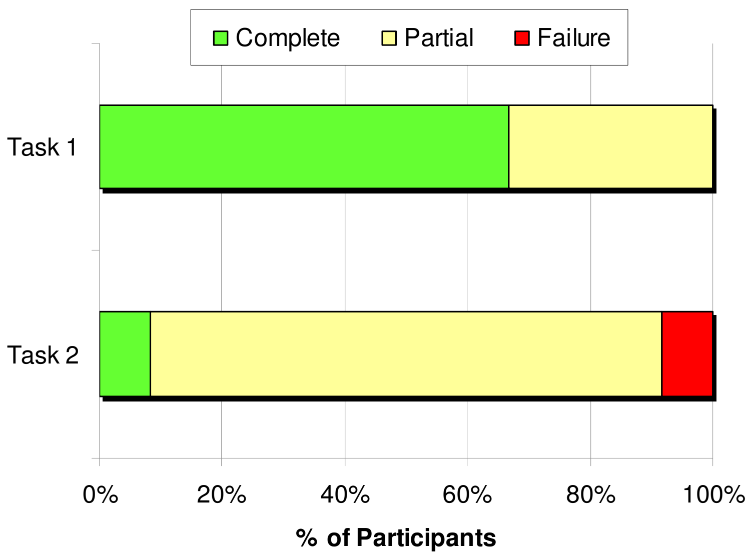


Figure 3.7: Stacked bar chart showing different levels of success based on task completion.

The measures of Satisfaction is analysed. Figure 3.8 shows that the lowest value of the Satisfaction is related to the clarity and effectiveness of the GUI (D8), while the highest value is related to the actual benefit perceived in the use of the GUI, during the improvement of the gearbox design (D5).

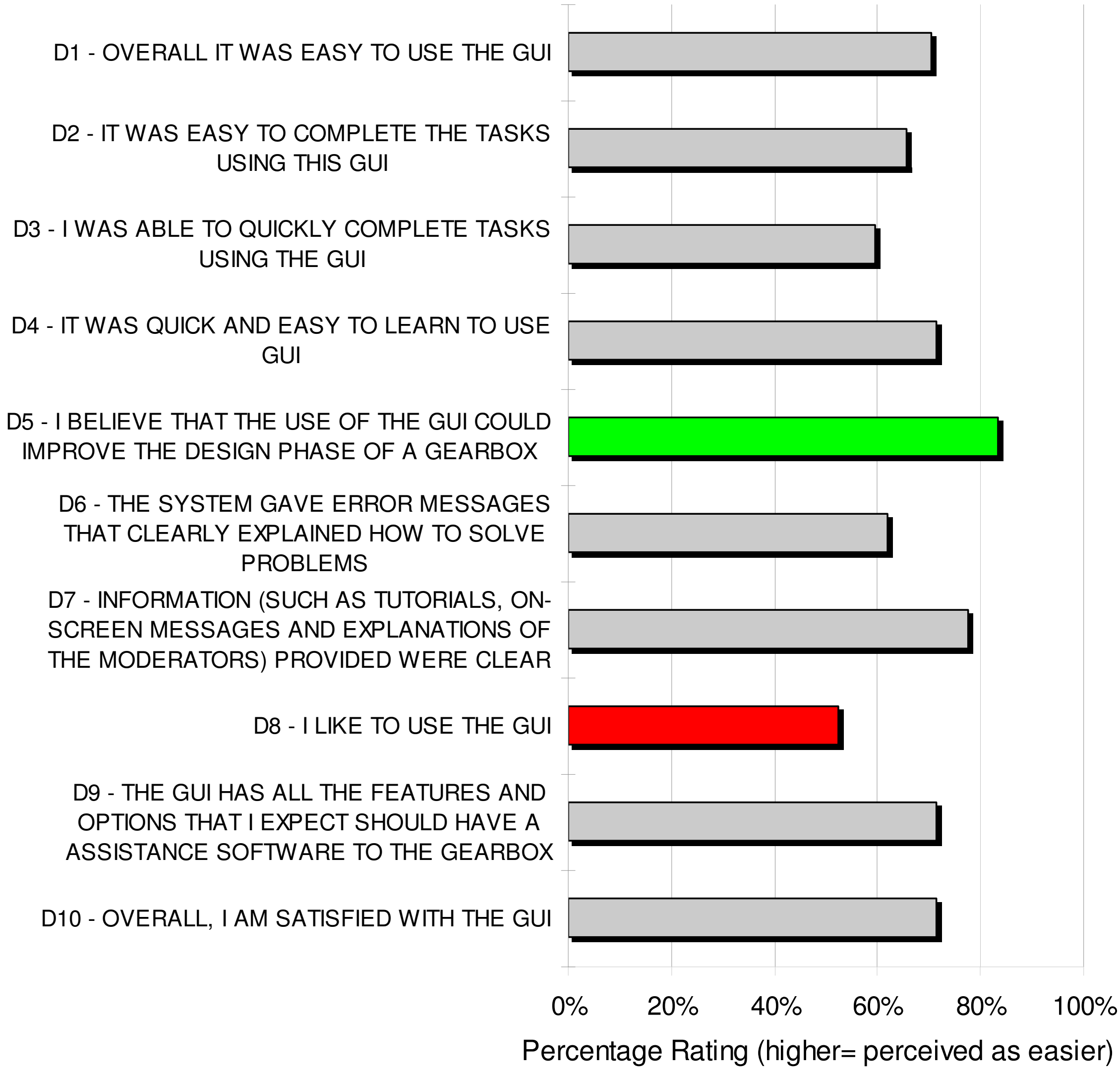


Figure 3.8: Average subjective ratings split by statement.

Some frequent critical issues are identified by analysing videos related to the tests (Table 3.6 on the next page). These problems involve difficulties in achieving the Tasks and they generally cause an increase of the operating time also due to a more than proportional increasing number of operations to be performed. In some cases, the user is confused and, then, she/he is led to an error or makes continuous action controls.

Table 3.6: Root causes and corrective actions of the critical issues identified.

#	Root Cause	Corrective Action	UD
1	The default fields are not empty	The default fields are empty	Efficiency
2	The interface does not allow you to overwrite the selected values in the fields, you are forced to cancel the existing	It's possible to overwrite the selected values	Efficiency
3	Poor visibility of function for exporting the file	A new visibility was given to the button to export the files	Efficiency
4	Poor functionality of the reset function	A new reset function was upgraded	Effectiveness
5	The user does not have a feedback on the correct setting of the parameters	The button on the control panel is divided and a section with the new (TEST) or upgraded (RESET) functions, and a new confirmation command (SET) are inserted. The latter turns on only when the input parameters are correct. In this way it provides an immediate feedback to the user.	Effectiveness
6	Poor visibility of the zoom function	A new visibility is given to the button aimed to Zoom	Efficiency
7	Inconsistency of the provision of the sections in the main window	A new provision of sections, with different background colours, that promotes the logical procedure for the input of project data is introduced	Effectiveness

Once these critical issues are identified, the GUI is re-designed. However, in order to avoid radical changes and with the aim to improve UD Effectiveness and Efficiency, the new GUI is developed but it keeps the initial sizes . So, all corrective actions listed in Table 3.6 are considered.

In order to confirm the validity of the new features a pairwise comparison between them is performed. For each upgrade, users are asked to rate the GUI with the new functions. They express their degree of preference on a scale of six points. The results of survey demonstrate a preference of the users for all new functions far higher than the initial ones.

Hence, all the new features are definitively implemented and a new release of GUI is developed (Figure 3.9 on the next page).

Finally, the validation test is accomplished in order to assess again the UI, according a continuous improvement loop.

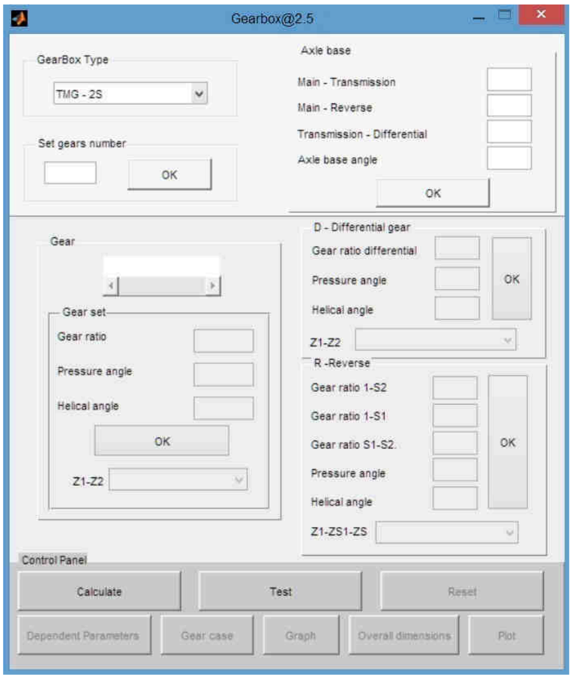


Figure 3.9: The new release of GUI.

Five users who tested the GUI in the first test are involved in the new validation test. Similarly to the first experimentation, the UI is calculated. In particular, the same weights of both UFs and UDs are used. In particular, the normalization measures of UFs are depicted (Table 3.7) and the UFs and UDs values are obtained (Tables 3.8-3.9 respectively).

Table 3.7: Normalized measures of UFs in the validation test.

	Effectiveness		Efficiency		Satisfaction
ID user	Number Errors (NE)	Task Completion (TC)	Number of Operations (NO)	Time (T)	Post Session Ratings (PSR)
VT.1	0.25	0.29	0.60	0.58	0.87
VT.2	0.20	0.67	0.88	0.57	1.00
VT.3	1.00	0.65	0.67	0.65	0.93
VT.4	1.00	0.76	0.87	0.76	0.94
VT.5	0.33	0.83	0.94	0.83	0.76

Table 3.8: Weights and values of UFs in the validation test.

	NE	TC	NO	T	PSR
w_i	0.74	0.26	0.25	0.75	1.00
um_i	0.56	0.57	0.79	0.68	0.90

Table 3.9: Weights and values of UDs in the validation test.

	Effectiveness	Efficiency	Satisfaction
w_i	0.59	0.31	0.10
UDI_i	0.56	0.76	0.90

The UI obtained is equal to 0.66. The overall GUI usability improvement is of 57%. For each usability dimension, the following percentage changes are registered: the Effectiveness UD increases of 81%, the Efficiency UD of 56% and the Satisfaction UD of 14%.

It's worthwhile to note the following improvements related to Effectiveness UD: the average NE decreases of 53%; the percentage of users able to complete with success the session test increases of 23% while those who are able to partially complete decreases of 28%. These results have a positive effects on TC measurement.

As regard the Efficiency UD, we highlight that the average T decreases of 22%. In particular, T decreases of 6% in Task 1, while decreases of 41% in Task 2.

Also the Satisfaction UD increases. A further investigation is carried out. Considering the paired data, matched samples, Wilcoxon-signed-rank test is used (Wilcoxon, 1945) to determine whether there is a significant difference between the average values of the PSR made under two different conditions (*i.e.* GUI before and after the changes). Both PSR measurements are made on each unit in a sample, and the test is based on the paired differences between these two values. The null hypothesis is the difference in the mean values is zero. Because the p-value is low (<7%), we can be assume that the changes have produced a significant effect on GUI usability.

3.6 Conclusions

The present study tackles the usability assessment of a GUI that is a part of a KBE system. To this aim, starting from a method for usability assessment successfully applied to a new product proposed by Di Gironimo *et al.* (2013), a new approach to evaluate the usability of a GUI is discussed. A new usability index (UI) is proposed based on AHP model and its use is validated thanks to experimental results. In particular, the experimental data, leading to a lower value for UI (0.42), are collected and discussed. Then, taking into account such experimental data, a new release of the GUI is proposed and a new set of experimentations are carried out in order to validate the new release. According to the validation test, the UI achieves the value of 0.66 *i.e.* it shows an increase equal to 57%. Such improvement induces to state that the use of the new release of the GUI could improve the KBE system and contribute to reduce the development time of gearboxes.

Further steps deal with evaluation and improvement of the new GUI. In fact, by using the present approach, new characteristics of GUI are discovered during the experimental sessions and could be introduced and evaluated, in iterative way.

3.7 References

Al-Awar, J., Chapanis, A., Ford, R., 1981. Tutorials for the first-time computer user. IEEE Transactions on Professional Communication, 24(1), 30–37.

Bevan, N., Barnum, C., Cockton, G., Nielsen, J., Spool, J., Wixon, D., 2003. The "magic number 5": is it enough for web testing?. In CHI '03 Extended Abstracts on Human Factors in Computing Systems (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). ACM, New York, NY, 698–699

Bourke, D., 2013. Software gives engineers quick access to product info from layers of company files. Mechanical Engineering, 44–47.

Caulton, D.A., 2001. Relaxing the homogeneity assumption in usability testing. Behaviour & Information Technology, 20(1), 1–7.

Di Gironimo, G., Matrone, G., Tarallo, A, Trotta, M, Lanzotti, A., 2013. A virtual reality approach for usability assessment: case study on a wheelchair-mounted robot manipulator. Engineering with Computers, 29(3), 359–373.

El Hani, M.A., Rivest, L., Maranzana, R., 2012. Product data reuse in product development: A practitioner's perspective. In PLM, proceeding of IFIP Advances in Information and Communication Technology, 388, 243–256.

Elgueder, J., Cochennec, F., Roucoules, L., Rouhaud, E., 2010. Product–process interface for manufacturing data management as a support for DFM and virtual manufacturing. International Journal on Interactive Design and Manufacturing, 4(4), 251–258.

Figueira, J., Greco, S., Ehrgott, M., 2005. Multiple Criteria Decision Analysis: State of the Art Surveys. Springer.

Ghosh, R., Dekhil, M., 2009. Discovering user profiles. In WWW '09: Proceedings of the 18th international conference on World wide web (New York, NY), ACM, 1233–1234.

Hertzum, M., Jacobsen, N.J., 2003 – corrected version (original published in 2001). The evaluator effect: A chilling fact about usability evaluation methods. International Journal of Human-Computer Interaction, 15(1), 183–204.

Hornbaek, K., 2006. Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer studies, 64(2), 79–102. doi:10.1016/j.ijhcs.2005.06.002

ISO 9241-11, 1998. Ergonomic requirements for office work with visual display terminals (VDTs) – Guidance on usability.

Kim, J., Han, S., 2008. A methodology for developing a usability index of consumer electronic products. International Journal of Industrial Ergonomics, 38(3-4), 333–345.

Lanzotti, A., Patalano, S., Vitolo, F., 2013. A graph-based approach to CAD modeling: a digital pattern application to the sizing and modeling of manual transverse gearboxes. In International Conference on Graphic Engineering, Madrid (Spain).

Lewis, J.R., 1982. Testing Small System Customer Setup. in Proceedings of the Human Factors Society 26th Annual Meeting, Santa Monica, CA, USA, SAGE Publications, vol. 26(8), 718–720.

Lewis, J.R., 1993. Problem discovery in usability studies: A model based on the binomial probability formula. In Proceedings of the Fifth International Conference on Human-Computer Interaction. Orlando, FL, USA, Elsevier, 666–671.

Lewis, J.R., 1994. Sample sizes for usability studies: Additional considerations. Human Factors, 36(2), 368–378.

Lewis, J.R., 2001. Evaluation of procedures for adjusting problem-discovery rates estimated from small samples. International Journal of Human-Computer Interaction, 13(4), 445–479.

Lewis, J.R., 2006. Sample sizes for usability tests: mostly math, not magic. Interactions, 13(6), 29–33.

Lindgaard, G., Chattratichart, J., 2007. Usability testing: what have we overlooked?. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, USA, April 28 - May 03, 2007. CHI '07. ACM, New York, NY, 1415–1424.

Madhavan, R., Alagarsamy, K., 2013. Usability issues in software development lifecycle. International Journal of Advanced Research in Computer Science and Software Engineering, 3(8), 1331–1335.

Nielsen, J., 1993. Usability Engineering. Morgan Kaufmann Publishers Inc., San Francisco, CA (USA). ISBN 0-12-518406-9.

Nielsen, J., 2000. Why You Only Need to Test with 5 Users. Jakob Nielsen's Alertbox. Available on line from www.useit.com/alertbox/20000319.html.

Nielsen, J., Landauer, T.K., 1993. A mathematical model of the finding of usability problems. In Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems, ACM, 206–213

Norman, D., 2004. Emotional Design. Perseus Publishing, New York, NY (USA).

Patalano, S., Vitolo, F., Lanzotti, A., 2013. A graph-based software tool for the CAD modelling of mechanical assemblies. In Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization, 60–69, Barcelona (Spain). doi:10.5220/0004299000600069

Perfetti, C., Landesman, L., 2001. Eight is not enough. User Interface Engineering.

Saaty, L., 2008. Decision making with the analytic hierarchy process. International Journal of Services Sciences, 1(1), 83–98.

Sandeberg, M., 2003. Knowledge based engineering – in product development. ISSN: 1402-1536.

Sauro, J., 2010. Why you only need to test with five users (explained). Measuring Usability.

Sauro, J., Kindlund, E., 2005a. A method to standardize usability metrics into single score. Proceedings of the SIGCHI conference on Human factors in computing system, Portland, 401–409.

Sauro, J., Kindlund, E., 2005b. How Long Should a Task Take? Identifying Specification Limits for Task Times in Usability Tests. Proceeding of the Human Computer Interaction Conference, Las Vegas (USA).

Schmettow, M., 2008. Heterogeneity in the Usability Evaluation Process. In Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction. ACM, Liverpool, UK, Vol. 1, 89–98.

Sharmin, M., Bailey, B.P., Coats C., Hamilton K., 2009. Understanding knowledge management practices for early design activity and its implications for reuse. In Proceedings of the 27th International Conference on Human Factors in Computing Systems, 2367–2376.

Sohaib, O., Khan, K., 2010. Integrating usability engineering and agile software development. In Proceedings of the International Conference on Computer Design and Application, volume 2, 32–38.

Spool, J., Schroeder, W., 2001. Testing web sites: five users is nowhere near enough, CHI '01 extended abstracts on Human factors in computing systems, March 31-April 05, Seattle, Washington, USA, ACM, 285–286.

Stenzel, I., Pourroy, F., 2008. Integration of experimental and computational analysis in the product development and proposals for the sharing of technical knowledge. International Journal on Interactive Design and Manufacturing, 2(1), 1–8.

Tullis, T., Albert, B., 2008. Measuring the user Experience. Morgan Kaufmann Publisher.

Turner, C.W., Lewis, J.R., Nielsen, J., 2002. Current issues in the determination of usability test sample size: how many users is enough?. In Usability Professionals' Association Conference Proceedings.

Turner, C.W., Lewis, J.R., Nielsen, J., 2006. Determining usability test sample size. In W. Karwowski (ed.), International Encyclopedia of Ergonomics and Human Factors, 3, 3084–3088

Virzi, R.A., 1990. Streamlining the design process: running fewer subjects. Proceedings of the Human Factors Society 34th Annual Meeting, Santa Monica, CA, USA, SAGE Publications, vol. 34(4), 291–294.

Virzi, R.A., 1992. Refining the test phase of usability evaluation: How many subjects is enough? Human Factors, 34(4), 457–468.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biometrics Bulletin, 1(6), 80–83.

Wixon, D., 2003. Evaluating usability methods: why the current literature fails the practitioner, interactions, 10(4), 28–34.

Woolrych, A., Cockton, G., 2001. Why and when five test users aren't enough. In Proceedings of IHM-HCI 2001 Conference, Toulouse, France: Cépadèus Éditions, Vol. 2, 105–108.

Wright, P.C., Monk, A.F., 1991. A cost-effective evaluation method for use by designers. International Journal of Man-Machine Studies, 35(6), 891–912.

Appended Papers

Paper A

Lanzotti A., Vanacore A., Del Giudice D.M., Proceedings of Joint Conference on Mechanical, Design Engineering & Advanced Manufacturing 2014, Paper n. 72, Toulouse (France), June 18th–20th 2014, 7 pp. It's waiting to be published on Research in Interactive Design Vol. 4 book by Springer Verlag (ISBN available soon)

Original Article Proceedings of Joint Conference on Mechanical, Design Engineering & Advanced Manufacturing, Toulouse, France, June 18th–20th, 2014

SEAT DESIGN IMPROVEMENT VIA COMFORT INDEXES BASED ON INTERFACE PRESSURE DATA

A. Lanzotti[1], A. Vanacore[1], D. M. Del Giudice[1]

(1) : University of Naples Federico II, Department of Industrial Engineering
P.le Tecchio, 80, 80125, Naples (NA), Italy
Phone +39-0817682506 – Fax +39-0817682187
E-mail : {antonio.lanzotti, amalia.vanacore, domenicomaria.delgiudice}@unina.it

Abstract: Literature on seat comfort recognizes that seat interface pressures are the objective comfort measures that most clearly relate to users' comfort perceptions about sitting experience. In this paper, the above relationship is quantitatively investigated by performing simple but effective explorative analyses on seat comfort data collected during experimental sessions involving 22 volunteers who tested 4 office chairs (differing in terms of cushion softness). Statistical data analyses show that subjective sitting comfort/discomfort ratings are significantly related to several combinations of pressure variables. The joint analysis of synthetic indexes based on seat interface pressures reveals to be a useful tool for comparative seat comfort assessment. Besides valuable suggestions for the definition of an effective strategy for seat comfort assessment, the results of data analyses provide useful information to support the product design phase. In fact, the sitting experience results to be significantly improved by: (1) a balancing of pressures between the bilateral buttocks; and (2) a balancing of contact areas between buttocks and thighs.

Key words: office chair; sitting comfort/discomfort assessment; interface pressure distribution.

1- Introduction

Research in the field of medicine and epidemiology has shown that, over the past decades, the incidence of backache has considerably increased (Harkness *et al.* 2005; Rubin 2007) due to sedentary lifestyle, closely related to prolonged period of sitting (Ehrlich 2003; Dul *et al.* 1987). More than 60% of people experience have at least one episode of lower back pain at work, in almost 45% of cases the first attack of lower back pain happens while working (Rezaee *et al.* 2011), with an incidence in the office workers of at least one episode backache every 3 years (Lengsfeld *et al.* 2000). The remedy that is most useful to prevent backache is the adoption of ergonomic chairs (Nelson *et al.* 1998, Herbert 2001, Loisel *et al.* 2001). Given the importance of ergonomic seat and prevention, this study aims to investigate the biomechanical aspects and the comfort evaluation of ergonomic office chairs, which are necessary steps to identify guidelines for designers to improve human wellness.

Specialized literature does not provide a universally recognized definition of comfort, but in recent years the assumption that comfort and discomfort are two distinct entities is winning broad respect (Vink 2012). Typically comfort assessment is realized on the basis of subjective evaluations and/or postural analysis.

Subjective evaluations are collected by surveying potential seat users who are asked to express their feelings of comfort/discomfort with the seat and/or compare, in terms of perceived comfort/discomfort, similar seats.

Postural analysis is realized by measuring one or more objective parameters, such as:

- the pattern of muscle activation measured through electromyography (EMG);
- the stress acting on the spine measured through pressure transducer and radio waves;
- the postural angles obtained using contact or non contact (like photogrammetric) techniques in real experiments or using virtual manikins in virtual experiments;
- the body–seat interface pressure measured through capacitative or resistive mats.

Anyway, subjective and objective methods are not alternative since they complement each other.

One of the main factors that affect seat comfort is seat-interface pressure distribution (Stinson *et al.* 2009). Moreover, pressure distribution is the objective measure with the clearest correlation with the subjective evaluation methods (De Looze *et al.* 2003, Kyung *et al.* 2008, Noro *et al.* 2012). In particular, in office chair design (Reed *et al.* 1993) pressure maps have been used to qualitatively verify the effectiveness on seat comfort of product features like, *e.g.*, cushion shape and materials (Kamijo *et al.* 1982, Park *et al.* 2000, Fujimaki *et al.* 2002) through correlation studies with the subjective user perceptions. Nevertheless the widespread use of pressure maps, just few authors have proposed synthetic indexes for the related multidimensional data, collected by performing real or virtual experiments involving a selected sample of potential users (Lanzotti *et al.*

2011). Furthermore, little effort has been made to highlight the usefulness of these pressure measures for specific purposes defined by designers, (*e.g*. design for a specific user or design for a generic user).

In this paper the main results of an extensive explorative analysis on seat interface pressure data are described. The explorative data analysis was aimed at investigating three critical aspects of seat comfort assessment: a) gender-based differences in seat interface pressure distribution; b) the relationship between subjective and objective measures of seat comfort; c) discriminant effectiveness of indexes based on seat interface pressure.

2- Methods

2.1 - Overview of experiment

The data analysed in this paper were obtained from a study aimed to validate a seat comfort index proposed by the authors. All details on the criteria for selection of participants and office chairs, the subjective rating scales and the experimental design have been provided in a previous paper (Lanzotti *et al*. 2011) only a synthetic description is given here.

Each of 22 volunteers [age 20-31, 14 males and 8 females, mean (SD) mass = 75.0 (12.2) kg, mean (SD) stature = 175.3 (11.6) cm] participated in short-term experimental sessions for the comfort evaluation of 4 office chairs.

Total duration of an experimental session (with 1 volunteer testing 1 seat) was 10-15 minutes including few minutes (≤5) for initial seat and posture adjustments and 10 minutes performing the task of reading a test on VDU. The choice of short-term experimental session is recommended when using pressure data for assessment of sitting comfort/discomfort (Helander *et al*. 1997; Kyung *et al*. 2008). In contrast, more extended durations are generally used when investigating sitting discomfort largely due to fatigue in seated postures not depending on chair design.

The 4 office chairs had a typical architecture of product (*i.e*., a five-pointed base, a backrest and two armrests) but differed in shapes and materials of the cushion.

In particular, different Seat Conditions were represented by the characteristic softness of the seat cushion (S) considered as a qualitative ordinal variable with four levels (0, soft; 1, medium; 2, compact; 3, semi-rigid). Each seat was representative of a specific Seat Condition level.

During each experimental session subjective measures of comfort perception as well as seat interface pressures were collected.

The comfort/discomfort ratings were based on a verbal numeric scale, with the comfort and discomfort at the extremes, thereby measuring a mixture of comfort and discomfort. In particular, for this assessment the scale Borg CR10 modified by Kyung *et al*. (2008) was used. This scale includes scores from 0 (no comfort, and maximum discomfort) to 10 (maximum comfort and minimum discomfort). Besides, in order to assess comfort and discomfort separately, another set of subjective measures was collect using two different scales. A verbal rating scale with four levels was used to collect data about perceived comfort, whereas to assess discomfort participants were asked to rank the chairs based on perceived discomfort.

Objective measures were obtained from pressure measured at the seat interface; these measures consisted of both overall and local pressures (Table 1).

Type	Name	Area to measure
Objective	PCP, Peak Contact Pressure (N/cm^2) CP, Contact Pressure (N/cm^2) CA, Contact Area (cm^2) UW, Unloaded Weight (kg)	• Left/right thighs (TL/TR) • Left/right buttocks (BL/BR)
	PCL, Press. Comfort Loss Index	• Sum of 4 local body part press.
Subjective	RT, Comfort/Discomfort Rating	• Whole body
	CR, Chairs Ranking	• Whole body
	CD, Comfort Degree	• Whole body

Table 1: Comfort Variables

2.2 - Experimental protocol

The experiments were performed at the Department of Industrial Engineering, University of Naples Federico II, in a suitable room cleared of furnishings.

For each session, a pressure mat was placed on the seat cushion and secured with masking tape to facilitate seat adjustments. The participants were instructed to sit carefully to minimize wrinkles on the pressure mat. In order to avoid the noise due to the sequence of the tested seats, participants tested the office chairs following a randomized experimental plan. Besides, in order to avoid that visual impact with the tested chair could affect comfort/discomfort assessments, each participant was introduced into the room blindfolded and made to sit. Subsequently, she/he was asked to take off the blindfold and adjust the chair in such a way that the legs were in rest conditions and the feet were comfortably on the floor so as to form an angle between the thigh and the leg equal to 90°. Thus, after an initial seat and posture adjustments, the participant had to read a text on a VDU for 10 minutes, after which it was blindfolded again and taken back out of the room.

2.3 - Data collection and processing

Subjective ratings were collected in a consistent order to minimize confusion: the *"Comfort/discomfort Rating (RT)"* and the *"Comfort Degree (CD)"* were obtained using a questionnaire immediately after each session; instead, the "*Chairs Ranking (CR)*" was obtained from a questionnaire submitted only after the participant had tested all four office chairs.

Pressure data were collected continuously during the reading text on VDU, using a Novel Gmbh (Munich, Germany) pressure mat (S2027 Pliance^{TM}). The pressure mat comprised 256 (16x16) thin (<1.2 mm) capacitive sensors that could easily conform to the contour of the seat, and measure pressures typically in a range from 0.2 N/cm^2 up to 6 N/cm^2. Thanks to its flexible structure the mat is a minimally invasive instrument, which does not interfere with user

Variable	Description	Measurement Unit
caTL (caTR)	Average contact area Thigh Left (Right)	cm^2
caBL (caBR)	Average contact area Buttock Left (Right)	cm^2
caSUM (caTL + caTR + caBL + caBR)	Sum of average contact areas	cm^2
caTL/caSUM, caTR/caSUM, caBL/caSUM, caBR/caSUM	Relative Average contact areas	
cpTL (cpTR)	Average contact pressure Thigh Left (Right)	N/cm^2
cpBL (cpBR)	Average contact pressure Buttock Left (Right)	N/cm^2
cpSUM (cpTL + cpTR + cpBL + cpBR)	Sum of average contact pressures	N/cm^2
cpTL/cpSUM, cpTR/cpSUM, cpBL/cpSUM, cpBR/cpSUM	Relative average contact pressures	
pcpTL (pcpTR)	Average peak contact pressure Thigh Left (Right)	N/cm^2
pcpBL (pcpBR)	Average peak contact pressure Buttock Left (Right)	N/cm^2
pcpSUM (pcpTL+pcpTR+pcpBL+pcpBR)	Sum of peak contact pressures	N/cm^2
pcpTL/pcpSUM, pcpTR/pcpSUM, pcpBL/pcpSUM, pcpBR/pcpSUM	Relative peak contact pressures	

Table 2: Contact area and pressure variables

perception of seat comfort. The mat had an active area of 392 mm x 392 mm, and sensor pitch was 24.5 mm (0.167 sensore/cm^2). Pressures were recorded at 50Hz. This sampling rate was considered sufficient to monitor the frequency of postural changes.

Contact area and contact pressure were calculated by including only data from sensors that were pressed (i.e. a positive value) at least once, and average (arithmetic mean) values were determined for the last 4 minutes of each session. Earlier data were excluded since they were transient due to settling into the chair (Reed et al., 1999).

2.4 – Data analysis

Data analysis aimed at deepening the following three aspects:

- effects of anthropometric variability and of differences in seat conditions on contact pressures;
- relationship between subjective evaluations and objective measurements of seat comfort/discomfort;
- discriminant effectiveness of indexes based on seat interface pressure in predicting comfort/discomfort.

The first aspect was investigated by building new pressure maps of the maximum Peak Contact Pressure (PCP) and analyzing the sampling distributions of the unloaded weight for male and female users.

The dependency of subjective ratings from contact area and pressure variables was investigated via Principal Component Regression (PCR) which develops into two steps: (1) a Principal Component Analysis (PCA, based on the correlation matrix) to reduce the number of explanatory variables (*i.e.* contact area and pressure variables) and (2) a multiple regression of each subjective rating on the factors from the PCA (obtained at step 1), which are in turn combinations of the contact area and pressure variables.

Following the data analysis strategy proposed by Kyung et al. (2008), all the collected contact area and pressure data were divided into four groups corresponding to four local body parts (*i.e.* right/left buttock and right /left thigh see Figure 1) and a total of 27 explanatory variables were derived (Table 2) to be used at step 1 of PCR. The first 9 variables were related to average contact areas and ratios; the second 9 variables described average contact pressures and ratios; and the last 9 variables indicated average peak contact pressures and ratios. Three overall pressure variables (caSUM, cpSUM, and pcpSUM) were only used to derive 12 ratio variables but they were not further analyzed.

The number of Principal Components (or Factors) was determined by two criteria, the size of the eigenvalue (>1) and the cumulative percentage (≈90%) of variance accounted for by the selected components. The selected factors were rotated by the varimax method. Statistical results were considered 'significant' or 'marginal' when $p \leq 0.05$ and $0.05 < p \leq 0.1$, respectively.

In order to select a robust response (*i.e.* a good proxy of perceived comfort/discomfort) to be used in step 2 of PCR and verify the consistency of the subjective data, the association among the three evaluation scales adopted was evaluated.

Finally, the discriminant effectiveness in predicting seat comfort/discomfort was evaluated for two indexes based on seat interface pressure: Peak of Contact Pressure (PCP) and Pressure Comfort Loss (PCL). The latter (Lanzotti et al., 2011) is based on a "Nominal is the Best" (NB) comfort loss function, standardized with respect to the nominal pressure. The formulation of this index takes into account the need to design for a specific target population through the introduction of a parameter θ related to the composition of the sample in terms of gender (eq. 1):

$$WPCL(\theta) = \theta \cdot WPCL_f + (1-\theta) \cdot WPCL_m \qquad (1)$$

with:

- $WPCL_f$, is comfort loss function (PCL) for the female population;
- $WPCL_m$, is comfort loss function (PCL) for the male population.

As regard the PCP index was considered the maximum PCP registered among the four groups of pressure corresponding to four local body parts (*i.e*. right/left buttock and right /left thigh in Figure 1).

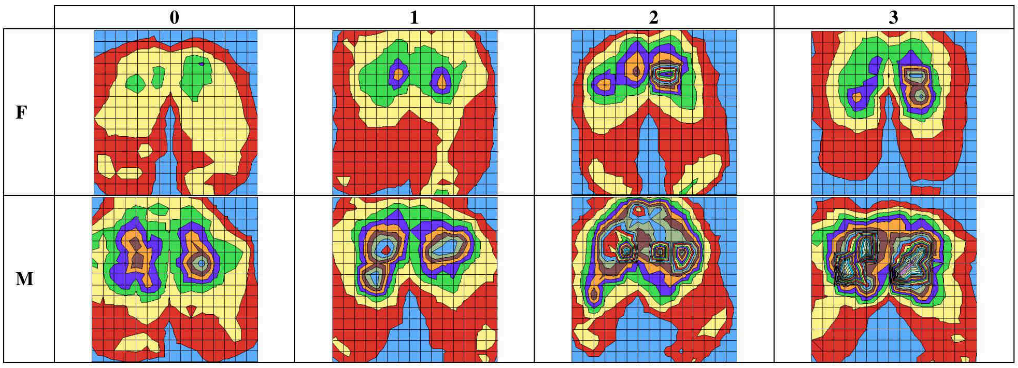


Table 3: Maps of the maximum PCP for the different sub-samples stratified by gender and softness of the cushion

3- Results

3.1 – Effects of anthropometric variability and seat conditions on interface pressures

In order to analyse the effects of anthropometric variability and of seat conditions on the contact pressures, the average peak contact pressure detected by each sensor mat was analysed. Specifically, new pressure maps stratified by gender and seat condition were developed, in which each map cell represents the PCP greater among all PCP sampled from a particular sensor (16 x 16), for a given seat condition (0, 1, 2, 3) and for a given gender (M, F). Thus the new maps were obtained (Table 3). The female maps (first row of the table) show PCP values lower than the corresponding male maps (second row of the Table 3). Similarly, the table was also examined by columns, in other words from the first column (soft cushion level) to the fourth (semi-rigid cushion level), PCP values shown gradually increase. It can be said that the Seat Condition 0 and 1, show pressure levels lower than the Seat Condition 2 and 3, for both males and females. One can assume that, independently of the anthropometric variability, the ideal contact pressure distribution of males differs from that of females. Hence, the seat designers could take this result into account when designing for a Target. Indeed, pressure levels and contact areas change significantly between males and females. It would seem that the males are more sensitive to changes in the seat condition and this could mean an amplification of discomfort effects in the long period.

In addition Figure 1 shows that, independently of the seat condition, the female users significantly differ from male users in terms of unloaded weight.

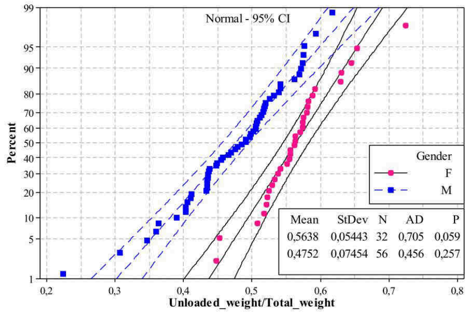


Figure 1: Probability Plots for the r.v. unloaded weight

3.2 – Relationship between subjective evaluations and objective measurements of seat comfort/ discomfort

The relationship between comfort degree and contact pressures was analysed from the statistical standpoint via PCR analysis. At step 1, a PCA was performed on the set of 27 variables listed and described in Table 2. Five principal components with an eigenvalue >1 accounted for 86.9% of the total variance (Table 4). After varimax rotation, each component appeared to have a more general interpretation (as indicated in Table 10). Indeed, a subset of (2 to 4) pressure variables was found in each principal component that predominantly determined the respective component level, as evidenced by high coefficients (>0.4). Further, these subsets of variables were mutually exclusive and distinguishable in terms of relevant body part, or type of pressure, and the principal components were termed accordingly to this (Table 4). Firstly, it's worthwhile to observe that for Factor 4, relating to contact area ratios, coefficients with opposite signs were found between the thigh and buttock average contact area ratios (i.e., caTL/caSUM vs. caBL/caSUM). In other words, there were negative associations between the thigh and buttock in terms of contact area ratio. Secondly for

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
	Left buttock (pressure)	Buttock (area)	Right buttock (pressure)	Left buttock vs thigh (area)	Right thigh (pressure)
caTL	0.024	-0.104	0.025	-0.422	0.028
caTR	-0.003	-0.247	-0.011	-0.501	0.014
caBL	0.101	-0.477	-0.007	0.025	0.021
caBR	-0.052	-0.568	0.050	-0.111	0.043
caTL/caSUM	0.035	0.257	0.029	-0.288	-0.014
caTR/caSUM	-0.020	0.070	-0.040	-0.445	-0.024
caBL/caSUM	0.057	-0.082	-0.030	0.443	0.019
caBR/caSUM	-0.092	-0.286	0.044	0.271	0.018
cpTL	0.001	0.092	0.067	0.042	0.380
cpTR	-0.004	-0.108	-0.031	-0.007	0.497
cpBL	-0.440	0.070	-0.151	-0.036	0.106
cpBR	-0.094	0.024	-0.453	-0.022	0.166
cpTL/cpSUM	0.161	0.113	0.260	0.014	0.108
cpTR/cpSUM	0.157	-0.128	0.151	-0.032	0.261
cpBL/cpSUM	-0.381	0.058	0.121	-0.007	-0.241
cpBR/cpSUM	0.084	-0.032	-0.413	0.020	-0.078
pcpTL	-0.126	0.253	0.037	0.010	0.280
pcpTR	-0.109	0.022	-0.121	0.005	0.499
pcpBL	-0.478	-0.019	-0.057	0.001	0.095
pcpBR	-0.025	0.087	-0.419	-0.003	0.033
pcpTL/pcpSUM	0.115	0.256	0.246	-0.024	0.042
pcpTR/pcpSUM	0.178	-0.052	0.121	-0.020	0.249
pcpBL/pcpSUM	-0.459	-0.135	0.184	0.025	-0.076
pcpBR/pcpSUM	0.235	-0.021	-0.423	0.006	-0.118
Eigenvalue	11.036	4.027	2.574	1.887	1.330
Cum percent	46.0	62.8	73.5	81.3	86.9

Table 4: Five principal components after varimax rotation (underlined values are >0.4 and maximal across factors in absolute value)

Factor 2, relating to contact areas of the bilateral buttocks, high coefficients were all negative.

The consistency of the subjective data was analysed via the Goodman and Kruskall's index, being all three adopted scales ordinal and polytomous. The Goodman and Kruskall's index was calculated for all possible combinations of binary association. Results (Table 5) show a substantial consistency of the scales.

The minimum value for Goodman and Kruskall's index in Table 5 is 0.653 which reveals a medium-high level of association between the scales ranking and rating. It is evident that the responses given on the scale "comfort degree" were highly associated with the other ones (0.984 e 0.860). So this scale was selected as a good proxy of perceived comfort/discomfort and set as a robust response function for step 2 of PCR analysis.

Rating (RT)	0.984	0.653
	Degree (CD)	0.860
		Ranking (CR)

Table 5: Results for association analysis on the evaluation scales

All three fitted regression models for comfort degree were significant ($p \leq 0.01$). As coefficients in table 6 show, increasing Factor 2 (significant for the mixed sample of users and male sub-sample) and decreasing Factor 1 (significant for the mixed sample of users and female sub-sample) and Factor 5 (marginal for the mixed sample of users and significant for the female sub-sample) would be effective at improving comfort degree. In particular, the coefficients for Factor 2 (-0.348 and -0.360, respectively for mixed sample and male sub-sample) indicated that increasing contact areas at the buttocks (specifically, caBL and caBR) would be the most effective method for improving comfort degree. Similarly, the coefficients for Factor 1 (0.174 and 0.350, respectively for the mixed sample and female sub-sample) suggest that decreasing average (peak) contact pressures and ratios relevant to the left buttock (specifically, cpBL, pcpBL e pcpBL/pcpSUM) would be the second most effective way of improving the subjective ratings. Finally, the coefficients of Factor 5 (-0206 and -0901, respectively for the mixed sample and female sub-sample) suggest a third strategy, the decrease of the contact pressure and peak at the right thigh (specifically, cpTR e pcpTR).

Term	Mixed		Males		Females	
	Coef	p	Coef	p	Coef	p
Intercept	2.602	0.000	2.648	0.000	2.449	0.000
Factor 1	0.174	0.024	0.132	0.209	0.350	0.004
Factor 2	-0.348	0.001	-0.360	0.003	-0.276	0.166
Factor 3	0.013	0.836	0.018	0.810	0.157	0.247
Factor 4	-0.037	0.575	-0.076	0.333	-0.054	0.784
Factor 5	-0.206	0.064	-0.139	0.273	-0.901	0.024

Table 6: Standard coefficients for regression models relating PCA factors to comfort degree expressed by mixed sample of users, male sub-sample and female sub-sample

3.3 – Discriminant effectiveness of indexes based on seat interface pressure in predicting comfort/ discomfort

Mean values of PCP and WPCL for the four tested chairs were compared to verify the consistency of discriminant information provided by these indexes.

The results, assuming WPCL as a response function, are shown in Figure 2 on the next page, for the sub-sample of female users, the mixed sample of users and for the sub-sample of male users. Level 0, corresponding to the level Soft, was the best one in terms of WPCL, whereas levels 2 and 3 got the worst results. Level 3 seems to be the most robust one against changes in the composition of the reference (sub-)sample. In fact, all other levels showed higher slopes for the mean effect plots. Assuming that the population were composed exclusively of males, level 3 would be better than level 2.

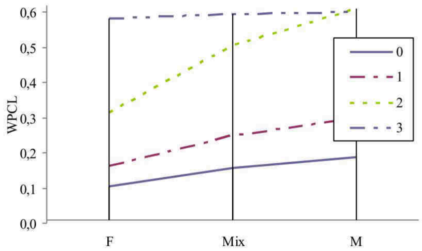


Figure 2: Mean effects assuming WPCL as response for the sample of MIXed users, the sub-sample of Male users, and the sub-sample of Female users

The same analysis was carried out, assuming the peak contact pressure, PCP, as a response function (Figure 3). The lowest values were recorded for level 0, which is the best one independently of the gender of users. Level 1 got comparable performance, whereas level 2 and 3 once again resulted to be the worst ones. Assuming that the (sub-)sample were composed exclusively of male users, level 3 would be better than level 2. Moreover, it's worthwhile to highlight that in the WPCL diagram the differences between the level 0 and 1 are clearer than in the PCP diagram. In order to verify if the two indexes significantly differ for capacity to discriminate among the four seat conditions, the Wilcoxon-Mann-Whitney test was performed. So for both indexes, three binary comparisons of seat condition (0 vs. 1, 1 vs. 2 and 2 vs. 3) were carried out for each composition of (sub-)sample (Mixed, Males and Females). Results (Table 7) show that, independently from the composition of the reference (sub-)sample, the WPCL is able to discriminate between the seat conditions 0 and 1, and between 1 and 2; whereas, the PCP only distinguishes between the seat conditions 1 and 2.

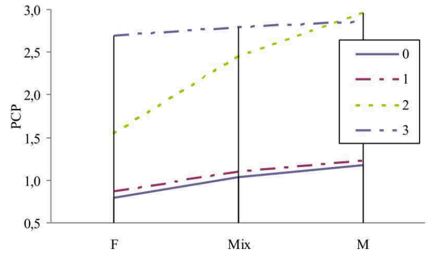


Figure 3: Mean effects assuming PCP as response function for the sample of MIXed users, the sub-sample of Male users, and the sub-sample of Female users.

The ranking of the four tested chairs shows substantial coherence of the results provided by PCP and WPCL, although WPCL seems to be more sensitive to the contact pressure distribution than PCP. This coherency in results does not mean that PCP and WPCL provide the same information. Indeed, for the sub-sample of male users, the maximum values of these indexes refer to different pressure maps and so identify different users (Figure 4).

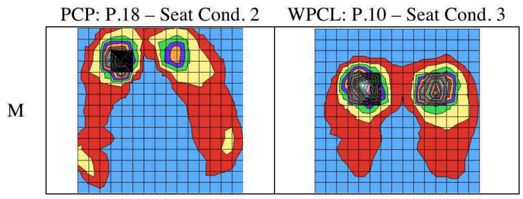


Figure 4: Pressure maps related to maximum values of WPCL and PCP for all seats

4- Discussion and conclusions

The purpose of this paper was to investigate some critical aspects of seat comfort/discomfort assessment. The illustrated experiment refers specifically to office chairs for use at VDT workstation, however the strategy of analysis can be easily adapted to assess seat comfort/discomfort in different context of use.

The results of data analysis provided satisfactory answers to the research questions.

Subjective evaluations resulted significantly correlated to several combinations of pressure variables. Consequently, these pressure variables, derived in terms of average contact area and average (peak) contact pressure, could be used

Variable	Mixed			Males			Females		
	0 vs 1	1 vs 2	2 vs 3	0 vs 1	1 vs 2	2 vs 3	0 vs 1	1 vs 2	2 vs 3
PCP		able to discr.			able to discr.			able to discr.	
WPCL	able to discr.	able to discr.		able to discr.	able to discr.		able to discr.	able to discr.	

Table 7: Results of non-parametric Wilcoxon-Mann-Whitney test

across anthropometric variability for the assessment of sitting comfort/discomfort.

An improved sitting experience is argued as requiring a balancing of pressures between the bilateral buttocks and a balancing of contact areas between buttocks and thighs. Indeed, asymmetrics of contact areas and pressure distributions can be considered undesirable as they appear to lead to lower subjective ratings.

However, use of pressure data is suggested as more appropriate for assessing short-term comfort/discomfort, reflecting seat support and the distribution of body load on it.

Interface pressures resulted significantly affected by anthropometric variability and gender, in particular. As shown in section 3.1, in fact, the pelvis of women is developed more in width, while in men the sacral and iliac bone is thicker and heavier, generating localized peaks of greater magnitude. Consequently, the analysis of pressure maps stratified by gender helps to take into account anthropometric variability and provides valuable suggestions to design the seat shape and choose materials. Moreover, the joint analysis of synthetic indexes (WPCL, PCP) could be a useful tool to discriminate among different seats conditions.

5- References

[DK1] De Looze, M., Kuijt Evers, L., Van Dieen, J., 2003. Sitting comfort and discomfort and the relationships with objective measures. Ergonomics, 46, (10), 985-997.

[DH1] Dul, J., Hilderbrandt, V. H., 1987. Ergonomic guidelines for the prevention of low back pain at the workplace. Ergonomics, 30, (2), 419-429.

[E1] Ehrlich, G. E., 2003. Low Back Pain. Bulletin of the World Health Organization, 81, (9), 671-676.

[FM1] Fujimaki, G., Mitsuya, R., 2002. Study of the seated posture for VDT work. Displays, 23, (1-2), 17-24.

[HM1] Harkness, E. F., Macfarlane, G. J., Silman, A. J., McBeth J., 2005. Is musculoskeletal pain more common now than 40 years ago?: Two population-based cross-sectional studies. Rheumatology, 44, (7), 890–895.

[HZ1] Helander, M., Zhang, L., 1997. Field studies of comfort and discomfortin sitting. Ergonomics, 40, (9), 895-915.

[HD1] Herbert, R., Dropkin, J., Warren, N., Sivin, D., Doucette, J., Kellogg, L., Bardin, J., Kass, D., Zoloth, S., 2001. Impact of a joint labor-management ergonomics program on upper extremity musculoskeletal symptoms among garment workers. Applied Ergonomics, 32, (5), 453-60.

[KT1] Kamijo, K., Tsujimara, H., Obara, H., Katsumatu, M., 1982. Evaluation of seating comfort. SAE Technical Paper Series 820761. Society of Automotive Engineers, Troy, MI.

[KN1] Kyung, G., Nussbaum, M. A., 2008. Driver sitting comfort and discomfort (part II): Relationships with an prediction from interface pressure. International Journal of Industrial Ergonomics 38, (5-6), 526-538.

[LV1] Lanzotti, A., Vanacore, A., Trotta, M., 2011. Validation of a new index for seat comfort assessment based on objective and subjective measurements. Proceedings of the IMProVe International conference on Innovative Methods in Product Design – ADM-INGEGRAF, June 15th – 17th, Venice, Italy.

[LF1] Lengsfeld, M., Frank, A., van Deursen, D.L., Griss, P., 2000. Lumbar spine curvature during office chair sitting. Medical Engineering and Physics, 22, (9), 665-669.

[LG1] Loisel, P., Gosselin, L., Durand, P., Lemaire, J., Poitras, S., Abenhaim, L., 2001. Implementation of a participatory ergonomics program in the rehabilitation of workers suffering from subacute back pain. Applied Ergonomics, 32, (1), 53-60.

[NS1] Nelson, N. A., Silverstein, B. A., 1998. Workplace changes associated with a reduction in musculoskeletal symptoms in office workers. Human Factors, 40, (2), 337-50.

[NN1] Noro, K., Naruse, T., Lueder, R., Nao-i N., Kozawa, M., 2012. Application of Zen sitting principles to microscopic surgery seating. Applied Ergonomics, 43, (2), 308-319.

[PK1] Park, M. Y., Kim, J. Y., Shin, J. H., 2000. Ergonomic design and evaluation of a new VDT workstation chair with keyboard-mouse support. International Journal of Industria.l Ergonomics, 26, (5), 537-548

[RG1] Reed, M., Grant, C., 1993. Development of a measurement protocol and analysis techniques for assessment of body pressure distributions on office chairs. TR. Ann Arbor, MI: Univ. of Michigan, Center for Ergonomics.

[RJ1] Rezaee, M., Jafari, N. J., Ghasemi, M., 2011. Low Back Pain and Related Factors among Iranian Office Workers. International Journal of Occupational Hygiene, 3, (1), 23-28.

[R1] Rubin, D., 2007. Epidemiology and risk factors for spine pain. Neurological Clinics, 25, (2), 353-371.

[SC1] Stinson, M., Crawford, S., 2009. Optimal Positioning: Wheelchair Seating Comfort and Pressure Mapping. In International Handbook of Occupational Therapy Interventions, Springer, 83-90.

[V1] Vink, P., 2012. Editorial: Comfort and discomfort studies demonstrate the need for a new model. Applied Ergonomics, 43, (2), 271-276.

Acknowledgements

The present work was developed with the contribution of the Italian Ministry of University and Research (MIUR) performing the activities of the PON01_01268 DIGIPAT Project, Digital Pattern Product Development: A Pattern Driven approach for industrial product design, simulation and manufacturing.

Paper B

Lanzotti A., Vanacore A., Del Giudice D.M., on the 24th of March in 2015 it was submitted to Applied Ergonomics, ISSN 18729126 and 00036870 (Q1 nel 2013), 12 pp.

Getting insight into Seating Discomfort via a comprehensive Statistical Data Analysis and new Diagnostic Tools

Antonio Lanzotti,
Fraunhofer JL IDEAS,
Department of Industrial Engineering,
University of Naples Federico II,
P.le Tecchio, 80,
Naples 80125, Italy
email: antonio.lanzotti@unina.it

Amalia Vanacore[1],
Fraunhofer JL IDEAS-SQUARE,
Department of Industrial Engineering,
University of Naples Federico II,
P.le Tecchio, 80,
Naples 80125, Italy
email: amalia.vanacore@unina.it
Phone +39-0817682506 – Fax +39-0817682187

Domenico Maria Del Giudice
Fraunhofer JL IDEAS,
Department of Industrial Engineering,
University of Naples Federico II,
P.le Tecchio, 80,
Naples 80125, Italy
email: domenicomaria.delgiudice@unina.it

(1) corresponding author

Abstract:

This paper provides new insights in the evaluation of seating discomfort with respect to three major concerns: 1) the relationship between subjective and objective measures of seat discomfort; 2) the gender-based differences in the distribution of seat-interface pressure; 3) the discriminant effectiveness of indexes based on seat-interface pressure.
Seating discomfort data (both subjective and objective measures) were collected performing a designed experiment involving 22 volunteers who tested 4 office chairs (differing in terms of cushion stiffness). Statistical data analyses showed that subjective sitting discomfort ratings were significantly related to several combinations of pressure variables. This result, together with the evidence of gender-based differences in the distribution of seat-interface pressure, pushes forward a better exploitation of all information available in a pressure map. For this purpose, two novel methods for both graphical (Maximum Peak Contact Pressure - MPCP map) and analytical (Weighted Pressure Comfort Loss - WPCL index) analysis of seat-interface pressure data are discussed. Their joint use can provide useful information to support the product design phase being effective for comparative seat discomfort assessment.
Though the paper focus is on the comparative assessment of *office* seating discomfort across a *gender* stratified population of *healthy* users, the proposed data analysis strategy can be easily adapted to other experimental seating contexts involving different target populations.

Key words:

office chair; sitting discomfort assessment; seat-interface pressure distribution.

1. Introduction

Research in the field of medicine and epidemiology has shown that, over the past decades, the incidence of work-related musculoskeletal disorders (WMSDs) has considerably increased (Harkness *et al.* 2005, Rubin 2007) due to sedentary modern lifestyle characterized by prolonged period of time spent in a seated position (Ehrlich 2003, Dul and Hilderbrandt 1987; Annetts *et al.* 2012). More than 60% of people experience at least one episode of lower back pain (LBP) at work, in almost 45% of cases the first attack of LBP happens while working, with an incidence in the office workers of at least one episode backache every 3 years (Lengsfeld *et al.* 2000, Rezaee *et al.* 2011).

Sitting on an ergonomic chair with a correct posture is undoubtedly one of the most useful remedy in preventing WMSDs, (Nelson and Silverstein 1998, Herbert *et al.* 2001, Loisel *et al.* 2001). The importance of good office seating design in improving human wellness greatly motivates the interest of specialized literature in topics related to the investigation of the biomechanical aspects of sitting and their effect on perceived (dis)comfort.

Typically, discomfort is associated to "an unpleasant state of the human body in reaction to its physical environment" (Vink 2012) and its assessment is realized on the basis of subjective evaluations and/or postural analysis. Subjective evaluations are collected by surveying potential seat users who are asked to express their feelings of discomfort with the seat and/or compare, in terms of perceived discomfort, similar seats.

Postural analysis is realized by measuring one or more objective parameters, several of which are listed in (Andreoni *et al* 2002):

- the pattern of muscle activation measured through electromyography (EMG) (Lueder 1986, van Dieen *et al.* 2001);
- the stress acting on the spine measured through pressure transducer and radio waves (Lueder 1986, Zenk *et al.* 2012);
- the postural angles obtained using contact or non-contact (like photogrammetric) techniques in real experiments (Dreyfuss 2002) or using virtual manikins in virtual experiments (Lanzotti 2008, Barone and Lanzotti 2009);
- these at-interface pressure measured through capacitative or resistive mats (Kyung and Nussbaum 2008).

Many researchers have tried to deepen the relationship between such measurements (Zhang *et al.* 1996). Among all objective parameters, pressure distribution results the objective measure with the clearest correlation with subjective evaluation (De Looze *et al.* 2003, Hamberg-van Reenen *et al.* 2008, Kyung and Nussbaum 2008, Stinson and Crawford 2009, Noro *et al.* 2012). In particular, in several studies on seating design (Kamijo *et al.* 1982, Reed and Grant 1993, Park *et al.* 2000, Fujimaki and Mitsuya 2002, Franz *et al.* 2012), the effects on seat (dis)comfort due to specific product features (*e.g.* cushion shape and materials)have been qualitatively verified by correlating the information obtained from pressure maps with users' (dis)comfort perceptions. In their recent review on the effectiveness of pressure measurements in the assessment of office chair comfort/discomfort, Zemp *et al.* (2015) highlight that investigations on the pressure-comfort/discomfort relationship are mainly based on seats other than the office one (*e.g.* car seats, wheelchairs, tractor seat and surgery seat); they call for further investigations in order to definitively answer whether pressure measurements are suitable for assessing the comfort/discomfort experienced while sitting in office investigating empirical chairs.

Independently from the specific investigation context, a further concern in studies on sitting comfort/discomfort assessment is that pressure measurements are not fully exploited being pressure distribution mostly described by the maximum (peak) pressure and/or the average pressure. Hitherto, little effort has been made to properly synthetize all the information provided by a pressure map and to highlight the usefulness of seat-interface pressure measures for specific purposes defined by designers (e.g. design for a specific user or design for a generic user).

In this paper the main results of an experiment aimed at deepening knowledge on office seat discomfort are described. In particular the experiment and the related explorative data analysis were aimed at investigating three critical aspects of seat discomfort assessment: 1) the relationship between subjective and objective measures of seat discomfort; 2) the gender-based differences in the distribution of seat-interface pressure; 3) the discriminant effectiveness of indexes based on seat-interface pressure.

The dependency of subjective discomfort ratings from contact area and pressure variables was explored via Principal Component Regression (PCR); gender-based differences in seat-interface pressure distribution were investigated by analysing the sampling distributions of the unloaded weight for male and female users and building new pressure maps of the Maximum Peak Contact Pressure (MPCP); finally, the discriminant effectiveness in predicting seat discomfort was evaluated for two indexes based on seat-interface pressure: the Peak of Contact Pressure (PCP) and the Weighted Pressure Comfort Loss (WPCL; Lanzotti *et al.* 2011).

The paper is arranged as follows: an overview of the experiment is provided in section 2; section 3 discusses the strategies for data analysis and the experimental results are illustrated and fully discussed in section 3 and section 4, respectively; finally, in section 5 conclusions are drawn.

2. Materials and methods

Data were obtained from an experiment performed at the Department of Industrial Engineering, University of Naples Federico II, in a suitable room cleared of furnishings and according to a well-defined experimental protocol. The whole experiment consisted of 88 experimental sessions during which22 volunteers tested four ergonomic office chairs performing a task of reading a text on a Visual Display Unit (VDU).

2.1–Seat Conditions

The four office seats have a typical architecture of market product (*i.e.* a five-pointed base, a backrest and two armrests), but differ for the stiffness of the seat pan foam. The seats were named with fantasy names (Oslo Chair, OC, Madrid Chair, MC, Chicago Chair, CC, and Toronto Chair, TC) so as to avoid any conditioning of the brand name or the model name on the evaluation (Table 1). The codes 0, 1, 2, 3 used to distinguish different Seat Conditions refer to increasing levels of cushion stiffness with extremes low (*i.e.* soft cushion) and high (*i.e.* rigid cushion).

Table 1: The tested seat conditions.

Office Seat	OC	MC	CC	TC
Seat Condition(Stiffness)	0 (Low)	1	2	3 (High)

2.2–Subjective and objective measures of seat discomfort

During each experimental session subjective measures of discomfort perception as well as seat-interface pressures were collected.

In order to collect users' evaluations about seat discomfort, three different scales were used: 1) the Discomfort Rating (DR) based on a 10-point ordinal scale with extremes 1 (no discomfort) and 10 (maximum discomfort); 2) the Discomfort Degree (DD) based on a4-level scale of agreement with the statement "I feel uncomfortable"; 3) the Chair Ranking (CR) based on ordinal ascending ranks assigned to chairs consistently with the level of perceived discomfort.

Objective measures were obtained from pressure measured at the seat-interface; these measures consisted of both overall and local pressures (Table 2).

Table 2: Subjective and objective measures.

Type	Name	Area to measure
Objective	PCP, Peak Contact Pressure (N/cm^2) CP, Contact Pressure (N/cm^2) CA, Contact Area (cm^2) UW, Unloaded Weight (kg) WPCL, Weighted Pressure Comfort Loss index	• Left/right thighs (TL/TR) • Left/right buttocks (BL/BR) • Sum of 4 local body part pressures
Subjective	DR, Discomfort Rating	Whole body
	DD, Discomfort Degree	Whole body
	CR, Chairs Ranking	Whole body

2.3–Participants

Twenty-two volunteers, including 8 Females (F) and 14 Males (M), participated in four short-term experimental sessions. Participants were recruited from a university student population. This population was deemed to be relevant to this study as university students tend to spend a large amount of time performing seated work. All participants were free of low back pain for 12 months prior to the testing period. Before experiment began, participants gave informed consent and their personal details (*viz.* gender, age and main occupation) as well as anthropometric data (*viz.* stature and weight) were collected and reported in Table 3.

Table 3: Anthropometric characteristics of participants.

Sex	Age	N	Anthropometric variable	Mean	SD	Min	Max
F	20-31	8	Stature (cm)	164	8	153	178
			Weight (kg)	67.2	13.3	52.8	96.1
M	20-31	14	Stature (cm)	182	8	170	198
			Weight (kg)	79.4	9.3	64.4	93.0

2.4–Experimental protocol

Participants tested the four office chairs in a random order, in order to prevent the disturbance due to the testing sequence. For each testing session, a pressure mat was put on the seat cushion and secured with masking tape to facilitate seat adjustments. Each participant was instructed to sit carefully to minimize wrinkles on the pressure mat. Besides, in order to avoid that discomfort assessments could be affected by the visual impact with the tested chair, the participant was introduced into the room blindfolded and made to sit. Subsequently, the participant was asked to take off the blindfold and adjust the chair in such a way that the legs were in rest conditions and the feet were comfortably on the floor so as to form an angle between the thigh and the leg equal to 90°. Few minutes (≤5) were devoted to initial seat and posture adjustments, then the test session started. In each session, participants performed the task of reading a text on VDU for 20 minutes. At the end of the testing session, the participant was blindfolded again and taken back out of the room.

The specific task of reading a text on VDU was chosen in order to minimize differences in postures among the participants due to the peculiarities in performing more complex task. Indeed, previous studies have shown that the type of computer workstation task performed has an effect on postural responses while sitting (van Dieen *et al.* 2001, Gregory *et al.* 2006, Dunk and Callaghan 2005, Moes 2005, Ellegast *et al.* 2012, Groenesteijn *et al.* 2012). The choice of short-term experimental session is recommended when using pressure mats in order to prevent the well-known effects of creep and/or hysteresis (Fay and Brienza 2000). Moreover, long-term sessions are generally suggested when investigating sitting discomfort due to fatigue resulting from sources other than chair design (Helander and Zhang 1997, Kyung and Nussbaum 2008).

2.5–Data collection and processing

The collection of subjective ratings was organised in such a way as to minimize confusion: the forms for the collection of the Discomfort Rating (DR) and the Discomfort Degree (DD) were administered to each participant immediately after each testing session; instead, the Chair Ranking (CR) was collected only after the participant had tested all four office chairs.

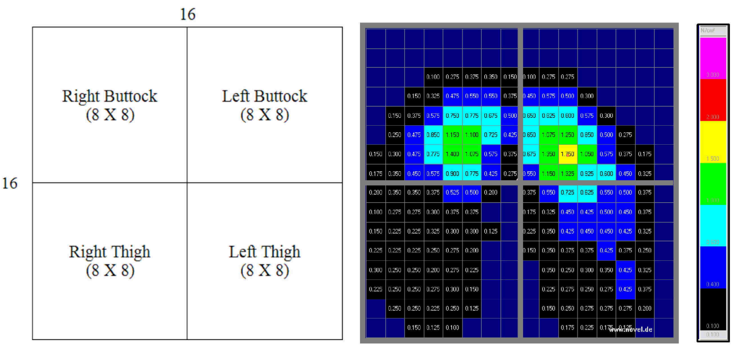
Pressure data were divided into four groups (Figure 1) and were collected continuously during the reading text on VDU, using a Novel Gmbh (Munich, Germany) pressure mat (S2027 Pliance^{TM}).

The above pressure mat comprises 256 (16x16) thin (<1.2 mm) capacitive sensors that could easily conform to the contour of the seat and measure pressures typically in a range from 0.2 N/cm^2 up to 6 N/cm^2. Thanks to its flexible structure the mat is a minimally invasive instrument, which does not interfere with user perception of seat discomfort. The mat has an active area of 392 mm x 392 mm, and sensor pitch is 24.5 mm (0.167 sensor/cm^2).

Pressures were recorded at 50Hz. This sampling rate was considered sufficient given the frequency of postural changes and resultant pressure changes (Kyung and Nussbaum 2008).

Contact area and contact pressure were calculated by including only data from sensors that were pressed at least once and average values were determined for the last 15 minutes of each session. Earlier data (5 min) were excluded since they were transient due to settling into the chair (Reed *et al.* 1999).

Figure 1: Division of pressure mat for four local body parts (left, number of sensors in parentheses) and exemplar pressure distribution (right, a higher peak pressure on left buttock).



3. Data analysis

Data analysis aimed at answering the following three research questions:

- Does a relationship between subjective evaluations and objective measurements of seat discomfort exist?
- How do anthropometric variability and differences in seat conditions affect contact pressures?
- Are indexes based on seat-interface pressure effective in predicting discomfort?

The first question was investigated by adopting a multivariate approach for the statistical analysis of collected data. The whole data analysis procedure developed into three steps: at the first step the association among the three adopted evaluation scales was evaluated via the Goodman and Kruskall's index in order to test the consistency of the subjective data and select the best proxy for perceived discomfort; at the second step pressure and contact variables were analysed via Principal Component Analysis (PCA) in order to reduce the number of explanatory variables; at third step a multiple regression of perceived discomfort on the PCA factors (obtained at step 2) was performed. The number of PCA factors was determined by two criteria, the size of the eigen value (>1) and the cumulative percentage (≈90%) of variance accounted for.

Following the data analysis strategy proposed by Kyung and Nussbaum (2008), all the collected contact area and pressure data were divided into four groups corresponding to four local body parts (*i.e.* right/left buttock and right/left thigh see Figure 1) and a total of 27 explanatory variables were derived (Table 4) to be used at step 2 of the data analysis procedure. The 1-9 variables were related to average contact areas and ratios; the 10-18 variables described average contact pressures and ratios; and the 19-27 variables indicated average peak contact pressures and ratios. The overall pressure variables (caSUM, cpSUM, and pcpSUM) were only used to derive the ratio variables but they were not further analysed. Statistical results were considered 'significant' or 'marginal' when $p \leq 0.05$ and $0.05 < p \leq 0.10$, respectively.

Table 4: Contact area and pressure variables.

Variable	Description	Measurement Unit
caTL (caTR)	Average contact area Thigh Left (Right)	cm^2
caBL (caBR)	Average contact area Buttock Left (Right)	cm^2
caSUM (caTL + caTR + caBL + caBR)	Sum of average contact areas	cm^2
caTL/caSUM, caTR/caSUM, caBL/caSUM, caBR/caSUM	Relative Average contact areas	
cpTL (cpTR)	Average contact pressure Thigh Left (Right)	N/cm^2
cpBL (cpBR)	Average contact pressure Buttock Left (Right)	N/cm^2
cpSUM (cpTL + cpTR + cpBL + cpBR)	Sum of average contact pressures	N/cm^2
cpTL/cpSUM, cpTR/cpSUM, cpBL/cpSUM, cpBR/cpSUM	Relative average contact pressures	
pcpTL (pcpTR)	Average peak contact pressure Thigh Left (Right)	N/cm^2
pcpBL (pcpBR)	Average peak contact pressure Buttock Left (Right)	N/cm^2
pcpSUM (pcpTL+pcpTR+pcpBL+pcpBR)	Sum of peak contact pressures	N/cm^2
pcpTL/pcpSUM, pcpTR/pcpSUM, pcpBL/pcpSUM, pcpBR/pcpSUM	Relative peak contact pressures	

The second question was investigated by building new pressure maps of the Maximum Peak Contact Pressure (MPCP) and by analyzing the sampling distributions of the unloaded weight for male and female users. Pressure data were stratified by gender and seat condition so as to obtain 8 (*i.e.* 2x4) strata. For each stratus a MPCPmap was built (Table 5) so that each map cell represents the greatest value among all (peak) contact pressures sampled from a particular sensor for a given stratus.

Finally, the third question was evaluated by analysing the discriminant effectiveness of two specific indexes based on seat-interface pressure: the Peak of Contact Pressure (PCP) and the Weighted Pressure Comfort Loss (WPCL). The PCP index is the overall maximum pressure value registered on the mat (De Looze *et al.* 2003, Dunk and Callaghan 2005, Hamberg-van Reenen *et al.*2008).

The Weighted Pressure Comfort Loss (WPCL; Lanzotti *et al.* 2011) is a discomfort index formulated under the assumptions that an ideal distribution of seat-interface pressures exists and that every deviation from it causes an increase in user's seat discomfort. Under the reasonable assumption that small deviation are not relevant and that larger deviations become increasingly important (*i.e.* the larger the deviation, the larger the increase in user's seat discomfort) the comfort loss is assumed to be a quadratic function of the deviation from the ideal pressure value. The existence of an ideal seat pressure distribution is accepted in the specialized literature (Kyung and Nussbaum 2008, Fujimaki and Mitsuya 2000) and it is generally believed that the ideal pattern of pressure distribution is obtained by uniformly distributing the body weight over the seating surface.

Coherently with this assumption, for each user, the ideal pressure (*i.e.* target pressure x_{0j}) can be defined as the mean pressure over the whole contact area (or any partition of it).

Let n_j be the number of activated cells in the pressure map for the *j-th* user and x_{ij} the pressure value measured by the *i-th* cell when the *j-th* user is seated, the target pressure is given by:

$$x_{0j} = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j} \qquad (1)$$

For each user and for each cell of the pressure map it is possible to evaluate the deviation of the observed pressure value, x_{ij}, from the target pressure, x_{0j} and thus identify the associated Pressure Comfort Loss (PCL) based on a (Nominal the Best) quadratic loss function.

For the *j-th* user the pressure comfort loss at the *i-th* activated cell of the contact surface can be defined as:

$$PCL_{ij}(x_{ij}) = k_{ij}\left(\frac{x_{ij} - x_{0j}}{x_{0j}}\right)^2 \qquad (2)$$

where k_{ij} is a proportionality coefficient that for each cell measures the loss corresponding to the maximum accepted deviation from the target pressure. In particular, let $Ä_{ij}$ be the maximum accepted relative deviation from ideal pressure at the *i-th* cell activated by the *j-th* user and let C_0 be the comfort loss due to uneven pressure, the proportionality coefficient k_{ij} can be defined as follows:

$$k_{ij} = \frac{C_0}{\Delta_{ij}^2} \qquad (3)$$

For the sake of simplicity and without loss of generality, hereafter C_0 can be assumed unitary and Δ_{ij} can be assumed constant over all activated cells and over all the users belonging to the same target population Ω (e.g. female or male users).

Assuming the hypothesis of additivity of comfort loss function, the PCL index for the *j-th* user belonging to Ω is given by:

$$PCL_{\Omega_j}(\vec{x}) = \Delta_{\Omega}^{-2}\sum_{i=1}^{n_j}\left(\frac{x_{ij} - x_{0j}}{x_{0j}}\right)^2 \qquad (4)$$

where $\vec{x}$ is a vector of dimension n_j with generic element x_{ij} and the proportionality coefficient Δ_{Ω} can be calculated by averaging the maximum relative deviations from ideal pressure over all pressure maps rated at the lowest level on the scale for perceived discomfort (*i.e.* no discomfort).

Starting from (4), the Weighted Pressure Comfort Loss for target population Ω can be defined as:

$$WPCL_{\Omega} = \sum_j w_{\Omega_j} \cdot PCL_{\Omega_j} \qquad (5)$$

Where $W_{\Omega j}$ is a weight that allows to account for the degree of anthropometrical representativeness of the *j-th* user inside the target population. The weights, $W_{\Omega j}$, via the discrete approximation of a continuous random variable (e.g. the stature of potential users) taken as representative of the population anthropometrical variability (for further details see Lanzotti and Vanacore 2007)

Moreover, the overall Weighted Pressure Comfort Loss for a mixture of sub-populations can be obtained as follows:

$$WPCL = \sum_k \theta_k \cdot WPCL_{\Omega_k}; \qquad \sum_k \theta_k = 1 \qquad (6)$$

being θ_k is the mixture parameter accounting for the representativeness of the (sub)population Ω_k inside the overall population.

Thus when dealing with an overall composed by female and male users, the WPCL index is obtained by summing up the gender specific WPCL indexes taking into account their mixture weight:

$$WPCL = \theta_F \cdot WPCL_F + (1 - \theta_F) \cdot WPCL_M; \qquad 0 \le \theta_F \le 1 \qquad (7)$$

4. Results and Discussion

4.1–Explaining the relationship between subjective evaluations and objective measurements of seat discomfort via multivariate data analysis

The Goodman and Kruskall's index was calculated for all possible combinations of binary association among the adopted subjective evaluation scales. Results (Table 5) show a substantial consistency of the tested evaluation scales. In fact, the minimum value for Goodman and Kruskall's index in Table 5 is 0.653 revealing a medium-high level of association between the CR and DR scales. Since responses given on the DD scale were highly associated with both CR and DR scales (0.984 and 0.860, respectively), this scale was selected as a good proxy of perceived discomfort and set as a robust response function for explorative data analysis via PCR.

Table 5: Results for association analysis on the subjective evaluation scales.

Rating (DR)	0.984	0.653
	Degree (DD)	0.860
		Ranking (CR)

PCR Analysis

From the PCA on the set of 27 variables listed and described in Table 4, resulted five principal components with an eigen value >1 accounted for 86.9% of the total variance (Table 6). After varimax rotation, principal components appeared to have a more general interpretation; indeed, for each of them a predominant subset of (2-4) pressure variables was found. Since these subsets of variables were mutually exclusive and distinguishable (*e.g.* in terms of body part) the components were termed accordingly to them. It is worthwhile to note that Factor 4 shows coefficients with opposite signs for thigh average contact area ratio and buttock average contact area ratio (*i.e.*, caTL/caSUM vs. caBL/caSUM), providing some evidence of negative association in terms of contact area ratio between thigh and buttock,.

Table 6: Five principal components after varimax rotation (underlined values are >0.4 and maximal across factors in absolute value).

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Variable	Left buttock (pressure)	Buttock (area)	Right buttock (pressure)	Left buttock vs. thigh (area)	Right thigh (pressure)
caTL	0.024	-0.104	0.025	-0.422	0.028
caTR	-0.003	-0.247	-0.011	-0.501	0.014
caBL	0.101	-0.477	-0.007	0.025	0.021
caBR	-0.052	-0.568	0.050	-0.111	0.043
caTL/caSUM	0.035	0.257	0.029	-0.288	-0.014
caTR/caSUM	-0.020	0.070	-0.040	-0.445	-0.024
caBL/caSUM	0.057	-0.082	-0.030	0.443	0.019
caBR/caSUM	-0.092	-0.286	0.044	0.271	0.018
cpTL	0.001	0.092	0.067	0.042	0.380
cpTR	-0.004	-0.108	-0.031	-0.007	0.497
cpBL	-0.440	0.070	-0.151	-0.036	0.106
cpBR	-0.094	0.024	-0.453	-0.022	0.166
cpTL/cpSUM	0.161	0.113	0.260	0.014	0.108
cpTR/cpSUM	0.157	-0.128	0.151	-0.032	0.261
cpBL/cpSUM	-0.381	0.058	0.121	-0.007	-0.241
cpBR/cpSUM	0.084	-0.032	-0.413	0.020	-0.078
pcpTL	-0.126	0.253	0.037	0.010	0.280
pcpTR	-0.109	0.022	-0.121	0.005	0.499
pcpBL	-0.478	-0.019	-0.057	0.001	0.095
pcpBR	-0.025	0.087	-0.419	-0.003	0.033
pcpTL/pcpSUM	0.115	0.256	0.246	-0.024	0.042
pcpTR/pcpSUM	0.178	-0.052	0.121	-0.020	0.249
pcpBL/pcpSUM	-0.459	-0.135	0.184	0.025	-0.076
pcpBR/pcpSUM	0.235	-0.021	-0.423	0.006	-0.118
Eigenvalue	11.036	4.027	2.574	1.887	1.330
Cum percent	46.0	62.8	73.5	81.3	86.9

Fitted DD regression models were significant ($p \leq 0.01$) for the whole sample of users (*i.e.* group of mixed users) as well as the two sub-samples obtained by stratifying by gender (*i.e.* group of male users and group of female users); however the five factors account for the DD of female users somewhat better than for the DD of male users and mixed users (R^2 equals to 52.9%, 23.7% and 25.7%, respectively).

As coefficients in Table 7 show, increasing Factor 2 (significant for the whole sample and for the sub-sample of male users) and decreasing Factor 1 (significant for the whole sample and for the sub-sample of female users) and Factor 5 (marginal for the whole sample and for the sub-sample of female users) would be effective at decreasing DD. In particular, the coefficients for Factor 2 (-0.348 and -0.360 for the whole sample and for the sub-sample of male users, respectively) indicated that increasing contact areas at the buttocks (specifically, caBL and caBR) would be the most effective method for decreasing DD in particular in the group of male users. Similarly, the coefficients for Factor 1 (0.174 and 0.350for the whole sample and for the sub-sample of female users, respectively) suggest that decreasing average (peak) contact pressures and ratios relevant to the left buttock (specifically, cpBL, pcpBL e pcpBL/pcpSUM) would be the second most effective way of decreasing subjective perception of discomfort especially in the group of female users. Finally, the coefficients of Factor 5 (-0.206 and -0.901,for the whole sample and for the sub-sample of female users, respectively) provides one more suggestion for seat design improvement consisting in decreasing contact pressure and peak at the right thigh (specifically, cpTR and pcpTR), this action will be particularly effective on the group of female users.

Table 7: Standard coefficients for regression models relating PCA factors to Discomfort Degree.

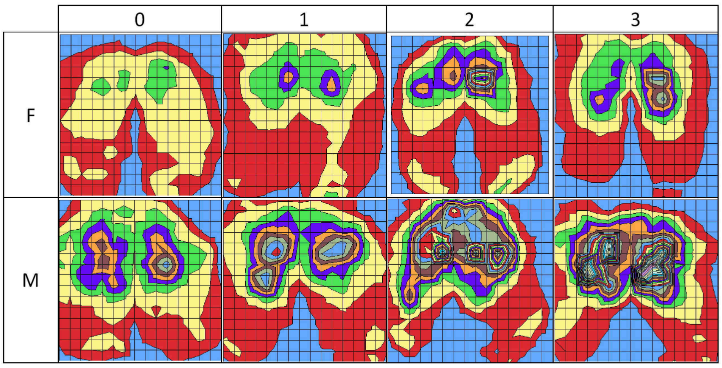
Term	Mixed		Males		Females	
	Coeff	p.value	Coeff	p.value	Coeff	p.value
Intercept	2.602	0.000	2.648	0.000	2.449	0.000
Factor 1 - Left buttock (pressure)	0.174	0.024	0.132	0.209	0.350	0.004
Factor 2 - Buttock (area)	-0.348	0.001	-0.360	0.003	-0.276	0.166
Factor 3 - Right buttock (pressure)	0.013	0.836	0.018	0.810	0.157	0.247
Factor 4 - Left buttock vs. thigh (area)	-0.037	0.575	-0.076	0.333	-0.054	0.784
Factor 5 - Right thigh (pressure)	-0.206	0.064	-0.139	0.273	-0.901	0.024

4.2–Effects of anthropometric variability and seat conditions on interface pressures

The analysis of MPCP maps shows that the contact pressure distribution of males is different from the contact pressure distribution of females. Table 8, reporting The MPCP maps can be arranged , can be read both by rows and by columns. In particular, the comparison by rows provides information on the effects of the anthropometric variability on seat-interface pressures. In fact, the female maps (first row of Table 8) show PCP values lower than the corresponding male maps (second row of Table 8). On the contrary, the comparison by columns provides information on the effects of the Seat Conditions. It's worthwhile to note that moving from the first column (low Stiffness) to the fourth one (high Stiffness), PCP values gradually increase. Thus, it can be said that the first two Seat Conditions show pressure levels lower than the last two Seat Conditions for both males and females.

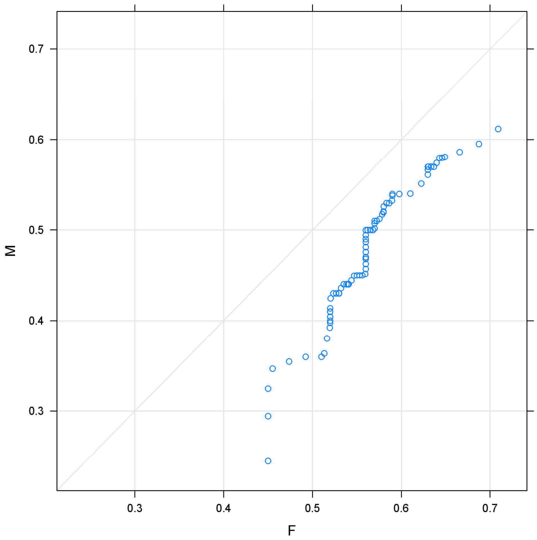
Briefly, the MPCP maps in Table 8 point out that: 1) pressure levels and contact areas vary between males and females and 2) it would seem that the males are more sensitive to changes in the seat condition and this could mean an amplification of discomfort effects in the long period.

Table 8: Maps of the MPCP for the different sub-samples stratified by gender and seat condition.



In addition the Quantile-Quantile plot in Figure 2 shows that, independently of the seat condition, the female users significantly differ from male users in terms of unloaded weight. In particular, quantiles of the unloaded weight are higher for female users than for male users. This result means that the location value of the unloaded weight is higher for female users than for male users; however, the non-linearity in the Quantile-Quantile plot implies that the difference between the two samples is not explained simply by a shift in location.

Figure 2: Quantile-Quantile plot for the unloaded weight

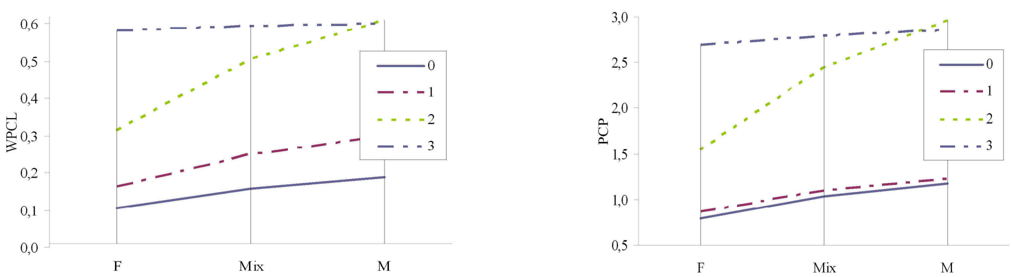


The obtained experimental results are consistent with the findings of previous studies (Dunk and Callaghan 2005; Gregory *et al.* 2006; Beach *et al.* 2008) investigating the influence of personal characteristics (*i.e.*, gender and flexibility) on postures adopted when performing seated computer work. These findings generally evidence that males and females react differently to seated exposures; in particular, the study of Dunk and Callaghan (2005) suggests that men tend to slouch against the back rest while females perch closer to the front of the seat pan. The above gender-based differences have been related to inter-individual variations in hip, hamstring, and low-back flexibility.

4.3 – Discriminant effectiveness of indexes based on seat-interface pressure in predicting discomfort

Mean values of PCP and WPCL were compared for the four tested chairs in order to verify the consistency of discriminant information provided by these indexes. The mean values of both indexes were calculated over the whole sample of users as well as over the two sub-samples of female users and male users. The results are shown in Figure 3(a) and Figure 3 (b) for WPCL and PCP, respectively.

Figure 3: Mean effect plots assuming WPCL (a) and PCP (b) as response.



(a) (b)

The two diagrams in Figure 3 show substantial coherence of the results provided by PCP and WPCL against increasing levels of Stiffness (S): the Seat Condition characterized by Low Stiffness (coded as 0) was the best one in terms of both WPCL and PCP, whereas the worst results were obtained for the Seat Condition characterized by High Stiffness (coded as 3).

For the Seat Condition 3, the mean effect plots of WPCL and PCP are constant against changes in the composition of the reference (sub-)sample. This result could be explained by a saturation effect due to the rigid cushion which produces very high pressure values which are comparable to those ones obtained in previous studies on a hard flat surface (Brienza and Karg 1998, Ragan *et al*. 2002). On the other hand, a similar effect is shown for the Seat Condition

0: the mean effect plots of both indexes, though not constant, show a lower slope compared to Seat Conditions 1 and 2 characterized by intermediate levels of Stiffness. This result could be explained as there were a point of diminishing returns beyond which decreasing the cushion stiffness is not really effective in further reducing the seat–interface pressure. A similar effect was found in a previous study with regard to cushion thickness (Regan *et al.*, 2002).

Though the mean effects plots of WPCL and PCP show similar patterns, it's worthwhile to highlight that, in the WPCL diagram the differences between the mean effect plots of Seat Condition 0 and 1 are clearer than in the PCP diagram.

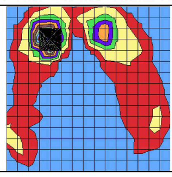
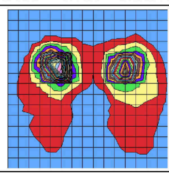
In order to verify if the two indexes significantly differ in discriminating among the four Seat Conditions, the Wilcoxon-Mann-Whitney test was performed. So for both indexes, three binary comparisons of Seat Conditions (0 vs. 1, 1 vs. 2 and 2 vs. 3) were carried out for the whole sample of users as well as the two sub-samples of Females and Males. Results (in Table 9) show that, independently from the composition of the reference (sub-)sample, the WPCL is able to discriminate between the Seat Conditions 0 and 1, and between 1 and 2; whereas, the PCP only distinguishes between the Seat Conditions 1 and 2.

Table 9: Results of non-parametric Wilcoxon-Mann-Whitney test.

	Mixed			Males			Females		
	0 vs. 1	1 vs. 2	2 vs. 3	0 vs. 1	1 vs. 2	2 vs. 3	0 vs. 1	1 vs. 2	2 vs. 3
PCP		X			X			X	
WPCL	X	X		X	X		X	X	

Finally, it is worthwhile to point out that, though their overall results are consistent, PCP and WPCL do not provide the same information. Indeed, for the sub-sample of male users, the maximum values of these indexes refer to different pressure maps and so identify different users (Table 10).

Table 10: Pressure maps related to maximum values of WPCL and PCP.

	PCP: P.18 – Seat Cond. 2	WPCL: P.10 – Seat Cond. 3
M		

5. Conclusions

This study provides satisfactory answers to some relevant issues related to the assessment of sitting discomfort due to office chairs.

Subjective discomfort evaluations resulted significantly correlated to several combinations of pressure variables, derived in terms of average contact area and average (peak) contact pressure. Consequently, these variables can be used across anthropometric variability for the assessment of static sitting discomfort in short-experimental sessions.

In particular, the perceived discomfort appears to be mainly due to the lack of pressure balance between the bilateral buttocks and the lack of balance in contact areas between buttocks and thighs. Thus, asymmetries in pressure distributions and in contact areas should be considered undesirable as they lead to increasing unpleasant state of human body.

The experimental results confirm the hypotheses that due to fundamental biomechanical differences in their sitting behaviours, males and females are exposed to different loading patterns and experience different discomfort pathways.

The adopted statistical approach can effectively support the designer in diagnosing seat discomfort (via the MPCP maps) and testing (via the WPCL index) alternative design solutions (e.g. in terms of both shape and materials) for specific purposes (e.g. design for a specific user or design for a generic user).

Though the paper focus is on the comparative assessment of office seating discomfort across a gender stratified population of healthy users, the proposed data analysis strategy can be easily adapted to other experimental seating contexts involving different target populations.

References

Andreoni, G., Santambrogio, G. C., Rabuffetti, M., Pedotti, A., 2002. Method for the analysis of posture and interface pressure of car drivers. Applied Ergonomics 33 (6), 511-522.

Annetts S, Coales P, Colville R, *et al.*, 2012. A pilot investigation into the effects of different office chairs on spinal angles. European Spine Journal 21 (Suppl 2), 165-170.

Barone, S., Lanzotti, A., 2009.Robust Ergonomic Virtual Design. In: Statistics for Innovation: statistical design of "continuous" product innovation. Springer, 61-83.

Beach, T.M., McDonald, K.A., Coke, S.K., Callaghan, J.P., 2008. Gender responses to automobile and office sitting — influence of hip, hamstring, and low-back flexibility on seated postures. The Ergonomics Open Journal, 1, 1-9. Retrieved January 29, 2014, from: http://benthamopen.com/toergj/articles

Brienza, D.M., Karg, P.E., 1998. Seat cushion optimization: a comparison of interface pressure and tissue stiffness characteristics for spinal cord injured and elderly patients. Archives of Physical Medicine and Rehabilitation, 79, 388-394.

De Looze, M., Kuijt Evers, L., van Dieen, J., 2003. Sitting comfort and discomfort and the relationships with objective measures. Ergonomics, 46 (10), 985-997.

Dreyfuss, H., 2002. The measure of man and woman. Wiley.

Dul, J., Hilderbrandt, V. H., 1987. Ergonomic guidelines for the prevention of low back pain at the workplace. Ergonomics, 30 (2), 419-429.

Dunk, N.M., Callaghan, J.P., 2005. Gender-based differences in postural responses to seated exposures. Clinical Biomechanics 20, 1101-1110.

Ehrlich, G. E., 2003. Low Back Pain. Bulletin of the World Health Organization, 81 (9), 671-676.

Ellegast, R.P., Keller, K., Groenesteijn, L., Krause, F., Berger, H., Vink, P., 2012. Comparison of four specific dynamic office chairs with a conventional office chair: impact upon muscle activation, physical activity and posture. Applied ergonomics, 43(2), 296-307.

Fay, B.T., Brienza, D.A., 2000. What is interface pressure? Paper presented at the 22nd Annual EMBS International Conference, Chicago, IL.

Franz, M.M., Durta, A., Zenk, R., Desmet, P.M.A., 2012. Comfort effects of a new car headrest with neck support Applied ergonomics, 43 (2), 336-343.

Fujimaki, G., Mitsuya, R., 2002. Study of the seated posture for VDT work. Displays, 23 (1-2), 17-24.

Gregory, D.E., Dunk, N.M., Callaghan, J.P., 2006. Stability ball versus office chair: Comparison of muscle activation and lumbar spine posture during prolonged sitting. Human Factors 48(1),142-153

Groenesteijn, L., Ellegast, R.P., Keller, K., Krause, F., Berger, H., De Looze, M.P., Office task effects on comfort and body dynamics in five dynamic office chairs. Applied ergonomics, 43 (2), 320-328.

Hamberg-van Reenen, H.H., Van der Beek, A.J., Blatter, B.M., Van der Grinten, M.P., Van Mechelen, W., Bongers, P.M., 2008. Does musculoskeletal discomfort at work predict future musculoskeletal pain? Ergonomics 51, 637-648.

Harkness, E. F., Macfarlane, G. J., Silman, A. J., McBeth J., 2005. Is musculoskeletal pain more common now than 40 years ago?: Two population-based cross-sectional studies. Rheumatology, 44 (7), 890–895.

Helander, M., Zhang, L., 1997. Field studies of comfort and discomfort in sitting. Ergonomics, 40 (9), 895-915.

Herbert, R., Dropkin, J., Warren, N., Sivin, D., Doucette, J., Kellogg, L., Bardin, J., Kass, D., Zoloth, S.,2001. Impact of a joint labor-management ergonomics program on upper extremity musculoskeletal symptoms among garment workers. Applied Ergonomics, 32 (5), 453-60.

Kamijo, K., Tsujimara, H., Obara, H., Katsumatu, M., 1982. Evaluation of seating comfort. SAE Technical Paper Series 820761.Society of Automotive Engineers, Troy, MI.

Kyung, G., Nussbaum, M. A., 2008. Driver sitting comfort and discomfort (part II): Relationships with an prediction from interface pressure. International Journal of Industrial Ergonomics 38 (5-6), 526-538.

Lanzotti, A., 2008. Robust design of car packaging in virtual environment. International Journal on Interactive Design and Manufacturing (IJIDeM) 2(1), 39-46.

Lanzotti, A., Vanacore, A., 2007. An efficient and easy discretizing method for the treatment of noise factors in robust design. Asian Journal on Quality, 8(3), 188-197.

Lanzotti, A., Vanacore, A., Trotta, M., 2011.Validation of a new index for seat comfort assessment based on objective and subjective measurements. Proceedings of the IMProVe International conference on Innovative Methods in Product Design – ADM-INGEGRAF, June 15^{th} – 17^{th}, Venice, Italy.

Lengsfeld, M., Frank, A., van Deursen, D.L., Griss, P., 2000. Lumbar spine curvature during office chair sitting. Medical Engineering and Physics, 22 (9), 665-669.

Loisel, P., Gosselin, L., Durand, P., Lemaire, J., Poitras, S., Abenhaim, L., 2001. Implementation of a participatory ergonomics program in the rehabilitation of workers suffering from sub acute back pain. Applied Ergonomics, 32 (1), 53-60.

Lueder, R. K., 1986. Seat comfort: A review of the Construct in the Office Environment. Human Factors 25 (6), 701-711.

Moes, N.C.C.M., 2005. Analysis of sitting discomfort, a review. In: Bust, P.D., McCabe, P.T. (Eds.), Contemporary Ergonomics 2005. Taylor & Francis, London, 200-204.

Nelson, N. A., Silverstein, B. A., 1998. Workplace changes associated with a reduction in musculoskeletal symptoms in office workers. Human Factors, 40 (2), 337-50.

Noro, K., Naruse, T., Lueder, R., Nao-i N., Kozawa, M., 2012. Application of Zen sitting principles to microscopic surgery seating. Applied Ergonomics, 43 (2), 308-319.

Park, M. Y., Kim, J. Y., Shin, J. H., 2000. Ergonomic design and evaluation of a new VDT workstation chair with keyboard-mouse support. International Journal of Industrial Ergonomics, 26 (5), 537-548

Ragan R., Kernozek T. W., Bidar M., Matheson J. W.,2002.Seat-interface pressures on various thicknesses of foam wheelchair cushions: a finite modeling approach. Archives of Physical Medicine and Rehabilitation, 83, 872-875

Reed, M., Grant, C., 1993. Development of a measurement protocol and analysis techniques for assessment of body pressure distributions on office chairs. TR. Ann Arbor, MI: Univ. of Michigan, Center for Ergonomics.

Reed, M.P., Manary, M.A., Schneider, L.W., 1999. Methods for measuring and representing automobile occupant posture. SAE Technical Paper Series 1999-01-0959.

Rezaee, M., Jafari, N. J., Ghasemi, M., 2011. Low Back Pain and Related Factors among Iranian Office Workers. International Journal of Occupational Hygiene, 3 (1), 23-28.

Rubin, D., 2007. Epidemiology and risk factors for spine pain. Neurological Clinics, 25 (2), 353-371.

Stinson, M., Crawford, S., 2009. Optimal Positioning: Wheelchair Seating Comfort and Pressure Mapping. In International Handbook of Occupational Therapy Interventions, Springer, 83-90.

van Dieen, J.H., de Looze, M.P., Hermans, V., 2001. Effects of dynamic office chairs on trunk kinematics, trunk extensor EMG and spinal shrinkage. Ergonomics 44, 739.750.

Vink, P., 2012. Editorial: Comfort and discomfort studies demonstrate the need for a new model. Applied Ergonomics, 43, (2), 271-276.

Zemp, R., Taylor, W. R., Lorenzetti, S., 2015. Are pressure measurements effective in the assessment of office chair comfort/discomfort? A review. Applied Ergonomics, 48, 273-282.

Zenk, R., Franz, M., Bubb, H., Vink, P., 2012. Technical note: Spine loading in automotive seating. Applied Ergonomics, 43 (2), 290-295.

Zhang, L., Helander, M., Drury, C., 1996. Identifying factors of comfort and discomfort. Human Factors 38, 377–389.

Acknowledgements

The present work was developed with the contribution of the Italian Ministry of University and Research (MIUR) performing the activities of the PON01_01268 DIGIPAT Project, Digital Pattern Product Development: A Pattern Driven approach for industrial product design, simulation and manufacturing.

Paper C

Patalano S., Del Giudice D.M., Gerbino S., Lanzotti A., Vitolo F., Proceedings of Joint Conference on Mechanical, Design Engineering & Advanced Manufacturing 2014, Paper n. 71, Toulouse (France), June 18th–20th 2014, 7 pp. It's waiting to be published on Research in Interactive Design Vol. 4 book by Springer Verlag (ISBN available soon). One of the best presented papers proposed to be published in the International Journal of Interactive Design and Manufacture (indexed in Scopus, Q1 in 2013; ISSN 1952513 and 19552505).

GUI USABILITY IMPROVEMENT FOR A NEW DIGITAL PATTERN TOOL TO ASSIST GEARBOX DESIGN

S. Patalano[1], D. M. Del Giudice[1], S. Gerbino[2], A. Lanzotti[1], F. Vitolo[1]

(1) : Department of Industrial Engineering, University of Naples Federico II
P.le Tecchio, 80, 80125, Naples (NA), Italy
Phone +39 0817682506 – Fax +39 0817682187
E-mail : { stanislao.patalano, domenicomaria.delgiudice, antonio.lanzotti}@unina.it; ferdy.vitolo@tiscali.it
(2) : DiBT Dept., University of Molise, Via Duca degli Abruzzi, Termoli (CB), Italy
Phone +39 0874 404593 - Fax +39 0874 404652
E-mail: salvatore.gerbino@unimol.it

Abstract: Design team can speed up the process of managing information related to gearbox design process by adopting digital pattern tools. These tools, as KBE systems, can assist engineers in re-using previous knowledge in order to improve time-consuming task as retrieval and selection of previous architectures and to modify and virtually test a new gearbox design. A critical point in the development of a KBE system is the interface usability to demonstrate effective reduction of development time and satisfaction in its use. In this paper, the authors face the problem of usability improvement of the Graphical User Interface (GUI) of the KBE system previously proposed. An approach based on Analytic Hierarchy Process (AHP) and Multiple-Criteria Decision Analysis (MCDA) has been used. A participatory test has been performed for evaluating the Usability Index (UI) of the GUI. Taking into account the data analysis some changes have been carried out and a new GUI release has been validated with new experimentations.

Key words: Usability assessment; *Graphical User Interface*; gearbox design; participatory design.

1- Introduction

Large and small companies adopt structured work teams to develop products [SB1]. These groups need a software toolset to help them to keep their design systems more and more effective. In such context, a "Digital Pattern" (DP) platform is recommended. A DP platform is a set of geometric and numeric data structures as well as models preconfigured and parametric which can be adapted to specific contexts. Therefore, a DP acts as a KBE system aimed to improve quality and reduce times and costs for product development through a massive use of knowledge and tools integration inside company. Such aims are accomplished through the fast and the best re-use of company knowledge, i.e. through technical and technological predefined solutions that quickly marry new projects, allowing fast performance evaluations and immediate checks. Such solutions should also be able to give feedback on designing and production costs. In [LP1] and [PV1] a Digital Pattern system, developed to assist gearbox design, was described.

In KBE system development the evaluation of interface usability, to demonstrate the effective reduction of development time and satisfaction in use, is a critical point. Indeed, usability can be defined as *the extent to which specific users, in a specific context of utilisation, can use a product to their satisfaction, in order to effectively and efficiently achieve specific goals* [MA1]. Sohaib and Khan [SK1] claim that agile projects needs to adopt aspects of *usability engineering* by incorporating user scenarios and including usability specialists in the team. During the designing of mechanical parts, designers need to verify the correctness of the hypotheses in use, especially in relation with multi-objective tasks [PV1]. Furthermore, the use of a KBE system is strategic for designers if we consider that they spend up to 30% of their time searching data [S2] and this percentage rises to 50 when we take into account the time spent validating the data [B1].

Following the approach in [DM1] this paper proposes the usability assessment of the gearbox software tool through a participatory testing, in order to evaluate the Usability Index (UI). The evaluation of the UI has been carried out with Saaty's *Analytic Hierarchy Process* (AHP), which implies the decomposition of the problem into several levels [S1]. In this case study, three dimensions of UI have been considered: effectiveness, efficiency and satisfaction. Taking into account the experimental data analysis, some changes have been carried out and a new release of the GUI has been validated with new experimentations.

The paper is arranged as follows. Section 2 summarizes the approach. Section 3 provides the experimental phases. Section 4 deals with results of experimentations. Finally,

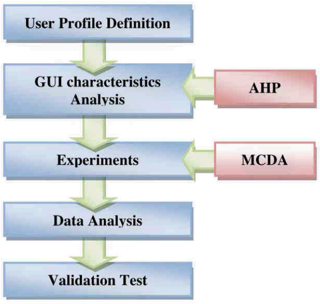


Figure 1: The logic flow chart of methodological approach

Section 5 draws conclusions.

2- The methodological approach

Usability evaluation is based on both objective aspects and subjective ones. Starting from [DM1], the adopted approach is participative, requiring the involvement of potential users during all phases of the usability evaluation process. In Figure 1 the logical flow chart of methodological approach is shown.

2.1 - User Profile Definition and GUI characteristics

A Graphical User Interface (GUI), developed in MatLAB® environment, was released in order to support the design of automotive manual transverse gearboxes (Figure 2). This tool has two main tasks: 1) - design gearboxes rapidly, reducing the risks of using incomplete information when making product development decisions; 2) - assist the designer in redesigning the gearboxes previously developed (for more details see [PV1][LP1]). It was developed for a user not very experienced in design, such as a junior designer.

The main window is divided in three fields (Figure 2): the upper field where the gearbox is pre-configured; the middle field where the gearbox is configured and the gears are characterised; the lower field where are located the post-processing and evaluating commands.

The designer can set: type of gearbox (among preconfigured types); number of gears (up to six); parameters layout controller (axle bases and angle between them); characteristic parameters for each gear (gear ratio, pressure angle, helical angle)(in viable range); gear teeth number (among those automatically generated). Specific panels to set reverse and differential gear were developed. By changing these parameters a new gearbox could be defined; otherwise a previous design could be edited.

As regards the post-processing, by clicking on "Plot" command, the surface models of toothing, i.e. head, sides and foot of helical-toothed gears, could be automatically generated and displayed on screen as mesh models; furthermore, it is possible to change the mesh degree in order to edit both representation and file export definition, ensuring the generation of the 3D solid model in any CAD environment. The designer could also evaluate the bounding volumes of CAD models. In a separate window, in fact, it is possible to evaluate the distance between two points and its components along assigned directions and, then, to easily assign different input values to gearbox parameters to fit, after the automatic re-generation of geometries, the desired bounding volume.

2.2 - The AHP model

The usability evaluation of the graphical user interface has been carried out by using Saaty's AHP [S1], essentially based on two steps: the first is to model the problem as a hierarchy to explore the factors of the problem at levels from general to detailed, then the scoring of the factors related to each level, by comparing them in pairs. In our case, the top level of the hierarchy is the usability of the GUI.

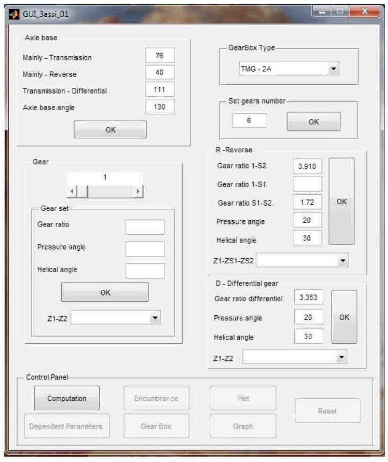


Figure 2: GUI for the gearbox CAD modeling software tool

The first decomposition can be made according to ISO 9241-11:1998 and, in *"Usability Dimensions"*, namely:

- *Effectiveness*: the measurement of the effectiveness related to the targets with the accuracy and completeness of the results achieved. The effectiveness value can be assigned in terms of accuracy related to the main task of the GUI (to assist the gearbox design);
- *Efficiency*: ratio between the effectiveness level and the use of resources, meant as time and computational operations;
- *Satisfaction*: user-perceived benefit and level of comfort felt during the use of the GUI. This dimension is strongly related to the subjective perception of user performance.

Starting from such considerations, the decomposition of the usability is shown in Figure 3. At the first level, there is the GUI's Usability (U) that is decomposed in Usability Dimensions (UDs) at second level. In turn, these dimensions are broken down at the next level in Usability Functions (UFs). Therefore, the last step in the model definition has been the definition of each *UF* that is strictly related to the design of the experimental task. The aim at this stage was to consider critical aspects in the usability assessment of the GUI. According to the hierarchical decomposition above

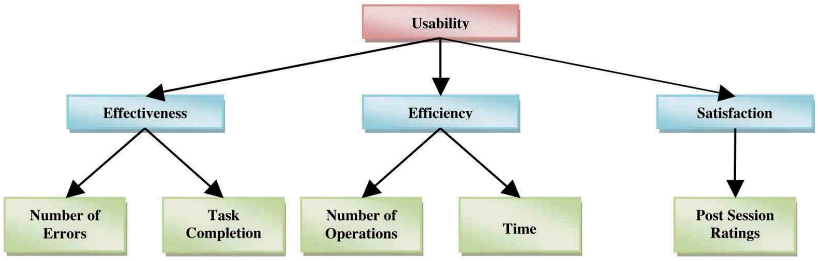


Figure 3: Usability hierarchical decomposition

described, the analysis of the characteristics provides the following *UFs*:

- *Number of Errors* (*E*), measure of effectiveness, is the number of error messages reported by the GUI during the task execution.
- *Task Completion (TC)*, measure of effectiveness, is the level of completion and accuracy in achieving the goals of the task.
- *Number of Operations* (*NO*), measure of efficiency, is defined as the number of operations used to complete a task in terms of mouse clicks and keystrokes.
- *Time* (*T*), measure of efficiency, is the effective time to perform a task or sub-activities.
- *Post Session Ratings* (*PSR*), measure of satisfaction, is a score, which expresses the feeling of users about the GUI use.

The *UFs* define the lowest level of the hierarchical model.

2.3 - The Usability Index definition

Starting from the assumption the factors of the hierarchy, for each level, are preferentially independent to each other, a simple linear additive evaluation model could be applied. By means of Multi-Criteria Decision Analysis (MCDA) all the measures corresponding to the factors could be combined into one overall value. In particular, the measure of each factor is multiplied by a weight based on a specific criterion, and then the weighted scores are summed up. The calculation of the index starts from the *UFs*, by using data collected during experiments. Being data of different nature and magnitude, a preliminary normalisation is required, in order to ensure the comparison between them.

The normalisation techniques, adopted for the specific UFs, are briefly described in the following:

- *0-Max* normalisation performs a linear transformation of the original data. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the equation (1):

$$e'_{ij} = \frac{e_{ij}}{\max_i} \quad (1)$$

- *Min/*e_{ij} normalisation performs a linear transformation of the original data that reverses the direction of preferences. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the equation (2):

$$e'_{ij} = \frac{\min_i}{e_{ij}} \quad (2)$$

- e_{target}/e_{ij} normalisation performs a linear transformation of the original data that reverses the direction of preferences and requires a target value. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the formula (3):

$$e'_{ij} = \frac{e_{target}}{e_{ij}} \quad (3)$$

The above techniques adopted for each UF are reported in Table 1.

Normalisation technique	Usability Functions
0-Max	*PSR*
Min/e_{ij}	*NE, TC*
e_{target}/e_{ij}	*NO, T*

Table 1: Normalisation techniques adopted for each UF

The outcomes of the normalisation procedure are the usability measures (um_i) that range from 0 to 1. Then, for each subgroup of usability measures, the usability dimension index (*UDI*) is defined (4):

$$UDI_i = \sum_{i=1}^{n} w_i \cdot um_i \quad (4)$$

where w_i is the weight of each usability measure, that could be different, based on the level of priority of usability measures in the specific application. The three usability dimension indexes are: 1- the effectiveness index; 2- the efficiency index; 3- the satisfaction index .

The weighted sum of these three indexes provides the overall results for the UI (5):

$$UI = \sum_{i=1}^{n} w_i \cdot UDI_i \quad (5)$$

In details, the AHP is applied in order to evaluate the relevance of the factors in the hierarchy, taking into account the analysis of GUI interaction. Starting from the hierarchy structure of the model, the matrix of weights has been defined. This matrix is accomplished for each level of the hierarchy and for each group (elements in the lower level hinge on the same element in the upper level), by placing the elements of the group both on matrix rows and columns. Hence, all the elements of the same group are compared in pairs. The generic matrix element a_{ij} is the result of the pairwise comparison between the attribute of the row *i-th* and the column *j-th*, with respect to a certain task, using the Saaty scale that is a 9-points scale anchored at the end with the terms *“Equivalent alternatives”* and *“The chosen alternative is absolutely better than the other one”*. Thus, the main diagonal of the matrix consists of unit elements only, while the values of the other cells are always positive, according to the reciprocity property (6):

$$a_{ij} = \frac{1}{a_{ij}} \quad (6)$$

Once the pairs comparison matrix has been defined, the weight of each element is assumed as (7):

$$w_i = \frac{\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}}; \quad i, j = [1,\ n] \quad (7)$$

where n is the dimension of the metrics related to the element at issue. In particular, the allocation of weights is done with a bottom-up logic, from the lowest level of the hierarchy (*UFs*) to the highest (Usability).

3- Experimental phase

3.1 – Overview of experiment

Preliminarily, a GUI *Tutorial* was defined to present the graphical interface, to explain the procedures for data entry and to discuss about the features of the interface. An example was also illustrated.

The first phase of testing involved the selection of a representative sample of the GUI user profiles. The minimum skills of the users were identified: 1) a good knowledge (at least theoretical) of a gearbox, 2) a good expertise with the use of graphical user interfaces, and more generally, with the use of specialized software. On this basis, participants were 12 newly graduated engineers (i.e. mechanical, electrical and management) selected, under the R&D project *Digital Pattern Development*, for highly specialized training in Computer-Aided Design.

To collect data an experimental session was prepared. In such session, each user had to complete two tasks: the first related to a new design activity, the second related to the modification of an existing gearbox model. As regards the first task, users were asked to design a new gearbox according to the specifications assigned, to assess the overall dimensions and, subsequently, to save the model. Then, for the second task, users were asked to modify the gearbox, designed in the previous task, according to new instructions, and, then, they were asked to carry out a new evaluation of the overall dimensions and to save again. At the end of the two previous tasks, the users were asked to fill a questionnaire; in particular, they were asked to express their agreement related to ten statements, all set in a positive sense, by using a seven-point scale, whose ends were the positions: *“strongly agree”* and *“strongly disagree”*.

3.2 - Experimental protocol

Several days before the test session, a preparatory meeting with the participation of users involved in the tests was accomplished. The purpose of the incoming experimentation and the functionality of GUI were presented. A detailed description of the *Tutorial* was given.

Then, on the day fixed for the tests and before starting the tests, users were informed, once again, that the aim of the experiment was to evaluate the GUI usability, and not the user’s ability to quickly perform a set of assigned tasks. In this way, we tried to minimize the “stress” that, generally, may affect the outcome of a proof. Furthermore, the inspectors of test session have shown the procedures of the experimental session, with particular attention to the rules of test performing. Then, they provided further details about the *Tutorial* and a short questionnaire for the personal details and for the informed agreement. The questionnaire was filled and returned before the start of the test. Finally, an ID code to each user was assigned.

Users tested the GUI individually and in random order to avoid noise factors. During the test, the inspectors recorded many details: the start and the ending time of the tasks, any notes on the bringing of the test, the number and kind of assistance provided. Specifically, if the user explicitly required the assistance of the inspector, then the inspector invited him to consult the *Tutorial* (classifying this as a *level 1* assistance). Otherwise, if the user was not able to continue the test, the inspector removed all doubts (classifying this as a *level 2* assistance).

3.3 - Data collection and processing

The experiments were performed at the Department of Industrial Engineering in a suitable room cleared furnishings and with a Visual Display Unit (VDU). The time limit for each test was set at 30 minutes after which the user was asked to suspend the operations. Furthermore, for each test was used a software that recorded all the user’s activities carried out on the monitor. Table 2 summarises the sources related to *UFs*.

As regards *UFs*, all the measures were collected. In the calculation phase, the total value of each *UF* was obtained as the sum of the measures/ratings respectively noted to perform both the Task 1 and Task 2 for the same user. This operation was repeated for all participant users.

Usability Dimensions	Usability Functions	Source
Effectiveness	Number of errors (NE) Task Completion (TC)	Video test Panel of experts
Efficiency	Number of operations (NO) Time (T)	Video test

Satisfaction	Post-session ratings (PSR)	Questionnaire

Table 2: The sources of *UFs*

Effectiveness: the *Number of Errors (*NE*)* was derived from the video by counting each time the GUI reported an error message; the *Task Completion (*TC*)* was measured using a rating given by a panel of experts who were asked to assess the completeness of the goals reached for all the activities performed in the test by using the following six-point scale: (1) Complete success without assistance, (2) Complete success with assistance, (3) Partial success without assistance, (4) Partial success with assistance, (5) Failure: the user didn't understand that the task is not complete, (6) Failure: the user does not completed the task despite the assistance. The references to determine the level of completion in task execution were decided beforehand.

Efficiency: the *Number of Operations (*NO*)* was derived from video by counting, from time to time, the operations that were performed to complete the task; the *Time* (T*)* was measured by the inspector as the difference between the ending and the beginning time of the session. This measure was subsequently validated by a comparison with the clock of VDU shown in video recordings.

Satisfaction: the *Post-Session Ratings (*PSR*)* were got from the post-session questionnaire that users filled out at the end of the test.

For the calculation of the UI the average values arithmetic mean) of all the aforementioned *UFs* were used.

4- Results

The following data were obtained. Table 3 shows the normalized measures of UFs split among users and subgroups.

ID user	Effectiveness: Number Errors (NE)	Effectiveness: Task Completion (TC)	Efficiency: Number of Operations (NO)	Efficiency: Time (T)	Satisfaction: PostSession Ratings (PSR)
1	0.50	0.25	0.84	0.43	0.56
2	0.20	0.33	0.80	0.57	0.95
3	0.10	0.33	0.47	0.36	1.00
4	0.17	0.33	0.75	0.37	0.87
5	0.25	0.40	0.77	0.49	0.82
6	0.25	0.40	0.70	0.46	0.93
7	0.11	0.40	0.53	0.45	0.46
8	0.08	0.33	0.49	0.24	0.46
9	0.20	0.40	0.63	0.32	0.90
10	0.10	0.50	0.78	0.43	0.72
11	0.33	1.00	0.85	0.58	0.92
12	1.00	0.40	0.64	0.32	0.84

Table 3: Normalized measures of *UFs*

According to the above UI definition (par. 2.3), the average values of UFs were obtained using the collected measurements. The weights of UFs were obtained submitting Saaty's questionnaire to a panel of experts and using the Equation 7, as summarised in Table 4.

UFs	NE	TC	NO	T	PSR
w_i	0.74	0.26	0.25	0.75	1
um_i	0.27	0.42	0.69	0.42	0.79

Table 4: Weights and values of *UFs*

Likewise, the weights and values of UDs were obtained (Table 5). Therefore the index of usability integrated obtained is equal to 0.42 (UI). This is a low value.

UD	$UDI_{effectiveness}$	$UDI_{efficiency}$	$UDI_{satisfaction}$
w_i	0.59	0.31	0.10
UDI_i	0.31	0.49	0.79

Table 5: Weights and values of *UDs*

The table 5 shows that the satisfaction index ($UDI_{satisfaction}$) is the highest (79%). However, this value could be smaller as the degree of satisfaction of the users could be influenced by the achievement of the goal (i.e., the effective gearbox modeling) rather than the difficulties they overcome in using interface. In any case, the satisfaction is the usability dimension that has a limited effect on the calculation of the global index. So, the results in Table 5 suggest that to improve the usability of the GUI it is necessary to increase the values of effectiveness and efficiency, improving the functions *Time* and the *Number of Errors*, respectively (Table 4). For a more detailed analysis, both Task 1 and Task 2 were divided into the following critical sub-activities: the choice of the gearbox architecture, the choice of the gears number, the setting of the dependent parameters, the setting of the wheel parameters, the setting of the reverse gear, the setting of the differential parameters, the overall dimensions, the procedure for file exporting. In this perspective, the measurement of efficiency was analysed.

Figure 4 shows the radar chart that highlights how the normalized average value, related to the number of operations due for each sub-task, is different from the normalized optimal value, equal to 1. For example, the value equal to 3, related to one of the axes, means that, on average, the number of operations necessary to accomplish that subtask, in order to obtain the required result, were equal to 3. Figure 4 shows that the more critical sub-tasks, involved in the first task, are (in descending order): the setting of the parameters of the differential (6.10), the setting of the reverse gear (5.29), the file exporting (4.42), the setting of the gear parameters (2.43). Further results for Task 2 are: the setting of the reverse gear (4.17) and setting of the differential parameters (2.35).

Also the Time was a critical UF. In fact, if we consider only the users who accomplished the Tasks, in a complete and without assistance way, the average Time was found to be almost double than the predetermined optimum value. More generally, the Time to complete the Task 1 was greater than the one related to the Task 2 (Fig. 5).

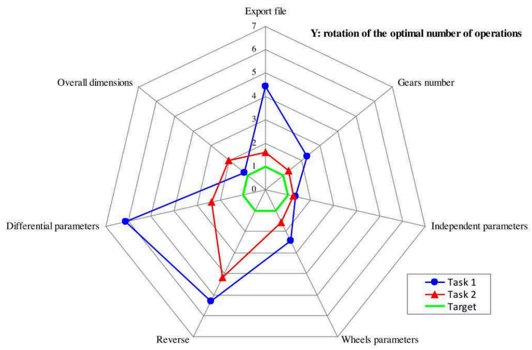


Figure 4: Rotation of the Number of Operations by sub-activity

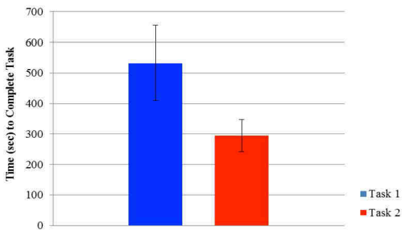


Figure 5: Average *Time* to complete Task 1 and Task 2

Tackling the measures of effectiveness, it is possible to highlight that there were no significant differences between the number of errors related to the Task 1 and Task 2, but the average values are not negligible (Fig. 6).

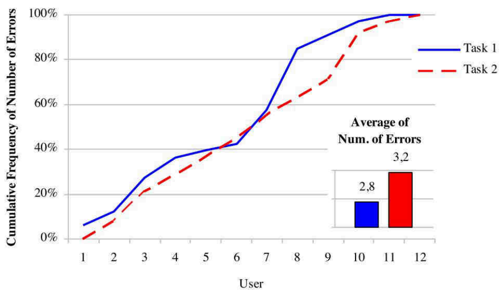


Figure 6: Cumulative frequency of number of errors related to each task and average values

The determinations of *Task Completion*, grouped using the level of success, are depicted in Figure 7. It should be noted that most users carried out the Task 1 in a complete success. Otherwise, in the accomplishment of Task 2, only 1 user had completely achieved the goals (complete success), while 10 users have had a partial success. There was also 1 user who had failed.

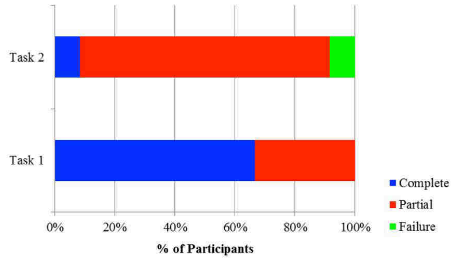


Figure 7: Stacked bar chart showing different levels of success based on task completion

The measures of satisfaction were analysed. Figure 8 shows that the lowest value of the satisfaction is related to the clarity and effectiveness of the GUI (D8), while the highest value is related to the actual benefit perceived in the use of the GUI, to improve the gearbox design (D5).

Based on what has been analysed, after careful viewing of the videos of the tests, the following critical issues were identified: 1) the default fields are not empty; 2) the interface does not allow users to overwrite the selected values in the fields, so the users are forced to cancel the existing; 3) the user does not have a feedback on the correct setting of the parameters; 4) the provision of the sections in the main window is not consistent; 5) the function for file exporting is poorly visible; 6) the reset command has an insufficient functionality.

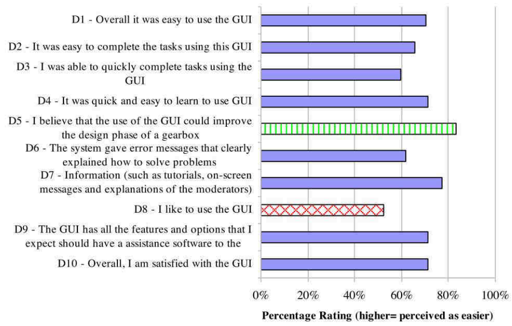


Figure 8: Average subjective ratings split by statement

These problems involve a difficulty in achieving the Tasks and, they generally cause an increase of the operating time also due to a more than proportional increasing number of operations to be performed. In some cases, the user is confused and, then, he is led to the error or makes continuous action controls.

Once these critical issues have been identified, it has been decided to design the GearBox 2.0 that was developed keeping unchanged the dimensions of initial GUI, avoiding radical changes and aiming to improve efficiency, and effectiveness. A series of new features were introduced (Figure 9). A new provision of sections, with different background colours, that promotes the logical procedure for the input of project data was introduced. A new visibility to the button aimed to file exporting was given. The buttons on the control panel were divided, and a section with the new *"TEST"* function, the upgraded *"RESET"* function and a new confirmation "SET" function were inserted. The latter turns on only when the input parameters are correct. In this way it provides an immediate feedback to the user. To assess the validity of the new features a pairwise comparison between the Gearbox 2.0 GUI and the initial version was performed. For each change users were asked if they preferred the original version or the GearBox 2.0. They expressed their degree of preference on a scale of six points. The survey results demonstrated a preference of the users for the new GUI far higher than the initial one. Therefore, all the features of Gearbox 2.0 GUI were definitively implemented and new tests were started for evaluating the UI, according a continuous improvement loop.

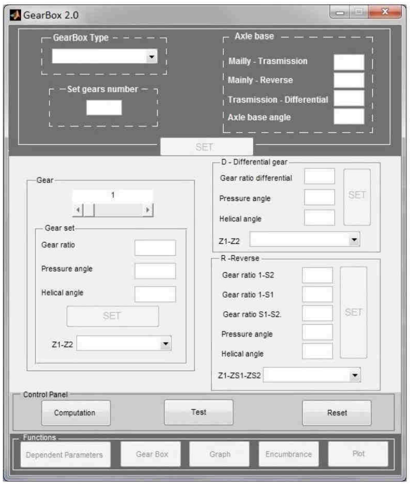


Figure 9: The Gearbox 2.0 GUI

5- Conclusions

In this paper the characteristics of the GUI belonging to a software tool for gearbox design were analysed. Then, an AHP model was proposed and an experimentation to calculate the integrated UI was performed. In particular, the experimental data, which have led to a lower UI, were collected and analysed. Therefore, a new release of the GUI was proposed. Further experimentations were carried out in order to validate the new release of the GUI and a positive evaluation of the new features was accomplished. This improvement induces to state that the use of the new release of the GUI could contribute to reduce the development time in gearbox design and, therefore, the overall *time to market* for new gearboxes. Further steps deal both with a further feature evaluation and improvement of the new GUI. In fact, by using the present approach, new characteristics of GUI could be introduced and evaluated.

6- Acknowledgements

The present work was developed with the contribution of the Italian Ministry of University and Research (MIUR) performing the activities of the PON01_01268 DIGIPAT Project, *Digital Pattern Product Development: A Pattern Driven approach for industrial product design, simulation and manufacturing*.

The authors also thank prof. Giuseppe Di Gironimo (Department of Industrial Engineering, University of Naples Federico II), coordinator of DIGIPAT Formation Project for selecting the GUI users, and the engineers who took part to the experiments. Finally, the authors thank eng. Alfonso Vitolo for his technical support.

7- References

[B1] Bourke D. Software gives engineers quick access to product info from layers of company files. Mechanical Engineering: 44-47, 2013.

[DM1] Di Gironimo G., Matrone G., Tarallo A, Trotta M, Lanzotti A. A virtual reality approach for usability assessment: case study on a wheelchair-mounted robot manipulator. Engineering with Computers 29 (3): 359-373, 2013.

[LP1] Lanzotti A., Patalano S., and Vitolo F. A graph-based approach to CAD modeling: a digital pattern application to the sizing and modeling of manual transverse gearboxes. In International Conference on Graphic Engineering, Madrid, 2013.

[MA1] Madhavan R., Alagarsamy K. Usability issues in software development lifecycle. International Journal of Advanced Research in Computer Science and Software Engineering, 3 (8): 1331-1335, 2013.

[PV1] Patalano S., Vitolo F., and Lanzotti A. A graph-based software tool for the CAD modelling of mechanical assemblies. In Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization: 60-69, Barcelona, DOI: 10.5220/0004299000600069, 2013.

[S1] Saaty L. Decision making with the analytic hierarchy process. International Journal of Services Sciences 1 (1): 83-98, 2008.

[S2] Sandeberg M. Knowledge based engineering – in product development. ISSN: 1402-1536, 2003.

[SB1] Sharmin M., Bailey B. P., Coats C., Hamilton K. Understanding knowledge management practices for early design activity and its implications for reuse. In Proceedings of the 27th International Conference on Human Factors in Computing Systems: 2367-2376, 2009.

[SK1] Sohaib O., Khan K. Integrating usability engineering and agile software development. In Proceedings of the International Conference on Computer Design and Application, volume 2: 32-38, 2010.

Paper D

Patalano S., Del Giudice D.M., Gerbino S., Lanzotti A., Vitolo F., to be submitted to International Journal of Interactive Design and Manufacture, ISSN 1952513 e 19552505 (Q1 nel 2013), 15 pp.

On the usability assessment of a new digital pattern tool graphical user interface

S Patalano (corresponding author)[1], A Lanzotti[1], D M Del Giudice[1], F Vitolo[1], S Gerbino[2]

[1] Fraunhofer JL IDEAS-COGITO
University of Naples "Federico II", Industrial Engineering Department (DII) Piazzale Tecchio 80 – 80125 Napoli, Italy

[2] DiBT Dept., University of Molise, Via Duca degli Abruzzi, Termoli (CB), Italy

Corresponding author details

Postal address: Unversità degli Studi di Napoli "Federico II", Dipartimento di Ingegneria Industriale, Piazzale Tecchio 80, 80125, Napoli, Italy
Tel: +39 0817682457
E-mail: stanislao.patalano@unina.it

Abstract

Design team spend up to 30% of their time to searching data and this percentage rises up to 50% if you take into account the time spent to their validating. To speed up these design processes, the use of a knowledge-based engineering (KBE) system is recommended. A critical point in the development of a KBE system is the interface usability for demonstrating an effective reduction of development time and satisfaction in its use. This work tackles the usability improvement of the KBE system's Graphical User Interface (GUI) which assists designing of automotive manual transverse gearboxes, through participatory tests. An approach based on iterative design has been used. A new usability index (UI) that uses methods as Analytic Hierarchy Process (AHP) and Multiple-Criteria Decision Analysis (MCDA) is presented. In the experimental phase, twelve newly graduated engineers specialised in Computer-Aided Design, have completed two specific tasks. Objective and subjective measures have been collected. Data analysis highlighted some critical issues. Then, a new release of the GUI has been developed and the validation test has been performed. Results have demonstrated an significant improvement.

Key words: Usability assessment; Graphical User Interface; gearbox design; participatory design.

1 Introduction

Large and small companies develop products through structured work teams supported by software toolsets aimed to keep up their design [1]. These tools are generally complex and require skilled users dealing with design, test and check activities. The main issue is that these users are geographically dispersed and interdepartmental [2], besides the design and manufacturing process often aren't concurrent but they turn in the loop. This induce a data flow loop which move through some division in the world. Systems, procedures and software to capture and manage design and manufacturing data are necessary to ride out these issues. Some authors [3-4] propose software tools to concurrent manage design and manufacturing process data. In such context, a "Digital Pattern" (DP) platform is recommended. A DP platform is a set of geometric and numeric data structures, as well as of preconfigured and parametric models, which can be adapted to specific contexts. Therefore, a DP acts as a knowledge-based expert (KBE) system aimed to improve quality and reduce times and costs for product development through a massive use of knowledge and tools integration inside company. Such aims are accomplished through the fast and the best re-use of company knowledge, i.e. through technical and technological predefined solutions that quickly marry new projects, allowing fast performance evaluations and immediate checks. Such solutions should also be able to give feedbacks on design and production costs. In [5-6] a DP system, developed to assist gearbox design, is described.

When developing a KBE system, the evaluation of interface usability, to demonstrate the effective reduction of development time and satisfaction in use, is a critical point. Indeed, usability can be defined as *the extent to which specific users, in a specific context of utilisation, can use a product to their satisfaction, in order to effectively and efficiently achieve specific goals* [7]. Sohaib and Khan [8] claim that agile projects needs to adopt aspects of *usability engineering* by incorporating user scenarios and including usability specialists in the team. During the designing of mechanical parts, designers need to verify the correctness of the hypotheses in use, especially in relation with multi-objective tasks [6]. Furthermore, the use of a KBE system is strategic for designers if we consider that they spend up to 30% of their time to searching data [9] and this percentage rises up to 50% when we take into account the time spent to validating the data [10].

Following the approach in [11] this paper proposes the usability improvement of the KBE system's Graphical User Interface (GUI) through a participatory testing. This GUI assists designing automotive manual transverse gearboxes. An approach based on Analytic Hierarchy Process (AHP) and Multiple-Criteria Decision Analysis (MCDA) is used. A participatory test was performed for evaluating the Usability Index (UI) of the GUI. The AHP approach implies the decomposition of the problem into several levels [12]. In the present work, three dimensions of UI are considered: Effectiveness, Efficiency and Satisfaction. The MCDA methodology implies that all the measures corresponding to the factors of the problem, being of different nature and magnitude, are first normalized and weighed, and then could be combined into one overall value through a bottom-up approach [16-17].

For the experimental phase, a set of selected end-users has to complete two specific tasks: 1) to design an automotive manual transverse gearbox; 2) to modify an existing gearbox model re-using previous knowledge. The goals of both Task 1 and Task 2 are clearly defined. Then, measures of subjective ratings and objective metrics are collected. So, the UI is calculated and the effects of the usability dimensions are analysed.

Taking into account the experimental data analysis, some frequent critical issues are identified. Before making any changes to the GUI, a questionnaire is administrated to the same users of the previous experiment to confirm the validity of new conceptual features proposed.

In this perspective, a new release of the GUI is developed and the validation test is again performed for a new assessment of the GUI, according to a continuous improvement loop.

The paper is arranged as follows: Section 2 summarises the methodological approach; Section 3 provides the experimental phases, Section 4 deals with results of experimentations; finally, Section 5 draws conclusions.

2 The Methodological approach

The traditional design process tends to favour the functional aspects of a product at the expense of the cognitive-emotional ones, not considering that a product can even have only an emotional function [13]. The usability assessment must take into account the analysis of both objective and subjective aspects that are closer to the emotional sphere of the individual. In this respect, the participation of the end-user into design process is crucial according to a User-Centred Design approach. Using such assumptions and starting from [11], the approach adopted to achieve the purpose of this paper requires the involvement of potential users during all phases of the usability evaluation process [14]. Figure 1 shows the logical flow chart of proposed methodological approach.

In summary, both the user profile and the GUI characteristics are identified. Given these requirements and the context of use, the develop design solutions are implemented and the usability testing are planned. Two specific tasks, devoted to translate the usability characteristics factors into measurable usability functions, are properly defined. In order to reduce the noise related to the user's skill, a training phase is conducted for all users. Then, the experimental data are collected and the GUI evaluation is settled.

As stated before, the GUI usability assessment is carried out by using Saaty's AHP [12] and MCDA methodology [15]. The analysis can be summarised as follows:

1) Decomposition of the problem into several hierarchical levels and factors;

2) Scoring of the factors related to each identified level by means of pairwise comparison.

In particular, MCDA methodology allows combining the values of the individual usability functions into a single usability index (UI) by means of a bottom-up approach [16-17].

Starting from results of experimental data, some changes are proposed. Finally, the validation experiment is performed to verify the goodness of these changes.

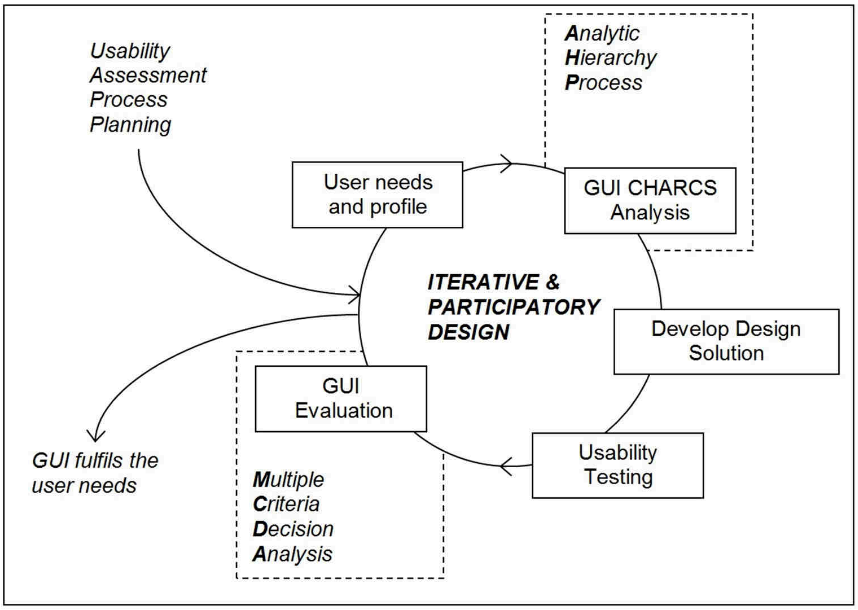


Figure 1: The logic flow chart of methodological approach

2.1 - User Profile Definition and GUI characteristics

A KBE system can assist engineers in re-using previous knowledge in order to improve time-consuming tasks, as retrieval and selection of previous architectures, and to modify and virtually test a new product design. A critical point in the development of a KBE system is the interface usability to demonstrate effective reduction of development time and satisfaction in its use. Specifically, the present work deals with a KBE system previously proposed and providing to assist within the design of automotive manual transverse gearboxes. Then, the GUI for the KBE system is released (Figure 2).

GUI interaction depends primarily on the kind of user and the context of use. All characteristics, which identify specific needs, desires and interests and even behaviours, contexts of use and personal preferences [18], define a specific user profile. The MatLAB®-based GUI is accomplished for junior designers belonging to automotive industry. Their minimum skills is defined as follows: 1) good knowledge (at least theoretical) of a gearbox, 2) good expertise with the use of graphical user interfaces, and more generally, with the use of specialised software. Hence, the GUI should easily satisfy user needs with no cognitive overload.

The GUI is designed to perform two major tasks: 1) to design gearboxes rapidly, reducing the risks of using incomplete information when making product development decisions; 2) to assist the designer in redesigning the gearboxes previously developed (for more details see [5-6]).

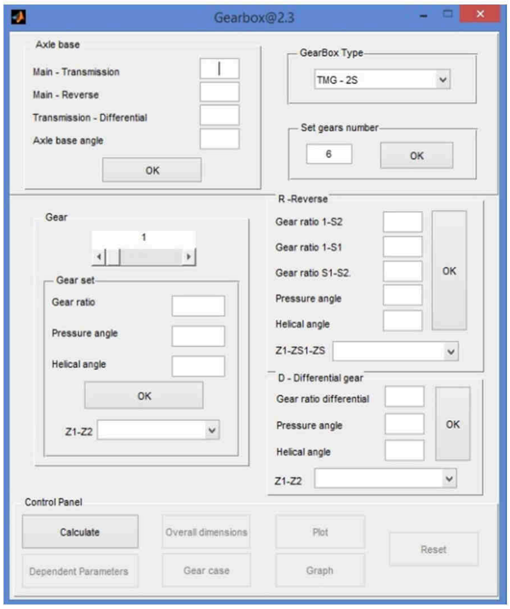


Figure 2: GUI for the gearbox CAD modelling software tool

The window is divided in three main fields (Figure 2): the upper field where the gearbox is pre-configured; the middle field where the gearbox is configured and the gears are characterised; the lower field where the post-processing and evaluating commands are located.

The designers could set: type of gearbox (selecting it from a list box); number of gears (up to six); layout parameters (axle bases and angle between them). Besides, designers could set three characteristic parameters for each gear: gear ratio, pressure angle and helical angle (all in viable range). Teeth numbers of the gear are automatically generated by means of an algorithm that pulls out ten set of teeth numbers that meet the three parameters of the gear.

Specific panels to set reverse and differential gear are developed. A new gearbox can be defined by setting these parameters, but a previous i.e. existing design can also be edited.

The computational structure is guided by a directed graph (digraph). The nodes are associated to parameters (dependent and independent), while directed edges represent the mathematical relationships among parameters. The “Graph” button displays the digraph in a new window where the designer is able to interact with the graph: for example removing relations and generating an isolated node, in order to set a constant value during the calculation step.

As regards the post-processing, the “Dependent Parameters” command displays all computed parameters useful to determine the geometry correctness (modules, pitch diameters, addendum (or outside) diameters, root diameters, base diameters, etc).

The “Plot” command and “Gear Case” command display the mesh representation of the gears and the gear case, respectively. The automatically generated models can be exported both as .txt and .stl files, so ensuring the generation of the corresponding 3D solid models in any CAD environment.

The bounding box of generated meshes can be shown and this option helps the designer to interactively set the gearbox parameters as to fit the whole gearbox within a desired volume.

2.2 - The AHP model for GUI usability

Figure 3 shows the decomposition of the usability according to the AHP approach. At first level we set the GUI usability (U) that is decomposed according to [17-18] in Usability Dimensions (UDs), whitin the second level. The UDs are defined as follows:

- Effectiveness, the level of accuracy and completeness with which users achieve a specified goal;
- Efficiency, the level of effectiveness achieved to the expenditure of resources;
- Satisfaction, the condition of freedom from discomfort and positive attitude towards the use of the product.

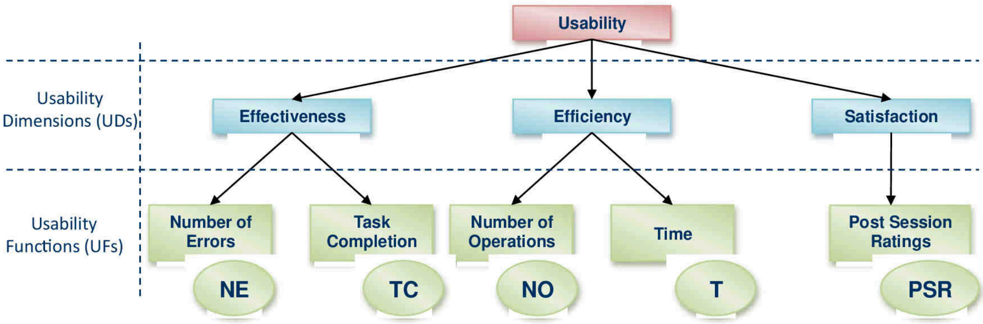


Figure 3: GUI Usability hierarchical decomposition

In turn, the UDs are broken down, at the third level, in Usability Functions (UFs) that are strictly related to the experimental tasks. These UFs, in fact, are accurately determined during the experimental phase and they tackle critical aspects for GUI usability assessment.

According to the hierarchical decomposition above described, the analysis of GUI characteristics provides the following UFs:

- *Number of Errors* (NE), measure of Effectiveness, is the number of error messages reported by the GUI during the task execution;
- *Task Completion* (TC), measure of Effectiveness, is the level of completion and accuracy in achieving the goals of the task;
- *Number of Operations* (NO), measure of efficiency, is defined as the number of operations used to complete a task in terms of mouse clicks and keystrokes;
- *Time* (T), measure of efficiency, is the effective time to perform a task or sub-activities;
- *Post Session Ratings* (PSR), measure of Satisfaction, is a score, which expresses the feeling of users about the GUI use.

2.3 - The usability Index definition

Starting from the assumption that the factors of the hierarchy, for each level, are preferentially independent to each other, a simple linear additive evaluation model could be applied. By means of Multi-Criteria Decision Analysis (MCDA) all the measures corresponding to the factors could be combined into one overall value [16-17]. In particular, the measure of each factor is multiplied by a weight based on a specific criterion, and then the weighted scores are summed up. The calculation of the index starts from the UFs, by using experimental data. Being data of different nature and magnitude, a preliminary normalisation is required, in order to ensure the comparison between them.

The normalisation techniques, adopted for the specific UFs, are briefly described in the following:

- *0-Max* normalisation performs a linear transformation of the original data. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the equation (1):

$$e'_{ij} = \frac{e_{ij}}{max_i} \qquad (1)$$

- *Min/*e_{ij} normalisation performs a linear transformation of the original data that reverses the direction of preferences. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the equation (2):

$$e'_{ij} = \frac{min_i}{e_{ij}} \qquad (2)$$

- e_{target}/e_{ij} normalisation performs a linear transformation of the original data that reverses the direction of preferences and requires a target value, lower than the minimal value. The considered value e_{ij} is transformed in a new value e'_{ij} ranged in the interval [0, 1] using the formula (3):

$$e'_{ij} = \frac{e_{target}}{e_{ij}} \qquad (3)$$

The above techniques adopted for each UF are reported in Table 1.

Table 1: Normalisation techniques adopted for UFs

Normalisation technique	Usability Functions
$0 - max$	PSR
min/e_{ij}	NE, TC
e_{target}/e_{ij}	NO, T

The outcomes of the normalisation procedure are the usability measures (um_i) that range from 0 to 1. Then, for each subgroup of usability measures, the Usability Dimension Index (UDI) is defined (4):

$$UDI_i = \sum_{i=1}^{n} w_i \cdot um_i \qquad 4)$$

where w_i is the weight of each usability measure, that could be different, based on the level of priority of usability measures in the specific application. The three usability dimension indexes are: 1) the Effectiveness index; 2) the Efficiency index; 3) the Satisfaction index.

The weighted sum of these three indexes provides the overall results for the UI (5):

$$UI = \sum_{i=1}^{n} w_i \cdot UDI_i \qquad (5)$$

In details, the AHP is applied in order to evaluate the relevance of the factors in the hierarchy, taking into account the analysis of GUI interaction. Starting from the hierarchy structure of the model, the matrix of weights is defined. Such matrix is accomplished for each level of the hierarchy and for each group (elements in the lower level hinge on the same element in the upper level), by placing the elements of the group both on matrix rows and columns. Hence, all the elements of the same group are compared in pairs. The generic matrix element a_{ij} is the result of the pairwise comparison between the attribute of the row *i-th* and the column *j-th*, with respect to a certain task, using the Saaty scale i.e. a 9-points scale anchored at the end with the terms *"Equivalent alternatives"* and *"The chosen alternative is absolutely better than the other one"*. Thus, the main diagonal of the matrix consists of unit elements only, while the values of the other cells are always positive, according to the reciprocity property (6):

$$a_{ji} = \frac{1}{a_{ij}} \qquad (6)$$

Once the pairs comparison matrix is defined, the weight of each element is assumed as (7):

$$w_i = \frac{\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}}; \quad i, j = [1,\ n] \qquad (7)$$

In (7), *n* is the dimension of the metrics related to the element at issue. In particular, the allocation of weights is done with a bottom-up logic, from the lowest level of the hierarchy (UFs) to the highest (Usability).

3 Experimental phase

3.1 – Overview of experiment

Based on of the requirements identified in section 2.1, participants are 12 newly graduated engineers (*i.e.* mechanical, electrical and management) aged between 27-32 years, attending a specialised course in Computer-Aided Design within a project named *Digital Pattern Product Development*.

The experiments are performed at the Fraunhofer JL IDEAS-COGITO laboratory, Department of Industrial Engineering of the University of Naples (IT) in a suitable room, with no furniture and equipped with a Visual Display

Unit (VDU). Preliminarily, a GUI *Tutorial* is defined to present the graphical interface, to explain the procedures for data entry and to discuss about the functions of the interface. An example is also illustrated.

An experimental session is performed. In such session, each user has to complete two specific tasks: 1) to design an automotive manual transverse gearbox; 2) to modify an existing gearbox model re-using previous knowledge.

The goals of the Task 1 are the follows:

- to design a new gearbox according to the specifications assigned (*i.e.* the parameters of six gears, of differential and of reverse);
- to plot the gearbox designed;
- to assess the overall dimensions;
- to save the model;
- to export the model.

Whereas, the goals of the Task 2 are the follows:

- to modify the gearbox designed in the first task, according to new instructions (*i.e.* it was asked to change the some parameters of the gears and of the layout controller);
- to plot the gearbox modified;
- to assess the overall dimensions;
- to save the model;
- to export the model.

Therefore, the measures of subjective ratings (TC and PSR) and objective metrics (NO, T and NE) are obtained.

3.2 - Experimental protocol

Several days before the test session, a preparatory meeting with the participation of users involved in the tests is accomplished. The purpose of the incoming experimentation and the functionality of GUI are presented. A detailed description of the *Tutorial* is given.

On the day fixed for the tests and before starting, users are informed, once again, that the aim of the experiment is to evaluate the GUI usability, and not the user's ability to quickly perform a set of assigned tasks. In this way, we try to minimize the "stress" that, generally, may affect the outcome of a proof. The inspectors show the procedures of the experimental session, with particular attention to the rules of test performing. Then, they provide further details about the *Tutorial* and they administrate the short questionnaire for the personal details and for the informed agreement to users. The questionnaire is filled and returned before the start of the test. Finally, an ID code to each user is assigned.

Users test the GUI individually and in random order to avoid noise factors. During the test, the inspectors record many details: the start and the ending time of the tasks, any notes on the bringing of the test, the number and kind of assistance provided. Specifically, if the user explicitly required the assistance, then the inspector invites him to consult the *Tutorial* (classifying this as a *level 1* assistance). Otherwise, if the user was not able to continue the test, the inspector removes all doubts (classifying this as a *level 2* assistance). The time limit for each test is set at 30 minutes after which the user is asked to suspend the operations [22]. After completing both tasks the questionnaire is administered to each user for detecting the PSR measurement.

The procedures described above is also applied to the validation test.

3.3 - Data collection and processing

All the UFs measures are collected. Table 2 summarises the sources related to UFs. In particular, an open source software is used to record all user's activities carried out during the experimental phase. Such tool is used to collect the NE, TC, NO and T metrics. The Effectiveness metrics are described in the following.

Table 2: The sources of UFs

UDs	UFs	Source
Effectiveness	Number of errors (NE)	Video test
	Task Completion (TC)	Panel of experts
Efficiency	Number of operations (NO)	Video test
	Time (T)	
Satisfaction	Post-Session Ratings (PSR)	Questionnaire

Effectiveness metrics:

- The *Number of Errors* (NE) are derived from the video by counting each time the GUI reports an error message.
- The *Task Completion* (TC) are measured using a rating given by a panel of experts who are asked to assess the completeness of the goals reached for all the activities performed in the test, by using the following six-point scale: (1) Complete success without assistance, (2) Complete success with assistance, (3) Partial success without assistance, (4) Partial success with assistance, (5) Failure: the user does not understand that the task is not complete, (6) Failure: the user does not complete the task despite the assistance. The references to determine the level of completion in task execution is decided beforehand [21].

Efficiency metrics:

- The *Number of Operations* (NO) are derived from video by counting, from time to time, the operations that are performed to complete the task.
- The *Time* (T) is measured by the inspector as the difference between the ending and the beginning time of the session. This measure is subsequently validated by a comparison with the clock of VDU shown in video recordings.

Satisfaction metrics:

- the *Post-Session Ratings* (PSR) are gained from the specific questionnaire that users filled out at the end of the session test (i.e. both Task 1 and Task 2). In particular, they are asked to express their agreement related to ten statements, all set in a positive sense, by using a seven-point scale, whose ends were the positions: *"strongly agree"* and *"strongly disagree"*.

In the calculation phase, the total value of each UF is obtained as the sum of the measures/ratings respectively noted to perform both the Task 1 and Task 2, for the same user. This operation is repeated for all users. For the calculation of the UI, the average values (arithmetic mean) of all aforementioned UFs are used.

The procedures described above to collect and process data are also applied to the validation test.

The experiment in numbers: 1 laboratory was used; 1 usability team consisting of 6 engineers, 1 panel of experts and 12 end-users were involved; 2 usability testing sessions were performed; over 4 hours of video footage and ca. 200 questions were examined.

4 Results and discussion

The following data are obtained. Table 3 shows the normalized measures of UFs for each user. According to the above UI definition (section 2.3), the average values of UFs are obtained using the collected measurements. The weights (w) of UFs are obtained submitting Saaty's questionnaire to a panel of experts and using the Equation 7, as summarised in Table 4.

Table 3: Normalized measures of UFs

	Effectiveness		Efficiency		Satisfaction
ID user	Number Errors (NE)	Task Completion (TC)	Number of Operations (NO)	Time (T)	Post Session Ratings (PSR)
S.1	0.50	0.25	0.84	0.43	0.56
S.2	0.20	0.33	0.80	0.57	0.95
S.3	0.10	0.33	0.47	0.36	1.00
S.4	0.17	0.33	0.75	0.37	0.87
S.5	0.25	0.40	0.77	0.49	0.82
S.6	0.25	0.40	0.70	0.46	0.93
S.7	0.11	0.40	0.53	0.45	0.46
S.8	0.08	0.33	0.49	0.24	0.46
S.9	0.20	0.40	0.63	0.32	0.90
S.10	0.10	0.50	0.78	0.43	0.72
S.11	0.33	1.00	0.85	0.58	0.92
S.12	1.00	0.40	0.64	0.32	0.84

Table 4: Weights and values of UFs

	NE	TC	NO	T	PSR
w_i	0.74	0.26	0.25	0.75	1.00
um_i	0.27	0.42	0.69	0.42	0.79

Likewise, the weights and values of UDs are obtained (Table 5). So, it is calculated the usability index: the value obtained is equal to 0.42, but this is not acceptable.

The Table 5 shows that the Satisfaction index ($UDI_{satisfaction}$) is the highest (79%). However, this value could be smaller as the degree of Satisfaction of the users could be influenced by the achievement of the goal (*i.e.* the effective gearbox modeling) rather than the difficulties they overcome in using interface. In this case, the Satisfaction is the usability dimension that has a "reduced" effect (10%, see Table 5) on the calculation of the global index. Hence, results in Table 5 suggest that the primary strategy for improving the usability of the GUI is to increase in Effectiveness values by acting mainly on the usability functions *Number of Errors* (NE). Whereas achieving a higher Efficiency value, by leveraging on the usability functions *Time*, may be the second strategy to improve the GUI usability.

Table 5: Weights and values of UDs

	Effectiveness	Efficiency	Satisfaction
w_i	0.59	0.31	0.10
UDI_i	0.31	0.49	0.79

Furthermore, both Task 1 and Task 2 are divided into the following critical sub-activities: the choice of the gearbox architecture, the choice of the number of gears, the setting of the dependent parameters, the setting of the wheel parameters, the setting of the reverse gear, the setting of the parameters of *differential*, the overall dimensions, the procedure for file exporting. In this perspective, the measurement of Efficiency is analysed.

Figure 4 shows the radar chart that highlights how the normalized average value, related to the number of operations due for each sub-task, departs from the normalized optimal value (equal to 1). For example the value 3, related to one of the axes, means that, on average, the number of operations necessary to accomplish that specific subtask is three times higher than the ideal value. Figure 4 points out that the more critical sub-tasks, involved in the Task 1, are (in descending order): the setting of the parameters of the differential (6.10), the setting of the reverse gear (5.29), the file exporting (4.42), the setting of the gear parameters (2.43). Further results for Task 2 are: the setting of the reverse gear (4.17) and setting of the differential parameters (2.35).

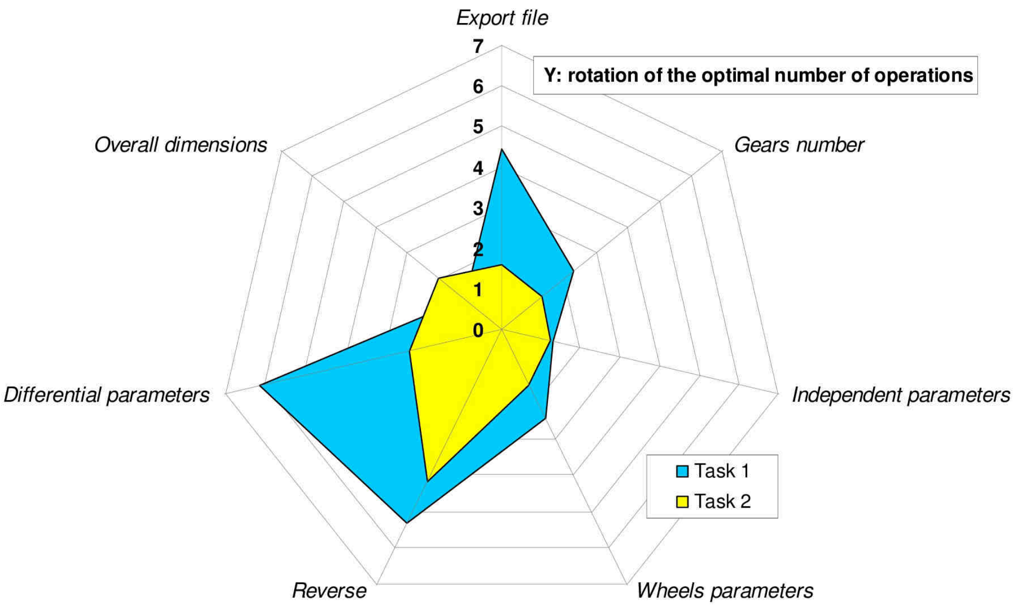


Figure 4: Number of Operations related to each sub-activity

Time is another critical UF. It's worthwhile to note that if we consider only the users who complete the tasks with success and without assistance, the average *Time* recorded is almost double than the predetermined optimum value. More generally, the average additional time to complete the Task 1 is much greater than the one related to the Task 2 as well as the variability of the measures (Figure 5). This may indicate a good level of learnability of the GUI.

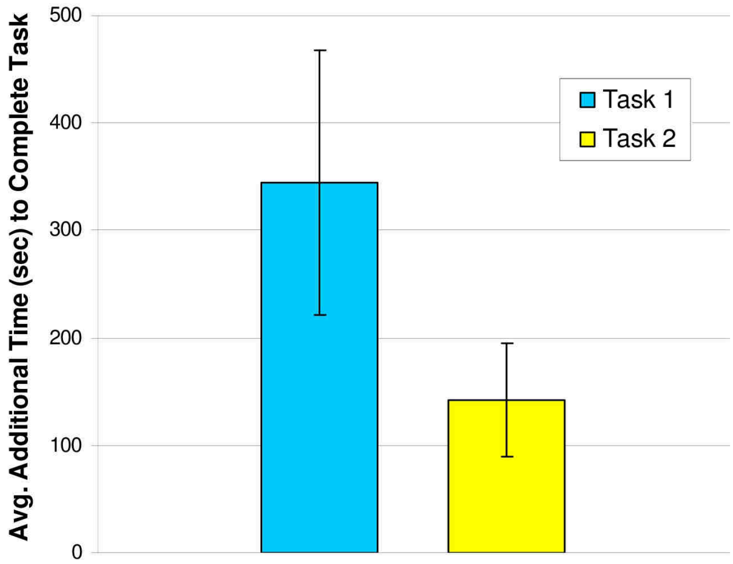


Figure 5: Average Additional Time to complete Task 1 and Task 2

Tackling the measures of Effectiveness, it's worthwhile to note that there are no significant differences between the number of errors related to the Task 1 and Task 2, but the average values are not negligible (Figure 6).

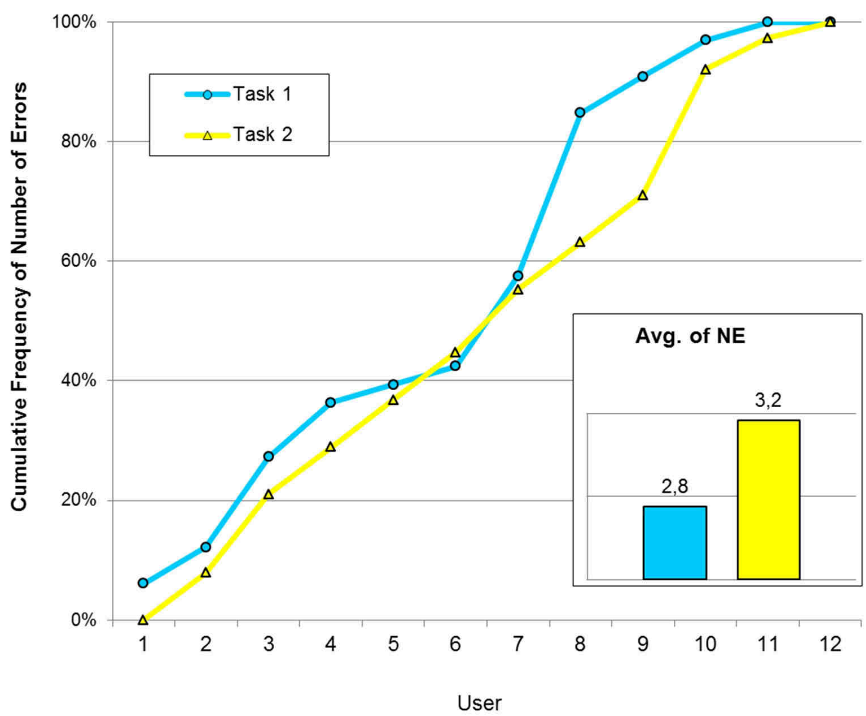


Figure 6: Cumulative frequency of number of errors related to each task and average values

The values of *Task Completion*, grouped using the level of success, are depicted in Figure 7. In particular, most users carry out the Task 1 in a complete success. Otherwise, in the accomplishment of Task 2, only 1 user completely achieves the goals (complete success), while 10 users get a partial success. There is also 1 user who fails.

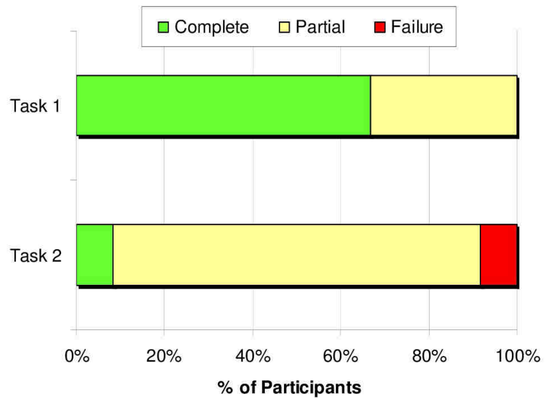


Figure 7: Stacked bar chart showing different levels of success based on task completion

The measures of Satisfaction is analysed. Figure 8 shows that the lowest value of the Satisfaction is related to the clarity and effectiveness of the GUI (D8), while the highest value is related to the actual benefit perceived in the use of the GUI, during the improvement of the gearbox design (D5).

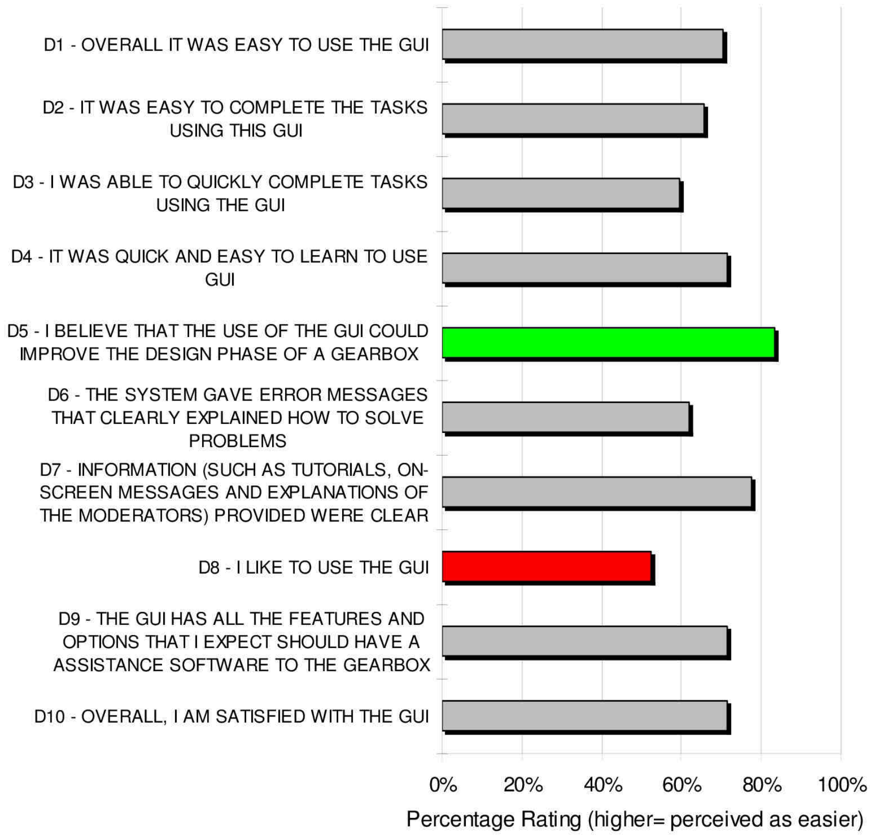


Figure 8: Average subjective ratings split by statement

Some frequent critical issues are identified by analysing videos related to the tests (Table 6). These problems involve difficulties in achieving the Tasks and they generally cause an increase of the operating time also due to a more than proportional increasing number of operations to be performed. In some cases, the user is confused and, then, he/she is led to an error or makes continuous action controls.

Table 6: Root causes and corrective actions of the critical issues identified

#	Root Cause	Corrective Action	UD
1	The default fields are not empty	The default fields are empty	Efficiency
2	The interface does not allow you to overwrite the selected values in the fields, you are forced to cancel the existing	It's possible to overwrite the selected values	Efficiency
3	Poor visibility of function for exporting the file	A new visibility was given to the button to export the files	Efficiency
4	Poor functionality of the reset function	A new reset function was upgraded	Effectiveness
5	The user does not have a feedback on the correct setting of the parameters	The button on the control panel is divided and a section with the new (TEST) or upgraded (RESET) functions, and a new confirmation command (SET) are inserted. The latter turns on only when the input parameters are correct. In this way it provides an immediate feedback to the user.	Effectiveness
6	Poor visibility of the zoom function	A new visibility is given to the button aimed to Zoom	Efficiency
7	Inconsistency of the provision of the sections in the main window	A new provision of sections, with different background colours, that promotes the logical procedure for the input of project data is introduced	Effectiveness

Once these critical issues are identified, the GUI is re-designed. However, in order to avoid radical changes and with the aim to improve UD Effectiveness and Efficiency, the new GUI is developed but it keeps the initial sizes . So, all corrective actions listed in Table 6 are considered.

In order to confirm the validity of the new features a pairwise comparison between them is performed. For each upgrade, users are asked to rate the GUI with the new functions. They express their degree of preference on a scale of six points. The results of survey demonstrate a preference of the users for all new functions far higher than the initial ones.

Hence, all the new features are definitively implemented and a new release of GUI is developed (Figure 9).

Finally, the validation test is accomplished in order to assess again the UI, according a continuous improvement loop.

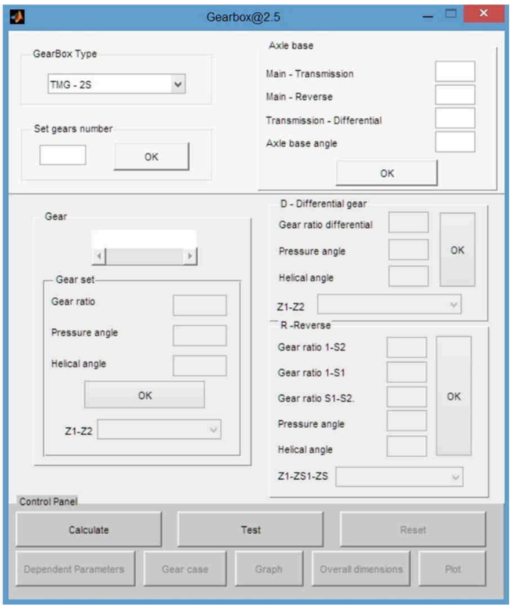


Figure 9: The new release of GUI

Five users who tested the GUI in the first test are involved in the new validation test. Similarly to the first experimentation, the UI is calculated. In particular, the same weights of both UFs and UDs are used. In particular, the normalization measures of UFs are depicted (Table7) and the UFs and UDs values are obtained (Tables 8-9 respectively).

Table 7: Normalized measures of UFs in the validation test

	Effectiveness		Efficiency		Satisfaction
ID user	Number Errors (NE)	Task Completion (TC)	Number of Operations (NO)	Time (T)	Post Session Ratings (PSR)
VT.1	0.25	0.29	0.60	0.58	0.87
VT.2	0.20	0.67	0.88	0.57	1.00
VT.3	1.00	0.65	0.67	0.65	0.93
VT.4	1.00	0.76	0.87	0.76	0.94
VT.5	0.33	0.83	0.94	0.83	0.76

Table 8: Weights and values of UFs in the validation test

	NE	TC	NO	T	PSR
w_i	0.74	0.26	0.25	0.75	1.00
um_i	0.56	0.57	0.79	0.68	0.90

Table 9: Weights and values of UDs in the validation test

	Effectiveness	Efficiency	Satisfaction
w_i	0.59	0.31	0.10
UDI_i	0.56	0.76	0.90

The UI obtained is equal to 0.66. The overall GUI usability improvement is of 57%. For each usability dimension, the following percentage changes are registered: the Effectiveness UD increases of 81%, the Efficiency UD of 56% and the Satisfaction UD of 14%.

It's worthwhile to note the following improvements related to Effectiveness UD: the average NE decreases of 53%; the percentage of users able to complete with success the session test increases of 23% while those who are able to partially complete decreases of 28%. These results have a positive effects on TC measurement.

As regard the Efficiency UD, we highlight that the average T decreases of 22%. In particular, T decreases of 6% in Task 1, while decreases of 41% in Task 2.

Also the Satisfaction UD increases. A further investigation is carried out. Considering the paired data, matched samples, Wilcoxon-signed-rank test is used [23] to determine whether there is a significant difference between the average values of the PSR made under two different conditions (*i.e.* GUI before and after the changes). Both PSR measurements are made on each unit in a sample, and the test is based on the paired differences between these two values. The null hypothesis is the difference in the mean values is zero. Because the p-value is low (< 7%), we can be assume that the changes have produced a significant effect on GUI usability.

5 Conclusions

The present paper tackles the usability assessment of a GUI that is a part of a KBE system. To this aim, starting from a method for usability assessment successfully applied to a new product proposed in [10], a new approach to evaluate the usability of a GUI is discussed. A new usability index (UI) is proposed based on AHP model and its use is validated thanks to experimental results. In particular, the experimental data, leading to a lower value for UI (0.42), are collected and discussed. Then, taking into account such experimental data, a new release of the GUI is proposed and a new set of experimentations are carried out in order to validate the new release. According to the validation test, the UI achieves the value of 0.66 i.e. it shows an increase equal to 57%. Such improvement induces to state that the use of the new release of the GUI could improve the KBE system and contribute to reduce the development time of gearboxes.

Further steps deal with evaluation and improvement of the new GUI. In fact, by using the present approach, new characteristics of GUI are discovered during the experimental sessions and could be introduced and evaluated, in iterative way.

Acknowledgements

The present work is developed with the contribution of the Italian Ministry of University and Research (MIUR) performing the activities of the PON01_01268 DIGIPAT Project, *Digital Pattern Product Development: A Pattern Driven approach for industrial product design, simulation and manufacturing*.

The authors also thank prof. Giuseppe Di Gironimo (Department of Industrial Engineering, University of Naples Federico II), coordinator of DIGIPAT Formation Project for selecting the users, and the engineers who took part to the experiments.

References

1. Sharmin M., Bailey B. P., Coats C., Hamilton K. Understanding knowledge management practices for early design activity and its implications for reuse. In Proceedings of the 27th International Conference on Human Factors in Computing Systems: 2367-2376 (2009).
2. Stenzel I., Pourroy F. Integration of experimental and computational analysis in the product development and proposals for the sharing of technical knowledge. International Journal on Interactive Design and Manufacturing, 2(1), 1-8 (2008).
3. Elgueder J., Cochennec F., Roucoules L., Rouhaud, E. Product–process interface for manufacturing data management as a support for DFM and virtual manufacturing. International Journal on Interactive Design and Manufacturing, 4(4), 251-258 (2010).
4. El Hani M. A., Rivest L., Maranzana R. Product data reuse in product development: A practitioner’s perspective. In PLM, proceeding of IFIP Advances in Information and Communication Technology, 388, 243-256 (2012).
5. Lanzotti A., Patalano S., and Vitolo F. A graph-based approach to CAD modeling: a digital pattern application to the sizing and modeling of manual transverse gearboxes. In International Conference on Graphic Engineering, Madrid (2013)
6. Patalano S., Vitolo F., and Lanzotti A. A graph-based software tool for the CAD modelling of mechanical assemblies. In Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization: 60-69, Barcelona, (2013). doi:10.5220/0004299000600069
7. Madhavan R., Alagarsamy K. Usability issues in software development lifecycle. International Journal of Advanced Research in Computer Science and Software Engineering, 3 (8): 1331-1335 (2013)
8. Sohaib O., Khan K. Integrating usability engineering and agile software development. In Proceedings of the International Conference on Computer Design and Application, volume 2: 32-38 (2010)
9. Sandeberg M. Knowledge based engineering – in product development. ISSN: 1402-1536 (2003)
10. Bourke D. Software gives engineers quick access to product info from layers of company files. Mechanical Engineering: 44-47 (2013)
11. Di Gironimo G., Matrone G., Tarallo A, Trotta M, Lanzotti A. A virtual reality approach for usability assessment: case study on a wheelchair-mounted robot manipulator. Engineering with Computers 29 (3): 359-373 (2013)
12. Saaty L. Decision making with the analytic hierarchy process. International Journal of Services Sciences 1 (1): 83-98 (2008)
13. Norman, D. Emotional Design. Perseus Publishing, New York (2004)
14. Nielsen, J. Usability Engineering. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1993) ISBN 0-12-518406-9
15. Figueira, J., Greco, S., Ehrgott, M. Multiple Criteria Decision Analysis: State of the Art Surveys. Springer (2005)
16. Sauro, J., Kindlund E. A method to standardize usability metrics into single score. Proceedings of the SIGCHI conference on Human factors in computing system, Portland, 401-409 (2005)
17. Kim, J., Han, S. A methodology for developing a usability index of consumer electronic products. International Journal of Industrial Ergonomics 38, 3-4, 333-345 (2008)
18. Ghosh, R., Dekhil, M. Discovering user profiles. In WWW ’09: Proceedings of the 18th international conference on World wide web (New York, NY), ACM, 1233-1234 (2009)

19. ISO 9241-11-1998: Ergonomic requirements for office work with visual display terminals (VDTs) – Guidance on usability (1998)
20. Hornbaek, K. Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer studies 64 (2), 79-102 (2006). doi:10.1016/j.ijhcs.2005.06.002
21. Tullis, T., Albert, B. Measuring the user Experience. Morgan Kaufmann Publisher (2008)
22. Sauro, J., Kindlund E. How Long Should a Task Take? Identifying Specification Limits for Task Times in Usability Tests. Proceeding of the Human Computer Interaction Conference, Las Vegas (2005)
23. Wilcoxon, F. Individual comparisons by ranking methods. Biometrics Bulletin 1 (6), 80-83 (1945)

Paper E

Malmo F., Del Giudice D.M., Sterle C., Proceedings of XIII International Symposium On Locational DEcision 2014 (ISOLDE), Naples/Capri (Italy), June 16th– 20th 2014, ISBN 9788898273072, abstract. ISOLDE is a triennial Symposium in conjunction with the XXI Meeting of EURO Working Group on Locational Analysis (EWGLA).

Computational procedure for location sensor network monitoring volcanic ash

Francesco Malmo[a], Domenico Maria Del Giudice[a], Claudio Sterle[b]

[a] *Department of Industrial Engineering University "Federico II" of Naples, P.le Tecchio, 80125, Naples, Italy – francesco.malmo, domenicomaria.delgiudice@unina.it*

[b] *Department of Electrical Engineering and Information Technology, University "Federico II" of Naples, Via Claudio 21, 80125 Naples, Italy - claudio.sterle@unina.it*

Keywords: volcanic ash, aviation safety, covering optimization, sensor location

Global air traffic is significantly affected by the volcanic ash especially when unfavorable weather conditions occur. About 500 active volcanos are in the world and the plume thrown up by the eruptions provoked several crisis. British Airways Flight 9 was, maybe, the first known case: the aircraft, on June 1982, flew into an ash cloud and this provided the power lost of all the four engines. The most known case, instead, is the eruption of the Eyjafjöll volcano, in Iceland on April 2010. European regulators acted the shutting down of the full airspace. Their decision cost $1.8 billion in lost revenues for the airlines and $5 billion for the global economy (IATA, 2011). In Italy, Mount Etna volcano eruptions, , repeatedly caused alarms requiring the restriction of local airspace. On February 2014, volcano Kelud brought to the closure of three international airports in the area. Therefore, managing the problem of volcanic ash is a new important challenge for civil aviation, which if neglected can cause significant damage to aircrafts and large economic loss. The main aim is to provide the competent civil and aviation authorities with real time information on the volcanic ash propagation. This allows to define no-flight levels, to re-route scheduled flights and to give warning massages to planes already on flight. This can be done using a permanent monitoring system composed of high range sensors to be installed on the land (Scollo *et al.,* 2013). Each sensor is characterized by its performance and its regional cost (Mallozzi, 2011).
The aim is to place the sensors optimizing an objective function which is a linear combination of cost and performance, guaranteeing that all the planes on flight can receive the warning with a sufficient advance along their route. Moreover a higher reliability of the system should be guaranteed covering each point of a route with more than one sensor. In this work we tackle this problem by the usage of covering optimization models (Berman et al., 2010). The following basic assumptions are required: the "Area of Interest"is discretized and divided in iso-risk sub-regions; a single flight level is considered; the air routes are linearized and discretized; the sources of volcanic ash are not in the "Area of Interest". The proposed model has been optimally solved by the usage of Xpress optimization software and tested on real test cases using the northern Italy as "Area of Interest".

Main references

Berman O., Drezner Z., Krass D. (2010). Generalized coverage: New developments in covering location models. *Computers and Operations Research*, 37: 1675–1687.

IATA (2011). Europe's Response to Grimsvotn - Formal Agreements Needed to Cement Progress. IATA Press release No. 28 – 24/05/2011.

Mallozzi L. (2011). Cooperative games in facility location situations with regional fixed costs. *Optimization Letters*, 5:173–181.

Scollo S., Azzopardi F., Boselli A., Coltelli M., Ellul R., Leto G., Pisani G., Prestifilippo M., Saliba M., Schiavone J., Spata G., Spinelli N., Wang X., Sanchez R. Z. (2013). Volcanic Ash Monitoring and FOrecaSting between Sicilia and Malta arEa and sharinG of the resUlts foR aviatiOn safety: the VAMOS SEGURO project, *Geophysical Research Abstracts, Vol. 15, EGU General Assembly*.

Paper F

Gerbino S., Del Giudice D.M., Staiano G., Martorelli M., Lanzotti A, on the 21th of January in 2015 it was submitted to International Journal of Advanced Manufacturing Technology, ISSN 14333015 and 02683768 (Q1, IF 1.78 in 2013), 19 pp.

On the influence of scanning parameters on the laser scanner based 3D inspection process

Salvatore Gerbino[(a)], Domenico Maria Del Giudice[(b)], Gabriele Staiano[(b)], Antonio Lanzotti[(b)], Massimo Martorelli[(b)]

[(a)] University of Molise - DiBT Dep't, Engineering Division, Via De Sanctis, 86100 Campobasso, Italy;
Email: salvatore.gerbino@unimol.it

[(b)] University of Naples Federico II - Dep't of Industrial Engineering, P.le Tecchio 80, 80125 Naples, Italy;
Email: {domenicomaria.delgiudice, gabriele.staiano, antonio.lanzotti, massimo.martorelli}@unina.it

Abstract

The quality of 3D scanned data is influenced by many factors both related to internal elements to the acquisition device, such as scanner resolution and accuracy, and external to it, such as proper selection of scanning parameters, ambient illumination and characteristics of the object surface being scanned (e.g. surface colour, glossiness, roughness, shape). Due to the recent developments in terms of accuracy, in particular for 3D laser scanners, today it is of great industrial interest to study and correctly setting the scanning parameters that allow to improve the quality of the 3D acquisitions so to increase the massive usage of these systems in the product inspection activities. In this paper the effects of some scanning parameters that may affect the measurement process, were analysed by using a commercial triangulation 3D laser scanner. The test geometry chosen for this study was a commercial sheet metal part more complex than the ones commonly used in laboratory and documented in the literature. Relative orientation, ambient illumination and scanner parameters were tested. The outcomes of the tests confirmed some results and suggestions documented in literature but also pointed out that among different conditions the most influencing factor are the relative orientation of the object with respect to the scanner, as well as its position of the measurement device within the field of view.

Key words

3D laser scanner; Design of Experiments; Statistical analysis; 3D inspection; Reverse Engineering.

1. Introduction

3D scanners are devices mainly addressed to the task of 3D digital reconstruction of real-world objects, even of complex shape, through principles codified in complete sets of procedures, specific to various applications. Non-contact active and passive 3D scanners are today widely used in numerous industrial applications, also for getting measurements for quality inspection. They mainly include systems based on the following optical technologies: laser triangulation, structured light, time-of-flight and photogrammetry. All these technologies have reached a good level of confidence and accuracy for Reverse Engineering applications and some of them (mainly those based on laser triangulation) are currently employed in a number of factories as quality inspection device [Martínez *et al.*, 2010]. They allow, in fact, to get a high speed in the acquisition of a large set of points over each surface, and the consequent time/cost reduction.

Coordinate Measuring Machine (CMM) systems based on touch probe still offer a higher accuracy in the inspection process for verifying dimensional and geometrical tolerance specifications of mechanical parts. The main drawback of such a technology is the long operational time to acquire also few points, that can become prohibitive when a complex shape needs to be accurately measured in many points [Martínez *et al.*, 2010].

The ability of 3D optical scanners to capture thousands of points in few seconds, together with the increased accuracy they recently offer, make possible to extend the quality control to the whole part's shape. In [Huang *et al.*, 2013 and Liu & Hu, 1997] authors documented researches aimed to check and model the statistical variational behaviour of automotive sheet metal components for virtual simulation of compliant assemblies by acquiring the whole geometry of each flexible component.

This need will become more and more common also in other industrial fields, so it is important to understand well the several conditions and parameters the user can chose to get the best from the 3D scanners, in particular the ones based on laser triangulation which are more common.

The testing analysis was conducted by designing specific experiments, by using Design of Experiments (DOE), according to [Box *et al.*, 2005, Montgomery, 2007 and Eriksson *et al.*, 2008], in order to discover how the laser scanner position with respect to the object, the ambient illumination and other internal settings of a laser scanner may influence the accuracy of the measurements. In the present paper the commercial 3D laser scanning system, VI 9i by Konica Minolta, was used for this purpose on a sheet metal car component with a quite complex shape.

The paper is arranged as follows: Section 2 introduces the main parameters to be considered that affect the measurements with a laser scanner; Section 3 presents materials and methods of the conducted investigation and introduces the statistical analysis made on the acquired measured data; in Section 4 the main outcomes of the analysis are described; finally, Section 5 draws the conclusions.

2. Main parameters in 3D data acquisition with a triangulation laser stripe system

The most common laser systems used in metrological applications are those based on principle of triangulation through a laser stripe. The reasons are the high precision, the relative lower cost with respect to other systems, and the possibility they offer to acquire objects on site.

In these non-contact active systems an incident laser beam, with known width, is projected onto a part to be scanned and the stripe generated on the surface is detected by a CCD sensor; then through the triangulation principle, 3D coordinates of the surface points are acquired.

The main parameters of these systems are depicted in Figure 1. We may distinguish parameters imposed by the system manufacturer and parameters depending on the user choices.

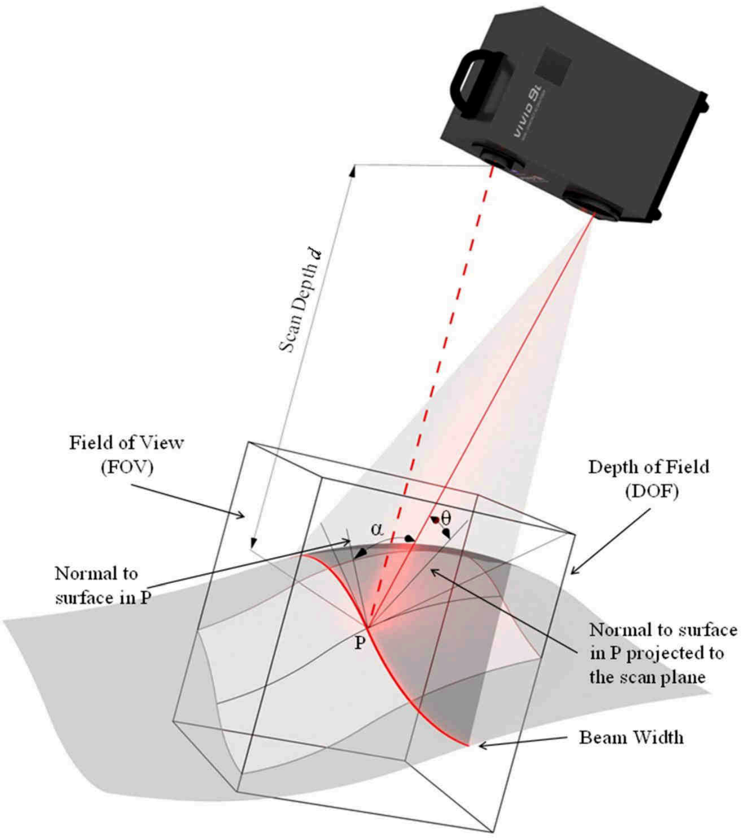


Figure 1: Main parameters of a laser stripe 3D scanner.

The main parameters imposed by the system manufacturer are:

- Field of view (FOV): 3D region within which the CCD sensor can acquire points on the scanned surface. It is defined by the depth of field and the scan width.
- Depth of Field (DOF): range of distance from the laser source within which the CCD sensor can acquire points on the scanned surface.
- Scan width: width of the laser beam measured in the half position of the depth of field.
- Stand-off distance: distance from the laser source to the reference surface located in the half zone of the field of view.

The main parameters not depending by the manufacturer, and that the user can (partially) control, are:

- Incident angle (θ): angle between the incident laser beam and the projected surface normal of the scanned point in the scanning plane [Prieto *et al*., 2003].
- Projected angle (α): angle between the scanning plane and the normal (minimum) at the surface in the scanned point [Prieto *et al*., 2003].
- Laser light intensity.
- Ambient light.

More factors influencing the accuracy of a measurement with a laser scanner must be added:

- Topology of the object to be acquired.
- Characteristics of the material object (e.g. surface colour, glossiness, roughness).
- Processing software.

The related orientations of the scanned surface with respect to the incident laser beams, represented by the incident angle θ (in the laser plane) and the projected angle α (measured out of the laser plane) in Figure 1, are the main scanning process parameters mostly influencing the measurement accuracy [Feng *et al*., 2001]. In fact, the intensity of the reflected laser beam and its impression on the optical CCD sensor are strictly related to the incident/projected angles. This dependency was empirically measured in [Mahmuda *et al*., 2011]. Increasing the incident angle causes more uncertainty in the measuring while keeping all other parameters constant.

In [Prieto *et al*., 2003] the following ranges were suggested for the angles θ and α: $-35° \leq \theta \leq +35°$ and $-15° \leq \alpha \leq +15°$, from which an algorithm was proposed for the optimal choice of the relative orientation laser-to-object_surface, starting from a planning based on the object's CAD model. This positioning error has a systematic nature and it can be partially controlled by the user thanks to the right positioning of the scanner with respect to the object, being usually the normal position laser-surface the best one for getting higher accuracy.

The ambient light plays a very important role in the optical measurements. In the recent version of laser devices, released by different manufacturers, it is pointed out their ability to work in a large range of ambient light conditions, as well as with different colours of the object's surface, but in practice many users still report a sensitivity to this effect. A technical solution more adopted in these cases is the use of specific light filters on CCD sensor [Forest & Salvi, 2004]. Practical solutions are the use of coating spray to cover the object with a matte white layer, as discussed later.

In [Blanco *et al*., 2009] the authors reported the results of a specific study on the influence of ambient light on the quality of laser scanner-based measurements on different typology of materials with the same surface treatment. Best results, in terms of less variability and more available data, were generated by using mercury vapour lamps instead of halogen illumination.

In [Cuesta *et al.*, 2009] the influence of the ambient light to the quality of captured signal by CCD sensor is confirmed and the dependency of the measurement from surface roughness of the component was pointed out. They suggested making measurements in absence of or limiting the ambient illumination. This condition is often difficult to apply in many real cases.

The intensity of laser source is strictly dependent from the surface's roughness. For surface with ISO roughness grade N1 (Ra=0.025µm) to N11 (Ra=25µm) the maximum laser intensity (whose correspond the highest amount of acquired data) is strongly related to the machining process generating that surface. The authors proposed a table with suggested ranges of laser intensity for getting the maximum results from the measurements for different roughness rate and machining processes. For sheet metal deep drawing components the recommended value for optimal laser intensity is about 20% of full power of laser source.

In [Feng *et al.*, 2001] the authors studied the random error which is strictly related to the surface's roughness causing the speckle effect: surface roughness during the scanning process produces a slight variation on the reflected light signal, and on CCD sensor the wave amplitude of the captured light can be cancelled or reinforced. This local interference between captured lights makes difficult to get the right centre point on CCD which is used to make the 3D measurement of the digitised point. This phenomenon is more evident when the laser wave length is close to the one of the scanned surface's roughness (around 1/10 µm).

Martínez et al. (2010) pointed out how difficult is the acquisition of dark, glossy and translucent surfaces. With very shiny surfaces a significant noise is added to the measurements, and points not belonging to the real surface can be generated.

Glossy surfaces affect the saturation of light captured by CCD sensor, whereas dark surfaces absorb too much light intensity. In these cases the CCD sensor is unable to acquire data locally (missing data) or the measurement is highly affected. In such conditions it is suggested using coating sprays to cover objects with a thin white matte layer, which is specifically helpful for scanning dark, glossy or translucent surfaces. In [Mahmuda *et al.*, 2011] it was evaluated in 5-15 µm the added thickness and the additional variation of about 45 µm by using very thin coating sprays.

Black coloured surfaces are harder to scan (they usually require the coating spray) but the new released devices provided by some scanner producers offer a limited solution to reduce measurement uncertainties by automatically tuning on the laser power and adopting specific filters for CCD on the captured signal.

3. Materials and Methods

The tested geometry chosen for this study is a commercial sheet metal part used in automotive field (Figure 2) with a general shape complexity that allows to test some aspects previously described.

A commercial laser scanner – VI 9i by Konica Minolta, whose measurement ranges for different interchangeable lenses (tele, middle and wide) are summarised in Table 1 – was used for the measurements.

The black original colour of the component resulted hard to acquire, so it was necessary to use a coating spray, as shown in Figure 2. Tests made with different operators demonstrated that the measurement result (also related to reference spheres located onto the part) was strongly dependent on the operator spraying the component, so we decided to paint the whole part with a matte light grey colour, very similar to the colour of the original calibration device of VI 9i scanner.

Figure 2: Tested sheet metal part partially sprayed with a white matte layer to allow acquisition with a laser scanner very sensitive to the original black colour of the part.

The part was firstly acquired with a CMM laser scanner to generate the whole reference model for the test as the original CAD model of the part was missing. This is acceptable for the higher accuracy of the used CMM laser scanner (Metris LC15 mounted on DEA Global Image Clima), equal to 8μm, than the one of the VI 9i scanner, equal to 50μm. Figure 3 depicts an intermediate step of the acquisition phase with the CMM laser scanner.

Table 1: Measurement ranges (X * Y * Z) for different interchangeable lens for the VI 9i laser scanner.

Distance (mm)	TELE (mm)	MIDDLE (mm)	WIDE (mm)
600	111 x 83 x 40	198 x 148 x 64	359 x 269 x 108
1000	185 x 139 x 110	329 x 247 x 176	598 x 449 x 284

All scan data were post-processed in Geomagic 2012 software to generate the final polygonal model (made by about 970k triangles, referred to the only component) shown in Figure 4.

The model in Figure 3 also shows the use of reference spheres to make easier and repeatable the alignment process of point data sets to the reference model (RPS method). We used spheres of diameter 1” and 1/2”.

Figure 3: Acquisition phase with CMM laser scanner to generate the reference 3D model.

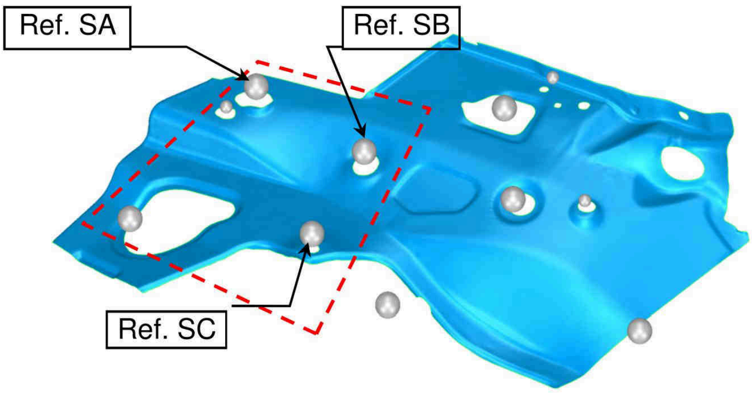


Figure 4: Full reference model for the case study. Reference spheres used to align point data sets are also shown.

According to a systematic approach to the DoE suggested in [Coleman & Montgomery, 1993], it follows the description of the pre-experimental phase that identifies all the factors of interest and their classification in control, held or noise factors. The classification is closely related to the specific objectives of the experiment.

In the context of the present work for the experimental tests the attention was focused on the acquisition of a limited portion of the component with more evident shape variation (highlighted in the left area in Figure 4). This allowed to work always with a single shot of the VI 9i scanner, having the possibility to capture a significant portion of the component by using the Middle lens whose measurement range is 329x247x176 mm at about 880 mm far away from the part (see Table 1), and the accuracy and precision of less than 0.20 mm and 0.048 mm, respectively.

In this way each acquired portion of scan data was aligned and compared with the reference model by using three sphere features always captured in the three different angles of the part: -30°, 0° and +30° (by using a rotating table mounted on a tripod). This "Angle" was classified as a control factor. Figure 5 shows how the part appears on the scanner's camera at different levels of the Angle, whereas Figure 6 shows the set-up of the experimental tests.

To test the effect of the ambient lighting (hereinafter "Lighting") the acquisitions were made in two very different environmental conditions: "light" with a white halogen lamp of 1000W not directly oriented to the object (so generating diffuse illumination), and "dark" condition with just a dim light.

-30° 0° +30°

Figure 5: Scanner's camera snapshots for the three Angle levels.

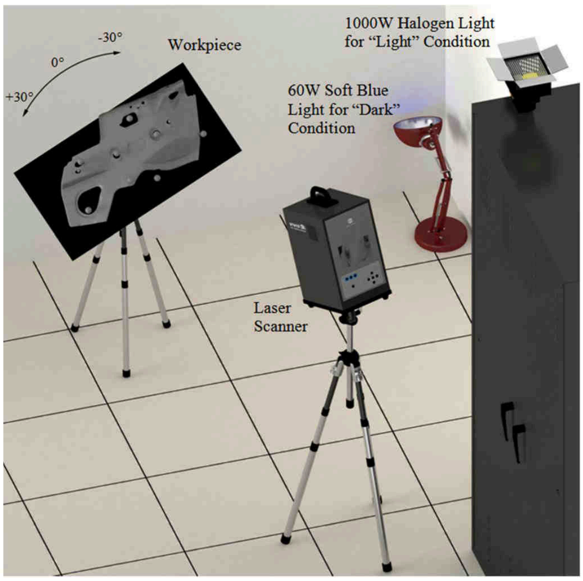


Figure 6: Set up of the experimental test.

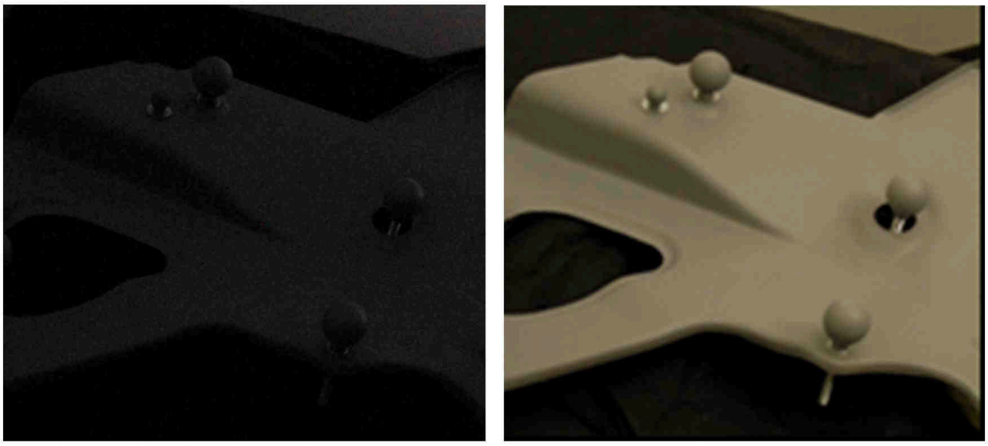


Figure 7: How the model appears on an external camera in dark (left) and light (right) Lighting levels.

The Lighting with a complete absence of external light source was not applicable to our test (as suggested in [Cuesta *et al.*, 2009] where a CMM laser scanner was used) as the laser scanner was unable to properly set up the camera lens focus in that condition. So a dim light with a blue filament lamp of 60W, located far away from the laser/object position, was used. Figure 7 shows two pictures captured in both Lighting levels.

Usually, in the industrial applications the Lighting is a noise factor as it is hard (sometimes impossible) to properly set it up for accurate measurements. In our lab tests, instead, Lighting was a control factor and its specific influence on the accuracy of the measurements was evaluated.

High quality filter is an option of the laser scan tool that, if ON, discards unreliable scan data, usually located close to the FOV boundary. This "Filter" was classified as a control factor.

There are some potential noise factors that may affect the experiment as the laser intensity and the stand-off distance. Noise factors are those that cannot be controlled from a technological or economic point of view.

VI 9i Minolta laser scanner offers an automatic setting both for the camera focus, AutoFocus, and the balance of laser intensity. These parameters may change a little when running consecutively more times the automatic setting in the same conditions, so an average value over five runs were set and kept constant for each repetition referred to Angle levels. In particular, the autofocus distance (taken at the "light" Lighting level) was used as reference stand-off distance and assigned to all scans related to both "light" and "dark" conditions for each Angle levels. Laser intensity, instead, was set differently for "light" and "dark" conditions, as previous tests have pointed out the influence of this parameter on the scan quality. Notice that the laser intensity automatically set by the scanner is about 20% of full power of laser source (250) according to [Cuesta *et al.*, 2009] for sheet metal deep drawing components.

Both laser intensity and stand-off distance can be classified as noise factors in the experimental tests. By means the set-up and the restrictive protocol described above the effects of these factors were controlled. This choice was then validated by running a restrictive experimental protocol.

The held-factors are controllable factors whose effects are not of interest in this phase, such as the white balance. The white balance option is suggested to set the internal level of Lighting and it can be carried out with a special matt white lens before starting the measurements. It was set once for all acquisitions in light condition, whereas for dark Lighting the default inner value of the scanner was used for the difficulty of setting this parameter in absence of enough ambient light.

To improve the alignment of measurements data with the CAD reference model and make this process repeatable, the three reference spheres shown in Figure 4 were used.

Based on the RPS alignment method with those three spherical features the Root Mean Square (RMS) error was used as response variable. RMS is an estimation parameter of the deviations of each scan data to the reference model. To make comparable the results in terms of RMS, a common subset of scan data (of about 350k polygons) was extracted and used in the next steps (see Figure 8).

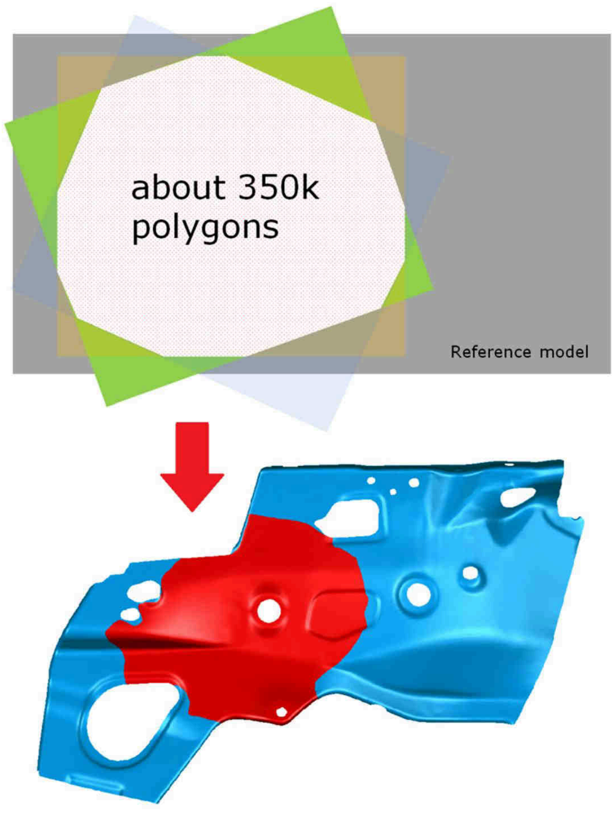


Figure 8: The common subset of scan data (of about 350k polygons).

The factors described above were taken into account for the experiment and are listed in Table 2.

Table 2: List of control, noise and held factors.

Factor	Type
Angle	Control
Lighting	Control
Filter*[1]	Control
White balance	Held
Laser intensity	Noise
Stand-off distance*[2]	Noise

*[1] It refers to the inner scanner filter.

*[2] It refers to the focus distance parameter of the camera.

A DoE study was carried out focusing the attention on the Angle, Lighting and Filter factors. As three factors were assumed important in the pre-experimental phase, a $3^1 \cdot 2^2$ mixed-level design matrix was adopted. This is a full resolution design, so no main effects either two-factor interaction or three-factor interaction are aliased with each others.

So, the experimental design enables reliable information to be obtained about main effects and two-factor interactions. Table 3 summarises the levels of control factors and their settings. The Lighting and Filter are qualitative factors while the Angle is a quantitative factor.

No smoothing was applied on the acquired data. Only outliers were detected and deleted when occurred (near edges). Data processing was conducted in Geomagic 2012. To take the post-processing analysis under control, we worked only with a single shot for each treatment so eliminating the additional error occurring during the registrations phase of more scan data sets.

Table 3: Control factors and their settings.

Control Factor	Low (-1)	Mid-level (0)	High (+1)	Unit
Angle	-30	0	+30	Deg
Lighting	Light	-	Dark	-
Filter	Off	-	On	-

Table 4 shows the $3^1 \cdot 2^2$ mixed-level design matrix. For each treatment, three replications were executed, giving a total of 36 experimental runs. Three replications, instead of two the ones usually used in a screening experiment, was here adopted to provide more consistent response repeatability, in particular for the Angle factor, during this first experimental phase. Assuming conservatively that the three replications for each treatment are not consistent with each other, it is necessary to group them in blocks and perform treatments in random order. In this way it is possible at same time to reduce the accidental variability and to increase the validity of inferences on the effects of the factors. In order to reduce time in resetting and recalibration the laser optical device, for each of treatment listed in Table 4 the three replications were made consecutively; so only the treatment order was randomised.

Table 4: Matrix for the $3^1 \cdot 2^2$ mixed-level design.

Treatment	Angle	Lighting	Filter
1	-1	-1	-1
2	-1	-1	+1
3	-1	+1	-1
4	-1	+1	+1
5	0	-1	-1
6	0	-1	+1
7	0	+1	-1
8	0	+1	+1
9	+1	-1	-1
10	+1	-1	+1
11	+1	+1	-1
12	+1	+1	+1

4. Results and discussion

Tests were firstly limited on spheres SA, SB and SC of Ø1" (see Figure 4) used as targets. Best fit spherical features were extracted from point data and the fitting error (deviation) was recorded. Figure 9 shows the mean deviations for each sphere in any combination of the control factors. Sphere SA in -30° reported the highest deviation as it was farer from the camera with respect to the FOV (see Figure 5). Lower deviations corresponded to sphere SC in "dark" Lighting. These variations affected of the alignment process of all data sets as shown later on.

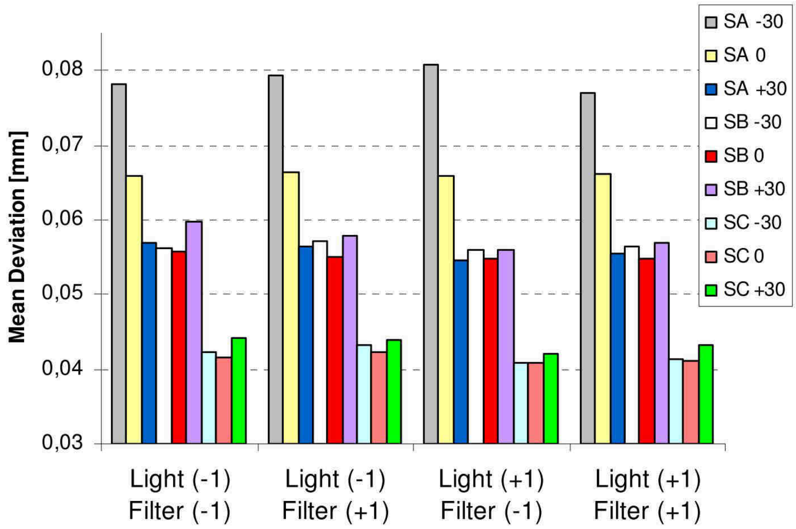


Figure 9: Mean deviations in the generation of spherical features.

Spherical features were used to look for the best sequence combination of reference features during the registration phase. Over the 36 runs, the experiments pointed out that the mating sequence SA-SC-SB was the one to which a general lower mating error occurred, so this combination was used for the subsequent steps of the analysis.

The ANOVA method was applied in order to test the statistical significance of the main effects, and the two- and three-factor interactions for Angle, Lighting and Filter factors. Diagnostic checking was successfully performed via graphical analysis of residuals. The experimental results are shown in Table 5, using ANOVA table (with $\alpha = 0.05$).

Table 5: Analysis of variance for RMS [mm] in the SA-SC-SB alignment sequence.

Source	DF	Adj MS	F	p-value
Blocks	2	0.0001444	2.42	0.11
Angle	2	0.0041194	69.12	0.00
Lighting	1	0.0001000	1.68	0.21
Filter	1	0.0001778	2.98	0.10
Angle*Lighting	2	0.0001083	1.82	0.19
Angle*Filter	2	0.0000028	0.05	0.96
Lighting*Filter	1	0.0001000	1.68	0.21
Angle*Lighting*Filter	2	0.0000250	0.42	0.66
Error	22	0.0000596		
Total	35			

Results show that the tested full regression model is statistically significant and has a high goodness of fit (R-Sq adj = 80%). A not significant p-value corresponds to blocks (p = 0.11) meaning that the mean values of the response variable are not different for the different blocks. For the response variable, the main effects (significant at $\alpha = 0.05$) are depicted in Figure 10; Figure 11 shows the interaction plots for RMS. An exhaustive technological interpretation of the results follows.

The first outcome of this screening was the main effect of the Angle control factor which played the most influent role on the accuracy of the over measurements. Whereas Lighting and Filter control factors had no significant effects (Figure 10). Also the two-interaction effects were not significant (Figure 11). In particular, lowest deviations were related to the 0° position (Figure 10). This seems to be in relation with the best "central" position (in terms of average distance) of all three spheres with respect to the FOV of the scanner, also valid for the +30° position.

On the other side, a lower accuracy corresponded to Angle at -30° where sphere SA was farer away from the FOV of the scanner. This result is in accordance with the mean deviations plot of the reference spheres in Figure 9.

From all the experimental campaign we may conclude that the best results correspond to combination of control factors: Angle = 0°, Lighting = Dark and Filter = On, having the only Angle factor a statistical significance on the accuracy of the measurements.

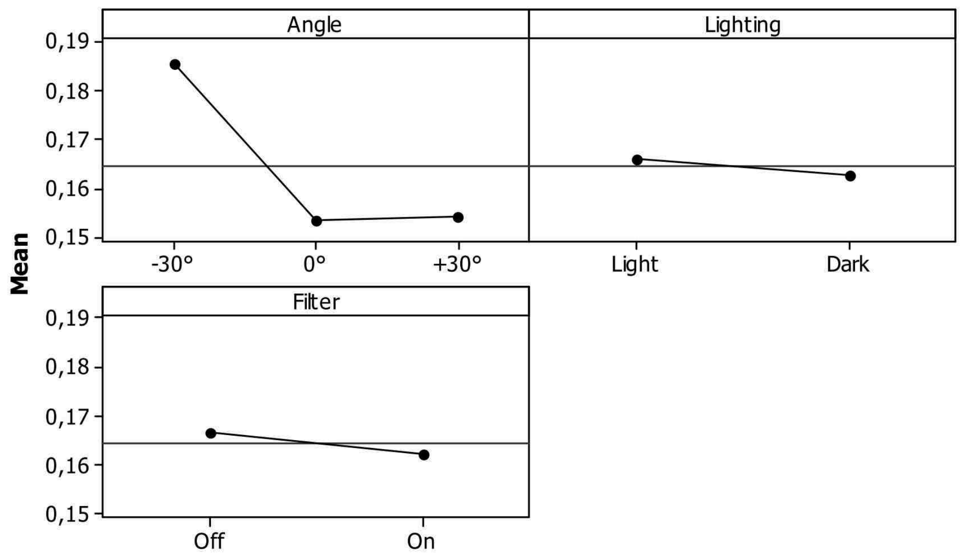


Figure 10: Main effects plot for RMS [mm] in the SA-SC-SB alignment sequence.

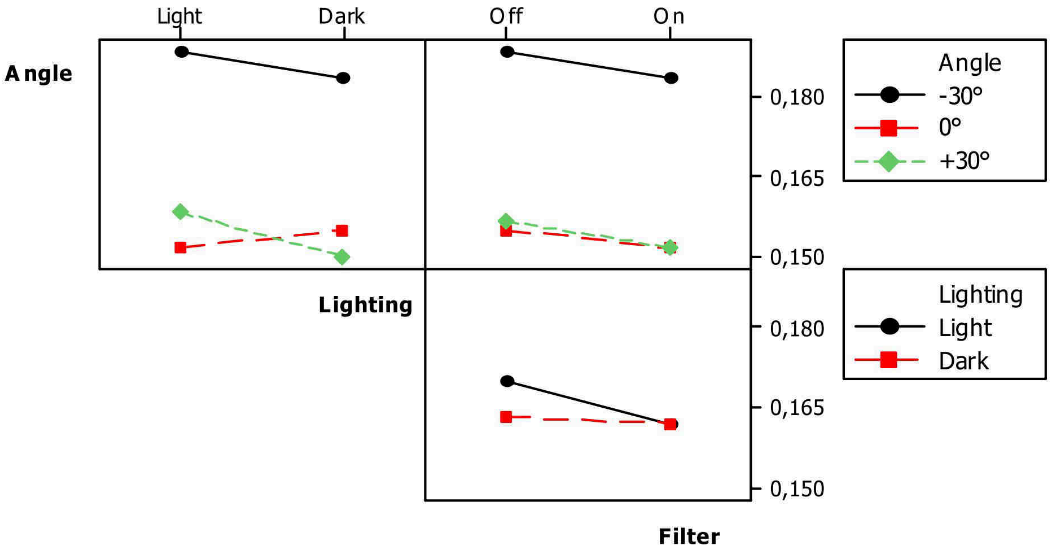


Figure 11: Interaction plot for RMS [mm] in the SA-SC-SB alignment sequence.

As the effects of both the blocks and the combinations of the factors with order higher than one were not significant, we repeated the analysis excluding them from the model. Analysing the new ANOVA test shown in Table 6, it appears that the simplified regression model is statistically significant and has a high goodness of fit (R-Sq adj = 79%). The Angle is confirmed to have a statistically significant effect.

Table 6: New analysis of variance for RMS [mm] in the SA-SC-SB alignment sequence.

Source	DF	Adj MS	F	p-value
Angle	2	0.0041194	64.75	0.00
Lighting	1	0.0001000	1.57	0.22
Filter	1	0.0001778	2.79	0.11
Error	31	0.0000636		
Total	35			

The previous analyses suggest that the control factor Angle could have covered the real effects of the secondary factors (Lighting, Filter), given its dominant influence on RMS. For this reason, a new analysis considering only those treatments with the Angle held at 0° (best level) was conducted. The design of experiment extracted from the initial one consists of four treatments (from 5 to 8 in the box of Table 4).

The results are shown, using Pareto charts of standardised effects (α = 0.05), in Figure 12. The secondary effects are not statistically significant. In Figure 12 the vertical dashed line marks the significant limit.

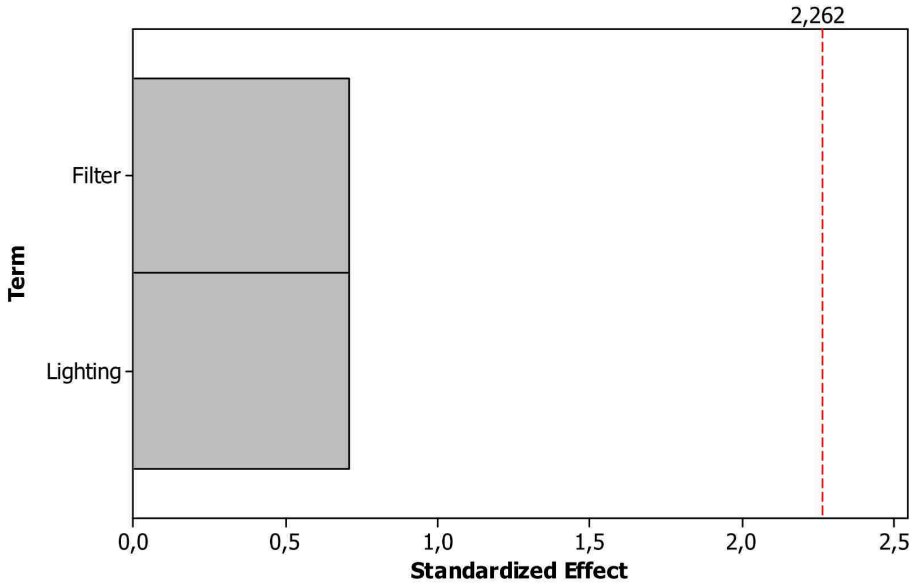


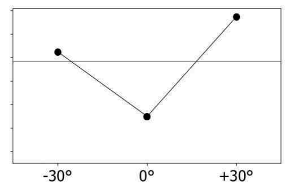
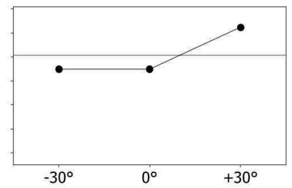
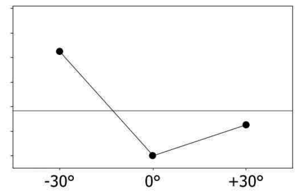
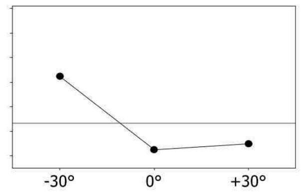
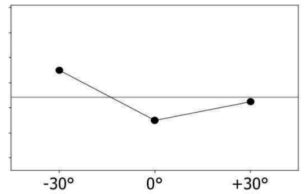
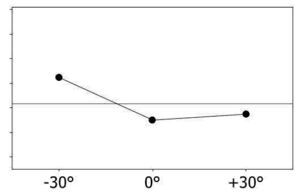
Figure 12: Pareto chart of the standardised effects using partial design for RMS in the SA-SC-SB alignment sequence (α = 0.05).

Looking at the Figure 9 we stated the best mating sequence was SA-SC-SB. The last experiment was oriented to confirm this. So the whole "assembly" process was repeated for each combination of mating sequence of spherical features and the related output in terms of ANOVA significant test were repeated as in Table 7. In particular, Table 8 shows the comparisons of the Angle main effects with respect of all alignment sequences.

Table 7: Comparison results of ANOVAs for RMS in all alignment sequences.

Alignment sequence	Angle	Lighting	Filter
SA-SB-SC	Significant	Not Significant	Not Significant
SB-SC-SA	Significant	Not Significant	Not Significant
SC-SA-SB	Significant	Not Significant	Not Significant
SC-SB-SA	Significant	Not Significant	Not Significant
SB-SA-SC	Significant	Not Significant	Not Significant
SA-SC-SB	Significant	Not Significant	Not Significant

Table 8: Comparison of the Angle main effect plots for RMS (in all alignment sequences).

SC-SB-**SA**			SB-SC-**SA**
SC-SA-**SB**			SA-SC-**SB**
SB-SA-**SC**			SA-SB-**SC**

We may state that the whole outcome of the analysis is robust with respect to the different order of spherical mating features, but at the same time the sequence SA-SC-SB was the one to which the best RMS results over the part measurements correspond, as shown in Figure 13. All sequences with SA as last feature in the mating sequence caused worst results. This finding is consistent with the results listed in Figure 9. A slightly better result came out for the sequence with SC as last feature.

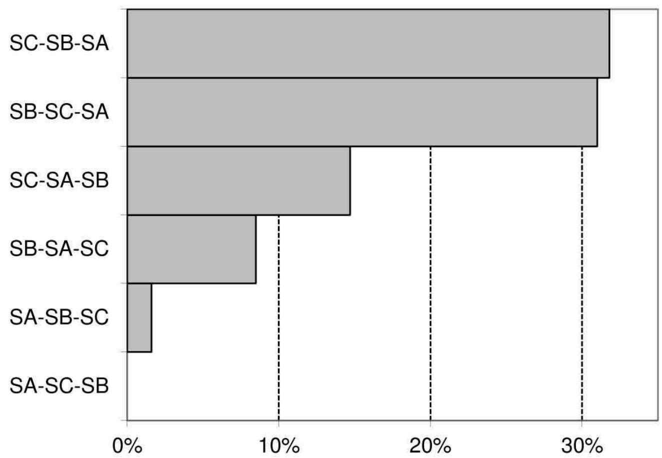


Figure 13: The comparison of relative variations in RMS with respect to best sequence alignment (SA-SC-SB).

5. Conclusions

The technical literature is worth of studies on measurement tests made with optical scanners (mainly based on laser triangulation) on simple workpieces and assisted by CMM device, aimed to highlight the factors most influencing the accuracy of the measurements.

A study related to a quite complex commercial part acquired by a mobile 3D laser scanner, VI 9i by Konica Minolta, was here presented. The DoE methodology was used with the aim to point out the influence of some external factors such as the relative position part-scanner (Angle factor) and the Lighting, and an internal factor such as the quality Filter of that specific scanner. Considering the working ranges of the scanner, a middle distance part-to-laser was adopted with the intent to capture a significant portion of the part without limiting so much the overall accuracy of the acquisitions. Reference spheres were included and used for the RPS alignment with a reference model. The study pointed out that part-scanner angular position has a statistically significant great effect on the measurement accuracy, while both the Lighting (if diffuse and not directly oriented to the part) and the Filter factor have no statistically significant effect (for the control factors selected at levels set) for the adopted VI 9i scanner.

In particular, the accuracy of the acquired reference spheres plays an important role: they should be in a central position with respect to the FOV of the scanner to get the best result.

The sphere SA, in a faraway position with respect to FOV in -30° Angle, presented a relative higher deviation and this affected the alignment of the data sets.

Closer acquisitions are required to improve accuracy even though this means that, when acquiring a large object, more scans have to be done, so introducing errors during the registration phase of such data patches. To make highly repeatable the measurements over a large batch of components, the use of well-defined features as references (see also fixtures) is highly recommended. Few separated tests highlighted that a direct illumination of the object may alter the ability of the scanner to locally detect points as the ambient light interferes with the reflected laser signal on the camera's CCD. In the presented test we used a part with a grey colour very close the one of the calibration device for that scanner. More investigation is needed to understand the behaviour of the laser scanner with different colour objects under multi-level light conditions.

Acknowledgment

This study was partially supported by the National Project PON01_01268 entitled ‘Digital pattern product development: a pattern driven approach for industrial product design’, and by the European research project EU-FP7 FoF-ICT-2011.7.4: Remote Laser Welding System Navigator for Eco and Resilient Automotive Factories - RLW Navigator - (grant agreement 285051).

References

Blanco D., Fenandez P., Cuesta E., Mateos S. and Beltran N., 2009. Influence of surface material on the quality of laser triangulation digitized point cloud for reverse engineering tasks. In ETFA2009, Palma de Mallorca (S).

Box. G.E.P., Hunter J.S. and Hunter W.G., 2005. Statistics for Experimenters, Wiley, New York.

Cuesta E., Rico J.C., Fernández P., Blanco D. and Valino G., 2009. Influence of roughness on surface scanning by means of a laser stripe system. In *Int J Adv Manuf Technol*, 43: 1157–1166.

Coleman, D.E. and Montgomery, D.C., 1993. A systematic approach to planning for a designed industrial experiment, *Technometrics* 35(1): 1–12.

Eriksson L., Johansson E., Kettanwh-Wold N., Wikström C., and Wold S., 2008. Design of experiments Principles and Applications, 3rd Ed., Umea: Umetrics.

Feng H.Y., Liu Y. and Xi F., 2001. Analysis of digitizing errors of a laser scanning system. *Precision Engineering*, 25: 185–191.

Forest J. and Salvi J., 2004. Laser stripe peak detector for 3D scanners. A FIR filter approach. In 17th Int. Conf. on Pattern Recognition (ICPR'04), Cambridge (UK), 3: 646-649.

Huang W., Liu, J., Chalivendra, V., Ceglarek, D., Kong, Z. and Zhou, Y., 2013. Statistical Modal Analysis (SMA) for Variation Characterization and Application in Manufacturing Quality Control. In *IIE Transactions*, Vol. 45.

Liu C. S. and Hu J.S., 1997. Variation Simulation for Deformable Sheet Metal Assemblies Using Finite Element Methods. In *ASME Journal of Manufacturing Science and Engineering*, 119: 368-374.

Lindau B., Lindkvist L., Andersson A. and Soderberg R., 2013. Statistical shape modeling in virtual assembly using PCA-technique. In *Journal of Manufacturing Systems*, 32: 456-463.

Mahmuda M., Joannic D., Royb M., Isheil A. and Fontaine J.F., 2011. 3D part inspection path planning of a laser scanner with control on the uncertainty. In *Computer-Aided Design*, 43: 345–355.

Martínez S., Cuesta E., Barreiro J. and Álvarez B., 2010. Analysis of laser scanning and strategies for dimensional and geometrical control. In *Int J Adv Manuf Technol*, 46: 621–629.

Montgomery, D.C., 2007. Design and Analysis of Experiments, 6th Ed. Set, Wiley, New York.

Prieto F., Lepage R., Boulanger P. and Redarce T., 2003. A CAD-based 3D data acquisition strategy for inspection. In *Machine Vision and Applications*, 15: 76–91.

Paper G

Lanzotti A., Del Giudice D.M., Staiano G., Martorelli M., On the Geometric Accuracy of RepRap Open-Source 3D Printer, on 15[th] of February in 2015 it was submitted to Journal of Mechanical Design – Transactions of the American Society of Mechanical Engineers (ASME), ISSN 10500472 (Q1, IF 1.17 in 2013), 13 pp.

On the Geometric Accuracy of RepRap Open-Source 3D Printer

Antonio Lanzotti, Domenico Maria Del Giudice, Gabriele Staiano, Massimo Martorelli*

Fraunhofer JL IDEAS-CREAMI, Department of Industrial Engineering
University of Naples Federico II, P.le Tecchio, 80, 80125 Naples, Italy
antonio.lanzotti@unina.it; domenicomaria.delgiudice@unina.it;
gabriele.staiano@unina.it;massimo.martorelli@unina.it

*Corresponding author. Tel./Fax: +39 0817682470.

ABSTRACT

In the field of Additive Manufacturing (AM) processes, there is a significant lack of scientific data on the performance of open-source 3D printers in relation to process parameter values. The purpose of this paper is to assess the impact of the main process parameters on the accuracy of a set of typical geometrical features, as obtained with an open-source 3D printer, the RepRap Prusa-Mendel I2.

For this purpose, a benchmarking part was set up, composed of elementary shapes, representing a series of different features. By means of a DoE approach, we were able to assess the effect of two process parameters - layer thickness and flow rate - on five geometrical features: cube, sphere, cylinder, cone and angled surfaces. A high resolution Laser Scanner was used to evaluate the variation between real features and nominal geometry.

On the basis of the experimental results, it was possible to analyze and discuss the main effects of the process parameters on each feature.

These results can help RepRap users in the correct selection of the process parameters with the aim of improving the quality of prototypes.

Keywords – Open-source 3D Printers, Geometric Features, Process Parameters, Laser Scanner.

1. INTRODUCTION

The term Additive Manufacturing (AM) is used to describe those technologies which allow the production of physical objects, made of various materials, through an additive process.The manufacturing process produces layer upon layer of the material, taking its input directly from a 3D data model.

In recent years, the new generation of AM techniques has rapidly become available to the masses thanks mainly to the expiration of some AM patents and to the open-source movement that allowed significant cost reductions.

One of the best known open source projects is the RepRap (Replicating Rapid prototyper) Project, developed by Adrian Bowyer at the University of Bath (UK) [1].

The extraordinary potential of these systems and the current increasing interest towards them, are demonstrated by the development of International Standards [2-7], as defined in a Partner Standards Development Organization (PSDO) cooperation agreement, signed in 2011, between on the one hand ISO Technical Committee 261 on Additive Manufacturing and on the other ASTM International Committee F42.

Although the partnership between ISO and ASTM International in the area of Additive Manufacturing represents a milestone for the Additive Manufacturing community, currently,

standard methods for the assessment of the accuracy of AM systems have not yet been defined.
Companies that sell AM systems use proprietary methods to define the accuracy of their systems. For open-source AM systems, such methods are not developed. As a result, there is a significant lack of scientific data on accuracy and repeatability with these systems.
Over the years, various benchmarking parts for evaluating the accuracy and repeatability of AM processes have been proposed: Kruth in 1991 [8], Lart in 1992 [9], Ippolito, Iuliano and de Filippi in 1994 [10], Juster and Childs in 1994 [11,12], Shellabear in 1999 [13], Mahesh, Wong, Fuh and Loh in 2004 [14], Sercombe and Hopkinson in 2006 [15].
However, none of these benchmarking parts comprehensively includes all the features necessary to establish the desired accuracy/repeatability related to AM process parameters. Recently the authors use the benchmarking part proposed by Fahad and Hopkinson [16], with the aim to evaluate the impact on system accuracy of the main process parameters in a RepRap 3D printer [17].
In a symmetrically repeatable sequence, the part includes elementary shapes representative of the main geometrical features: cube, cylinder, sphere, cone and angled surfaces.
Features have been used widely and successfully in traditional manufacturing processes [18-21]. For example, the feature-based approach allows for the evaluation of a design and for its modification or redesign into one that is functionally acceptable and compatible with a selected manufacturing process. In Additive Manufacturing processes the parts, even of complex shapes, are obtained layer by layer. Therefore, compared to traditional manufacturing, features would seem less useful for these processes. However, for AM to be accepted as a mainstream manufacturing process, parts created by these systems will have to consistently satisfy critical geometric tolerances specifications for various features of the part. Therefore, there is a need to study AM process parameters that influence the accuracy on critical features of the manufactured part [22-24].
The purpose of the paper is to assess the effect of the process parameters on the accuracy of five typical geometric features - cube, sphere, cylinder, cone and angled surfaces – using an open-source 3D printer.

2. MATERIALS AND METHODS

In this study the open-source RepRap Prusa-Mendel I2 3D printer was used. The 3D printer was calibrated using a dial indicator with magnetic base, Mitutoyo 2046-08 (Mitutoyo, Japan) with an accuracy of ±10 μm. Marlin and Cura open-source software were used to get the final 3D objects printed in 2.85 mm PLA biodegradable material.

2.1 Geometric Features

The benchmarking part proposed by Fahad and Hopkinson [16] (Fig. 1) was taken into account for this study. Five geometric features - cube, sphere, vertical cylinder, cone and angled surfaces - (Fig. 1), were selected. In the paper for each feature, it was evaluated the Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} d_i^2} \qquad (1)$$

where d_i is the distance between the correlated points of manufactured real and CAD nominal part and n is the number of points.

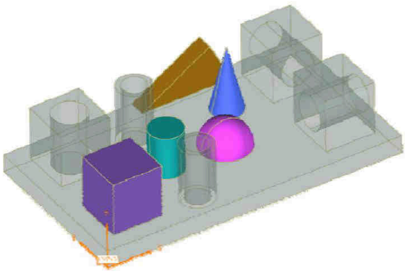


Fig. 1 – Five features selected in the test part

2.2 PROCESS PARAMETERS

Several process parameters are involved in the experimentation. Starting from previous results [17] that take into account the main slicing software (CURA, KISSLICER, SLIC3R), two process parameters were selected:

- Layer thickness is the height of each slice that deposits the 3D printer that is the feed rate along the vertical axis taken before extruding a new layer atop previous one.
- Flow rate is a measure of the material quantity that comes out from the extruder; it is expressed as a percentage of the revolution numbers that the electrical motor has to fulfill for extruding 1 mm of filament.

Layer thickness and flow rate were adopted as control factors, whereas the other process parameters as held factors. The factors are listed in Table 1.

Table 1: List of control and held factors

Factor	Type	Value
Layer thickness (mm)	Control	various
Flow rate %	Control	various
Deposition speed (mm/s)	Held	30
Wall thickness (mm)	Held	0.7
Bottom/top thickness (mm)	Held	0.6
Fill density %	Held	20
Bed temperature (°C)	Held	80
Printing temperature (°C)	Held	200

Based on the aim of the study the range of variation of the control factors was selected. For the layer thickness, the authors, in the previous experience, carried out tests with a layer thickness value of 0.05 mm. These tests showed the poor quality of the parts manufactured using this level. Hence, the minimum value is set to 0.10 mm with increasing in steps of 0.05 mm.
For the flow rate, based on the experience of the RepRap experts the three values 100%, 105%, and 110% were chosen.
Table 2 summarizes the levels of control factors and their settings. Both control factors are quantitative variables.

Table 2: Three level chosen for the control factors

Control Factor	Label	Low (-1)	Mid-level (0)	High (+1)	Unit
Layer thickness	Lt	0.10	0.15	0.20	mm
Flow rate	Fr	100	105	110	%

It's worthwhile to highlight that in [17] a Design of Experiment (DoE) study was carried out focusing the attention on the Layer thickness (Lt), Deposition speed (Ds) and Flow rate (Fr) factors. As three factors were assumed important in the pre-experimental phase, a 3^3 full factorial design was adopted. In this paper the assessments are carried out at the same deposition speed that has been set to the optimal level (30 mm/s).
Table 3 shows the 3^2 full factorial design adopted. For each treatment 3 replicates were planned. 27 PLA benchmark parts (viz. 9 treatments * 3 replicates) were fabricated and 135 geometric features (27 benchmark parts * 5 typical geometric features) were examined. To reduce the effects of variability, treatments were performed in random order. So, the accuracy of geometric test parts, described in Sec. 2.1, was evaluated.

Table 3: 3^2 full factorial design

Treatment	Lt	Fr
1	-1	-1
2	-1	0
3	-1	1
4	0	-1
5	0	0
6	0	1
7	1	-1
8	1	0
9	1	1

2.3 Laser Scanner Acquisition

A high resolution Laser Scanner, D700 Scanner – 3Shape, Denmark was used to get the cloud points of each feature (Fig. 2). The accuracy of this non-contact Reverse Engineering (RE) system is of ±20 µm.
Based on the strict protocol all printed parts were acquired. The acquisitions were conducted by the same operator.
Using an Iterative Closest Point (ICP) algorithm [25, 26], in Geomagic® software by 3D Systems, the points cloud of each feature was aligned with the reference CAD model. So, the analysis of deviations was carried out.

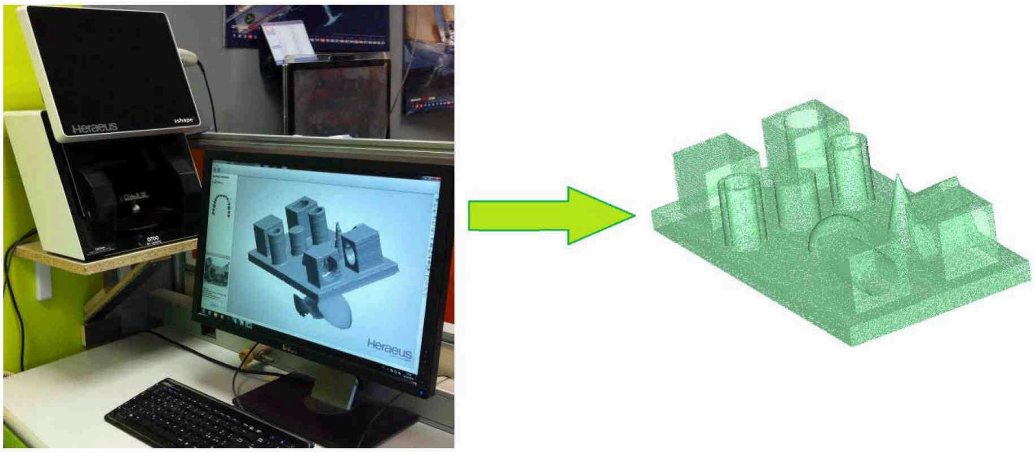


Fig. 2 - Benchmarking part RE acquisition by D700 Laser Scanner

To avoid the noise due by laser scanner acquisition and to make comparable results, a common subset of scan data for each feature was extracted and used in the subsequent analysis (Fig. 3).

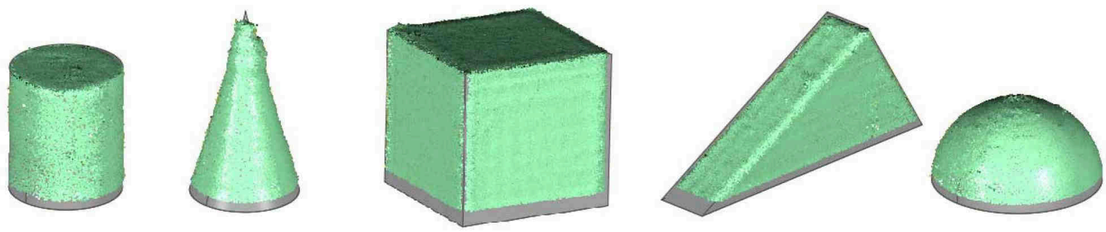


Fig. 3 - Common subset of scan data extracted for each feature

3. RESULTS

Root Mean Square Error (RMSE), as defined in eq. (1), was considered as response variable. This response is a good measure of the accuracy and can be easily obtained applying the procedure described in Sec. 2.3. Using the best fit alignment, all distances between each point clouds and the 3D CAD nominal model were recorded. Figure 4 shows, for example for the first replicate, the colored map of distances between real geometric feature and nominal part that is correlated to a value of RMSE.

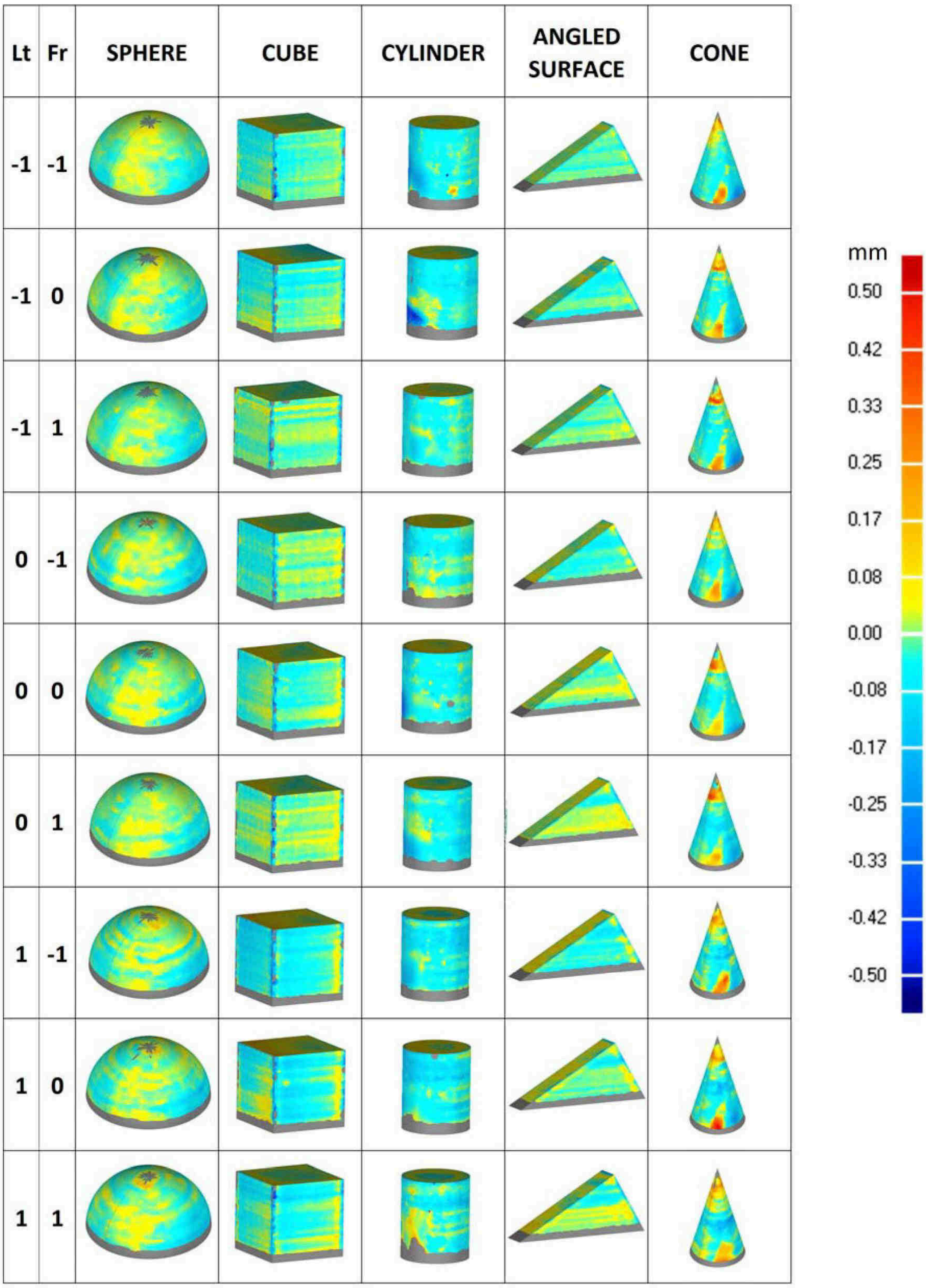


Fig. 4 - RMSE colored maps for the first replicate

For each geometric feature, experimental data and the RMSE responses were collected (Table 4).

Table 4: Full factorial design for geometric features RMSE responses (the dotted line divides two replicates)

		Factor		Response:RMSE (mm)				
Treatment	Replicate	Lt	Fr	Sphere	Angled surface	Cylinder	Cube	Cone
1	1	-1	-1	0.03	0.04	0.09	0.04	0.09
2	1	-1	0	0.03	0.05	0.08	0.05	0.09
3	1	-1	1	0.03	0.04	0.05	0.04	0.09
4	1	0	-1	0.04	0.04	0.06	0.04	0.08
5	1	0	0	0.03	0.04	0.08	0.04	0.08
6	1	0	1	0.03	0.04	0.07	0.04	0.11
7	1	1	-1	0.04	0.06	0.08	0.07	0.10
8	1	1	0	0.04	0.06	0.08	0.07	0.10
9	1	1	1	0.05	0.06	0.06	0.05	0.10
1	2	-1	-1	0.03	0.04	0.07	0.05	0.09
2	2	-1	0	0.03	0.04	0.05	0.04	0.08
3	2	-1	1	0.03	0.04	0.06	0.04	0.09
4	2	0	-1	0.05	0.04	0.12	0.04	0.09
5	2	0	0	0.04	0.04	0.06	0.05	0.08
6	2	0	1	0.03	0.04	0.05	0.05	0.10
7	2	1	-1	0.05	0.06	0.09	0.06	0.10
8	2	1	0	0.05	0.05	0.06	0.05	0.11
9	2	1	1	0.04	0.04	0.06	0.05	0.12
1	3	-1	-1	0.03	0.05	0.06	0.04	0.09
2	3	-1	0	0.04	0.04	0.07	0.04	0.09
3	3	-1	1	0.03	0.04	0.06	0.03	0.08
4	3	0	-1	0.04	0.04	0.07	0.05	0.09
5	3	0	0	0.04	0.04	0.08	0.04	0.07
6	3	0	1	0.04	0.04	0.06	0.04	0.10
7	3	1	-1	0.06	0.06	0.08	0.06	0.12
8	3	1	0	0.05	0.06	0.11	0.05	0.12
9	3	1	1	0.04	0.05	0.06	0.06	0.12

Figure 5 shows the box-plots of RMSEstratified for geometric feature. It's worthwhile that the coneand cylinder RMSE distributionsare not comparable with the others nor between them. Sphere, angled surface and cube RMSE distributions appear comparable, although the sphere obtains the lowest values.

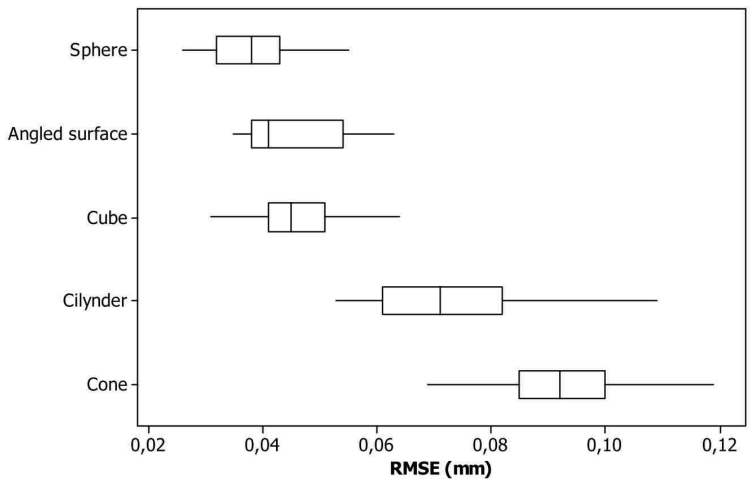


Fig. 5 - Box-plots of RMSE stratified for geometric features

In order to verify if the means of the RMSE distributions of the five geometric features significantly differ, the T-Test of difference (one-tail) was performed. According to the Box-plots shown in Fig. 5, four pair comparisons of geometric featureswere carried out:

1. Sphere vs. Angled surface,
2. Angled surface vs. Cube,
3. Cube vs. Cylinder,
4. Cylinder vs. Cone).

Results show that test 1, 3 and 4 are significant with α equal to 0.05, while test 2 is not significant. In other words in terms of geometric accuracy:

- Sphere s significantly better than angled surface;
- angled surface is not better than cube,
- cube is better than vertical cylinder;
- cylinder is significantly better than cone.

ANOVA table for RMSE shows that the mean effect of factor Lt is significant with α equal to 0.05 for sphere, angled surface, cube and cone; the factor Fr is significant for cylinder, while the interaction (Lt, Fr) is significant for cone (Table 5).

Table 5: ANOVA table for the RMSE (α = 0.05)

		Sphere		Angled surface		Cilynder		Cube		Cone	
Source	DF	Seq SS	p-value	Seq SS	p-value	Seq SS	p-value	Seq SS	p-value	Seq SS	p-value
Lt	2	0.00110	**0.00**	0.00127	**0.00**	0.00047	0.44	0.00147	**0.00**	0.00282	**0.00**
Fr	2	0.00014	0.15	0.00010	0.14	0.00216	**0.04**	0.00014	0.23	0.00047	**0.04**
Lt*Fr	4	0.00010	0.55	0.00008	0.48	0.00018	0.95	0.00010	0.68	0.00071	**0.05**
Error	18	0.00060		0.00040		0.00487		0.00080		0.00107	
Total	26	0.00194		0.00185		0.00767		0.00252		0.00507	

Fig. 6 shows the main effect plots for the 5 geometric features and is useful to evaluate the optimal expected combination of levels that maximise geometric accuracy for each feature (viz. minimize RMSE):

- the level -1 of factor Lt (0.10 mm) for sphere andcube;
- the level 0 of factor Lt (0.15 mm) for angled surface;
- the level 1 of Factor Fr (105%) for cylinder
- the levels 0 of factor Lt (0.15 mm) and 0 of Factor Fr (105%) for cone.

From these results, it is recommended to choose the combination (-1,1) for Lt and Fr factors due to the small difference between the mean effect on angled surface in choosing -1 instead of level 0. The alternative is (0,0) when conic surface are predominant in a part.

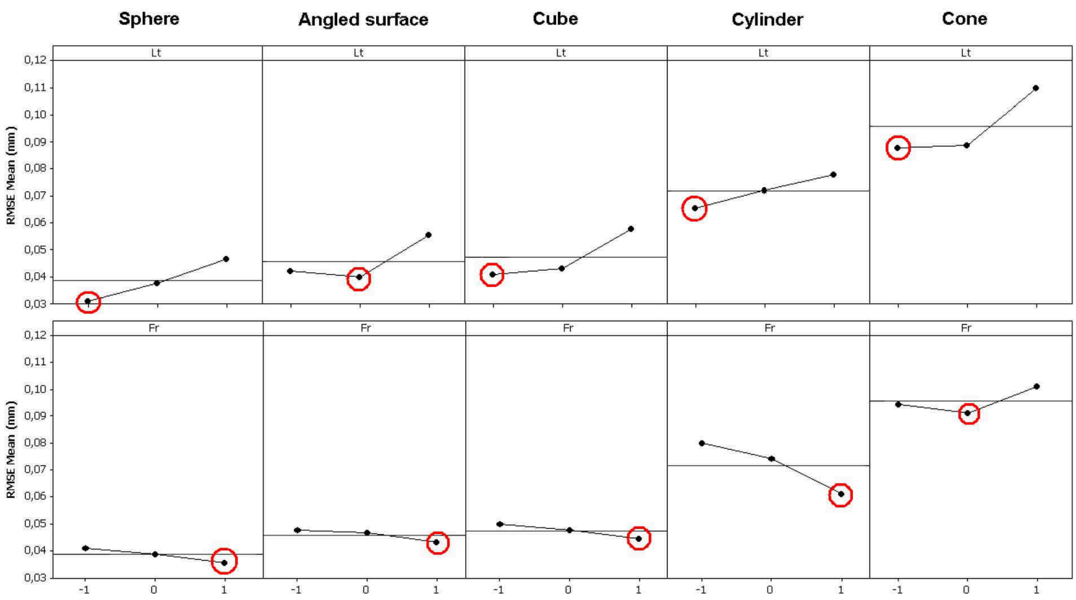


Fig. 6 –Main effect plotsobtained for each geometric feature for RMSE

The interaction plot (Fig. 7) for Cone shows that the levels 0 (105%) of Factor Fr and 0 (0.15 mm) of Factor Lt are the optimal choice. The increment of RMSE due to the mean effect from -1 to 0 is counter balanced by the decrement due to interaction.

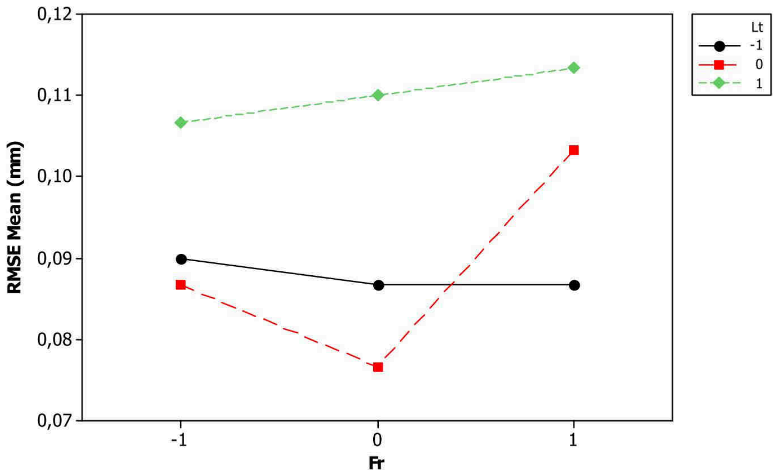


Fig. 7 -Interaction plot for RMSE Cone

Table 6 shows for each treatment the marginal means and signal-to-noise values evaluated for each treatment. The treatment 3 is the most robust combination of control factors*i.e.* it is the combination that improves the geometric accuracy of the test part.
The treatment 5 is the second choice in terms of S/N.

Table 6: Marginal means and Signal to Noise (S/N)ratio based on RMSE data

	Factor		Response:mean RMSE (mm)					Marginal response	
Treatment	Lt	Fr	Sphere	Angled surface	Cylinder	Cube	Cone	Mean	S/N
1	-1	-1	0.03	0.04	0.07	0.05	0.09	0.06	24.33
2	-1	0	0.03	0.04	0.07	0.04	0.09	0.06	24.10
3	-1	1	0.03	0.04	0.06	0.05	0.09	0.05	24.78
4	0	-1	0.04	0.04	0.08	0.05	0.08	0.06	24.06
5	0	0	0.04	0.04	0.07	0.04	0.08	0.06	24.53
6	0	1	0.03	0.04	0.06	0.04	0.10	0.06	23.52
7	1	-1	0.05	0.06	0.08	0.07	0.10	0.07	22.60
8	1	0	0.04	0.06	0.09	0.06	0.11	0.07	22.49
9	1	1	0.04	0.05	0.06	0.05	0.11	0.07	23.02

Tab. 7 shows the results of the ANOVA test: the factor Lt and the interaction are significant with α equal to 0.05. So, the variation of the marginal meanfor the five geometric features is most influenced by the factor Lt.

Table 7: ANOVA for Marginal Mean (α = 0.05)

		Mean	
Source	DF	Seq SS	p-value
Lt	2	0.00103	**0.00**
Fr	2	0.00003	0.53
Lt*Fr	4	0.00028	**0.04**
Error	18	0.00040	
Total	26	0.00174	

The main effects plot for the mean (Fig. 8) and the interaction plot (Fig. 9) provides the expected optimal combination:the levels -1 of Factor Lt (0.10 mm) and 1 of Factor Fr (110%) identifies the most robust treatment.

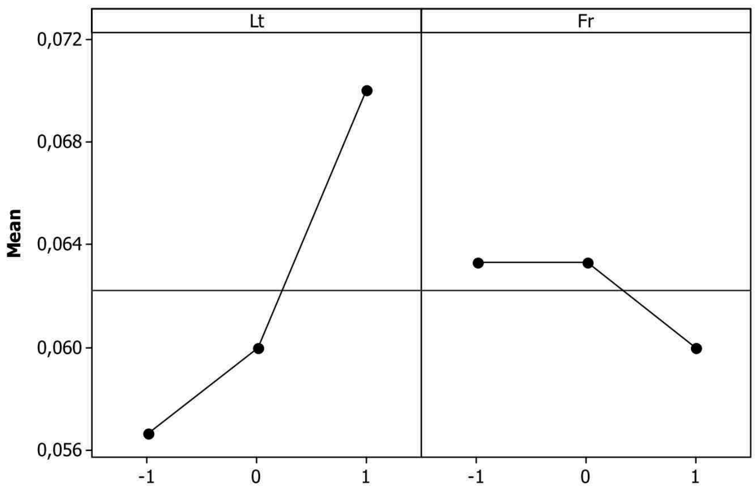


Fig. 8 - Main effects plot for RMSE Overall Mean

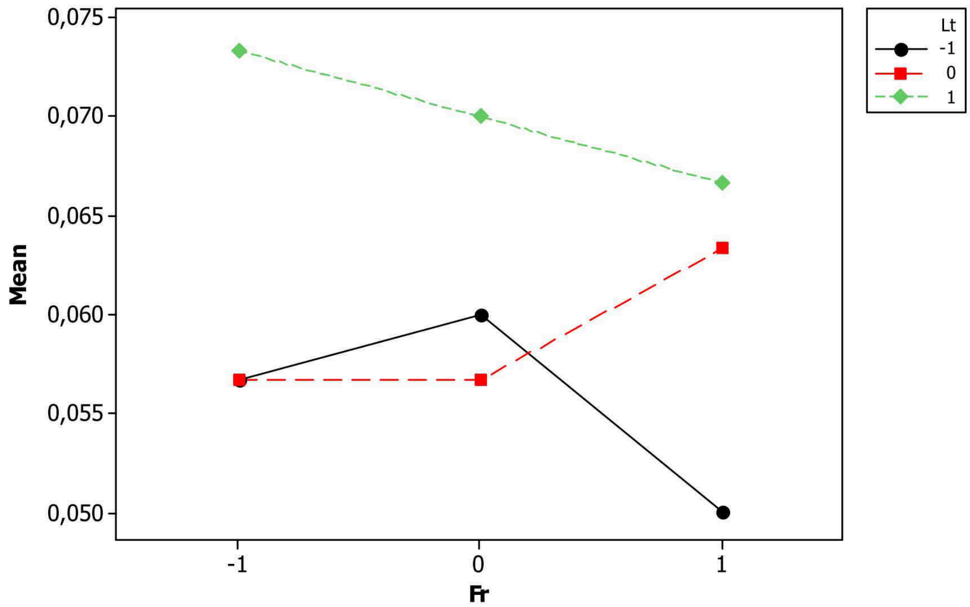


Fig. 9 –Interaction plot shows synergic and anti-synergic effects on all features

4. DISCUSSION

The present study concerns the effects of two process parameters on the accuracy of geometric features manufactured using an open-source 3D printer. Starting from the results collected in the experimental phase, it can be settled out that:

- sphere, angled surface, cube and cone are significantly affected by changes of Layer thickness. The level 0 (0.15 mm) maximizes the accuracy of angled surface, whereas the level -1 (0.10 mm) is the best choice for the other ones. This results agrees with a “rule of thumb” that empirically suggest a value of layer thickness equal to one-fourth of the nozzle as optimal choice (in this test the diameter is 0.35 mm). The choice of level -1 for angled surface is acceptable being the RMSE differences very small if compared to the optimal level 0;
- cylinder and cone are significantly affected by changes of Flow rate control factor. The level 1 (110%) maximizes the accuracy of cylinder, while the level 0 (105%) is the optimal one for thecone.So the practical suggestion to increase the flow rate over 100% is correct and the effects on the accuracy are robust, being the improvement obtained in a wide range (from 105% to 110%). The level 1 of Fr is a robust choice even for the sphere, angled surface and cube geometric features for which the flow rate is not significant (see the maximum S/N value in Table 7);
- cone shows the highest RMSE values, whereas the sphere achieved a greater level of accuracy, in any condition of the process parameters. These results agree with expectations of experts.

Besides, in order to improvethe accuracy whenall geometric features have to be manufactured as in the case of the test part, some practical suggestions are proposed:

- to maximize the mean accuracy, the layer thickness could be set to level -1. This choice is the most robust, even if it involves aquite small loss of accuracy for the cone.

- the flow rate could be set to level 1 and this is the optimal choice taking into account interaction effects.

5. CONCLUDING REMARKS

ISO and ASTM Technical Committees are working on the development of International Standards for Additive Manufacturing to be adopted worldwide. However, currently standard methods for the assessment of the accuracy of AM systems have not yet been defined. The benchmarking part used in this paper can be suitably fabricated by different AM processes, with a view to assessing the accuracy and repeatability of the system.
In the paper, the benchmarking part was used to analyze the impact of the layer thickness and flow rate process parameters on the accuracy of the production of five typical geometric features in an open-source 3D printer.
Taking into account the limits of the present investigation, it can be concluded that with respect to accuracy, the results show that the layer thickness is a significant parameter for 4 out of 5 geometric features considered whereas the flow rate for 2 out of 5, and that the best results are obtained with the lowest values for layer thickness (0.10 mm for this study) and the highest one for flow rate (110%). This is true except for the cone.
The highest values for layer thickness (0.20 mm in this study) and the lowest values for flow rate (100%) produce loss in accuracy for geometric features (except the cone).
If you consider the interactions of different geometric features you cannot exclude some potential non-linear effects of process parameters (see the significant effect of interaction for cone).
The results discussed in the paper provide the AM community with additional scientific data on the impact of process parameters on the quality of parts obtained using a Rep-Rap 3D printer. Being quality, a critical factor for the industrial successful application of AM processes, future work should be carried out to understand the effect of process parameters on test parts with complex shape to satisfy geometric tolerance specifications and to validate predictive models useful to anticipate expected accuracy.

ACKNOWLEDGMENTS

This work has been supported by Project POR FSE 2007/13 entitled “WISCH, Work Into Shaping Campania’s Home”.

REFERENCES

[1] Jones R., Haufe P., Sells E., Iravani P., Olliver V., Palmer C., Bowyer A., RepRap-the replicating rapid prototyper. Robotica, 2011, Vol. 29, Issue 1, pp. 177-191.

[2] ISO/ASTM 52915: 2013, Standard specification for Additive Manufacturing file Format (AMF) Version 1.1.

[3] ISO/ASTM 52921: 2013, Standard terminology for Additive Manufacturing-Coordinate systems and test methodologies.

[4] ISO 17296-1:2014, Additive manufacturing - General - Part 1: Terminology.

[5] ISO 17296-4:2014, Additive manufacturing - General principles - Part 4: Overview of data processing technologies, ASTM Fact Sheet.

[6] ISO 17296-3:2014, Additive manufacturing - General principles - Part 3: Main characteristics and corresponding test methods.

[7] ISO 17296-2:2015,Additive manufacturing - General principles - Part 2: Overview of process categories and feedstock.

[8] Kruth J. P., Material incress manufacturing by rapid prototyping techniques CIRP Annals, 1991, 40(2), pp. 1603-1615.

[9] Lart G., Comparison of rapid prototyping systems, Proceedings of First European Conference on Rapid Prototyping, 1992, University of Nottingham, pp. 243-254.

[10] Ippolito NR.,Iuliano L., de Filippi A., A new user part for performance evaluation of rapid prototyping systems, Proc. of Third European Conference on Rapid Prototyping and Manufacturing, 1994, University of Nottingham, pp. 327-339.

[11] Juster N.P., Childs T.H.C., Linear and geometric accuracies from layer manufacturing, CIRP annals, 1994, 43(1), pp. 163-166.

[12] Juster N.P., Childs T.H.C., A comparison of rapid prototyping processes, Proc. of Third European Conference on Rapid Prototyping and Manufacturing, 1994, University of Nottingham, pp. 35-52.

[13] Shellabear M., Benchmarking study of accuracy and surface quality in RP models, RAPTEC, 1999, Task 4.2, Report 2.

[14] Mahesh, M., Wong, Y. S., Fuh, Y. H., and Loh, H. T. Benchmarking for comparative evaluation of RP systems and processes. Rapid Prototyping J., 2004, 10(2), 123-35.

[15] Sercombe T. B., Hopkinson N., Process shrinkage and accuracy during indirect laser sintering of Aluminum, Advanced Engineering Materials, 2006, Vol. 8, Number 4, pp. 260-264.

[16] Fahad M., Hopkinson N., A new benchmarking part for evaluating the accuracy and repeatability of Additive Manufacturing (AM) processes, 2nd International Conference on Mechanical, Production and Automobile Engineering (ICMPAE 2012), 2012, Singapore April 28-29.

[17] Lanzotti A., Martorelli M., Staiano G., Understanding Process Parameter Effects of RepRap Open-Source Three-Dimensional Printers Through a Design of Experiments Approach, Journal of Manufacturing Science and Engineering, 2015, 137(1), pp. 1-7, ISSN: 1087-1357, Transactions of the ASME.

[18] Shah J., Mantyla M., Parametric and Feature-based CAD/CAM: Concepts, Techniques, and Applications, Wiley-inter-science, 1995, ISBN 0-471-00214-3.

[19] Chen Y.M. et al., Extraction of geometric characteristics for manufacturability assessment, Robotics and Computer Integrated Manufacturing, 2003, 19, pp. 371–385.

[20] Gayretli A., Abdalla H.S., A feature-based prototype system for the evaluation and optimisation of manufacturing processes, Computers & Industrial Engineering,1999, 37, pp. 481-484.

[21] Ip C.Y., Regli W.C., A 3D object classifier for discriminating manufacturing processes, Computers & Graphics, 2006, 30, pp. 903–916.

[22] Zhang Y., Bernard A., Using AM feature and multi-attribute decision making to orientate part in Additive Manufacturing, in High Value Manufacturing: Advanced Research in Virtual and Rapid Prototyping, Bàrtolo et al. (Editors), Taylor & Francis Group, London, 2014, ISBN 978-1-138-00137-4.

[23] Rosen D.W., Computer-Aided Design for Additive Manufacturing of Cellular structures, International CAD Conference and Exhibition, Honolulu, 2007, June 25-29.

[24] Qian X., Dutta D.,Feature Based Fabrication in Layered Manufacturing, Journal of Mechanical Design, 2001, 123 (3), pp. 337-345.

[25] Besl PJ, McKay ND. A method for registration of 3-D shapes.IEEE Transactions on Pattern Analysis and Machine Intelligence1992;14: pp. 239-56.

[26] Pottmann H, Leopoldseder S, Hofer M. Simultaneousregistration of multiple views of a 3D object. ProceedingsPCV’02, Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 3A, CommissionIII. 2002.