Development and validation of a product creativity evaluation framework for the assessment of functional consumer products

by

Kamyar Hazeri

Imperial College London, Dyson School of Design Engineering

A thesis submitted for the degree of Doctor of Philosophy in Design Engineering

February 2019

Declaration of Originality

I hereby certify that the intellectual content of this thesis is prepared and developed by my own except where explicitly stated otherwise by reference or acknowledgment. I declare that neither any part of this thesis nor the whole of the thesis has been submitted for any degree to any other University or Institution.

Kamyar Hazeri

February 2019

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work

Abstract

A reliable method for creativity assessment of functional consumer products has been developed. In literature, creativity is explored in four main aspects of person, process, press, and product known as 4P's of creativity. In this research, the product creativity measurement is studied aiming at assisting designers to better understand creativity and to be able to implement it in the design process. The research begins with defining the role of creativity in engineering design and continues with reviewing existing approaches for measuring product creativity in three main categories.

In this research, a goal is to investigate the ability of the CAT for the assessment of consumer products produced under non-experimental conditions. The CAT was used together with domain experts to assess three categories of consumer products for creativity (home-accessories, electric-vehicles, and smoke-alarms). The results of inter-judge reliabilities indicated that CAT can reliably evaluate these products, although a larger number of raters (10 to 15) compared to those in artistic domains is recommended. Factor analysis revealed that for products with complex structures (electric-vehicles), experts are unable to distinguish the border between creativity and other influential factors such as product's technical performance and or aesthetic appeal. As a result, the assessment of complex engineering products should be performed with caution.

A theoretical framework to enable researchers to develop their own criterion-based instrument based on the product's functionality is presented. The framework comprises 65 indicators in nine broad categories. A case study using 5 lounge-chairs and 288 university students was conducted to test the applicability and validity of the framework. A new internally consistent and factorially valid model was achieved. Overall, this project contributed to the field by extending the usage of CAT into functional products and by proposing a new framework to evaluate creativity. Accordingly, researchers/designers can get insights and enhance product creativity.

Table of Contents

Declaration of Originality ii

Abstract iii

Table of Contents iv

List of Figures vi

List of Tables viii

Acknowledgement ix

List of Abbreviations x

1. Introduction 13
   - 1.1. Research aim 17
   - 1.2. Research structure 18
2. Introduction to engineering design and creativity 23
   - 2.1. Overview of engineering design 23
   - 2.2. Design process 24
   - 2.3. Engineering problem-solving 28
   - 2.4. Creativity in Engineering Design 31
   - 2.5. Measuring creativity in engineering domain 38
   - 2.6. Chapter conclusion 45
3. Product Creativity Assessment in Engineering Design 47
   - 3.1. Objective methods 48
   - 3.2. Indirect measurement methods 53
   - 3.3. Global judgment 55
   - 3.4. Criterion-based measurement methods 60
   - 3.5. Comparison of the product creativity measurement methods 68

3.6. Chapter Conclusion 69

4. Investigation of the Extension of Consensual Assessment Technique into Consumer products: three case studies 72

4.1. Research aim and objectives 74

4.2. Methodology 75

4.3. Results of pilot study 82

4.4. Results and Discussion 82

4.5. Chapter Conclusion 90

5. Development of a Theoretical Framework for Evaluation of Functional Consumer Product Creativity 92

5.1. Scope of measurement 93

5.2. Proposal of a new theoretical model for the evaluation of functional consumer product creativity 95

5.3. Research Methodology 105

5.4. Quantitative data analysis 108

5.5. Chapter conclusion 114

6. Confirming the Validity of the Consumer Product Creativity Evaluation Framework 117

6.1. Research aim and objectives 119

6.2. Methodology 120

6.3. Results 124

6.4. Discussion 130

6.5. Research limitation 134

6.6. Future work 135

6.7. Chapter conclusion 136

7. Overall Discussion and Conclusion 138

7.1. Summary of the main findings 139

7.2. Guidelines for implementing the measurement approaches 147

7.3. Directions for future research 152

7.4. Chapter conclusion 153

References 155

Appendix: Questionnaire 180

List of Figures

Figure 1-1 NASA's BEST Engineering Design Model (NASA, 2018) 14
Figure 2-1 Total design process (Pugh, 1991) 25
Figure 2-2 Design process steps (Pahl & Beitz, 1996) 26
Figure 2-3 General problem-solving process 28
Figure 2-4 Creative Problem Solving (CPS) schematic 30
Figure 2-5 Wallas's creative process model 34
Figure 2-6 Relationship between design process and creative design process (Adapted and updated from Wong & Siu, 2012) 35
Figure 3-1 An example of measuring novelty using proposed method by Chakrabarti and Khadilkar (2003) 50
Figure 3-2 SAPPhIRE model of causality (Adapted and updated from Sarkar and Chakrabarti, 2011) 52
Figure 3-3 Method of assessing novelty (Adapted and updated from Sarkar and Chakrabarti, 2011) 53
Figure 3-4 Revised Creative Product Analysis Matrix (CPAM) (Besemer & O'Quin, 1986) 62
Figure 3-5 Revised CSDS Structure (Cropley & Kaufman, 2012) 64
Figure 3-6 Product Creativity Measurement Instrument (PCMI) (Horn & Salvendy, 2009) 67
Figure 4-1 Three groups of experts independently assessed the products of their corresponding group. Expert judges were first told to observe and compare the products to one another and then asked to separately evaluate each product for creativity and additional nine dimensions using a 5-point Likert scale 78
Figure 4-2 Screenshots of the product presentation videos for electric-vehicles group. Images show different product feature at a time (e.g. charging time, technology, safety, and practicality) 79
Figure 5-1 New framework for measuring the creativity of functional consumer products 97

Figure 6-1 Standardized estimates of the five-factor model (x^2=520.923, df=199, p<0.001, RMSEA=0.035, CFI=0.981)..... 128

List of Tables

Table 1-1 Thesis Overview 20
Table 2-1 Associated dimensions of creativity based on 4Ps of creativity (Adapted and updated from Horn & Salvendy, 2006) 38
Table 3-1 Levels of creativity in products (Adapted and updated from Cropley and Kaufman, 2012) 63
Table 4-1 Boundary of performance for selected consumer products 77
Table 4-2 Dimensions of product creativity 80
Table 4-3 Factor analysis on 10 dimensions of judgment for all groups of products 83
Table 4-4 AVE and squared correlations between the constructs for home accessory and smoke alarm groups 85
Table 4-5 Significant predictors of creativity per product category 86
Table 4-6 Inter-judge reliabilities for group of 5, 10, 15, 20, and 25 randomly selected raters per product category 87
Table 6-1 Inter-judge reliability for each product as well as the entire dataset 124
Table 6-2 The five-factor model extracted from the EFA analysis 125
Table 6-3 The degree of internal consistency between the items of each constructs 126
Table 6-4 The AVE, Composite Reliability (CR), and squared correlations between the constructs 129
Table 6-5 Factor score predictions of creativity ratings 129
Table 6-6 Significant predictors of product creativity for the lounge chairs 130
Table 6-7 The unique criterion-based instrument obtained for creativity measurement of lounge chairs 132
Table 7-1 The theoretical framework for evaluating functional product creativity 144

Acknowledgement

I would like to thank my supervisor, Professor Peter Childs who has provided guidance and support during this research. It was his enthusiasm that encouraged me to research and develop my knowledge in this field, and his patience and wisdom which kept me motivated and inspired throughout this journey. I would like also to thank my secondary supervisor Dr Marco Aurisicchio for his support and guidance.

None of this would have been possible without the support of my family who have persuaded me to understand my goals and encouraged me to work for them. I am fortunate to have parents who gave me peace in my most stressful circumstances, who have also supported me financially.

Special thanks to Professor David Cropley for his tremendous help and advice on the project. For his insightful comments and recommendations through the collaborative work done to accomplish the first phase of the research.

I am grateful for all the guidance and assistance I have received from the faculty of Dyson School of Design Engineering, in particular Dr Nejra Van Zalk and Dr Samuel J Cooper. Thanks also goes to all my colleagues and friends in the Design Engineering group at Imperial College London for their contribution in my thinking, and all of whom made research more enjoyable.

List of Abbreviations

AVE _ Average Variance Extracted

BICB _ Biographical Inventory of Creative Behaviours

CAQ _ Creativity Achievement Questionnaire

CAT _ Consensual Assessment Technique

CBI _ Creative Behaviour Inventory

CCQ _ Creative Climate Questionnaire

CDQ _ Creativity Domain Questionnaire

CFA _ Confirmatory Factor Analysis

CFI _ Comparative Fit Index

CI _ Confidence Interval

CMIN _ chi-square

CPAM _ Creative Product Analysis Matrix

CPS _ Creative Personality Scale

CPS _ Creative Problem Solving

CPSS _ Creative Product Semantic Scale

CR _ Composite Reliability

CSDD _ Creativity Scale for Diverse Domains

CSDS _ Creative Solution Diagnosis Scale

DE _ Design Engineering

DF _ Degree of Freedom

DFA _ Design for Assembly

DFM _ Design for Manufacturing

DT _ Divergent Thinking

EDP _ Engineering Design Process

EFA _ Exploratory Factor Analysis

EV _ Electric Vehicle

FFM _ Five-Factor Model

GFI _ Goodness of Fit Index

GIFT _ Group Inventory for Finding Creative Talents

K-DOCS _ Kaufman Domain of Creativity Scale

KMO _ Kaiser-Meyer-Olkin

MLR _ Multiple Linear Regression

NFI _ Normed Fit Index

NNFI _ Non-Normed Fit Index

PAF _ Principle Axis Factoring

PCMI _ Product Creativity Measurement Instrument

RCFI _ Robust Comparative Fix Index

RMSEA _ Root Mean Square Error of Approximation

SD _ Standard Deviation

SSM _ Soft System Methodology

SSSI _ Siegel Scale of Support for Innovation

SUS _ System Usability Scale

TCI _ Team Climate Inventory

TLI _ Tucker-Lewis Index

TRIZ _ Theory of Inventive Problem Solving

TTCT _ Torrance Tests of Creative Thinking

1. Introduction

Creativity is often cited as the major factor capable of differentiating the products of small and medium-sized enterprises and is essential for their success and survival (Weerasiri, Zhengang, & Perera, 2012). It is among the factors that positively influence the revenue growth of innovative organisations while giving them the opportunity to enhance their profit growth, market share, and customer satisfaction (Csikszentmihalyi & Sawyer, 2014; Von Nordenflycht, 2007). Creativity provokes consumer's mindset and encourages them to use new products that are performing well at the same time. The results of interviewing over 1500 CEOs worldwide indicated that creativity is the most crucial quality for leadership (IBM, 2010). The study revealed that despite the importance of other qualities such as global thinking or influence, it is the creativity that inspires managers as it persuades execution of new ideas and let firms to become leaders of domains that themselves once generated it. Creativity gives companies their competitive edge and convinces innovation managers to promote new and useful ideas instead of replicating others. Records from the UK treasury indicate 75% of the revenue of the top innovative companies is generated from products or services that did not exist a few years ago (Cox, 2005). All these explanations emphasize the importance of creativity for businesses.

The BCG consulting group has recently listed over 50 companies as the most innovative firms of the year 2018 and among them were Apple, Netflix, and BMW (BCG, 2018). From an innovation perspective an arising question warranting attention is 'what factors discriminant Apple's products from the others?' and/or 'what elements have caused the vehicles made by BMW group to be seen as unique and original?'. It appears that creativity up to some extent is able to answer these questions. Indeed, every creative product is the result of visualizing the common sense that exists between the consumer and the manufacturer. On the one hand, creativity is the main driver for innovation and profitability of the firms, on the other hand, it acts as an instrument to fulfil customer requirements in a way they want to see it. Firms use creativity as a tool to generate

solutions that are both novel and useful. This ability of creativity later founds the basis for a potential innovative output (Cropley & Cropley, 2005).

The drive to continuously discover deficiencies in products and the ability to replace them with alternative thoughts that convert unproductive actions into ones that are new and effective is what inspires creative thinking. However, understanding creativity within the context of engineering design requires much deeper investigations. Design Engineering (DE) is a practice to develop new ideas and elaborate them into shapes and forms that are capable of enriching people's lives. In other words, engineering design is a systematic intellectual process that designers utilise to detect problems, create novel ideas and determine appropriate solutions that can meet the consumer needs (Dym, 2006). Within the engineering design discipline, the design process is a segment that is responsible to guide designers in various stages of the product development process and aims to enhance work efficiency by integrating engineering knowledge and applying them in the right place. In overall, the process can be summarized into a series of investigations and collaborations with major design sources to improve customer satisfaction as well as company profit (Ulrich & Eppinger, 2004). NASA (2018) describes the Engineering Design Process (EDP) as an iterative process that constitutes a series of fundamental actions (see Figure 1-1). Usually, the process starts with the identification of the needs and constraints and continues to the development of potential new concepts. It then leads the design team toward the selection of appropriate solutions and prepares them for prototyping their idea. The procedure ends

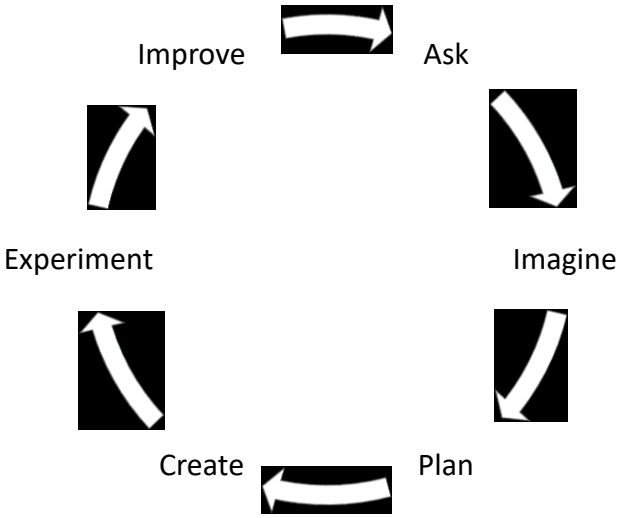


Figure 1-1 NASA's BEST Engineering Design Model (NASA, 2018)

typically by applying additional alterations and improvements to the design or re-starts from the beginning if designers discover further problems.

The role of creativity as a quality in the design process is essential (Christiaans & Venselaar, 2005). Creativity offers the ability to create something novel and useful which founds the basis for a potential innovative output (Cropley & Cropley, 2005). The emergence of creativity is the first stage of the innovative process as the quality of the initial idea directly influences the characteristics of the final work. Creativity rises from individuals and what happens inside their head (Jones, 1992). However, in creativity research, there is a distinction between a creative person and creative product (M. H. Chen & Kaufmann, 2008; Rhodes, 1961). Understanding the difference between the two latter is essential at the conceptual design stage where the employment of creative person does not necessarily lead to creative outcome. Indeed, the quality of the final work depends on further factors such as the thinking mechanism and/or the environment to which these interactions take place (Cropley, 2016).

Understanding how creativity takes place at the conceptual design stage can help designers to generate large number of ideas. Nonetheless, the assessment of creative ideas (or products) does not necessarily require consideration of the process (Amabile, 1982). In fact, the assessment in this context should be done separately as it then would allow companies to recognise creativity more precisely at any stage that it occurs. In addition, the independent study of creative products can help to explain inconsistencies in other components of creativity (Horn & Salvendy, 2006a). Other researchers have also referred to product evaluation as the most practical assessment of creativity (Baer, Kaufman, & Gentile, 2004).

Creativity is a dynamic and continuous phenomenon (Piffer, 2012), which means any predetermined definition for creativity that attempts to identify it within a closed boundary is indeed contradicts with the essence of it. This to some extent explains the reasons for using subjective rather than objective methods for the assessment of product creativity. Since the beginning of the creativity research, scholars have emphasised that a subjective measurement of product creativity is something relatively feasible and reliable (Amabile, 1982; Baer & McKool, 2009a; C. Chen et al., 2002; Freeman, Son, & McRoberts, 2015; B. Hennessey, Gia, Guomin, &

Weiwei, 2008; O'Quin & Besemer, 1989). Researchers over decades have attempted to develop reliable assessment tools to measure product creativity, and as a result, a number of methods have been emerged (Amabile, 1996; Cropley, Kaufman, & Cropley, 2011). It is widely believed that for measurements to be reliable, the presence of domain relevant experts is essential (Amabile, 1996). The complex structure of creativity, and the fact that it can appear at distinctive shapes for different products makes it difficult for novice raters to evaluate creativity with a high degree of consistency, and as a result, their validity for such measurement remains obscure.

Researchers have employed experts to measure creativity of artefacts in various fields such as art, poetry and music (Baer, 1994; Brinkman, 1999; B. Hennessey & Amabile, 1988; Hickey, 1995; Runco, 1989). In addition, most of the products where their creativity was judged by domain experts were produced under experimental conditions where individuals generated tangible objects using pre-identified equipment and boundary conditions (B. Hennessey et al., 2008). However, less attention has been paid to the creativity assessment of functional consumer products. In this thesis, the term "functional product" represents products that serve as solutions to existing problems. They have specific functions and are considered as practical, useful and beneficial to the consumer and in this sense, they are distinctive from those that are purely aesthetic. Cropley and Cropley (2005), referred to functional creativity as objects, systems, or processes that are capable of performing tasks or solving problems. Functional products comprise of hardware, software and service support system (Lindstrom, Sas, Lideskog, Lofstrand, & Karlsson, 2015). The hardware represents any tangible assets in the product, the software refers to the operating tool and system, and the service support system denotes the maintenance of the product, operations planning and educating and training the usage of it.

This research further undertook the investigation of the applicability and reliability of domain experts to assess the creativity of products produced under non-experimental conditions (non-parallel products). The terms "non-parallel" and "non-experimental" products in this thesis refer to products sharing the same functionality, but they are not produced with exact same materials, resources, equipment, procedures, and not designed within the same time period, environment, and conditions. For instance, the main purpose of two comparable cars is to provide convenient transportation for consumers; however, they can follow distinctive mindsets and attitudes (e.g.

sport, luxury and executive), offer different features, and even consume different types of fuel. Baer, Kaufman, and Gentile (2004) referred to these types of products as those that are "... not created in response to the same assignments or prompts" (p. 116).

Another issue that has been less addressed in the relevant literature is to find out how the measurement of product creativity could be improved to allow designers to better understand creativity as a result of measurement, and to be able to implement it in the design process. This is important since the first step after performing any measurement is to undertake actions to resolve the shortage of creativity. Thereby, the outcomes of measuring product creativity should also include essential information that can support designers to realise exactly what areas of the product is currently missing the necessary quality.

1.1. Research aim

This research aims to identify a reliable method for assessing the creativity of functional consumer products. In addition, the measurement of product creativity is explored aiming at assisting designers to better understand creativity and to be able to implement it in the design process. The main questions that this research intends to answer are as follow:

1) How to assess the creativity of functional consumer products?
2) How reliable is this assessment? What elements influence assessment reliability?
3) How the assessment of product creativity could be improved to make it possible for designers to better recognize creativity and to be able to implement it in the product development process?

Below is the list of objectives that need to be addressed in order to fulfil the research aim.

- Objective I:

To comprehend the role of creativity in engineering design and to explore different components of creativity. In addition, to identify the definition of creativity within engineering design discipline and to find out how to assess creativity in general.

- Objective II:

To review and compare already existing methods for the assessment of product creativity. To determine what measurement approaches can be employed for the application of functional consumer products.

- Objective III:

To investigate the use consensual assessment technique to assess the creativity of functional consumer products produced under non-experimental conditions. To find out whether the reliability of creativity assessment differs between functional and artistic products. To understand whether the number of raters required for a reliable assessment is equivalent to those of reported in artistic domain.

- Objective IV:

To develop a theoretical framework allowing the formation of unique measurement tools with respect to product functionality as well as assisting with understanding consumers' perception of creativity. Also, to support researchers and designers in the research and development, marketing, ideation and conceptualization departments to better understand creativity and to enable them to get insights from the list of indicators provided in the framework to enhance creativity.

- Objective V:

To exemplify the use of framework for generating a unique measurement instrument for a specific category of products. To test the reliability and validity of the theoretical framework.

1.2. Research structure

To address the research aims and to be able to respond to questions arising from each research objective, this thesis makes use of both theoretical approaches and empirical studies to investigate creativity in products. The overall research path within this thesis includes a review of the relevant literature, administering surveys to understand how certain viewers perceive creativity for specific products, and proposition of a theoretical model to further improve the measurement of functional product creativity.

The current literature on creativity and methods for assessing it is explored focusing on understanding the factors which have an effect on creativity measurement as well as to specify the appropriate measurement approaches for the application of functional consumer products. Several empirical studies are conducted to evaluate the reliability of utilising the panel of experts for the assessment of functional products and to compare the results with those reported in the artistic domain. In addition, to further improve the measurement precision, a theoretical framework containing several creativity indicators and factors is developed based on the synthesis of the literature. The new framework attempts to support researchers and practitioners in the research and development, marketing, ideation and conceptualization departments by facilitating them with the right tool to recognise creativity and assisting with implementation of creativity within product development process. In addition, any researcher who intends to better understand how consumers of certain types of products perceive creativity could benefit from this framework and get effective insight from the list of indicators provided in this instrument. A case study is conducted to assess the capability and validity of the theoretical framework as well as to exemplify the procedure of utilising it to generate a unique measurement instrument for a particular category of products. Finally, based on the outcomes of the thesis conclusions are drawn and new research directions are presented for future study. Table 1-1 presents the structure of this thesis.

Table 1-1 Thesis Overview

Section	Chapter	Objective	Key Contribution(s)
Introduction	One	Introduce the concept of creativity and its importance for the design process	Establish research orientation; Emphasise the need for and extending the creativity assessment to functional consumer products
Role of creativity in engineering design	Two	Understand creativity within engineering design discipline	Understand creativity in engineering design; Define various components of creativity and understand how creativity is measured in general (Review 1)
Product creativity assessment	Three	Investigate the assessment of product creativity	Review of existing methods for product creativity assessment; Understand the applicability of the current approaches for functional consumer products (Review 2)
Extension of the CAT to functional consumer products	Four	Investigate the capability of the domain experts for evaluating functional consumer products	Identify the degree of reliability of the consensual assessment technique for evaluating non-parallel functional products (Study 1, 2, 3)
Proposal of a theoretical framework for evaluating functional consumer products	Five	Develop a theoretical framework to evaluate the creativity of functional consumer products	Theoretical framework for enabling the formation of unique measurement instruments based upon product's functionality
Reliability and validity of the theoretical framework	Six	Demonstrate the procedure of making use of the framework to generate a distinct measurement instrument for a particular category of products	Confirm the reliability and validity of the theoretical framework (Study 4)
Discussion and conclusion	Seven	Present conclusions from the thesis and suggest future directions for the research	Insights from the studies and directions for future work

Chapter one introduces the concept of creativity and reviews its importance in engineering and business. The chapter attempts to establish the overall research orientation and emphasise the need for and extending the creativity assessment to functional consumer products.

Chapter two discusses the role of creativity in the engineering design discipline. The chapter begins with an exploration of the engineering design as well as various parts of the design process with more attention to creativity. Main components that have an impact on the emergence of creativity are described and further explanation is provided to indicate how creativity is measured in general. According to each aspect of creativity, several measurement methods are briefly introduced, with particular attention to measurement of product creativity. The literature for the definition of (product) creativity in engineering domain is presented while primary methods for the assessment of product creativity are introduced.

Chapter three reviews the current approaches for the measurement of product creativity in three broad categories including indirect measurement, global-judgments, and criterion-based instruments. The chapter also includes a brief explanation of the present objective practices for the assessment of creativity and novelty. Based upon each form of measurement, several techniques and strategies to reliably assess product creativity in various domains are presented, with specific attention to consensual assessment technique and criterion-based instruments. Relevant measurement methods are compared to one another and a conclusion is drawn to signify which of the methods are more appropriate for the application of functional consumer products.

Chapter four investigates the capability of the consensual assessment technique for assessing three categories of non-parallel consumer products. The chapter presents the results of creativity ratings generated by domain relevant experts for each category of products. The chapter also explains the methodology used for data gathering including participants, material, judging, and overall operational procedure. It gives insights upon the degree of reliability of the consensual technique for functional consumer products and further discusses the considerations for conducting a valid assessment.

Chapter five introduces a theoretical framework consisting of relevant criteria for functional products to allow the formation of unique criterion-based instruments dependent on a products' functionality. The chapter reviews the relationship between creativity and each of the criteria proposed. To confirm the reliability and validity of the theoretical framework, a case study is conducted. This chapter explains the methods that have been considered for data gathering and analysis of the data for conducting a quality case study. Certain factors affecting the outcomes of the study has also been reviewed such as sample size and survey design.

Chapter six demonstrates a case study to test the applicability and capability of the theoretical framework. It employs the new framework to generate a unique criterion-based instrument for a specific category of consumer products (i.e., a series of lounge chairs). The chapter reports the results of surveying 288 university students rating the products using the indicators proposed in this thesis. It also includes an explanation for methods used for data gathering including participants, material, and judging procedure.

Chapter seven summarises the research findings, draws conclusions from the research, and suggests new directions for future work. Guidelines for implementing the new framework in the engineering design domain is reported in this chapter.

2. Introduction to engineering design and creativity

Innovation depends greatly but not limited to the emergence of creative thoughts (Amabile, 1988; Thompson & Lordan, 1999). Creativity provides the basis for innovation; however, the emergence of innovation depends on several other factors, one of which is the ability of the firm to effectively implement the creative idea into the market (Scott & A, 1994; Van de Ven, 1986). New ideas increase a firm's profitability and improve customer satisfaction (Sethi, Smith, & Park, 2001). Companies need to understand the process of creativity emergence in order to consistently produce innovative outputs. Engineering design is a discipline that enables researchers to concentrate on studying the creativity process and specifying how to recognise creativity within the design process.

2.1. Overview of engineering design

In the past, engineering professions were split into four main disciplines including Mechanical, Civil, Electrical, and Chemical engineering. Engineering disciplines emerged to provide comfortable and sustainable environment for people. Engineers apply scientific and mathematical principles to problems and needs requested by society, commercial enterprises, or organizations to develop and deliver well-crafted products and solutions that are capable of solving the complication (P. R. N. Childs, 2014). For instance, Mechanical engineering concerns with the design, analysis, manufacturing and operating of a wide range of mechanical components, devices, and systems (IMechE, 2018). Electrical engineering addresses everything associated with electricity such as electrical power, signal processing and control, multimedia and telecommunications (IEEE, 2018). Previous advancements in engineering knowledge improved societies and as a result, new demands and concerns have raised. Engineering design is a practice that intends to provide a comprehensive platform for designers and engineers to systematically identify problems and deliver creative solutions to efficiently overcome existing concerns in wide range of fields. Pugh (Pugh, 1991) explained that the term design includes all tasks that designers

perform over the creation of a product, artefact or a service in order to fulfil the market demand. Childs (2018) defined that “Design engineering is the fusion of design thinking, engineering thinking and practice, within a culture of innovation and enterprise”. Overall, engineering design consists of analytical tools and frameworks that assist designers to test and analyse their solutions using scientific methods. It also optimises the designer’s thinking process by supressing unnecessary thoughts and increases productivity through tools that intentionally developed for each step (e.g. creativity tools). It is an intellectual procedure that considers all necessary steps in the product development process and offers engineered instruments to effectively deal with the problem. Typically, the end version of any newly produced product is the result of many trials and reconfigurations that the design team accomplished in order to safely present the outcome to society. Design engineering provides a process to better understand consumer needs and stimulates designers to come up with new and useful concepts. It then leads the design team to improve and upgrade their initial idea based on engineering approaches. In addition, it suggests how to select the right material and manufacturing process and indicates innovation managers how to implement the output into the market.

2.2. Design process

The design process controls all actions that the design team should perform in order to better perceive customer needs and create something new of value. Understanding of this process provides engineering companies with the knowledge to optimise their use of resources and time management. There are several notable existing design process models for engineers to follow for enhancing the quality of outcome (Baxter, 1995; Design Council, 2018; French, 1985; Pahl & Beitz, 1984; Svensson, 1974). Figure 2-1 shows Pugh’s total design model. Pugh (Pugh, 1991) believed that a successful design requires a comprehensive market assessment together with the consideration of both customer and competitor’s viewpoints. Once market data gathered, the design team should classify and specify the design requirements and transform them into a form of meaningful concept. At this stage, detailed product characteristics such as functionality, list of components, shapes and forms must be visualised and documented. This includes final engineering drawings, selection of material and manufacturing process. Subsequently, marketing

executives should determine the most useful market strategy based on product's features and capabilities.

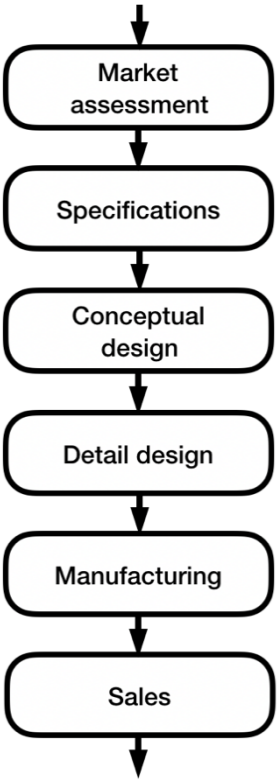


Figure 2-1 Total design process (Pugh, 1991)

One of the comprehensive systematic design process models has been presented by Pahl and Beitz (1996). Their model composes of four main steps including a) clarification of the task, b) conceptual design, c) embodiment design and d) detail design (see Figure 2-2). In general, the design task acts as in the form of a development proposal for the design department. It describes the market needs and specifies the product's functionality, features and performance characteristics (Pahl, Beitz, Feldhusen, & Grote, 2013). The task contains key information about product attributes and holds objectives that can satisfy future consumers. The outcome of this phase is a list of initial requirements which the design group should then inspect and refine in order to identify the design boundaries.

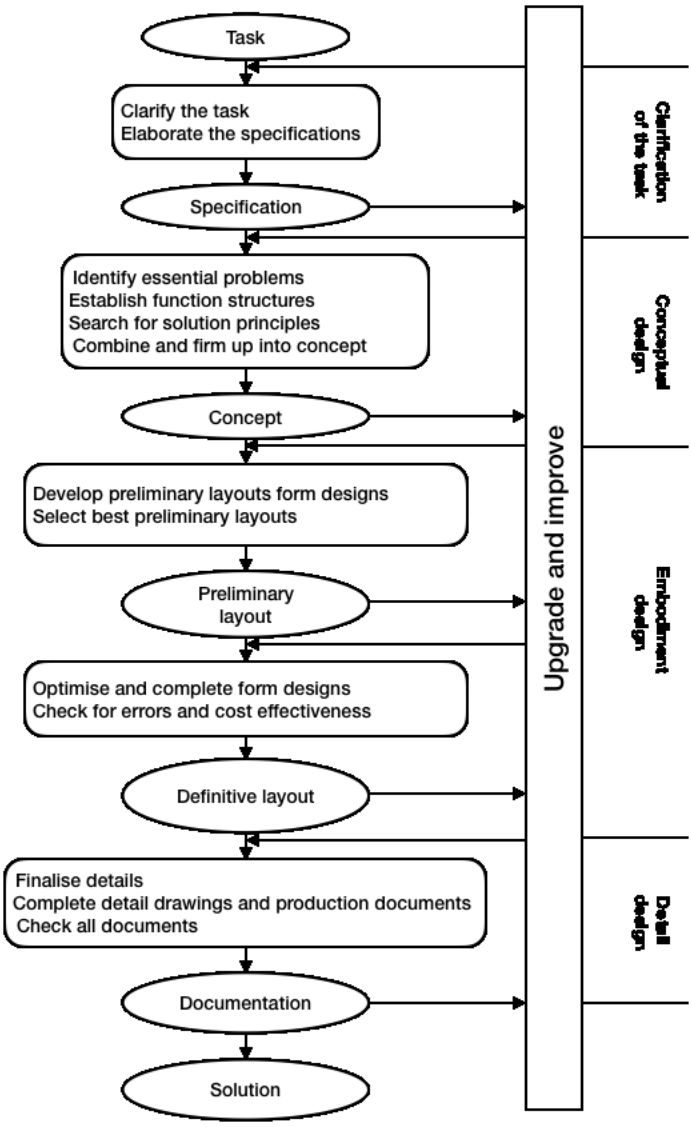


Figure 2-2 Design process steps (Pahl & Beitz, 1996)

Once designers have decided on major specifications, concepts can be generated to represent how the product will perform and look. The conceptual design contains a rough understanding of all components that require to make up the whole system. At this stage, a series of potential solutions are developed, and their practicality and performance are evaluated. Creativity tools are of great value in this phase as they can significantly increase the number of ideas which then leads to an increase in the probability of extracting the most useful solution. Shah et al. (2003) categorize creativity tools in two major classes including intuitive and logical. Intuitive tools as indicated by their name aim to guide individuals to produce novel ideas quickly without paying

attention to their effectiveness. Several methods have been developed in this class among them are SCAMPER, Morphological analysis and mind mapping. On the other hand, logical tools are built to assist engineers in solving problems systematically and being consistently inventive in the long term. One of the most well-known techniques in this class is the theory of inventive problem solving (TRIZ). TRIZ offers a step-by-step guideline to identify and eliminate contradictions in the design process of technical systems (Altshuller, 1999). This comprises several strategies that have been carefully captured from the study of past inventions and are capable of providing the right solution for the given problem. Recently, analysts have reported a number of successful applications of creativity tools in industry (Ekong, Long, & Childs, 2013). In the design process, engineers and managers also need an evaluation tool to assess the quality of ideas generated. Having the ability to recognize creativity and being able to measure the degree of it would enable firms to choose the solution with greatest potential and be efficient in using the resources. This thesis will cover an in-depth review of product/idea creativity measurement approaches in the next chapter.

Embodiment design is a part of the design process that inspects the product from all angles and intends to prepare the initial concept for final production. In this phase the design team shifts from a macro to a micro perspective where all components and design elements are investigated with engineering principles together with consideration of their economic effectiveness. The phase consists of analytical iterations to define the general arrangement of components and assemblies in the product, identifying how components interfere and defining geometric limitations (Sobek, 2005). At the detailed design stage, the components in the case of a physical product get to be refined, and their properties are founded such as dimensions, tolerances and surface roughness. During the detailed design, the design team selects appropriate material and specifies what manufacturing approach should be utilised at the production stage (P. R. N. Childs, 2014).

Prototypes are often created after completion of the detailed design stage to allow designers to verify the practicality of their design and detect the shortcomings of it. Typically, a prototype looks and functions similar to the final product and is intended to reveal how users will feel and interact with the product in the real-world scenario. The use of prototypes in the design process

is of great value as they can inform designers early about costly errors in the product and let firms launch their products more confidently. The manufacture stage impacts all other phases as not all designs in earlier phases are actually feasible for manufacturing. Some parts might be too expensive to manufacture, and some might be too complicated. To prevent contradictions in this phase, a number of design systems have developed such as design for assembly (DFA) and design for manufacturing (DFM) (Molloy, Tilley, & Warman, 1998).

2.3. Engineering problem-solving

Engineers deal with many problems during the process of product development. Seldom it is the case that problems are well-structured, and usually, a straightforward answer exists in another domain, while sometimes problems are difficult to be recognised at first glance before being able to find out what solution is appropriate. By definition, problems are described as "any situation in which a gap is perceived to exist between what is and what should be" (Arthur & VanGundy, 1988, p.3). In other words, problems occur when the required efficiency is not attained and the reason for its ineffectiveness is still undefined. Design engineers need problem-solving techniques in order to be consistently efficient and productive in resolving the difficulties. Typically, a problem-solving process consists of a number of steps which begin with clarifying the problem, continue with breaking down the problem into smaller sections and analysing them with the aim to detect the main cause, followed by generating and evaluating alternative solutions, and eventually end with selecting and implementing the most useful solution (see Figure 2-3). Afterwards, the potential solution must be refined in order to be standardised for further usage.

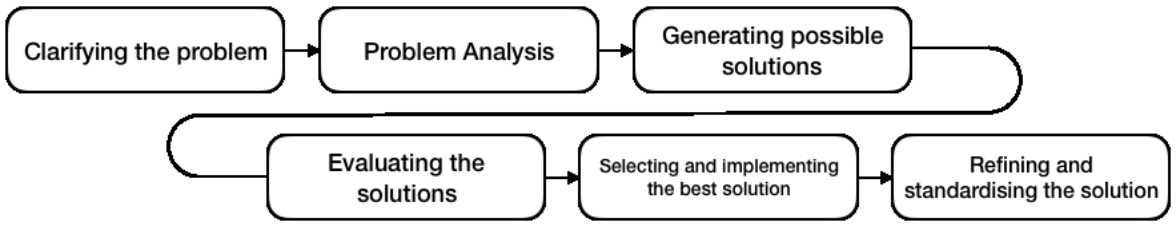


Figure 2-3 General problem-solving process

In engineering design, several authors have established approaches for problem solving. Hicks (1991) categorised them into a) Soft System, b) Rational, and c) Creativity approaches. The Soft

System methodology (SSM) is convenient for highly sophisticated and interrelated problem situations. The SSM considers the problem as a whole and attempts to solve it by initially rendering a rich picture of the overall problem and then providing conceptual models based on the study of other relevant systems. Rational approaches imply solving problems systematically and usually are most utilised for situations in which problems are well-structured. These methods commonly include decomposing the system into constituent elements in order to understand their nature and analyse their function separately. Then existing engineering principles and standards are employed to solve the problem. On the other hand, creative approaches are useful when problems are not well-defined, and relevant systems are inaccessible or not exist.

Osborn (1963) and Parnes (1977) developed the Creative Problem Solving (CPS) approach based on a combination of divergent and convergent thinking. They believe that the process of solving problem is engaged with divergent and convergent thinking as novel ideas usually emerge in the divergent phase whereas examination of ideas take place in convergent phase. Divergent thinking enables the design team to find dilemma, confront with issues, explore new solutions, and expand their range without being judged too early. On the other hand, at the convergent stage designers use critical thinking to organise and analyse the potential solutions and plan to improve their usefulness. The CPS comprises of six main phases where each split into two sub-phases of divergent and convergent thinking as shown in Figure 2-4. The steps include I) Objective finding, II) Fact finding, III) Problem finding, IV) Idea finding, V) Solution finding, VI) Acceptance finding.

Objective finding includes specification of the goal(s), and in general, finding the answer to the question ‘what needs to be accomplished, improved, and/or replaced?’. The origin of problems usually commences from difficulties that individuals encounter relative to each service or product they utilise. Occasionally objectives can arise from past knowledge or develop from non-human sources. During fact finding period, designers extend their understanding of the issue by searching the background of the problem and specifying more detailed information from its corresponding environment. Once the design team gathered sufficient data, they can determine the main concern that requires the most attention, and thus clarify the problem. New ideas should emerge in response to the requirements, and their effectiveness in the system must be

evaluated subsequently. In the solution finding phase, engineers need to inspect the potential solutions and choose the most promising idea that has the capacity to fulfil the original demand.

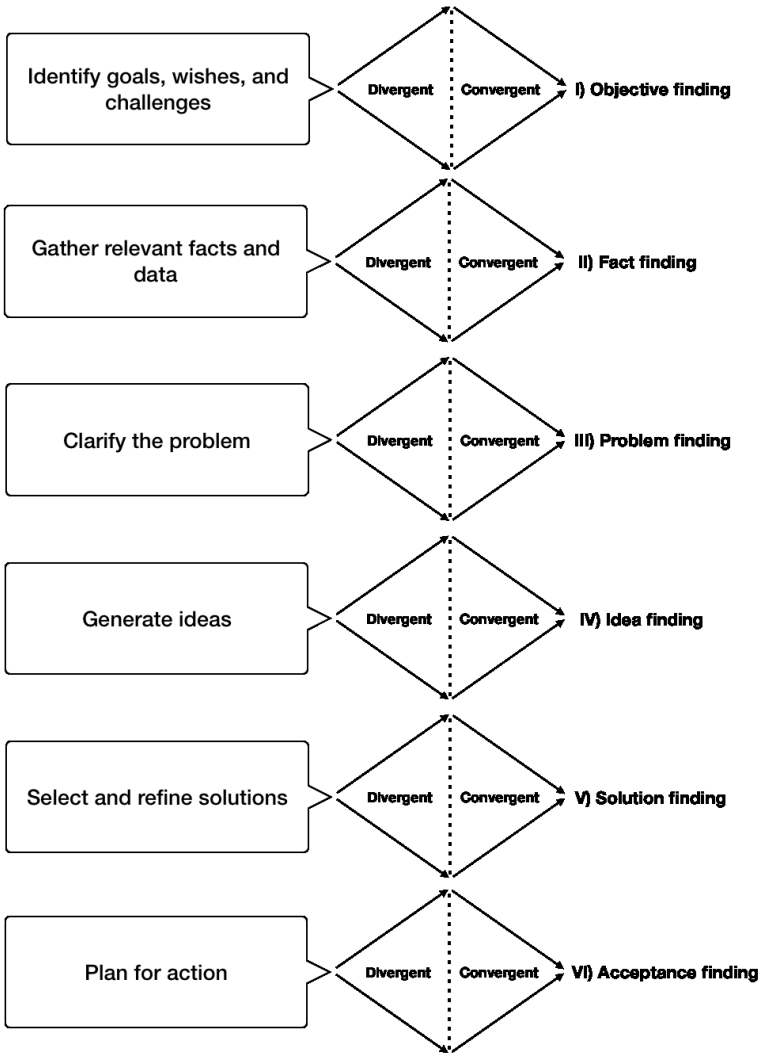


Figure 2-4 Creative Problem Solving (CPS) schematic

In the final step, the degree of acknowledgement of the solution among clients, consumers, internal and external members of the community must be assessed before asserting the solution into the field. Over the years, researchers based on Osborn's original model have developed improved versions of it. Recently, (Miller, Vehar, Firestein, Thurber, & Nielsen, 2011; Puccio, Mance, & Murdock, 2011) introduced a new CPS model containing four main steps including clarification, ideation, development, and implementation. Similar to the original model each stage uses several divergent and convergent thinking approaches to simplify the process and to

improve the quality of the final outcome. In this model, the clarification phase intends to recognize the problem by means of asking questions from different viewpoints to further assist with the exploration of the vision, data gathering, and formulating the challenge. Stage 2 (ideation) facilitates the development of new and practical concepts. Stages 3 and 4 are to help to move from initial ideas to viable solutions by further evaluating and elaborating the concepts and providing guidance on identifying the future steps as well as the resources required.

2.4. Creativity in Engineering Design

The interpretation of creativity spans several subjects including personality traits, social, cultural as well as environmental effects (Csikszentmihalyi, 1996; Sawyer, 2006). Recent investigations have also shown that there is neuroanatomical evidence in the DNA structure that indicates a portion of ideational fluency is due to genetics (Runco, Noble, Reiter-Palmon, Acar, & Ritchire, 2011). Although fluency is not strongly associated with creativity, yet it leads to the emergence of a large number of ideas from which a few can be original or creative. The study of creativity in engineering domain is categorised into four main facets of *Person*, *Process*, *Press* (or environment) and *Product*. The presence of all elements for the occurrence of creativity is essential as indeed the 4Ps implies the creativity process, from the emergence of the idea in the person's mind to the creation of a tangible outcome.

Studying creativity from four unique aspects was first presented by Rhodes with the following statement: "The word creativity is a noun naming the phenomenon in which a person communicates a new concept (which is the product). Mental activity (or mental process) is implicit in the definition and of course no one could conceive of a person living or operating in a vacuum, so the term *press* is also implicit. The definition begs the questions as to how new the concept must be and to whom it must be new" (Rhodes, 1961). Since then the hierarchy is considered as the basis for creativity research while many scholars have attempted to further enhance the statement by adding to the definition that the concept must also be effective and useful for domain users (Gardner, 1989; Runco & Jaeger, 2012). Likewise, a few studies have intended to improve the original model. Simonton (1995) suggested the notion of persuasion and discussed that creative person in addition to having the ability to produce something novel; they

must also possess the skill to convince others and influence their thinking. Furthermore, Runco (2008) proposed a hierarchy that divides creativity into two pieces including performance and potential at the highest level. The author argued that creative performance signifies the properties of product and persuasion whereas the creative potential represents the capacity that comes across the person, process and place.

Creative Person

Without doubt individuals are influential on the outcome creativity. Many scientists have studied the relationship between personality traits and creativity and believe that creative people have specific characteristics by which they are distinguishable from others (Costa & McCrae, 1992; Guastello, 2009). Cattell (1947) introduced a taxonomy of 16 bipolar traits as indicators of creativity in individuals. Among them were qualities such as dominant, socially board, inattentive to rules, open to experience, and intelligent which according to Cattell et al. (1970) is essential before even planning to describe person's potential by other factors. Barron (1963) investigated the association between creativity and intelligent and concluded that the two are correlated up until an IQ of 120, while no sensible relationship was detected between people with an IQ greater than 120. One of the well-known models of personality traits is Five-Factor Model (FFM) (Costa & McCrae, 1992). The Big Five comprises of openness, extraversion, neuroticism, agreeableness, and conscientiousness.

Openness is defined as the desire to explore new ideas and to wonder how differently the outside world can be observed. Openness also has a positive correlation with divergent thinking ability which further influences the originality of thoughts (Wolfradt & Pretz, 2001). The link between creativity and openness to experience has been supported with substantial works of analysts (Costa & McCrae, 1985; Soldz & Vaillant, 1999). Many scientists have argued that extraversion can positively influence creativity and described extraverts as active, highly energetic, emotionally bland and passionate individuals who are willing to take risks in order to convert their thought into reality (McCrae, Costa, & Busch, 1986). Agreeableness is another factor that may beneficially impact the creativity (Walker, Koestner, & Hum, 1995). Agreeableness refers to the ability of an individual as being friendly and enthusiastic to cooperate and avoid conflict. The

relationship between creativity and neuroticism has not well established in the field though. Several authors have reported positive correlations between the two (Götz & Götz, 1979), while some have recorded contrary or no relation (Martindale & Dailey, 1996).

Neuroticism stands for the tendency of an individual to experience 'negative' emotions, anxiety, depression, and anger. In overall, it has been shown that a low level of neuroticism is preferred as opposed to the high level, as individuals with the latter quality may experience helplessness and anxiety when they encountered with complex tasks which may lead to low creativity (Raja & Johns, 2010). Conscientiousness denotes responsibility and tendency to persistence and hard work. Although the negative relationship between creativity and conscientiousness has seldom been reported in personality studies (King, McKee Walker, & Broyles, 1996), the general perception is that high conscientiousness can positively affect creativity (Runco, 2002). Other characteristics that researchers have studied their impact on creativity are knowledge, cognition, biological aspects (i.e. genetics) and motivation (B. A. Hennessey, 2010; Kozbelt, Beghetto, & Runco, 2010; Ward & Kolomyts, 2010; Weisberg, 1999).

Creative Process

The study of the creative process has its roots in engineering design and cognitive psychology. Howard et al. (2008) inspected the differences and similarities between the two using an in-depth review of 23 engineering design processes and 19 creative processes in cognitive psychology. They have indicated that the common steps of the creative process are 'analysis phase', 'generation phase', 'evaluation phase', and 'communication/ implementation phase'. The authors then argued that the first three phases are often perceptible in the conceptual and embodiment phases of the engineering design process. However, the main difference between the former two methods was seen in the communication/ implementation phase. This step involves less communication with manufacture, and mostly focuses on accepting and approving the idea, while in the design process, engineers produce official documents for production and implementation of their original concept.

In cognitive psychology, the creative process is an intellectual procedure in which individuals attempt to explore new ideas and transform concepts into viable solutions (Boden, 1991). Wallas

(Wallas, 1926) presented one of the earliest cognitive models comprising four stages: preparation, incubation, insight, elaboration and evaluation (see Figure 2-5). In the preparation stage, the design team seek and organize information to better understand the context and determine the requirements for the rest of the process. The incubation stage refers to the unconscious cognitive process of problem-solving especially when individuals switch to another assignment or make a distance with the problem in question. In literature, researchers have investigated the effect of incubation period on problem-solving using two groups of participants; one group was experimentally controlled while the other party was distracted from the problem for a while and then asked to return and try to solve the problem they were first presented (Dodds, Ward, & Smith, 2003; S. Smith & Dodds, 1999). Results of their study showed that the group with an incubation period could solve the problem more effectively than the group without. They have identified that factors such as incubation length and preparatory activities can efficiently enhance the outcome of the problem-solving process. Insight (or illumination) represents the 'Aha' moment when a thought that is previously planned and planted at the unconscious level comes along as a potential solution at the conscious level. Recently, Scholars have characterised the necessary elements of the illumination stage. According to Pols (Pols, 2002), insights typically occur unexpectedly and result in a substantial alteration in the process of solving the problem. The new realisation is more comprehensive, and capable of significantly impacting the problem even if they are indirectly related to the problem.

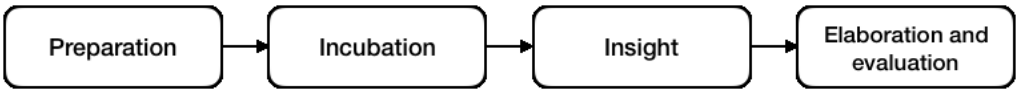


Figure 2-5 Wallas's creative process model

The elaboration and evaluation stage aim to assess the practicality of the final outcome and test whether the features of the solution are in line with what have been set during the preparation phase. Most of the recently developed creative models have shifted from considering the incubation stage in the process and put more focus on adding more conscious steps that can produce creativity systematically. For instance, Amabile (1983) has constructed the componential framework of the creative process and specified that the level of product (or response) creativity depends on a number components including task motivation, domain

relevant skills, and creative relevant skills. Amabile's model consists of five stages: problem or task presentation, preparation, response generation, response evaluation, and creative outcome. During the first phase, individuals introduced to the problem, then use their knowledge and skills to formulate the problem and attempt to detect the most useful solution. Having a high level of motivation during the first phases is necessary as it maintains individuals' performance at a sufficient level and encourages them to continue the exploration. In the response generation, creativity skills along with the motivation of individuals are of great value as they directly influence the quality of the final output. In the evaluation stage, the assessment of ideas relies mostly on domain relevant skills.

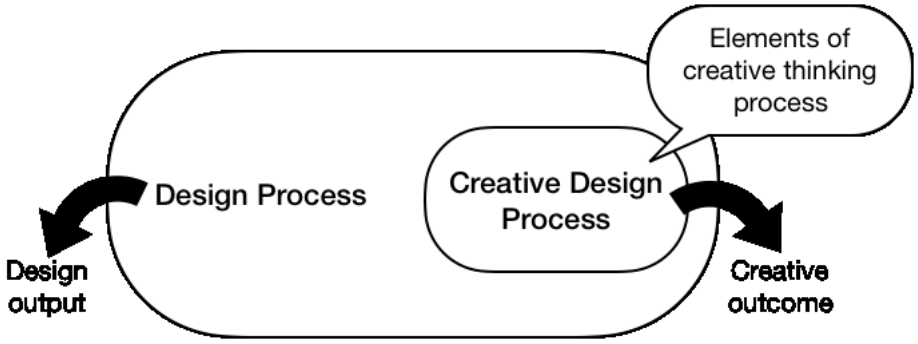


Figure 2-6 Relationship between design process and creative design process (Adapted and updated from Wong & Siu, 2012)

Other researchers have also adopted the Amabile's attitude toward paying more attention to conscious actions that are capable of generating creativity more consistently (Couger, Higgins, & McIntyre, 1993; Shneiderman, 2000). Alternatively, a group of scientists believe that creativity is resulting mostly from divergent thinking; however, it is more effective particularly if employed conjointly with convergent thinking (Eric, 2003; Osborn, 1963; Runco, 2003). In conclusion, the creative design process belongs to larger mainstream of the design process, where it is responsible for generating novel and useful ideas that can fulfil designers and managers requirements at further stages of the design process (see Figure 2-6) (Wong & Siu, 2012).

Creative place (press)

Apart from individuals and their thinking process, the workplace is another foundation that creativity relies on it. Scholars believe that a portion of organisational performance depends on

the existence of a creative climate (Baer & Frese, 2003; Ekvall, 1996). Indeed, the environment in which designers and engineers interact with each other, the managerial system that is in place, the set of equipment and the amount of pressure that are given to the design team are all capable of significantly influencing the quality of the final work. Studies have indicated that creative climates also positively impact the creative thinking of the employees (Moghimi & Subramaniam, 2013). Tsai and Kao (2004) investigated over 250 employees of the Taiwanese firms and showed that the higher innovative behaviours was captured for enterprises where their employees have better perceived and benefited the organisational creativity climate.

In creativity research, the study of creative climates is mainly focused on recognising attributes that enable organisations to promote creativity across their team and specifying the key elements in organisational atmospheres that the development of novel and useful outcomes depends on them. For instance, Krause (2004) refers to freedom as one the characteristics of innovative process and points out that such property is effective for individuals as it allows them to decide without restraint the way they wish to accomplish tasks. According to Shalley and Gilson (2004), the development of innovation in firms demand an internal catalyst to inspire individuals to deal with challenges and maintain their motivation during the creative work. Organisations must also be supportive of their employees and be attentive and thoughtful when dealing with new ideas (Kanter, 1983). In a creative organisational climate, managers and peers should pay attention to other group members and make an effort to analyse ideas wisely before trying to accept or reject them. Meanwhile, Lin and Liu (2012) believe that the two factors of organisational impediments and workload pressure may act as obstacles for the occurrence of creativity inside organisations.

Organisational creativity can be described as a complex social system in which individuals work together to create novel ideas, artefacts, processes, and/or systems that are also appropriate for the consumer (Griffin, 1983). Kanter (1988) report that for organisations to advocate creativity, they first need to establish a unique culture and create a climate in which individuals are encouraged to collaborate while paying close attention to diversity. In line with this, Amabile et al. (1996) describes ten elements that differentiate creative climates with the contrary such as organisational encouragement, supervisory encouragement, and sufficient resources. Therefore,

it can be realised that in addition to inventive people and efficient intellectual processes, creativity also hinges to the way enterprises coordinate the team members and the method they choose to execute ideas.

Creative Product

Creativity is a crucial factor for innovative firms as it enables the design team to solve their problems more differently and efficiently (P. Childs & Fountain, 2011). However, the key element in the process of ideation is to first recognise creative thoughts. In engineering science, the capacity to recognize creativity is hooked up with the ability to first define it. Creativity despite its simple sense consists of a much more complex structure. In general, the terms such as original, unusual, practical and useful are utilized when people intend to explain a novel work of an individual. The level of variations in describing creativity is due to the uncertainties that exist in the characterisation of creativity. In overall, creativity can occur for everyone depending on their environment or tradition. Scientists have classified creativity into two categories of the "big C" and the "little C" (Kaufman, Beghetto, & Ronald, 2009). The "big C" represents the noteworthy works of genius while the "little C" signifies the everyday creativity. Creativity can take place to any quantity and quality but the definition of it is built based on one uniform basis. The research on creativity was ignited by the president of American Psychological Association, J. P. Guilford (1950). Since then many scientists over decades have attempted to identify a consistent definition for (product) creativity (Mayer, 1999; Runco & Jaeger, 2012). Stein (1974) stated that creativity is non-existed novel work that the field accepts its usefulness. Another definition was given by Mednick (1962) where creativity characterised as the original formation of solutions that are able to fulfil the minimum requirements of the user. Gardner (1989) believed that creativity is the ability to generate new answers for existing problems that are also satisfactory for the domain. Other researchers have reported similar definitions for creativity. Plucker et al. (2004) explained that creativity is the result of an interaction between individual aptitude, intellectual process, and environment that is new and practical for the consumer. In another study, Childs et al. (2006) specified that “creativity is the ability to imagine or invent something new of value. It is not the ability to create out of nothing, but the ability to generate new ideas by combining, changing, or reapplying existing ideas” (p. 1). Overall, creativity defines as an idea that can be

formed in the shape of speech, behaviour, finished product, process or system that never ever occurred before along with complete acceptance of people who are observing, understanding or utilising it. The solution must be effective, useful and appropriate to the user.

2.5. Measuring creativity in engineering domain

Creativity is a complex and dynamic phenomenon meaning that the shape of it continuously alters from field to field and/or application to application. Although the basis of creativity is simple to perceive as this thesis reviewed over the previous sections, the definition is still insufficient in that it brings complications and confusions regarding the measurement of creativity. Indeed, any attempt to measure creativity must first be able to supply a comprehensive definition that can describe creativity in all forms and facets. However, the associated dimensions of creativity per aspect are themselves extensive which make it impossible to measure them all at once (see Table 2-1). As mentioned earlier, creativity occurs as a result of a creative process that include four main components: person, process, press, and product (Rhodes, 1961).

Table 2-1 Associated dimensions of creativity based on 4Ps of creativity (Adapted and updated from Horn & Salvendy, 2006)

Forms of creativity	Associated dimensions
Person	Active imagination, curiosity, independence, acceptance of own difference, trust in own senses, openness to experience, ability to restructure problems
Process	Fluency of ideas, problem recognition and construction, combination of ideas, recognising ideas, transformation and restructuring ideas, elaborating and expanding ideas, self-directed evaluation
Press	Superiority encouragement, workload pressure, freedom of choice, sufficient resources, organisational impediments
Product	Novelty, originality, usefulness, impact, elegance, pleasantness

Creative people must possess a specific set of skills and traits (e.g. openness to experience, active imagination) and also be able to conform to certain cognitive processes in addition to being in climates where divergent thinking and freedom of choice are appreciated in order to be able to produce creative outputs. As a result, researchers have attempted to separate constituent

components of creativity and developed tools and instruments that are capable of measuring creativity and associated dimensions based on situations that creativity occurs within them (Amabile, 1982; Amabile et al., 1996; McCrae et al., 1986; Torrance, 1974).

Several methods have been established for assessing creativity including psychometric, biographical, biological, computational, and contextual tests (Mayer, 1999). For instance, biographical methodologies aim to specify the effective characteristics of creative people by studying the case histories of individuals that have been recognised as creative in a specific domain (Wallace & Gruber, 1989). These methods typically intend to provide an in-depth report of the case history, then attempt to identify whether there is a relationship between the occurrence of a particular event in the past and creativity of individuals who are subjected to the same or similar instances. Computational methods attempt to formulate and simulate the creative thinking using artificial intelligence programming techniques. The computational approach presents high level of consistency as they construct the process of creative thinking by means of mathematical equations and computer coding, although they do not consider the non-cognitive aspects of creative thinking yet (Mayer, 1999). To assess creativity with the computational computer programming, researchers usually attempt to quantify the design process, then analyse outputs to understand which of the factors of the design process has the most impact on the creativity. Once the influential factors are recognised, then analysts define an equation that best describes the creativity for the given factors (Yuan & Lee, 2014).

On the other hand, psychometric tests consider creativity as an intellectual ability that can be measured numerically using relevant instruments. In creativity research, psychometric tools are currently the predominant methods for the assessment of creativity, and in general are categorised into four main types including: investigation of creative personality traits, creative processes, environmental attributes that promote creativity, indicators and distinguishing factors of creative products (Plucker & Makel, 2010). Many researchers have attempted to understand creativity and identify the influential factors that foster creativity across various disciplines, this thesis will review some of the most contemporary academic researches on the application of psychometric tests in creativity domain.

Measuring creativity of person, process, and press

One approach to distinguish creative individuals from the contrary is through self-report inventories, in which persons assess themselves on numerous factors. Typically, these factors are obtained by studying the personality characteristics of individuals who already identified as creative, and then specifying the qualities that were most common across all individuals. Over the years, researchers have developed several psychometric inventory tools for the measurement of personality attributes, among which the NEO Five-Factor Inventory (Costa & McCrae, 1992), Group Inventory for Finding Creative Talents (GIFT)(Rimm & Davis, 1979), and Gough's Creative Personality Scale (CPS)(Gough, 1979) are notable.

The CPS is an adjective checklist consisting of 30 items including 18 indicative and 12 contraindicative (negative) of the creative potential. In the CPS, individuals are asked to choose the adjectives that best describe themselves, and then the overall creativity is calculated based on the sum of all positive and negative selections. The instructor considers a positive point each time the person picks from one of the indicative adjectives and gives a negative point for the opposite condition. Researchers have reported a high consistency among the items of the CPS and many believe that the tool is accurately capable of predicting creativity (Gough, 1979; Oldham & Cummings, 1996). The NEO Five-Factor Inventory tool is another psychometric personality assessment instrument that is widely used in the field. The Five-Factor evaluates personality based on 60 questions that altogether represent five key aspects of the creative personality including openness to experience, extraversion, neuroticism, agreeableness, and conscientiousness. The validity of the Five-Factor has been repeatedly confirmed (Costa & McCrae, 1992; McCrae & Costa, 2004), and researchers reported a high test-retest reliability ratio (Robins, Fraley, Roberts, & Trzesniewski, 2001).

Apart from assessing creativity by personality traits, researchers have also developed alternative instruments that intend to separate creative persons by taking into account their previous behaviours and achievements in certain domains. Assessment of creativity under this method follows the assumption that “the best predictor of future creative behaviour may be past creative behaviour” (Colangelo, Kerr, Hallowell, Huesman, & Gaeth, 1992, p.158). Creativity Achievement

Questionnaire (CAQ) is a self-report tool that measures the individual's significant accomplishments on 10 different domains including music, drama, humour, writing, visual arts, discovery, culinary arts, dance, architecture, invention, and science (Carson, Peterson, & Higgins, 2005). In the CAQ, persons are asked to rate themselves on 10 domains from 0 representing "I have no training or recognised talent" to 7 representing "my work has been recognised by a national publication". The reliability, as well as the convergent and discriminant validity of the CAQ, have been examined and confirmed by the researchers (Carson et al., 2005). Other domain-based assessment tools are Creativity Scale for Diverse Domains (CSDD) (Kaufman & Baer, 2004), Creativity Domain Questionnaire (CDQ) (Kaufman, Cole, & Baer, 2009), and Kaufman Domain of Creativity Scale (K-DOCS) (Kaufman, 2012).

The creativity of intellectual processes has been commonly investigated using Divergent Thinking (DT) approaches. The divergent thinking tests demand participants to generate numerous solutions to a prompt, then the quality and quantity of the ideas are measured and results usually presented as indication of creativity. Guilford (1950) offered a divergent thinking test that evaluates creativity of responses to a divergent thinking task based on four batteries including fluency, flexibility, originality, and elaboration. Guilford describes *fluency* as the quantity of solutions provided for the subject, *flexibility* as the number of variations among the solutions (or the extent to which category of responses are different from one another), *originality* as how uncommon the solution is in compare to other responses, and *elaboration* as the tendency to provide further details in the solutions. Another and by far the most frequently employed test of divergent thinking is presented by Torrance (1974).

The Torrance Tests of Creative Thinking (TTCT) was founded on Guilford's divergent thinking batteries and consists of both verbal and figural tests (Clapham, 2004; Cropley, 2000). The verbal tests comprise six activities (including Asking, Guessing Causes, Guessing Consequences, Product Improvement, Unusual Uses, and Just Suppose) in which participants, for example, are instructed to ask questions around a picture that is selected as a stimulus, then to make guesses about the causes and consequences of actions associated to the drawing. After completion of all activities, the creative performance of the participants is determined based on divergent thinking batteries. On the other hand, the figural tests (i.e. Picture Construction, Picture Completion, and

Line/Circles) intend to measure creativity according to the ability of individuals in providing different, detailed, and original ideas around the given subjects. Torrance and Ball (1984) revised the TTCT and added two more dimensions including Abstractness of Titles which defines as the ability to abstract thoughts, and Resistance to Premature Closure that represents the degree of openness to explore new information. These revisions has caused the TTCT to get beyond divergent thinking test and become an effective creativity test (Kim, 2006).

Later, Hebert et al. (2002) reported a high correlation between the *flexibility* and *fluency* batteries and as result *flexibility* had been eliminated from the instrument. The validity of the TTCT has been examined and constructed by many researchers in the field (Cramond, Matthews-Morgan, Bandalos, & Zuo, 2005; Krumm, Lemos, & Filippetti, 2014). In a longitudinal study, Runco et al. (2010) indicated that over fifty years, the composite scores of the TTCT have successfully predicted the creative accomplishments of the individuals who were previously participated in the study. In addition to DT tests, several researchers have developed and validated tools and instruments for the assessment of creative ideation including Runco Ideational Behaviour Scale (RIBS) (Runco, Plucker, & Lim, 2001), Grit scale (Duckworth & Quinn, 2009), and Instances and Uses tests (Wallach & Kogan, 1965).

The study of creative climates intends to determine the key factors in working environments that significantly influence the creative performance of the team members. Scientists to this point have recognised a number of elements and believe that the existence of these factors in organisations highly correlates with creative achievements, productivity and quality of the employees (Amabile et al., 1996). One of the earliest attempts to evaluate creative potentials within organisations is the Swedish Creative Climate Questionnaire (CCQ), introduced by Ekvall (1996). The tool consists of 50 statements that have been factorially structured into 10 sub-scales including challenge, freedom, idea support, trust/openness, dynamism/liveliness, playfulness/ humour, debate, risk-taking, idea time, and conflicts. The questionnaire is a self-report tool that aims to assess the team members' perceptions of the environment. Researchers applied the CCQ in various climates and have reported a high level of internal consistency and goodness-of-fit (S. G. Isaksen & Ekvall, 2007, 2010).

Amabile et al. (1996) have developed the KEYS questions (known as Assessing the Climate for Creativity) based on the assumption that the perception of individuals from their organisational climate, unit, and co-workers is more dominant than the source of management when it comes to enhancing creativity within the organisational context. For instance, if organisational managers want their employees to take risks in their work, the team must first be convinced and perceived that they are allowed to do as such. The KEYS comprise 78 self-report questions that have been further categorised into 10 sub-scales: organisational encouragement, supervisory encouragement, work group supports, sufficient resources, challenging work and freedom, organisational impediments, workload pressure, creativity, and productivity. In the KEYS, employees are given the 78 items and asked to rate their perceptions about various aspects of their working environment as well as other team members. After collecting the results, the managers will understand their employees' judgments of the organisation while several suggestions can be provided to enhance the overall creativity. The KEYS is a reliable tool, and its convergent and discriminant validity have been confirmed by researchers in the field (Amabile et al., 1996). Other organisational creativity climate assessment tools are the Siegel Scale of Support for Innovation (SSSI)(Siegel & Kaemmerer, 1978), and Team Climate Inventory (TCI)(N. R. Anderson & West, 1998).

Measuring creativity of product

Researchers believe that the measurement of product is among the most influential aspects of creativity measurement (Plucker & Makel, 2010). The assessment in this context can assist analysts to better diagnose the problems and deficiencies in other forms of creativity measurement. The ability to accurately measure product creativity is also essential for innovative firms as it facilitates them to gain useful insights over the success rate of their new design before delivering it to the market. Organisations would be able to assess the efficiency of their design team as well as their thinking processes, if they are being able to systematically measure product creativity. However, creativity in its underlying structure is a dynamic concept meaning that it changes its shape continuously from domain to domain, which makes it difficult to evaluate creativity objectively. The objective assessment of creativity requires a clear definition for creativity that can distinguish it from other influential elements at any condition. Also, it is

needed a reference system (or unit) to which all other measurement tools can be calibrated and/or validated.

In recent years, many subjective and criteria specific tools have been emerged and their validity have been examined by the researchers in the field (Amabile, 1982; Besemer & O'Quin, 1999; Cropley & Cropley, 2005). The main assumption of the subjective approaches is that as creative products are made for specific users, they must also be perceived as creative by the same observers. In other words, a product's creativity can be assessed subjectively as similar to person, process, and place by appropriate judges who independently observe and agree with a high internal consistency that the outcome is creative. This is the main logic behind the well-known Consensual Assessment Technique (CAT)(Amabile, 1982). Amabile's CAT considers no pre-definition for creativity and only relies on domain experts as judges for the assessment. In the CAT, raters (experts) are given a set of products that have been produced by a number of participants in an experimental condition and asked to independently rate all the products for creativity. The CAT has been extensively used in the creativity measurement of a wide range of products including poems, music, and narratives, and its reliability and validity is reported by several studies (Amabile, 1983; Baer, Kaufman, & Riggs, 2009; Kaufman, Baer, Cole, & Sexton, 2008).

Nevertheless, finding appropriate domain experts for the assessment is sometimes problematic and even if they can be found the level of accessibility can also be limited. In addition, the outcomes of the subjective methods do not provide useful information about the areas that due to them the product is rated uncreative or less creative. To overcome this issue, researchers developed the rating scales that are typically designed for a specific domain and able to be utilised by non-experts as they are cheap and easily accessible (Cropley & Kaufman, 2012; Horn & Salvendy, 2009; O'Quin & Besemer, 1989). In contrast to the CAT, these scales have been built based upon a theory that decomposes creativity into its indicators that have been previously identified as gauges of creativity in domain that is in question. In the Criterion-based approach, inspectors usually ask a large number of novices to rate a couple of products of the same functionality on several criteria. The validity of such method is gained by ensuring that the ratings are reliable at the first place and verifying that the raters are able to distinguish the differences

between the latent factors which embedded in the model. If the instrument lacks the former conditions, then it implies the convergent and discriminant validity of the tool is under question.

Other researchers have used measures which are indirectly assessing the products creativity. These approaches usually focus on the study of creative people, the behaviour and past achievements. One of the commonly known methods under this category is the attempt to measure creativity by asking experts in a particular domain to appoint the most creative members of their domains. The theory is the basis for the *Peer and teacher nomination and awards* approach, where teachers intend to judge and nominate the child's creativity by considering to a certain degree the quality and quantity of the child's work (O'Quin & Besemer, 2011). Additional indirect techniques are the *Measures of eminence* (D. W. MacKinnon, 1968), and *Self-reported creative activities and achievements* (Daniel Hocevar & Bachelor, 1989). This thesis will cover an in-depth review of the existing psychometric product creativity measurement tools in the next chapter.

2.6. Chapter conclusion

This chapter has considered the role of engineering design in the process of product development. Also, it emphasised the importance of creativity in the design process and reviewed the definition of creativity. Engineering design is a discipline that intends to educate engineers the tools and strategies which can be used to simplify the process of understanding design problems while boosting their productivity. The design process instructs designers and managers to clarify consumer requirements systematically, and better realise what functions should be applied in order to better fulfil clients' needs. The creative design process supplies useful knowledge for designers to employ in order to produce creative solutions while facilitates them with reliable approaches to assess their effectiveness. In overall, the design model consists of a set of techniques that allow managers to prevent costly mistakes in the early phases of the design and leads innovative companies toward the production of solutions that are capable of satisfying both user and firm's demands. Creativity can be described as a combination of both novelty and useful dimensions.

This chapter reviewed different forms of creativity (known as 4Ps of creativity) in the engineering design and covered the general approaches for the evaluation of creativity. Investigations of the contemporary psychological researches on creativity assessment indicated that the field is noticeably advanced and according to many studies, a subjective measurement of creativity using psychometric approaches is systematically practicable. In addition, it is indicated that the use of psychometric tools for the assessment of creativity is reliable and most of these methods possess convergent and discriminant validity. This chapter highlighted the importance of product in the creative process and emphasized the point that an accurate measurement of product creativity can facilitate researchers in effectively diagnosing the lack of validity and/or calibration in other forms of creativity measurement. However, the current trend of research on the assessment of product creativity is focused on artistic domains (e.g. music, art, writing, and collages) and less attention has been paid to the assessment of functional consumer products in which the element of *usefulness* is of great value for the user. This thesis is an attempt to fill the gap in the knowledge by exploring the applicability and reliability of the current measurement techniques in assessing creativity of non-parallel engineering products. Also, the research aims to develop a product creativity evaluation framework that enables the designers to accurately measure creativity of engineering products with a high degree of reliability.

3. Product Creativity Assessment in Engineering Design

Creativity is among the influential factors affecting positively on companies' revenue growth, profit growth, and overall market share (Csikszentmihalyi & Sawyer, 2014; Von Nordenflycht, 2007). Results of interviewing more than 1500 CEOs worldwide indicated that creativity aside from other important factors is the most significant element impacting the leadership and success of the firms (IBM, 2010). Every creative product indeed results from an interaction among creative individuals (personality traits), and the intellectual processes utilised by those individuals (see chapter 2, Creativity in Engineering Design). The creation of creative ideas/products also depends on environments where individuals are encouraged to produce such output, the place that provides appropriate resources, time, and conditions (e.g. team working support).

Another benefit of studying creative products is that it allows researchers to better understand how creativity forms in the mind of an individual and/or what environmental factors are affecting the creation of novel ideas. From the beginning of the research on creativity, scholars made efforts toward the development of various tools and instruments in order to measure idea/product creativity. Horn and Salvendy (2006a) have divided the measurement methods into three major categories including indirect measurement, global judgment, and criterion-based instruments. The indirect assessment of creativity is mostly based on the study of background, and previous great works of inventive individuals who are acknowledged by experts in their domain. Indirect measurement of creativity is based on the assumption that creative individuals are more likely to produce creative outcomes. The Measures of Eminence, as well as the Self-reported creative activities and achievements, are among the most well-known techniques in this category.

On the other hand, the global judgment approaches neither consider any definition for creativity nor limit it into exceptional people who are seldom judged as creative after their death. The global judgment of creativity instead focuses on the notion that a product is creative as long as a

group of knowledgeable and experienced people in a particular agree that the outcome is creative. The Consensual Assessment Technique (CAT) is the most recognised method in this class. The viewpoint of the global judgment methods is exactly in contrast with the concept that most of the criterion-based instruments are mainly based upon it. The criterion-based models assess creativity through a set of criteria and indicators that are commonly utilised when people are asked to rate the creativity of a product (e.g. novelty, technical performance of the product). Creative Product Analysis Matrix (CPAM) and Creative Product Semantic Scale (CPSS) are among the most well-known instruments in this division.

It appears that a common point among all these tools and instruments is that their measurement mostly relies on the judgment of a group of appropriate judges. Possibly the main reason for such subjective approach is due to the absence of a reference product (a genuinely creative product) that its creativity has been universally accepted by the majority of people worldwide, whereas its creativity level is being calculated using specific mathematical algorithms (i.e. its value is repeatedly measurable and obtainable). The existence of a referral point (product) is essential for any objective assessment as it then enables the existing tools to be validated while their accuracy and precision can be further calibrated. However, because creativity is a dynamic concept and factors that are illustrating it are continuously changing their shape from product to product, an objective measurement of creativity at this period does not seem feasible. Nonetheless, several investigations have been made for objectively assessing the degree of novelty in new products. This chapter reviews the majority of the existing creativity measurement tools and covers the advantages and disadvantages of each type of measurement.

3.1. Objective methods

Shah and Vargas-Hernandez (2003) suggested a method for assessing the novelty of ideas produced in response to a specific task. In this approach, the key performance characteristics of the problem are first defined, and then the overall novelty is mathematically computed through the equation proposed (see equation 1). The authors believe that novelty is a function of the sum of the weights of the key performance attributes of the product multiplied by the sum of novelty scores for the idea generated per attribute. In general, the performance attributes include those

elements that together form a framework (or boundary) for any potential solutions that can be emerged for the given problem. For instance, in order to move from point A to point B, one of the involving attributes is the medium of travel (air, water, land, and space), and any potential solution for this task must indeed use a type of medium in addition to further ideas that belong to other attributes in the system (e.g. propulsion or motion type).

$$M = \sum_{j=1}^{m} f_j \sum_{k=1}^{n} S_{1jk} P_k \qquad (1)$$

In this equation, the overall novelty (M) is equivalent to the sum of the weights (coefficients) of the performance attributes (f_j) multiplied by sum of the novelty scores (S_{1jk}) of the idea generated per each attribute. The attributes and their corresponding coefficients are defined based on the previous knowledge and experience, while the novelty score for each solution is calculated mathematically by subtracting the total number of solutions generated for that specific attribute from the number of solutions that are similar (or equivalent) to the one presented in the idea (e.g. all solutions with air selected as their medium of travel). Equation (2) shows how novelty scores are calculated. Where T_{jk} denotes the total number of ideas and C_{jk} represents the number of equivalent ideas generated for that attribute. The notion P_k refers to the weights of the stages at which the idea is provided, for example, conceptual and/or embodiment design stage.

$$S_{1jk} = \frac{T_{jk} - C_{jk}}{T_{jk}} \times 10 \qquad (2)$$

Chakrabarti and Khadilkar (Chakrabarti & Khadilkar, 2003) stated that novelty is a relative concept in the sense that the degree of it is comparatively defined according to other products that are selected as reference in a certain time period. The authors proposed and validated a measurement system for assessing novelty by introducing two levels at which newness can occur in the product, vertical and horizontal levels. Vertical level refers to the hierarchy of the product which starts from the *need* (the primary motivation for creating the product), and goes to *task* (*ta*), *principle (p)*, *technology (te)*, *sub-technology (ste)*, and ends with *implementation*, i.e., the way the internal and external components in the system are arranged in order to conduct the

task. On the other hand, the horizontal level denotes the type of function (*main (m)*, *supplementary (s)*, and *additional (a)*), as well as how crucial each function is for the whole system. The originators also labelled each criterion with a weightage (*W*) to ease the process of measurement. In this method, novelty is measured by multiplying the weightages of those parts where product novelty is detected (see Figure 3-1). For instance, if novelty is detected at the technology level where it serves as the main function, then novelty is equivalent to *[(Wte*Wm)*Wm]*. In this equation, the second weightage (*Wm*) signifies the importance of the function of the previous section in the product hierarchy which in this example is sub-system level.

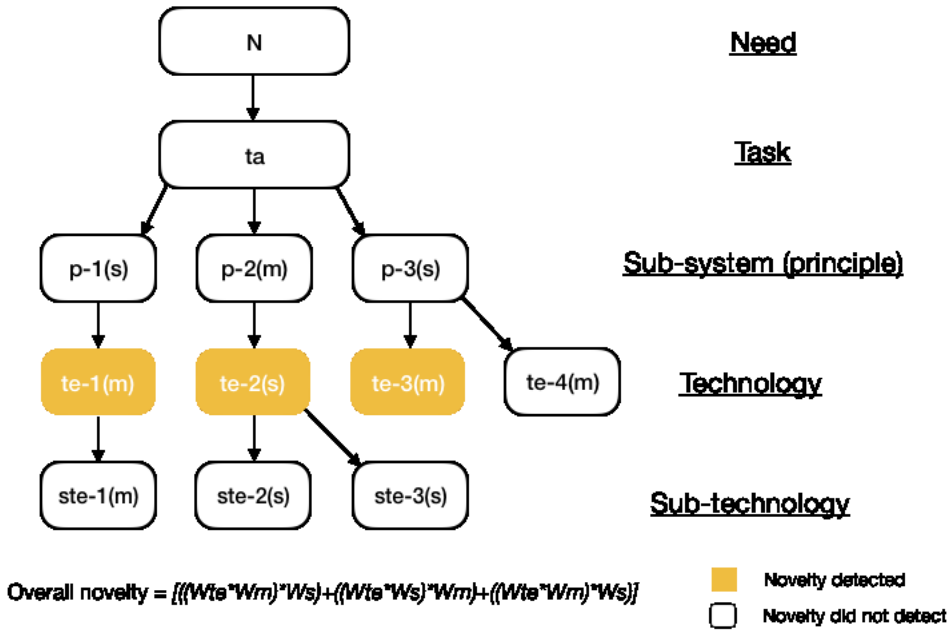


Figure 3-1 An example of measuring novelty using proposed method by Chakrabarti and Khadilkar (2003)

The authors validated the method through comparing the results of assessing the novelty of four computer mice once using their method and once using the subjective judgment of 13 experts (engineers and architects). In this study, one of the products was used as the reference to allow comparison. The experiment showed that the outcomes of measuring novelty using the proposed method is equivalent to those of expert judges and hence indicated the reliability of the method.

In another study, Sarkar and Chakrabarti (2011) proposed a method for assessing novelty based on the causality model of the given system (product). They explained that a systematic

measurement of novelty is possible through a methodical comparison of the features and characteristics of the given product with those of other products in the domain. For instance, if the function of the product is different from the others, then it can be concluded that the product is novel. However, in order to understand the degree of novelty researchers utilised the SAPPhIRE causality model. The SAPPhIRE model was first introduced by Chakrabarti et al. (2005) with the aim to illustrate the behavioural relationships between function and the underlying structure in both natural and engineering systems. The acronym SAPPhIRE stands for *state change*, *action*, *parts*, *phenomenon*, *input*, *organ*, and *effect*. Figure 3-2 presents the general structure and relationships among the sections of the model.

In this model, *parts* represent the physical belongings of the product as well as the links between the components and their surroundings that all together form a system in a closed environment. The *organ* refers to all required platforms, programs, conditions, and/or structures (apart from input) that enables a physical effect to stay operative for the desired duration. The input of the system is usually a quantitative element that is derived from an external source into the system, and it is something crucial as it allows the system to interact with its environment. The input of the system can be in the form of information, material, or energy. The physical *effect* implies all physical principles that are responsible to explain and control the collaborations and communications among the components (e.g. Newton's equation of motion is needed to clarify the displacement of a body). The physical *phenomenon* refers to the higher abstraction of the interaction between organ(s) and input(s) that is ruled by physical law (e.g. heat transfer phenomenon). The *state of change* represents the overall alterations that appeared as a result of previous stages and indeed defines the system's property. Action denotes the overall description of all interactions and communications that take place within the system or with the environment outside of it (e.g. movement of the body).

The SAPPhIRE causality model contains three types of relationships including creation, activation, and interpretation. In this model, parts are forming organs (creation), and organs together with inputs trigger physical effects (activation). Subsequently, the physical law generates physical phenomenon which then produces state of change (creation).

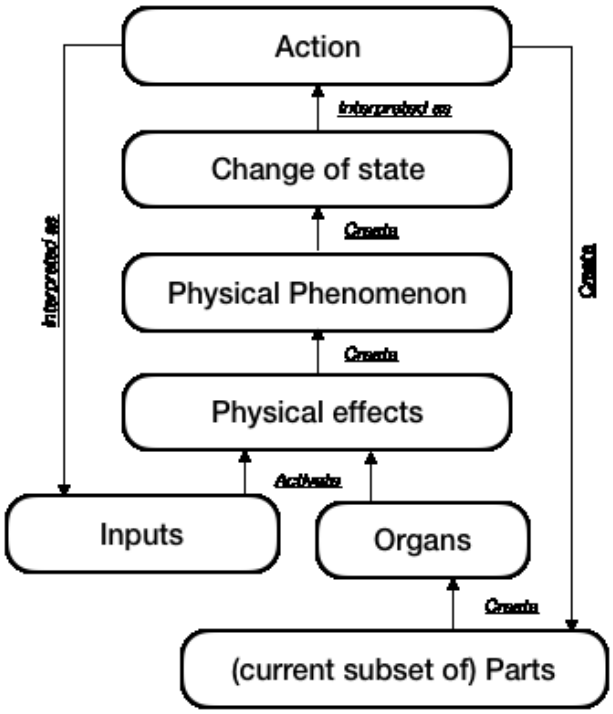


Figure 3-2 SAPPhIRE model of causality (Adapted and updated from Sarkar and Chakrabarti, 2011)

Finally, the changes in the system are interpreted as action, or new inputs, and/or lead to the creation of new parts. To assess novelty under this algorithm, researchers (Sarkar & Chakrabarti, 2011) proposed a flow-chart that is able to indicate the degree of novelty in products (see Figure 3-3). In this chart, the most novel product (very high novel) is recognised as the one that has an original function where no other product in the domain is in possession of such function. If the function exists, the structure of the product should be compared with those of other products, and if it is the same, then it indicates that the product is not novel. However, if the newness is detected at any other constructs of the SAPPhIRE causality model, then it means that the product is to some extent novel. The authors stated that highly novel products should possess new (and original) characteristics at the construct level of *state of change* or *input*, while moderately novel products (medium novelty) should have newness only in terms of physical *phenomenon* or *effects*. The products are diagnosed with low novelty if the distinctions are detected in the form of new parts and/or organs that are just arranged in a different way.

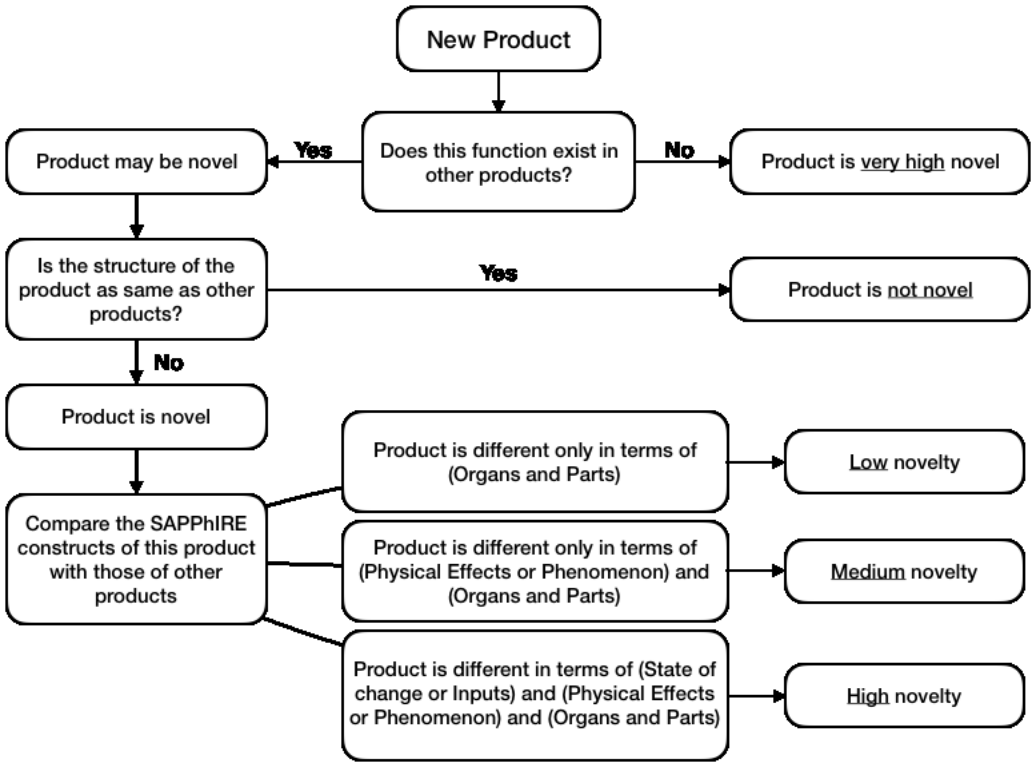


Figure 3-3 Method of assessing novelty (Adapted and updated from Sarkar and Chakrabarti, 2011)

In addition to assessing novelty, researchers have proposed a method for measuring usefulness of the product. In their system, usefulness (U) is described as the common effects between *level of importance* (L), *rate of popularity of use* (R), *frequency of usage* (F), and *duration of use* (D) (see equation 3).

$$U = L \times R \times F \times D \qquad (3)$$

For instance, the level of usefulness of a pen can be described as how important it appears for the user multiplied by how widely it is used and accepted by the people in the domain, and how frequent as well as for how long it is employed by the users.

3.2. Indirect measurement methods

Indirect measurement of creativity refers to utilising the assessment of product creativity as a technique to investigate and identify the most creative individuals in a particular domain. In other words, a different way to recognise eminent people is to study the creativity of their works in the past. In this area, researchers have presented several approaches for evaluating creativity

including peer and teacher nomination, measures of eminence, and self-reported creative activities and achievements.

Peer and teacher nomination and awards

The peer nomination methodology involves asking experts in a specific domain to nominate or rank the most creative individuals of their field. The use of experts enhances the reliability of the judgment as experienced people typically take into account those essential characteristics that are most common and well-known to the other experts in the field. Likewise, when teachers decide to identify the most creative student of their class, they also consider a variety of quantitative and qualitative aspects of the students' works. Often researchers use the peer and teacher nomination approach to award the most creative outcomes of a group of participants that take part in some creative competition. Examples of such events can be seen in several institutional contests (e.g. Red dot design awards), national and international creative competitions (e.g. young writers creative writing competition, UK).

Measures of eminence

Another indirect approach to measure (product) creativity is to investigate the history of the eminent individuals. Typically, the measure of eminence does not suggest a specific rule or methodology for performing the assessment, and researchers often develop their own strategy depending on the field that is in question. For example, Mackinnon (1962) identified the 40 most creative architects in the United States using a panel of judges, where five professors of architecture are asked to nominate the creative architects of their field. In another study, Simonton (1977) attempted to generate a structural model for eminence using historical records for over 700 classical composers. In this study, the author employed several variables for establishing the model such as individual-level factors (e.g., the amount of time dedicated to the composer's music in a particular occasion), creative productivity (the total number of products), creative longevity (the time interval between each notable outcome), and the birthplace of the composer. In some art cases, eminence is measured by analysing the total number of illustrations as well as the price fluctuation of the work of an artist throughout the time (Galenson, 2001). The

measure of eminence is similar to other indirect methods as it mostly utilises product assessment as a tool to validly evaluating and recognising eminent people of a specific domain.

Self-reported creative activities and achievements

Self-reported scales intend to measure creativity by asking individuals to consider their past activities and rate themselves on a series of questions that are commonly regarded as creative. One of the earliest attempts in this field was made by Hocevar (1979) for developing the Creative Behaviour Inventory (CBI). The CBI is an inventory containing a list of 90 items that together assess creativity from six main sub-scales including creativity in arts, crafts, literature, music, performing arts, and math/science. In this method, the overall creativity score is computed by counting the number of activities checked by the individual, although the authenticity of the selected items cannot be examined. Since then the CBI has been used in several studies, and its reliability and validity has been established by the researchers in the field (D Hocevar, 1980). Other well-known self-rating scales are Creativity Domain Questionnaire (CDQ) (Kaufman, Cole, et al., 2009), Biographical Inventory of Creative Behaviours (BICB) (Batey, 2007), and Creative Achievement Questionnaire (CAQ) (Carson et al., 2005).

3.3. Global judgment

The global judgment of creativity involves an overall assessment of creativity and does not aim to identify the factors that are forming creativity. These methods do not even consider any definition of creativity, and their assessment of creativity mostly relies on the judgment of a group of experts who are independently agree that the product is creative.

Consensual Assessment Technique (CAT)

The Consensual Assessment Technique (CAT) is the most well-known approach to evaluate product creativity. CAT overcomes the problem of 'does the constructed model genuinely represent creativity?'. Amabile (1982) established CAT based on judgment of experts in the domain and stated that if a group of experienced people in a particular domain agree that the product is creative then it must be accepted that the product is indeed creative. It is worth mentioning that, the assessment of creativity in this context is a relative measure meaning that

the highest score for creativity in a study can still be of moderate creativity in the domain using another group of products. Indeed, utilizing the CAT does not lead to determination of the absolute value of creativity in a domain but rather it is used to understand the level of creativity for a group of products selected as stimuli in a study with respect to one another. In other words, the outcomes of the CAT cannot be generalised unless all the products in a domain have been identified and already included in the study. In addition, creativity is a dynamic concept and its definition changes over the time and based on the place of measurement. In Amabile's CAT, the definition for creativity is determined based on the subject in question as well as the perception of experts that selected as judges. Since CAT assumes no definition for creativity, the validity of it is typically justified by measuring the inter-rater reliability between judges (Kaufman, Lee, Baer, & Lee, 2007). However, in order to eliminate potential systematic errors from the assessment Amabile (1982) suggested five requirements to consider before taking the measurement. These are (a) the experts in the domain must be selected as judges, (b) Judges must assess the products independently, (c) products must be judged relative to one another, (d) each judge must observe the products in different random order, and (e) judges must be asked to rate the products on other dimensions in addition to creativity. The measurement of additional items is to indicate that the judges are able to distinguish the differences between creativity and other factors, and that their judgment of creativity is not solely based on, for example, technical merit of the product. Moreover, it allows to determine the relationships between the selected factors and creativity. For instance, Tsai (2013) examined the creativity of 46 collages using 6 experts and found that the overall expert's perception of creativity contains two dimensions including technical and aesthetic aspects. Valgeirsdottir and Onarheim (2015) as well as Hennessey, Amabile, and Mueller (2011) used these two factors in assessing product creativity and also added extra items to investigate their associations with creativity.

In CAT, participants are first provided a certain amount of time and asked to create an artefact using a set of pre-assigned instructions, materials, and equipment. Then, experts in the domain are requested to judge the creativity of all products created by those participants (Baer & McKool, 2009a). The consensual technique has been extensively utilised on a wide range of products such as art, music, poems, business products, and even mathematical equations

(Amabile, 1983; Baer, 1993; Lee, Lee, & Youn, 2005; Robert Sternberg & Lubart, 1995). However, the process of assessing creativity using the consensual technique is not without limitations. For instance, a crucial question to answer before starting the evaluation is 'who is the most appropriate judge for assessing creativity?'. Hickey (2001) compared the inter-judge reliabilities between different panels of judges measuring the creativity of 12 musical compositions using the consensual technique. They concluded that choral teachers and music theorists are the most appropriate judges to assess musical compositions with the inter-judge reliabilities of 0.81 and 0.73 respectively. In the field of music, Byrne, MacDonald, and Carlton (2003) also reported similar outcomes and noted that for reliability and robustness, the use of experts is a must for any assessment. In general, CAT has shown sufficient validity as long as relevant and well-experienced judges are in place. CAT has become known as the "gold standard" of creativity measurement due to the ability to obey no specific theory of creativity (Carson, 2006). Kaufman, Plucker, and Baer (2008) stated that the validity of the consensual technique is subjective in the use of experts in the domain as judges. Therefore, the most important item in creativity evaluation under this category is the selection of an appropriate panel of judges. However, experts might not always be accessible, and they are not usually cost-effective. Baer, Kaufman, and Riggs (2009) inspected the intra-class correlations between expert and quasi-expert ratings for poems, stories, and personal narratives generated by eighth-grade students.

For personal narratives, they have reported a high correlation ($r = 0.88$) between the ratings of experts and English teachers (quasi-experts). Their analysis of correlations among other experts and quasi-experts has also resulted in similar outcomes for other cases (e.g., $r = 0.95$ for creativity ratings of stories). They concluded that in domains where quasi-experts have considerable experience, their ratings to some extent can reflect the expert's opinion. In a separate study by Tsai (2016), the reliability of the CAT was evaluated for measuring the creativity of 10 comic strips using 84 quasi-experts (design-college students). Tsai assessed creativity using five criteria consisting of originality, quality, attractiveness, technical quality, and elegance. For all dimensions, the ratio of reliability (coefficient alpha) was attained at a good ratio (ranging from 0.859 to 0.891), revealing the applicability of quasi-experts in creativity assessment.

Further investigations on identifying appropriate judges have resulted in convincing evidence that prevents the use of novices in creativity measurement. For instance, Kaufman et al. (2008) examined the reliability ratios between experts (10 poets) and non-experts (106 college students) rating 205 writing samples gathered from a separate group of students. For both cases, the results of Cronbach's alpha reliability test were obtained at an exceptional range, i.e., 0.83 for experts and .93 for novices. However, after considering the ratio of the number of experts to novices and standardising the outcomes to 10 raters per group, the value of coefficient alpha for non-experts experienced a sharp decrease to 0.575. This significant drop in the value of inter-rater reliability revealed the inconsistency among the novice raters. Even though the use of novices in creativity assessment has been shown to be unreliable, some researchers have shown that combined ratings of experts and novice judges can be reliable.

Freeman, Son, and McRoberts (2015) measured the creativity of 30 fashion illustrations by six judges consisting of three experts and three undergraduate students (novices). They achieved an excellent internal consistency coefficient (alpha) of 0.96 for both creativity and technical quality indicators. Their analysis of the variance (one-way ANOVA) with creativity and technical qualities as dependent factors and the judges as between subject variable revealed an insignificant difference between the ratings of experts and non-experts. However, most of the researchers in the domain (Kaufman & Baer, 2012; Kaufman, Baer, & Cole, 2009) believe that the use of novices in the measurement of creativity is in general unreliable. Nevertheless, for conditions in which the access to experts is restricted, those quasi-experts with appropriate knowledge in the field can represent the experts' judgment, while attention is needed when employing quasi-experts in domains where no previous study has been done to show their validity.

Apart from the experience of the judges, the number of judges who are asked to perform the assessment also influences the CAT reliability. In the literature, unfortunately there is no clear guideline specifying an appropriate number of raters for the assessment, however, different scholars have suggested that for most studies (on average) a panel consisting of between 5 to 10 judges should be adequate and it would probably be able to assist researchers in achieving an acceptable level of reliability (Kaufman, Plucker, et al., 2008). In the past, researchers were able to capture reliable assessments from using only two raters (Daly, Seifert, Yilmaz, & Gonzalez,

2016) up to over 130 raters (Valgeirsdottir & Onarheim, 2015). Overall, the larger the panel of judges the greater inter-rater reliability ratio between the raters (Baer & McKool, 2009a). However, a larger panel of judges may artificially inflate the value of Cronbach's alpha as mentioned by the researchers in the field (Cseh & Jeffries, 2019). It appears that the process of determining an appropriate size for the panel of judges depends on many factors that have yet to be researched, however, researchers may still be able to assemble the correct number of raters by studying the previous studies conducted by other researchers in the domain of their interest.

Other factors that influence creativity measurement are the effects of gender, culture, and ethnicity. Hennessey et al. (2008) examined the reliability of creativity measurement in non-western societies such as China and South Korea. They reported that CAT is capable of achieving a high degree of reliability across distinct cultures for assessing the creativity of collages and stories made by children. In another study by Chen et al. (2002), the consensual technique was employed to evaluate the creativity of 294 drawings produced by students from diverse cultures including European American and Chinese. Two groups of undergraduate students (8 Chinese, and 6 European American) judged the drawings using a 5-point scale. High degree of consistency was obtained for both groups which showed that CAT is indeed a reliable assessment across different cultures. Moreover, the analysis of correlations verified high levels of agreement between the two groups of raters. Researchers also inspected the effects of different ethnicities and gender on the outcomes of creativity measurement.

Kaufman et al. (2010) investigated the reliability of the consensual technique in measuring creativity of 205 poems and stories through 108 college students. The participants and judges consisted of mixed gender with various ethnicities such as African American, European American, and Asian American. Excellent value of inter-rater reliability (ICC) were reported for both writing tasks, i.e., 0.92 for poems and 0.91 for stories. They showed that regardless of the panel of African American judges, the rest of the raters preferred to select the writings of their own ethnic as the most creative. Their study revealed that creativity can be reliably assessed across different ethnicities, but the final scores can be dissimilar. Kaufman, Baer, and Gentile, (2004) employed the consensual technique with 13 experts for assessing the creativity of 103 poems, 104 fictional

stories, and 103 personal narratives generated by Caucasian, African American, Latino/a, and Asian eighth-grade students. According to their study, there were no significant gender or ethnic differences among the creativity of most participants in terms of reliability. However, results indicate that the value of creativity can be different based on the cultural viewpoints of the judges. However, again it must be noted that the assessment of creativity using CAT only produces relative results and these outcomes are not the absolute value for creativity in any domain. The outcomes of the CAT cannot be generalised to any domain unless all the products in that domain have been carefully identified and included in the study. Otherwise, these outputs only indicate the degree of creativity between the subjects in question, that is any high score for creativity in one study may still be of moderate creativity in another study using a different set of artefacts.

3.4. Criterion-based measurement methods

Indirect and global judgment approaches mostly dominate the field of product creativity measurement. However, recent studies on creativity try to decompose it into several indicators that are arranged in a simple and meaningful configuration. The use of indicators allows analysts to have a uniform definition of creativity for a specific domain while also enabling designers to implant the creativity characteristics in their design. In the criterion-based techniques, researchers first make an effort to identify the influential attributes of creativity and then ask a group of untrained judges to compare and rate a number of products of the same functionality on those attributes. Evaluating creativity by criteria, in general, has higher feasibility as untrained raters are considerably cheap and accessible compared to expert judges. In addition, the outcomes are usually informative and can support the design team in recognising the area that require the most attention in order to enhance creativity.

Creative Product Analysis Matrix (CPAM)

Criterion-based models are constructed for a specific category of products but recently researchers have attempted to develop general criteria of creativity that can cover a broader range of products. One of the first efforts was made by Besemer and Treffinger (1981), where

they proposed a three-dimensional theoretical model known as the Creative Product Analysis Matrix (CPAM). The constituent components of the CPAM are Novelty, Resolution, and Style. The authors described *novelty* as the degree of originality and ability of the product to present surprising (unexpected) information to the user in the shape of a new approach, new idea, and/or new technology. The *resolution* is described as the extent to which the product serves useful functions that are capable of fulfilling user demands together with the level of benefits it can provide for the consumer (either physically, financially, socially, or psychologically). The *resolution* also refers to the ability of product in being effortlessly understandable and the degree to which it follows the accepted principles of the domain in question. The authors explained that the *style* (or Elaboration and Synthesis) denotes the level of personality in the product, and further characterised it as the measures of how well components fit together (organic), how well the product crafted, and how elegant it looks in the eyes of the viewer.

Besemer and colleagues (Besemer, 1984; Besemer & O'Quin, 1986) later developed the Creative Product Semantic Scale (CPSS) based on their previous theoretical model (CPAM). The CPSS comprises 11 sub-scales where each scale further defines with five bipolar semantic questions (i.e., 55 items in total) that altogether represent the original three-factor model (see Figure 3-1). In the CPSS, participants are given a few products to examine and then asked to judge each on 55 items using a 7-point Likert type scales (e.g., Overused 1-2-3-4-5-6-7 Fresh). The level of creativity is then computed by calculating the sum of average of the total ratings. The reliability and validity of the CPSS have been established through a decade of investigations utilising numerous types of products and judges (Besemer, 1998; Besemer & O'Quin, 1987). For instance, Besemer and O'Quin (1999) tested the construct validity of the model using a set of 3 novel Norwegian chairs and one traditional American chair. In their study, 185 psychology students were instructed to evaluate all the chairs on 55 bipolar items independently. The results of internal consistency (Cronbach's alpha) were reported at an acceptable range for all factors per product ($0.69 < \alpha < 0.86$). The authors also conducted a Confirmatory Factor Analysis (CFA) for each chair and results verified the validity of the initial three-factor model of creativity (e.g., the value of Robust Comparative Fix Index (RCFI) was captured in an excellent range for all chairs indicating the goodness-of-fit across all models ($0.91 < RCFI < 0.94$)).

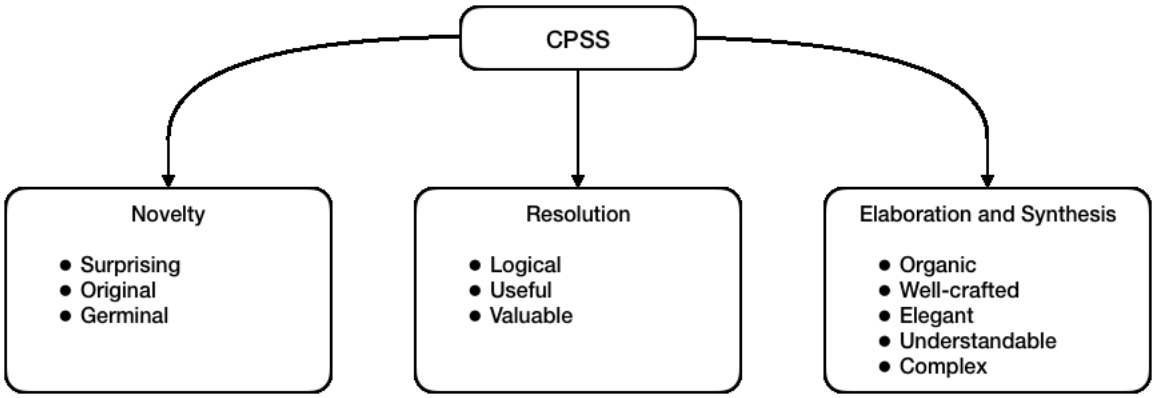


Figure 3-4 Revised Creative Product Analysis Matrix (CPAM) (Besemer & O'Quin, 1986)

Over the years, scholars have extensively used the CPSS to assess or improve the creativity of both tangible or intangible products (B. Smith, 1993; White & Smith, 2001). White, Shen, and Smith (2002) employed the CPSS to measure the creativity of 15 print advertisements using three groups of judges; college students, advertising professionals, and the general public. In their study, the authors; however, measured creativity utilising only three (out of 11) sub-scales of the model including originality, logic, and well-crafted. They indicated that the judgments of the originality and logic of ads were comparable among the three groups of judges, whereas the ratings of advertising professionals on the craft quality dimension differed significantly from those of students and the general public. In another study, Besemer (2000) evaluated the capacity of the CPSS in predicting the willingness to purchase for a cluster of consumer products. Contrary to what was expected, results revealed that the main reasons for purchasing a product are value and elegance dimensions and that novelty has the minimum impact.

The successful applications of CPSS in the product design and product improvement courses have been reported by the researchers. For example, Kristensson, Gustafsson, and Archer (2004) examined the creativity of three clusters of users (professional product developers, advanced and ordinary users) on a particular design task using an altered version of the CPSS. Outcomes of this study showed that ordinary users are able to produce more unique and meaningful products than the former two groups. Christiaans (2002) employed the semantic scale for examining the judgments of senior male students on creativity, prototypical value, attractiveness, interest, and technical qualities for rating three design models (telephone booths, displays, and drill holder).

Results showed that while creativity is highly associated to attractiveness and interest qualities, senior raters perceived them as separate dimensions. In this study, items such as original, surprising, germinal, valuable, and elegant were reported to have positive relationships with creativity.

Creative Solution Diagnosis Scale (CSDS)

Without a doubt, the active attitude of creativity does not allow it to be identified only with a specific set of indicators for all products. Typically, novelty is a must for creative products, yet in some cases, it is even more important for the product to be able to serve some useful functionality. Cropley and Cropley (2005) presented the concept of functional creativity and referred to product effectiveness as the main factor that provokes the sense of satisfaction in the user. The functional creativity aimed to assess those products in which the elements of effectiveness (usefulness) is of greater value to the user than the novelty aspect, such as manufactured consumer goods, scientific and/or technological products. In addition to novelty and effectiveness elements, the authors added two more dimensions of *elegance* and *genesis* in order to better explain the product creativity. The *elegance* factor pertains to the notion of grace in products and further characterised by how complete and pleasant the product appears in the eyes of the beholder. The *genesis* element on the other hand is the higher version of the novelty and specifically refers to situations in which the solution is capable of opening up new ways to existing problems and/or when the solution suggests new approaches for solving unrelated problems in other domains. The originators later employed the functional creativity definition to introduce the hierarchical organisation of creative products (see Table 3-1) (Cropley & Cropley, 2008).

Table 3-1 Levels of creativity in products (Adapted and updated from Cropley and Kaufman, 2012)

	Kind of Product				
	Routine	Original	Elegant	Innovation	Aesthetic
Effectiveness	✓	✓	✓	✓	✗
Novelty	✗	✓	✓	✓	✓
Elegance	✗	✗	✓	✓	?
Genesis	✗	✗	✗	✓	?

This hierarchy emphasises that creativity of a product can have different levels as well as kinds, and that creativity is not a concept that either exists or does not exist. The entry level of this hierarchy includes products that only serve some useful functionality and do not provide new features to the domain. At this level, products are recognised as uncreative (or routine) as they are only offering the ability to solve the problem that is in question without eliciting the admiration of the user. The next level (original) refers to the basic level of creativity where products contain both novel and useful attributes. Products are generally recognised as more creative if in addition to previous factors they can also offer to some extent a degree of completeness, gracefulness, and pleasantness (elegant), and they become even more creative if, for example, they possess the ability to provide solutions to other problems and/or attract user's attention into previously unnoticed ideas (innovative).

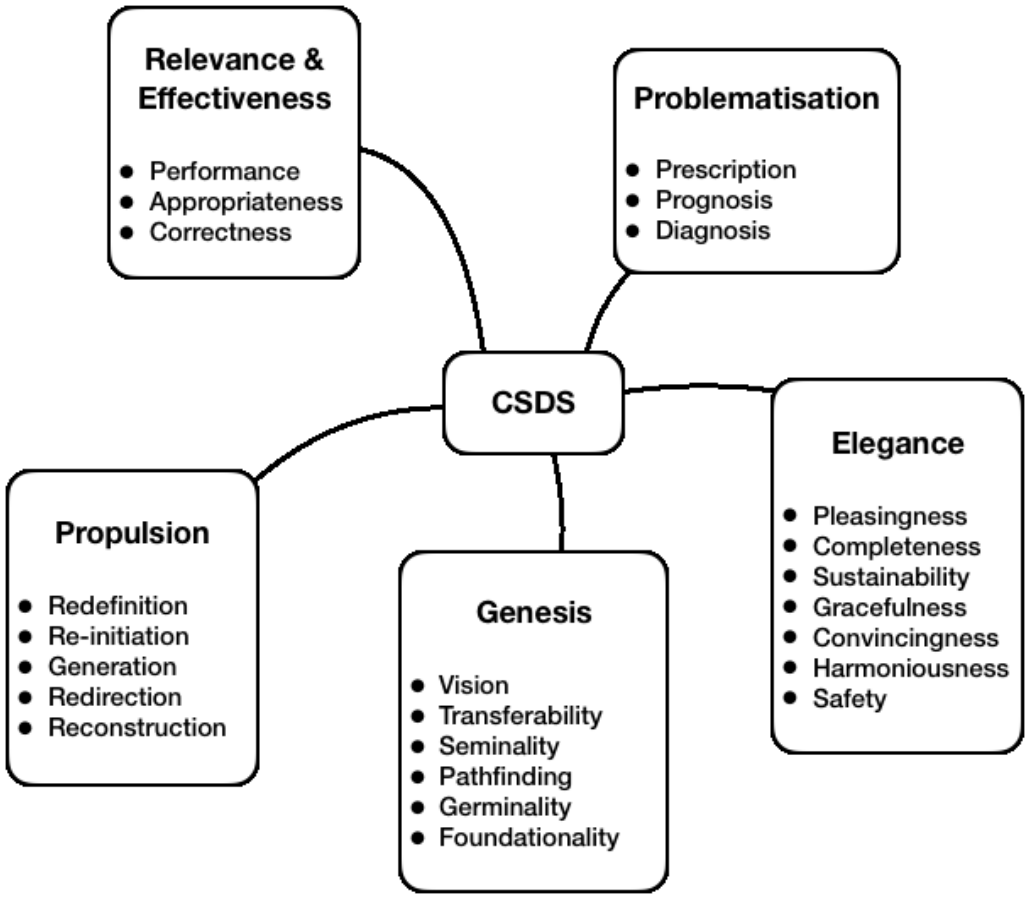


Figure 3-5 Revised CSDS Structure (Cropley & Kaufman, 2012)

Cropley and Cropley (2005, 2008) introduced the Creative Solution Diagnosis Scale (CSDS) consisting of internal and external indicators that were derived based on the previous hierarchical system and works of other researchers in the field (Besemer & O'Quin, 1986; Robert Sternberg & Lubart, 1999; A. Taylor, 1975). Initially the CSDS comprised 30 items that categorised into four factors: Relevance and effectiveness, Generation of novelty, Elegance, and Genesis.

The first factor refers to the extent to which the product does what it is designed to do (*performance*), and the degree to which the solution fits within task limitations (*appropriateness*). Moreover, it reflects the level of conventional knowledge and techniques that the product holds (*correctness*). The elegance factor relates to aesthetic aspect of the product and in general, signifies the details that represent overall gracefulness in products. Indeed, if components of a product are arranged in a consistent way (*harmoniousness*), and if the overall form of the product possesses the required proportion (*gracefulness*), and also if beholders perceive proficiency in the way product performs its tasks (*convincingness*), then it can be argued that the product contains the elements of elegance. The genesis factor on the other hand is indeed the extended version of the novelty factor. It assesses the product from further perspectives such as the ability of the product to offer new approaches to view and solve the current problem (*germinality*), the extent to which the solution introduces new standards and norms for evaluating other products in the domain (*vision*), and the degree to which the product causes a revolutionary change in the field and sets up a new path for other solutions (*pathfinding*).

Nevertheless, researchers had to perform several revisions to test the reliability and confirm the validity of the proposed structure. Cropley, Kaufman, and Cropley (2011) asked 323 college students to judge five mousetraps using the original 30 items of the CSDS instrument. The consistency among the raters (alpha) was obtained at a good to excellent range for all products indicating that CSDS is a reliable tool ($0.851 < \alpha < 0.961$). The outcomes of factor analysis (CFA) resulted in a five-factor construct as well as elimination of three items from the model as they did not load high enough on any factor. Researchers then split the novelty factor into two factors of *problematisation* and *propulsion.* The *problematisation* refers to a type of novelty in which, for example, the solution provides the user(s) with the ability to understand existing problems in the current solutions that they would not know otherwise (*diagnosis*), or when the idea causes a

movement in a particular field and makes other creators to consider it as the potential future of that field (*prescription*). The *propulsion* factor denotes another form of novelty where, for example, the product re-defines the way of accomplishing the task or utilising the product (*redefinition*), or even when solution offers a fundamentally new concept and perception to the current trend (*generation*).

In another study, Cropley and Kaufman (2012) examined the construct validity of the original CSDS instrument using 203 college students rating five different mousetraps of varying designs. Results of the inter-judge reliability (alpha) was reported at an excellent range for all five stimuli ($0.947 < \alpha < 0.968$). In this study, the outcomes of the confirmatory factor analysis confirmed the previous five-factor structure; however, resulted in the removal of 6 indicators from the original model due to either low loadings or cross-loadings higher than the cut-off value (0.3) (see Figure 3-5).

Product Creativity Measurement Instrument (PCMI)

Other researchers have suggested the consideration of additional factors to include as criteria in the measurement of product creativity. Horn and Salvendy (Horn & Salvendy, 2006b) proposed the Product Creativity Measurement Instrument (PCMI) containing the elements of the CPAM model as well as the consumer's perception of creativity. In their model, measurement of creativity is defined as the subjective assessment of a set of pre-identified criteria by a group of judges (consumers). Initially, the PCMI was introduced with seven factors consisting of Novelty, Resolution, Style, Pleasure, Arousal, Centrality, and Applicability. The first three factors are taken from the CPAM model of creativity which have been discussed previously (see p.55), while the rest of criteria are selected from the works of other researchers in the field (Bloch, Brunel, & Arnold, 2003; Christiaans, 2002; Demirbilek & Sener, 2003). The factors of arousal and pleasure represent the extent to which the product is able to excite the user(s) on an emotional level. While the terms centrality and applicability refer to how important (*applicability*) as well as how interesting (*centrality*) is creativity for the individual. The authors further discussed the consequence of consumer preferences in the judgment of creativity and argued that such

assessment can be greatly influenced by the degree to which the solution perceived as significant and engaging to the task by the rater.

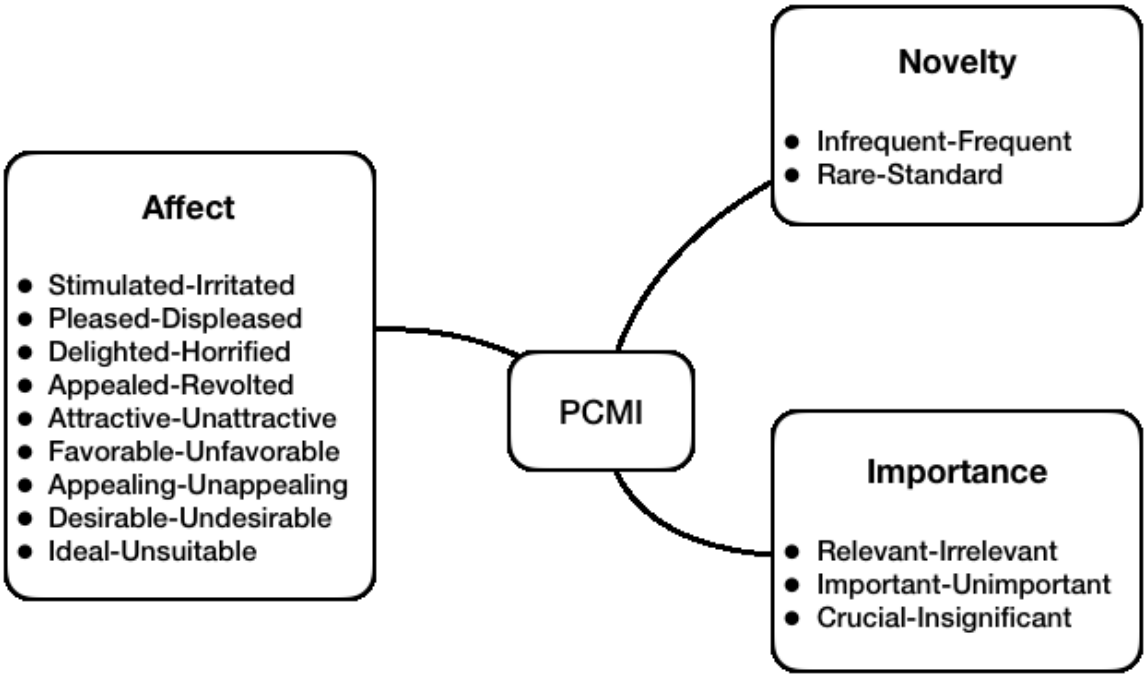


Figure 3-6 Product Creativity Measurement Instrument (PCMI) (Horn & Salvendy, 2009)

Horn and Salvendy (Horn & Salvendy, 2006b) examined the construct validity of the proposed model by surveying 205 undergraduate students and asking them to rate 90 items based on their previous experience with creative products. Results of the Explanatory Factor Analysis (EFA) revealed six significant factors in total that explained 64.8 % of the variance in the dataset (Novelty, Resolution, Emotion, Centrality, Importance, and Desire). In line with this, researchers have also conducted a confirmatory factor analysis to test the goodness-of-fit and long-term reliability of their criterion structure. The model was reported to be capable of achieving a good fit in all statistical indexes indicating the validity of the construct (e.g. GFI = 0.92, NFI = 0.93, RMSEA = 0.044). In this study, the originators also examined the degree of purchase intention as well as consumer satisfaction with respect to consumer's perception of the product creativity. The study indicated that both willingness to purchase and consumer satisfaction can be significantly predicted by three factors of Desire, Importance, and Centrality. It was shown that Novelty, Emotion, and Resolution factors are ineffectual to these important elements of marketing.

Subsequently, Horn and Salvendy (2009) refined their previous structure with 208 general product consumers evaluating four chairs and lamps. In this study, the results of three EFA analyses (chairs, lamps, chairs and lamps combined) all indicated a three-factor structure explaining 68.2%, 69.5%, and 67.8% of the total variance respectively. The authors labelled the three factors as Affect, Importance, and Novelty (see Figure 3-6). They characterised the *affect* dimension as the ability of the product to excite, please, and attract the attention of users, whereas the description of the other factors remained unchanged as before. Novelty considered as the extent to which the solution appears as infrequent and rare in eyes of the observers, and Importance is explained as how relevant and crucial the solution is for the problem. In overall, the PCMI indicated a good inter-rater reliability coefficient (alpha = 0.81).

3.5. Comparison of the product creativity measurement methods

In creativity assessment, CAT dominates the field as its validity (reliability) has been extensively studied and accepted by researchers (Amabile, 1982; B. Hennessey et al., 2008; Kaufman et al., 2010). Nevertheless, CAT is a time-consuming process and often involves expensive operating costs. The process of searching and defining appropriate panels of judges for measurement is challenging, and even when the panel is formed, the availability of each individual is usually constrained. Also, CAT does not suggest additional information for understanding the causes of why a product is being judged as low creative or uncreative. In addition, the application of the consensual technique has been less investigated for non-parallel consumer products. Recently, a few studies have been conducted using stimuli being generated under non-experimental conditions where final results revealed that CAT is indeed capable of achieving high reliability for such products (Baer et al., 2004; Christiaans & Venselaar, 2005). Nonetheless, further investigations are required to fully comprehend the consensual assessment technique's abilities for achieving valid outcomes using non-parallel products as well as to explore potential changes to CAT guidelines in order to maintain the assessment efficiency at high level.

On the other hand, criterion-based models define creativity using a set of pre-identified criteria and sub-criteria and hence they are effective in providing useful information about how creativity can be improved in the product. However, it appears that not all criterion-based techniques

define creativity in the same way. The review of the existing methods reveals that while some factors are shared across all instruments (e.g., novelty), there are certain factors that are distinctive. For instance, the *genesis* factor in the CSDS model refers to a higher level of novelty and has been less considered in other instruments. Likewise, the *affect* factor in the PCMI model contains items representing both aesthetic and emotional impact of the product on the consumer attributes, whereas in other instruments these elements either are not considered, or they are not measured under one factor. More contradictions between criterion-based models are evident by studying their factorial structure and corresponding items that form the latent factors. Although a portion of this distinction between instruments is due to the idea they follow, a part of it is due to the product functionality used for creating the model. Typically, researchers select a group of products of the same functionality as stimuli in order to create the model, while the outcomes of the factor analysis vary depending on what product is selected for the study. Sometimes factors are positively or negatively correlated with each other precisely because of the product that is in question. For example, when people are asked to give their opinion about the practicality of a product where practicality is described as a combination of human and performance factors, then there is a high chance for the factorial model to consider those interfering items as one latent factor which indeed might not be correct for another product or situation. In other words, apart from raters and psychometric questions (items) in the questionnaire, the functionality of the selected products also influences the outcomes of the factor analysis.

3.6. Chapter Conclusion

In this chapter, the majority of existing methodologies for the measurement of product creativity are explained. An objective assessment of creativity requires a reference standard to allow other measurement tools to be validated (and calibrated). As result, almost all explorations in this class are devoted to the assessment of product novelty. The objective measurement of novelty is achievable by first decomposing the product into its constituent elements and then searching for originality and/or uniqueness at different layers of the product structure. If novelty occurred at higher levels of the product hierarchy, then it can be concluded that the product has a higher degree of novelty than those to which newness took place at lower levels of the product

hierarchy (see objective methods). The vast majority of the creativity measurement tools and instruments are subjective and typically rely on appropriate panels of judges as well as the judging process. Scholars have divided these methods into three categories including indirect measurement, global-judgment, and criterion-based techniques. The indirect methods utilise the measurement of product creativity as a means to evaluate a person's creativity. Under this methodology, for example, peers determine the most creative person of their fields by considering the past works of the successful individuals of their domain. On the other hand, the global-judgment approaches neither consider any definition for creativity nor attempt to limit it into a set of creativity norms. In fact, this type of measurement assumes that the product is creative to the extent that appropriate expert judges agree that the product is creative. The global judgment methods have by far dominated the assessment of (product) creativity in the arts, music, poetry, and writing industries; and this chapter has reviewed some of the important studies in this area. It was indicated that for almost all type of products, the use of expert is a must for any measurement completed making use of global-judgment approach, although in some cases that the access to experts is restricted, those quasi-experts with appropriate knowledge in the field can indeed represent the expert opinion. Judgment of creativity is independent to gender or ethnicity; however, expert judges from a particular culture may rate the products differently compared to those of other cultures, but their outcome is often reliable.

Contrary to global-judgments, the criterion-based techniques distinguish creativity using a set of pre-identified criteria and sub-criteria. In general, almost all criterion-based techniques link creativity to novelty and discuss that no product can be creative without being novel at the first place. All other factors are indeed representing the product *effectiveness* and some of them are including Resolution, Affect, Importance, Elegance, Elaboration and Synthesis. However, not all models look at product effectiveness from the same point of view. Defining criteria for measuring *effectiveness* (appropriateness) is by far the most challenging aspect of the product creativity assessment. Usefulness can be generally described as technical performance and/or sometimes as aesthetic factors, but in order to measure it accurately one must understand and consider the relevant scientific variables, principles, and technologies that altogether represent performance for that product (e.g., understanding of horsepower, acceleration, and torque are required for

describing the engine performance of the consumer vehicles). The study of the current instruments indicate that the underlying structure of creativity is different depending on the functionality of the product that is used, raters, and other experimental conditions that have potential to alter the outcomes of factor analysis.

4. Investigation of the Extension of Consensual Assessment Technique into Consumer products: three case studies

The existing approaches for measuring product creativity have been reviewed in the previous chapter. In general, the main assessment methods include indirect measurements, criterion-based techniques, and global judgments (Horn & Salvendy, 2006a). Indirect measurements include considering product creativity assessment as a means to determine a person's creativity. For instance, in self-reporting scales, individuals are asked to specify their previous achievements and awards in various fields, and then creativity is measured by considering the number and degree of each accomplishment (Carson et al., 2005).

In criterion-based techniques, creativity is decomposed into several criteria and sub-criteria where it then becomes equivalent to the average value of the scores for each item given by raters who assessed the product. However, one problem with criterion-based techniques is that they are typically validated for a particular domain and employing them for other fields where they have never been used before is usually questionable. The reason for this inconsistency lies in the structure of creativity. Creativity cannot be defined equally for all products in the universe using a set of constant factors, and this makes it difficult for general criterion-based models which intend to assess the creativity of a wide range of products. There is always a risk that the constructed model does not genuinely represent creativity in the real-world scenario as criteria that define product usefulness (effectiveness) usually vary from product to product.

Global-judgment methods neither define creativity nor product creativity. They are only based on the overall assessment of expert judges. The Consensual Assessment Technique is among the most popular methods in this category and has been extensively utilised in artistic fields such as music, poetry, and writing. However, in the research literature, less attention has been given to the assessment of creativity of functional consumer products which usually possess complex structure and interactive elements. Consumer products form an important part of everyday life

and promoting creativity in this domain can significantly affect the success of innovative companies (IBM, 2010). Consumer products can be defined as products that result from a production process and typically purchased by the general public for their personal (and/or private) consumption and often in the absence of direct supervision (Mital, Desai, Subramanian, & Mital, 2014). Functional products, on the other hand, are products comprising of hardware, software and service support system (Lindstrom et al., 2015). The hardware represents any tangible assets in the product, the software refers to the operating tool and system, and the service support system denotes the maintenance of the product, operations planning and educating and training the usage of it. Those functional consumer products which are competing in a market are often manufactured under distinct conditions using contrasting designers and resources as well as within different time intervals. Therefore, it is important to understand whether the CAT is capable of assessing such products for creativity under its original conditions and protocols (see Amabile, 1982). In addition, the process of identifying the appropriate size of the panel of judges is among the essential elements affecting the inter-judge reliability. Therefore, for domains with limited backgrounds it might be worthy to conduct the assessments with a larger panel of judges at the beginning and then make an attempt to see if the number of judges can be decreased in any way. For instance, more studies can be done using the same or similar products but different panels of judges where the outcomes can be utilised to establish the minimum number of raters for a reliable assessment in that domain.

Nevertheless, there is not that much information about the ability of CAT in assessing the creativity of non-parallel functional consumer products, i.e., products produced under dissimilar conditions. The process of making creative products requires not only creative individuals but also a creative working climate. A creative person must have adequate capabilities and be able to stay motivated during the process. Whereas, environmental factors such as having freedom in deciding what to do, sufficient resources and time also play important roles. As a result, if creativity is associated with individuals and their interaction with the environment, one can say that a practical measurement tool should be capable of assessing the creativity of products produced under such conditions.

If CAT indicates a valid measurement for non-parallel functional consumer products, it then can be employed to inspect the relationships between other forms of creativity. For instance, the performance of individuals who are subjected to specific thinking processes can be compared and analysed to one another, and it can be determined that which thinking procedure is more responsible for creating creative thoughts (or outcomes). Also, researchers can investigate the effects of different environmental conditions on creativity and specify the most influential elements affecting the performance and technical efficiency of creative thinking. If CAT shows a reliable assessment for non-parallel functional consumer products, it can also be useful for businesses to not only improve their leadership in the market but also enhance their market size and long-term profit.

4.1. Research aim and objectives

The primary aim of this study is to investigate the validity of the consensual technique in assessing creativity of three categories of functional consumer products, including home accessories, electric vehicles, and fire alarms. Also, to better understand the constraints and provide recommendations of utilising the CAT for assessing creativity of products produced under dissimilar environments utilising distinctive equipment and material.

Objective 1: to examine the creativity of three distinctive categories of functional products using the consensual technique, as well as to measure inter-judge reliabilities and compare the results with consistency ratios obtained in artistic domains such as art, writing, and music.

Objective 2: to investigate whether expert judges are able to distinguish the differences between creativity and other dimensions for all product categories.

Objective 3: to examine the minimum number of experts required for a reliable assessment comparing to the conventional number of experts (Baer & McKool, 2009a; Hickey, 2001; Kaufman, Baer, et al., 2009) used for a reliable assessment in other artistic domains.

Objective 4: to understand how experts are preserving creativity across three categories of products. In other words, to compare the influential factors of creativity across all three categories of products and to find out whether they are constant for all categories of products.

4.2. Methodology

In this study, a non-experimental research method was used to investigate the assessment of product creativity using the consensual technique. Quantitative techniques were employed to numerically collect experts' opinions on creativity as well as to analyse and determine relationships between the influential factors. The data was collected by administrating surveys containing Likert type (rating scales) questions. In this study, creativity is treated as a continuous variable as defined in creativity research meaning that it exists on a spectrum containing two endpoints called (highly) creative and uncreative (Amabile, 1996; Piffer, 2012; Shalley, 1995). For instance, sometimes a significant contribution in technology can alter the future of that field, and hence it can be considered highly creative. On the other hand, one can re-establish an already existing idea and suggest it for a new application. While this action is creative, it is at a relatively low level. Therefore, creativity occurs on a continuum with some outputs are begin judged as more or less creative than others, and some as in-between. The study was conducted in two phases including a pilot study and the main study. The pilot study was performed to find out information about the overall performance and feasibility of the research including implementation, readability of questions, and average time required for completing the survey. The main study was conducted to assess the reliability and validity of consensual technique in assessing creativity products produced under dissimilar conditions.

Pilot Study

The aim of this study is to resemble the main study using a smaller population in order to specify those administrative problems and systematic errors that need to be addressed prior to the main study. Consequently, 11 design experts were invited to assess three cordless drywall screwdrivers for creativity. The survey was administrated online where raters read and agreed to a consent form and then observed a short video containing technical specifications and appearance of the products. Then, they were instructed to compare and assess each product independently for creativity in addition to nine more items. The panel of experts included only male raters mostly ranging from 45 to 64 years old with a standard deviation of {SD} = 1.453. The result of this study can be seen in detail in "results of pilot study" section of this chapter. In overall, the outcomes of

the descriptive statistical analysis as well as reliability tests did not indicate any administrative or design problems.

Main study

i. Material

For this study, three independent sets of consumer goods were selected, including (a) Home accessories (wall-clocks), (b) Electric vehicles (EVs), and (c) Smoke alarms. Under each group, there were three products coded as A, B, and C as demonstrated in Figure 4-1. The products of each group had identical functionality but featured distinctive characteristics to allow comparison such as dissimilar components, appearance, structure, and technology. In this study, the number of products is relatively low compared to most studies using the CAT (Baer et al., 2004; Kaufman et al., 2010; K. Tsai, 2016). This is mainly because in a well-established market, the main characteristics of functional products are defined based on consumer demands and in contrast to those in artistic domains they are very limited in numbers when it comes to comparing them against each other. In other words, the number of products that can be put against each other in a fair comparison is usually limited to a few items given the distinctive characteristics of each product. For instance, current automotive market is divided into several small areas where each has its own unique characteristics. Manufacturers often classify their range of products based on different elements such as body type (e.g. SUV, sport, and sedan), engine size, gearbox (automatic/manual), number of seats, and fuel type to not only distinguish them from others but also to emphasize on the fact that each class is different and they cannot be compared to one another in a fair comparison. Since, the fairness of comparison is an important element in the CAT, then it is sensible to have a limited number of stimuli but pay more attention to product selection to ensure comparison is fair. In this research, products of the home accessory and smoke alarm groups were carefully chosen from competing items in the specific departments of a reliable online shopping website. The EVs considered in this study were classified in the A-segment of the European Commission classification (i.e., a small car designed to be used primarily in urban areas). The boundary of performance for the selected products is presented in Table 4-1.

Table 4-1 Boundary of performance for selected consumer products

Performance parameters	Quality and quantity
Electric vehicles (EVs)	
Classification	A-segment (city car)
Electric range	160 to 300 km (according to NEDC*)
Fast DC/ AC charging time	30 to 40 mins (DC) _ 3.5 to 5.5 hrs (AC)
Output power	81 to 170 hp
Practicality	5 doors _ 4 to 5 seats
Home accessories	
Style	Modern/Classic
Material	Plastic/wood
Colour	White/Orange/Black
Number of components	1 to 13 components
Smoke alarms	
Technology	Ionization/Photoelectric/Dual sensing
Average time to detect fast flaming fire	105 to 200 s
Average time to detect smouldering fire	60 to 110 mins
Error (%)	4% to 19%

ii. Judges

Three groups of experts were recruited to assess the products. To achieve sufficient number of experts, a total of 360 invitations were sent to specialists in both industry and academic sectors. For electric vehicles, 34 complete r esponses (out of 120 invitations) were collected. Among the judges were academic lecturers and researchers who study electric-vehicles, electrification of vehicles, electric mobility, and vehicle ergonomics. Raters were 94% male with an average age interval of 45 to 54 years old and the standard deviation of {SD} = 1.17. Similarly, a group of 36 expert judges composed of 31 industrial designers and 5 design engineers evaluated the wall-clocks. There were six females and 30 males with an average age interval of 35 to 44 years old, {SD} = 0.95. Finally, 34 individuals with more than 10 years of experience in fire safety agreed to assess the creativity of three fire alarms (26 fire safety advisors and officers, 4 fire protection engineers, and 4 health and safety officers). Among them were 2 females and 32 males, and on average they were ranged from 55 to 64 years old, {SD} = 1.2. In CAT, the number of judges is an important element as it is capable of influencing the assessment's reliability. In the literature, there is no clear statement about an ideal number of raters, however, many researchers believe that for most domains a panel consisting of between 5 to 10 judges is an appropriate size (Kaufman, Plucker, et al., 2008).

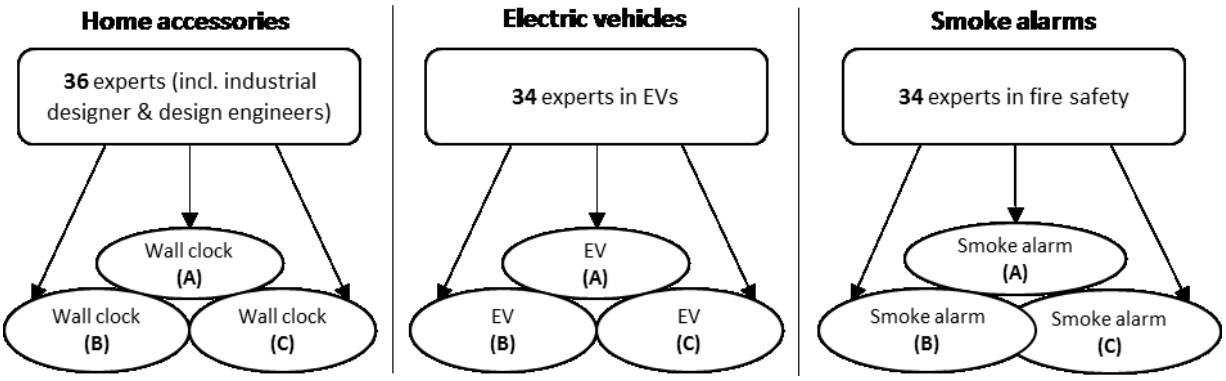


Figure 4-1 Three groups of experts independently assessed the products of their corresponding group. Expert judges were first told to observe and compare the products to one another and then asked to separately evaluate each product for creativity and additional nine dimensions using a 5-point Likert scale

In the past, researchers were able to capture reliable assessments by using only 2 judges (Daly et al., 2016) as well as they reported with over 130 judges (Valgeirsdottir & Onarheim, 2015). In practice, a small panel may not achieve acceptable reliability while a large panel may mislead the ratio of Cronbach's alpha (Cseh & Jeffries, 2019). In this study, a larger panel of judges was selected mainly for two reasons. First, to improve statistical power of the assessment and second to ensure that the correct number of raters for having an acceptable level of reliability could be detected given that a new category of product (i.e. functional consumer product) is being investigated.

iii. Judging Procedure

The process of data collection was conducted online, where judges read and agreed to consent conditions, and then were given instruction for the task. A novel approach for product presentation was used in the sense that each group of experts were presented with three short videos of the products of their corresponding group to visually assess their appearance and become familiar with their technical specifications. Videos are preferred because of their ability to provide a large volume of information in a short amount of time while giving raters the opportunity to observe the product from all perspectives without being misled in any direction. All videos last just under one minute. Initially, each video provided an introduction to the product highlighting its corresponding code (name) in the study plus a few 2D images to help judges to

become familiar with the aesthetic appeal of the product. Then, product specifications were given using both text and 2D pictures (animation where possible).

To prevent confusion and fatigue, only one feature of the product was presented at a time. Throughout the video, judges were presented with all essential information about the product, for example, with electric-vehicles, information such as performance characteristics, lithium-ion battery details, range and charging time, information about materials (whether they are recyclable), and vehicle's safety factor (Euro NCAP stars) were provided to the observers (see Figure 4-2).

Figure 4-2 Screenshots of the product presentation videos for electric-vehicles group. Images show different product feature at a time (e.g. charging time, technology, safety, and practicality)

To be consistent with the CAT procedure, the experts were asked to work independently without consultation about their ratings. Judges did not ask to explain or provide any support for their ratings. In addition, the order of presenting the products as well as questions were randomised to limit systematic errors in the study. After judges familiarized themselves with the products, they were told to compare them to one another and assess each for creativity using their own

comprehension of the concept. A 5-point Likert scale with 1 representing 'Not at all' and 5 representing 'Extremely' was used for data collection. According to Preston and Colman (2000), an optimum degree of reliability, validity, and discriminatory during the assessment can be achieved with rating scales of between 5 to 7 points while any fewer or more points could negatively influence the outcomes. In addition, judges were also asked to compare the products on nine dimensions which are given in Table 4-2. These dimensions expose a wide range of aspects and their relationships with creativity have been already established by the researchers in the field (Besemer & O'Quin, 1986; Cropley & Kaufman, 2012; Harold, Balaban, & Pecht, 2009; Horn & Salvendy, 2006a; Podolyakina, 2017). During the evaluation process, judges were allowed to change their opinions about different attributes. After finishing the survey, they had no access to observe or make further changes to the ratings.

iv. Reliability

The ratio of inter-rater reliability has been calculated using Cronbach's coefficient alpha. The coefficient alpha is a standard measure of internal consistency (Cronbach, 1951). In creativity research, scholars have employed the alpha in analysing the levels of inter-judge reliability by transposing the items with raters in the data matrix (Cropley et al., 2011; Kaufman, Baer, et al., 2008). Researchers usually report an excellent consistency if the (α) coefficient is equal or greater than 0.9 (Bland & Altman, 1997; Nunnally & Bernstein, 1994). The alpha of $0.8 \leq \alpha < 0.9$ illustrates a 'good' level of internal consistency whereas any value between $0.7 \leq \alpha < 0.8$ shows an 'acceptable' level of agreement among raters. A coefficient alpha below 0.7 is often interpreted as unacceptable consistency or indicates the presence of systematic errors in the study.

Table 4-2 Dimensions of product creativity

Novelty (originality)	Reliability and Maintainability
Technical performance	Emotional impact on the consumer
Safety and security	Attractiveness
Human factors	Importance (relevancy to task)
Sustainability	

v. Factor Analysis (Exploratory Factor Analysis)

Factor analysis is a statistical approach to simplify a large set of variables into smaller, easy to interpret, and meaningful factor structure. The procedure is used to discover patterns from a group of observed variables and reduce the number of variables to fewer (latent) factors that share common variance as the original observation (Bartholomew, Knott, & Moustaki, 2011). Exploratory Factor Analysis (EFA) is a technique within the factor analysis family that aims to determine the optimum underlying relationships between variables to which data can be described with a minimum number of latent factors while presenting the same concept together with the maximum portion of the variance of the original observation (DeCoster, 1998).

In the Consensual Assessment Technique, EFA is employed to indicate the degree of relatedness or dissociation between creativity and other dimensions used in the study. The procedure illustrates the extent to which judges are able to psychologically separate creativity from other factors so that their ratings are truly representing creativity ratios. This is an important step to indicate the levels of discriminant validity of the consensual assessment (Amabile, 1982; Baer & McKool, 2009a). In this study, factor analysis was conducted on the mean ratings of 10 dimensions of judgment per product category. To extract the factor loadings, Principal Axis Factoring (PAF) was used to numerically estimate the variable's communalities (i.e. the variance of variable accounted by the common factors plus residual) and generate a reduced correlation matrix to compute loadings that best represent the original observation. Afterwards, loadings were rotated under oblique rotation algorithm to indicate the best possible factor structure. The oblimin rotation system was selected over orthogonal approach for its assumption that correlations exist between the factors. A cut-off point of 0.4 was used as a criterion to accept loadings on to one factor (Stevens, 1992).

vi. Regression Analysis

Multiple linear regression analysis (MLR) is a parametric (statistical) approach to specify the significant predictors of the target variable (dependent) out of a set of predictive variables (independent). In creativity research, linear regression is used to determine the significant predictors of creativity in different domains (Besemer & O'Quin, 1999; Cropley & Kaufman, 2012). In this study, the influence of nine additional dimensions (see Table 4-2) as independent

variables on creativity as the dependent variable was inspected across three product categories using multiple linear regression analysis. The IBM-SPSS (version 24) software was used to perform all statistical computations.

4.3. Results of pilot study

Eleven design engineers and product design academics were asked to assess the creativity of three cordless screwdrivers. The result of inter-judge reliability was attained at an acceptable rate ($\alpha = 0.73$) indicating a sufficient consistency between expert judges. Also, the ratio of inter-item reliability (Cronbach's alpha) was calculated to examine whether all items in the survey measure the same concept or construct. In this dataset, the value of internal consistency among the items was obtained at a good ratio ($\alpha = 0.894$). In addition, the data was examined for normality using the Skewness and Kurtosis tests and results did not indicate statistically significant difference from normality. Overall, the results of this study revealed the adequacy and acceptability of the survey to participants. The survey indicated reliable outcomes, all performance characteristics were statistically acceptable, and the time required for completing the survey was appropriate.

4.4. Results and Discussion

Domain relevant experts measured the creativity of three sets of products (home accessories, EVs, and smoke alarms). The inter-judge reliability of the creativity ratings for all product categories was achieved at an excellent range, 0.90 for home accessories, 0.95 for EVs, and 0.97 for smoke alarms. Moreover, there was consistency in the judgment of judges on nine additional dimensions. For home accessories, 9 out of 10 items had inter-rater reliability greater than 0.7 with the median (reliability) of 0.88. For the EVs, the median value dropped to 0.86, where 7 dimensions had reliabilities above the 0.7 threshold. Similarly, seven items had inter-judge reliabilities larger than 0.7 for the group with smoke alarm products. The median reliability for this category was 0.8. The data were subjected to factor analysis to examine whether the judgements of creativity by experts was essentially separated from the assessment of technical goodness or aesthetic aspects of the products.

Exploratory factor analysis was conducted utilising PAF and Oblique rotation algorithm. The factorability of the items was evaluated prior to analysis. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was achieved above the frequently suggested threshold of 0.6 for all product categories (Coakes & Steed, 1999); 0.77 for home accessories, 0.85 for EVs, and 0.79 for smoke alarms. The Bartlett's test of sphericity was also significant ($x^2(45) = 514.997,\ p < .001$ for home accessories; $x^2(45) = 453.288,\ p < .001$ for EVs; and $x^2(45) = 357.816,\ p < .001$ for smoke alarms).

Table 4-3 Factor analysis on 10 dimensions of judgment for all groups of products

Home accessory group		
	Creativity cluster	Technical goodness cluster
Creativity	**0.940**	- 0.046
Novelty (originality)	**0.773**	- 0.083
Attractiveness	**0.727**	0.131
Emotional impact on the consumer	**0.766**	0.04
Performance	0.101	**0.695**
Sustainability	- 0.09	**0.886**
Reliability and Maintainability	0.049	**0.651**
Human factors	- 0.02	**0.839**
Importance	- 0.003	**0.642**
Smoke alarm group		
	Creativity cluster	Aesthetic appeal cluster
Creativity	**0.961**	0.012
Novelty (originality)	**0.775**	0.034
Performance	**0.597**	0.183
Reliability and Maintainability	**0.781**	- 0.026
Safety	**0.534**	- 0.154
Sustainability	0.116	**0.578**
Attractiveness	- 0.018	**0.403**
Emotional impact on the consumer	- 0.148	**0.803**
Electric vehicle group		
	Factor 1	Factor 2
Creativity	1.021	- 0.188
Novelty (originality)	0.878	- 0.009
Attractiveness	0.665	0.187
Emotional impact on the consumer	0.558	0.339
Safety	- 0.045	0.295
Performance	0.628	0.224
Human factors	0.737	- 0.416
Sustainability	0.058	- 0.474
Reliability and Maintainability	0.510	- 0.032
Importance	0.406	0.386

For home accessories, initially, a solution with three factors was extracted based on the eigenvalue cut-off point of 1.0. After examination of the pattern matrix, one factor was dropped due to the insufficient number of primary loadings. As a result, a factor analysis using PAF and Oblimin rotation settings was repeated with two factors explaining 60% of the variance. All items under the main factors had strong loadings (greater than 0.5) without any substantial cross-loadings (see Table 4-3). Only one item (safety) was removed as it failed to load greater than 0.4 on any of the factors. The first latent factor appeared to be Creativity which consisted of *creativity*, *novelty*, *attractiveness*, and *emotional impact on the consumer*, whereas the second factor labelled as technical goodness composing of *sustainability*, *reliability and maintainability*, *technical performance*, and *human factors*. The internal consistency of the items for each cluster of variables was in the acceptable range (0.87 for creativity, 0.86 for technical merit). Since the oblique rotation algorithm allows factors to be correlated, the value obtained 0.182 between the extracted factors in this category indicates that factors are distinct and uncorrelated.

Similarly, a factor analysis utilising the same settings was performed for the products of smoke alarms group and results presented in Table 4-3. A solution with two factors was achieved that explained 49.9% of the variance. The factors seem to be Creativity and Aesthetic appeals of the products. In this category, the items of *importance* and *human factors* loaded low on both factors (i.e. smaller than 0.4) and hence were removed from the final model. The rest of items loaded strongly under at least one factor and there were no cross-loadings greater than 0.4. Overall, CAT was indicated as a reliable assessment method for the products in the home accessory and smoke alarm groups. The results of factor analysis illustrated that the ratings of creativity by experts were psychologically detached from the evaluation of technical merit and aesthetic aspects of the products. The items that loaded highly on Creativity factor were different across the two product categories indicating that the perception of creativity is a dynamic concept and that the influential elements would change depending on product functionality. Discriminant validity was attained between the constructs of the two models.

Table 4-4 gives the values of Average Variance Extracted (AVE) as well as the correlation between factors per product category. According to the criteria suggested by Fornell and Larcker (1981), to establish discriminant validity the AVE for the two factors should be larger than the square

correlation (i.e. shared variance) between the factors. In this study, the AVE values for both creativity, and technical merit factors for the home accessories group were obtained at larger ratios in comparison with the squared correlation between the factors (0.64 > 0.038 & 0.56 > 0.038 respectively). For the smoke alarms, the AVE values of the obtained latent factors (0.55 for creativity; 0.38 for aesthetic appeal) were also larger than the squared correlation between the factors (0.32). The outcomes indicated that discriminant validity exists within the constructs.

For the EVs, a factor analysis (using oblimin rotation) on the mean ratings of the 10 dimensions resulted initially in a solution with two factors. However, the analysis of pattern matrix did not show a meaningful separation between the clusters of variables. Many of the items cross loaded on both factors and some did not load high enough on any of the factors (see Table 4-3). For this category, discriminant validity was not achieved as there was a high correlation between the two constructs.

Table 4-4 AVE and squared correlations between the constructs for home accessory and smoke alarm groups

Home accessory group			
	AVE	Creativity factor	Technical goodness factor
Creativity factor	0.64	1	
Technical goodness factor	0.56	0.038	1
Smoke alarm group			
	AVE	Creativity factor	Aesthetic appeals factor
Creativity factor	0.55	1	
Aesthetic appeal factor	0.38	0.32	1

This indicated that the judgments of creativity by domain experts is not psychologically separated from the ratings of other dimensions such as technical goodness and/or aesthetic appeals of the products. One reason for this untidiness in factor loadings is probably due to the complexity that exist in structure of electric-vehicles. A typical electric vehicle consists of multiple systems and sub-systems, which makes it difficult to perceive the distinctions between various product attributes without examining them in the real situation. For instance, in order to describe a vehicle's overall performance, aside from its dynamic and engine characteristics, several other

factors are contributing, including chassis design, aerodynamic, suspension, and associated electronics. This issue is applied to many other functional products that contain a complex structure where numerous advanced technologies and concepts are involved. One potential way to overcome this issue is to either ask more specific questions or break down the product into its constituent elements to ensure raters can effortlessly differentiate the distinctions.

This study has indicated that creativity measurement of non-parallel consumer goods by domain relevant experts is reliable as long as the product is simple to understand. Simple to understand in this context means any product that is composed of multiple systems (products) where specific knowledge is required to understand each section and to be able to assess them for creativity. As the level of complexity increases, the ability of the judges to perceive creativity at a glance reduces. Therefore, the use of CAT under its current condition is not recommended for complex products produced under dissimilar environments.

Table 4-5 Significant predictors of creativity per product category

Significant predictors	Standardized coefficients (Beta)	95% CI		*p*-value
Home accessories				
Novelty	0.489	0.347	0.669	.000
Attractiveness	0.325	0.218	0.559	.000
Emotional impact on the consumer	0.188	0.039	0.385	.017
Electric vehicles				
Novelty	0.415	0.264	0.554	.000
Technical performance	0.244	0.115	0.346	.000
Human Factors	0.211	0.121	0.361	.000
Attractiveness	0.154	0.027	0.271	.018
Smoke alarms				
Novelty	0.450	0.377	0.653	.000
Technical performance	0.312	0.218	0.464	.000
Reliability and Maintainability	0.258	0.176	0.463	.000

Table 4-5 indicates the outcomes of three linear regression analyses. Results revealed that the influential elements of creativity are dissimilar across different product categories. In the home accessories group, factors such as *novelty*, *attractiveness*, and *emotional impact on the consumer* statistically significantly predicted the overall creativity. These indicators accounted for R^2 =

0.728 (*F* (9, 106) = 28.889, *p* < .001) of the variation in this dataset with adjusted R^2 = 0.70, a large size effect according to Cohen (1988). For electric-vehicles, a model was attained with four significant predictors including *novelty*, *human factors*, *technical performance*, and *attractiveness*. These items were responsible to explain R^2 = 0.834 (*F* (9, 99) = 50.246, *p* < .001) of the variance in this category (adj. R^2 = 0.817, a large size effect). Finally, three of the nine criteria significantly predicted creativity for the smoke alarms group. These were *novelty*, *technical performance*, *reliability and maintainability*. They accounted for more than 84% (*F* (9,98) = 53.497, *p* < .001) of the variance in this cluster (with adj. R^2 = .828, a large size effect). The outcomes of the regression analysis revealed that factors affecting product creativity are varying depending on product functionality, which further verifies that creativity is a dynamic concept.

The inter-judge reliabilities obtained in this study are exceptionally high, and they are in a similar range to studies in artistic domains where artefacts were gathered under tightly controlled conditions (Baer et al., 2009; Kaufman, Baer, et al., 2008; K. Tsai, 2016). These extremely high ratios of reliability can be due to the use of larger panel of judges. Researchers often assess creativity using a smaller group of raters (3 to 10), as larger panels are difficult to access and involve a higher cost (Baer & McKool, 2009a; Hickey, 2001; Kaufman, Baer, et al., 2009).

Table 4-6 Inter-judge reliabilities for group of 5, 10, 15, 20, and 25 randomly selected raters per product category

	Home accessory group	Smoke alarm group	Electric vehicle group
α for 5 judges	0.52 - 0.76	0.54 - 0.65	0.12 - 0.59
α for 10 judges	0.57 - 0.79	0.75 - 0.87	0.47 - 0.68
α for 15 judges	0.75 - 0.79	0.90 - 0.95	0.67 - 0.88
α for 20 judges	0.78 - 0.81	0.93 - 0.96	0.86 - 0.90
α for 25 judges	0.84 - 0.91	0.96 - 0.96	0.93 - 0.95

The current study (on average) employed 35 skilful individuals to assess the creativity. Since this value is considerably larger than the frequently used number of raters, the effect of the number of raters on reliability ratios was inspected using the modified version of the approach introduced by Baer et al. (2004). The experts per product category were split into 5 clusters of randomly

selected judges, including a cluster of 5, 10, 15, 20, and 25 experts. Each cluster contained three randomly selected groups of experts who assessed the products of the same category. The ratio of coefficient alpha was then computed per group of raters and results presented in Table 4-6. For the home accessories, the coefficient alpha was obtained in an acceptable range for the clusters with 15, 20, and 25 experts (0.75-0.79, 0.78-0.81, 0.84-0.91 respectively). Whereas, the value of alpha was ranged insufficiently for the groups with less than 15 raters (0.52-0.76 for groups with 5 experts, and 0.57-0.79 for groups with 10 experts). For the smoke alarms, the alpha was attained adequately for almost all clusters of judges except the groups with 5 experts (0.54-0.65). The level of agreement between the judgments of creativity in the electric vehicles group was captured only appropriately for the clusters with 20, and 25 judges (0.86-0.90, 0.93-0.95 respectively). Considering that the acceptable range of alpha is higher than 0.7 (Nunnally & Bernstein, 1994), the results of the analysis with fewer judges indicated that in general a greater number of experts is required for the creativity assessment of consumer goods produced under uncontrolled conditions. In the current investigation, the minimum number of experts for a reliable assessment of creativity were 10 for smoke alarms, 15 for home accessories, and 20 for electric vehicles. The outcomes demonstrated that the creativity judgments of relatively complex products may not be valid unless an adequate number of raters are presented. When the degree of complexity in products passes a certain level to which evaluation is unreasonable without testing the features of the product, the assessment of creativity using CAT under its original condition needs further investigations. Complexity in this context includes a series of characteristics in a system that make it complicated to understand its constituent components, the connectivity between different elements, and how these elements are controlled and managed. Main characteristics of a complex system consist of a large number of components and technologies within the system, high degrees of diversity between components, a non-linear and complicated level of interaction between the components. Also, factors such as higher autonomy (degree of independence of the elements in the system) could contribute to complexity by making it difficult to control the system from several regions rather than more centrally (Systemsinnovation, n.d.).

The analysis of EFA and Cronbach's alpha with fewer judges refer to this issue. Among the products, the agreement between the randomly selected judges of the electric-vehicles group were lowest compared to other product categories. The outcomes of EFA analysis revealed that expert judges in this category (electric-vehicles) were unable to perceive creativity and that there was no clear distinction between their ratings of creativity and technical goodness and/or aesthetic appeals of the products. The difficulty in characterising creativity and inability to psychologically separating it from other critical factors can be due to the complex structure of the internal components of the electric vehicles. One solution to overcome this problem is to first decompose the product into its constituent elements and then measure creativity of smaller sections. Under this condition, it can be a straightforward task for raters to perceive the influential attributes of creativity as the factors that causing confusion are eliminated. Having said that, another issue to be responsible for a lower reliability could be due to instructions given to the raters. This study followed the instructions suggested by Amabile (1982), however, the way the products presented to the judges might be also affected the reliability. Given the complex structure of the products in this category, it might be worthy (if possible) to allow judges to have physical access to the products in order to enhance consistency. Baer et al. (2004) assessed the creativity of writing samples that were written in response to variety of assignments using 13 domain experts. They concluded that CAT is a reliable tool for evaluating the creativity of artefacts that are created under uncontrolled conditions. In general, the outcome of this study is in line with the findings of other researchers and adds the fact that creativity measurement using CAT is reliable as long as it is evident for experts to perceive creativity within products. When the product in question contains complex structure (see p. 88), it becomes complicated to determine the performance characteristics of the product, and as result it becomes difficult even for experts to differentiate the border between creativity and other influential factors. Consequently, the measurement of creativity using CAT under its current condition is not recommended for complex functional products. Despite the complexity of products, the way the raters are instructed to judge creativity can also influence the results and may contribute to lowering the ratio of reliability. To overcome this problem, (complex) products can be decomposed into their constituent sub-systems where creativity of each section can be assessed

individually. The total creativity then becomes equal to the average value of creativity ratios obtained for all sub-systems.

The results of this study revealed that CAT can be now utilized for the creativity assessment of functional products created under uncontrolled conditions. However, caution should be taken with assessing products that are relatively complex (complexity means a product consists of too many components with a high degree of diversity. Also, the connectivity between the components is complicated (see page 88 for more information)). The outcomes of this study also allow researchers to investigate and determine the influential factors affecting the development of product creativity in various organizational environments. For instance, one can ask a group of designers to develop an artefact for the same assignment while defining a specific set of equipment and experimental conditions for each individual. Then, the effects of each environmental condition on the overall creativity performance can be examined. In addition, CAT can be used as a reference to validate and calibrate other measurement tools.

4.5. Chapter Conclusion

The consensual assessment technique is regarded as the "gold standard" of creativity measurement (Baer & McKool, 2009b). CAT has been widely utilised for identifying the creativity in various fields such as writing, poetry, art and music. In the consensual technique, typically a group of participants are initially asked to create an artefact using a series of identical equipment in a limited time interval. Then domain-relevant experts are invited to judge the creativity of those creations independently. However, the current arrangement of the CAT makes it inadequate for the assessment of products produced under different conditions using distinctive tools and resources. In this study, the creativity of three categories of functional consumer goods was evaluated using the recommended settings of the CAT except that products did not produce under tightly controlled conditions. On average 35 experts rated the products of each category for creativity. Also, the raters judged each product on nine further dimensions in addition to creativity to examine the discriminant validity of the measurement. Results of inter-judge reliabilities indicated that CAT is a reliable method for the assessment of artefacts created under uncontrolled conditions. The reliability ratios obtained in this study are in the same range as obtained in other artistic domains using artefacts produced under tightly controlled

environments. Considering that the number of judges impacts the reliability, the test of inter-judge reliability was replicated using smaller groups of randomly selected raters per product category. Results revealed that, in general, a larger panel of rater is required for the assessment of products produced under distinctive environments.

Exploratory factor analysis of ratings to 10 items for both home accessories and fire alarms group revealed that experts are able to psychologically separate creativity from other factors, implying that their creativity ratings are not due to their ratings of technical or aesthetic aspects of the product. However, the judges of the electric vehicles were unable to distinguish the differences between creativity and other factors. It appears that the complex structure (see page 88 for more information) of electric vehicles resulted in confusion in ratings of creativity and caused the judges to become psychologically incapable of recognising creativity as a single factor for this group of products. Also, the outcomes of linear regression analyses indicated that the significant predictors of creativity are varying from product to product which further suggests that creativity is a dynamic concept and limiting it into a set of constant variables is indeed in contrast with the main idea behind creativity.

5. Development of a Theoretical Framework for Evaluation of Functional Consumer Product Creativity

In the previous chapter, the ability of the consensual assessment technique for assessing the creativity of non-parallel functional consumer products was investigated. Initial research outcomes indicated that CAT is a reliable technique for measurement of such products; however, further factor analysis revealed that even expert judges cannot psychologically separate creativity from other technical or aesthetic aspects for complex products. In addition, the use of CAT is typically time-consuming and operationally expensive (Cropley et al., 2011). Also, CAT does not provide useful information to support designers to better understand how creativity is perceived by observers. From this viewpoint, criterion-based instruments have advantage over consensual technique. These methods assess creativity using relevant indicators and hence are able to provide notable insights to aid product designers to maximise creativity. Criterion-based approaches allow novice raters to judge creativity, making the overall measurement quicker than the consensual technique.

Review of existing approaches for assessing creativity revealed that creativity is a multidimensional variable and is interconnected to a broad range attributes (see chapter 3, section 3.4). Creativity is a dynamic concept because it must be seen, felt, and understood by appropriate observers in a particular time period rather than being directly measured in the laboratory. In this context, Rubera, Ordanini, and Griffith (2011) state that "… creativity is not an objective property of the product, but rather is determined by the interaction of three elements: the product, the perceiver (e.g., the consumer), and the context"; where *context* reflects other

domain products in a particular time period (Amabile et al., 1996). Therefore, contrary to other criterion-based techniques, the goal of this research is to fill the gap in knowledge by proposing a theoretical framework to evaluate product creativity using novice raters. The new model does not aim to suggest a single construct for evaluating creativity, but rather presents a general platform containing the most common and internally consistent creativity indicators. This model supports practitioners and researchers to develop their own unique and reliable assessment instrument for the given product by employing the indicators proposed in this research. In order to enhance assessment accuracy as well as to allow measurements to be performed using more specific dimensions, the scope of measurement is reduced from all products to mostly (but not limited to) functional consumer products in the shopping goods category.

5.1. Scope of measurement

Consumer goods are among the well-known categories of products. The SAGE encyclopaedia of business ethics and society describes the consumer good as: "… any product or service purchased primarily for personal, family, or household uses. Consumer goods such as clothing, foodstuffs, or toys are intended to satisfy human wants and needs through their direct consumption or use. Industrial goods, in contrast are purchased by individuals or organizations to produce other products and services that are sold to, provided for, other individuals or organisations. According to their usage, many goods (e.g., cars, printers, or personal computers) can be categorised either as consumer goods or as industrial goods."(Grabner-Kräuter, 2018, p. 2).

Typically, there is no universal classification for consumer goods, but rather they are categorised based on the application that is in question. In marketing sector, consumer goods are classified based on the amount of effort customers expended on searching and selecting the product; including (1) convenience goods, (2) shopping goods, and (3) speciality goods (Copeland, 1924). A convenience good is a product that is accessible to a wide extent and generally purchased regularly with a minimum amount of effort. Considering that convenience goods are needed frequently, the process of selecting them usually do not involve a notable decision-making procedure and/or risk. Examples of products in this category are groceries and office supplies.

The process of selecting and purchasing a shopping good, on the other hand, involves spending considerable amounts of time and money for deciding between available products by comparing their attributes, quality, price, and style (American Marketing Association, 1948). Contrary to convenience goods, shopping goods are not purchased routinely and the process of selecting them includes significant levels of risk. As a result, consumers spend time and effort for collecting information required for the evaluation. Furniture, clothing, and used-cars are examples of products in this class. Speciality goods are described as items or services in which consumers select or purchase them because of their distinctive features or brand identification. Customers of this category usually do not search for other alternatives and instead willing to spend the highest amounts of time and effort on finding the product that is best fitting their specific demand. The process of selecting and purchasing speciality goods comprises higher amounts of risk compared to the other two classes of goods. Some of the products in this category are premium cars and clothes.

Among the three classes of goods, creativity is most influential to products of the shopping goods category. Customers of this class consider the product's attributes including novelty and usefulness more than any other customer. Whereas, in convenience goods, consumers usually pay less attention to the levels of novelty and mainly select the product because of its usefulness. Likewise, customers of speciality goods are usually choosing the right product either because of the brand or their previous experience with the product and sometimes because of their specific demand. Creativity can give the products in the shopping goods category the ability to not only differentiate themselves from the contrary, but also to better satisfy consumers which further increases the likelihood of being successful in the market.

As a result, in this research the scope of measurement is reduced to functional consumer products (i.e., products that serve some practical function) that are classified in the shopping goods category. Limiting measurement scope allows to select more relevant creativity dimensions which is required for enhancing the assessment accuracy.

5.2. Proposal of a new theoretical model for the evaluation of functional consumer product creativity

Product creativity is associated with novelty and usefulness (Boden, 2003), and thus assessments in this context must be in line with this definition. This only applies if the evaluation is intended to be done based on pre-identified characteristics, however, if the evaluation is meant to be done without previous knowledge about creativity then this rule does not apply. In this regard, the assessment of creativity using the consensual assessment technique heavily relies on the judgements of a group of appropriate judges and it does not consider any definition for creativity. Overall, the process of breaking creativity down into its constituent elements is a complex task because parameters that represent usefulness (effectiveness) relate extensively to product functionality. When functionality changes, all variables that define usefulness in the product also change. This issue was evident in the regression analysis between creativity and nine dimensions for three categories of products in the previous study (see chapter 4, section 4.4). Results indicated that the significant predictors of creativity are varying from product to product as their functionality change. As a result, the idea of assessing creativity using constant and general variables might not be useful. In contrast, the process of searching and specifying the precise creativity indicators for a particular product can be extremely time-consuming.

This research aims to propose a framework (platform) for evaluating the creativity of functional consumer products especially those in the shopping goods category. The factors proposed in this instrument do not suggest the actual factorial model for all products in the universe. They are instead exemplifying an ideal factorial construct for creativity assessment. However, in the real-world scenario, some factors may be combined while some others may be perceived as irrelevant and hence eliminated. This model aims to support researchers by presenting the most common and relevant creativity indicators for functional consumer products. Using this instrument, the need for searching and selecting creativity indicators for designing new instruments for specific products is eliminated. This model enables practitioners to quickly perform factor-analysis for the desired product and as a result understand how consumers perceive creativity for that product. For example, in order to find out how computer mouse's consumers are recognising

creativity, a factor analysis using indicators introduced in this model can be employed and hence a unique construct can be generated that is explicitly dedicated for computer mouse.

This theoretical model contains a set of nine priori dimensions for measurement of product creativity including *novelty*, *performance*, *safety*, *human-factors*, reliability, *sustainability*, *attractiveness*, *emotion*, and *relevance* (see Fig. 5-1). To follow the rules of the psychometric tests, each criterion must be further explained with several items (indicators) to allow judges to have a common understanding of the selected dimension (criterion) with respect to the given situation (J. Anderson & Gerbing, 1988). As a result, a pool of items containing on average 7 indicators per each dimension (65 in total) was generated (see Table 5-1). The use of multiple items is to explain each dimension from different viewpoints as well as to enhance the overall (inter-item) consistency of the constructs (Nunnally & Bernstein, 1994). In general, researchers do not recommend the use of a single item as the analysis of such a construct does not provide useful information about the psychometric properties of the measurement (Netemeyer, Bearden, & Sharma, 2003).

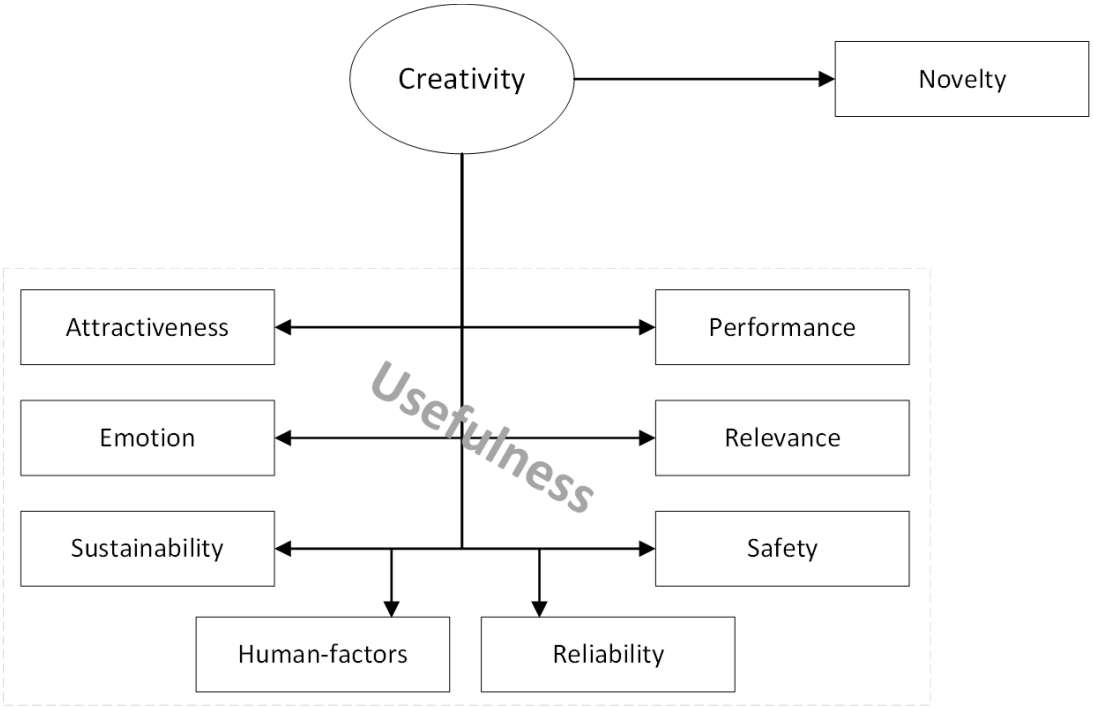


Figure 5-1 New framework for measuring the creativity of functional consumer products

In creativity research literature, novelty is the crucial factor but not sufficient (Cropley et al., 2011; Maher, 2010). The existence of novelty in creative products is a must, meaning that even the most practical and/or aesthetically efficient products cannot be deemed creative unless they comprise new and original features. However, the novelty on its own does not describe creativity. The essential factor for any consumer is the ability of the product in completing the task that it is supposed to do effectively. For instance, a computer monitor aims to display a bright and high-resolution pictorial information to the user. For any reason that it fails to fulfil the user's demand, it implies that the product lacks usefulness and hence cannot be creative regardless of how many new features have already been implanted in the product. On the other hand, usefulness defined as the extent to which a product is able to successfully and efficiently perform its intended task. The eight dimensions that are selected to describe product usefulness include *performance* (product is capable, efficient, and flexible), *relevance* (solution is relevant to task and important to the user), *safety* (product is safe and secure to use), *reliability* (product is trustworthy and durable), *human-factors* (product is ergonomically comfortable and intuitive to use), *sustainability* (product is environmentally friendly), *attractiveness* (product is beautiful, elegant, and proportional), and *emotion* (product is desirable and pleasant to use). The literature of the relationships between each dimension and creativity is reviewed in the next section.

Creativity and Novelty

Novelty gives the product the ability to differentiate itself from other competing products (Rodríguez, Perez, & Gutierrez, 2007), and hence is crucial from the consumer point of view. Previous investigations revealed that customers are interested in selecting products that comprise new and original features in addition to having adequate levels of usefulness compared to those without (Danneels & Kleinschmidt, 2001; Sethi & Sethi, 2009). In the majority of explorations about creativity, novelty is referred to as a necessary quality for creativity (Amabile, 1982; Csikszentmihalyi & Wolfe, 2000; Runco, 2007). Novelty is defined as the extent to which the product includes unique and original characteristics. A novel product must not be similar or equivalent to any previously existed artefact (Robert Sternberg & Lubart, 1999).

In the relevant research literature, the relationship between novelty and creativity has been widely investigated and result always verified the positive association between the two variables (Besemer & O'Quin, 1999). Almost all criterion-based models agree that the measurement of novelty to some extent correlates with the assessment of creativity. Cropley and Cropley (2005) introduced an instrument to measure product creativity consisting of four dimensions including *novelty*, *relevance and effectiveness*, *elegance*, and *genesis*. In another research study, Goldenberg and Mazursky (Goldenberg & Mazursky, 2002) discussed the importance of product novelty and referred to it as an essential characteristic for the creative product. Other analysts evaluated the creativity of ideas in engineering design domain using a series of dimensions including novelty factor (Oman & Tumer, 2009). In Horn and Salvendy's (2009) three-dimensional (creativity assessment) instrument, creativity is measured by *novelty*, *affect*, and *importance* factors. Besemer and O'Quin's Creative Product Semantic Scale (1986) has defined creativity using factors including *novelty*, *resolution*, and *style*. Taylor (1975) presented the Creative Product Inventory consisting of dimensions such as *originality*, *relevancy*, and *complexity*. A generic product's novelty has been measured with several indicators such as *original*, *unusual*, *unique*, *infrequent*, and *revolutionary* (Besemer & O'Quin, 1987; Cropley, 2016; Horn & Salvendy, 2009).

Creativity and product effectiveness

Despite the necessity of novelty in creativity, the ability of the product to effectively meet user's expectations is also important from the consumer perspective. Researchers believe that a novel, but unfitting product has a higher chance of being detected as an unreliable and meaningless by the consumer (Ford, 1996). In fact, usefulness forms the basis of creativity, and almost all creativity definitions emphasize the need for appropriateness in product. Rothenberg and Hausman (1976) described creativity with novelty and value. Bruner (1962) characterised creativity as "an act that produces effective surprise". Other researchers link creativity to originality and appropriateness while interpreting it (creativity) as an original response that is also capable of solving a problem as well as fitting a situation (D. MacKinnon, 1965).

Product usefulness (appropriateness) includes all characteristics that enable the product to successfully fulfil user's demands. In this regard, Cagan and Vogel (2002) proposed a classification system for describing product value comprise of seven components including Emotion (the ability of the product to provide freedom, excitement, luxuriousness, feeling of safety and authority), Ergonomics (ease of use and comfort), Aesthetics (the ability of the product to stimulate users through either visual, tactile, auditory, olfactory, and gustatory senses), Identity (the ability of the product to differentiate itself from the others), Impact (social and environmental impacts), Core-Technology (sufficient features and reliability), and Quality (craftsmanship). In creativity research, product usefulness is measured as part of creativity assessment. Besemer and O'Quin (1986) developed and validated the Creative Product Semantic Scale consisting of three dimensions (novelty, resolution, and elaboration and synthesis) and fifty-five pairs of semantic items. In this model, for example, *resolution* is further described using indicators such as *value* (product is important for the solution), *logic* (product is adequate and relevant to the solution), and *usefulness* (product is functional and operable). MacKinnon (1968) presented a five-factor model for assessing product creativity including following factors: *originality*, *adaptiveness*, *elegance* (aesthetic quality), *transcendence*, and *realisation*. In Cropley and Kaufman's Creative Solution Diagnosis Scale (2012), creativity is evaluated based on four elements comprise of *relevance and effectiveness*, *novelty*, *elegance*, and *genesis* (also see chapter 3, section 3.4).

i. Product performance

One of the areas influencing (product) effectiveness is the level of technical performance and technology that the product provides. Technology on its own does not meet consumer expectations; however, it contributes to overall performance, consistency, and reliability of the product in the long term. Technology also offers the product the opportunity to differentiate itself from the other products that are missing this service. Another feature that adds value to the product is the ability of product in adapting (or adjusting) itself with the given condition or working environment. For instance, a sports car is more capable on a racetrack than a sport-utility vehicle (SUV), while an SUV is more useful for off-road conditions. However, any vehicle capable of adapting itself to both road conditions is perceived as more valuable to consumer than the former two vehicles. In the literature of criterion-based techniques, many scholars have

either directly or indirectly pointed out the importance of having an adequate level of technical performance for creative products and included in their instrument indicators that resemble product's technical performance characteristics. For instance, Besemer and O'Quin (1999) have measured product performance with *resolution* factor containing items such as *functional* and *effective*. Cropley et al. (2012) assessed product effectiveness using *performance* indicator.

ii. Product safety

Product safety is an essential part of product development process, and in general is referred to as a mandatory requirement by laws, regulations, and/or standards. The EU general product safety directive (EC & EP, 2002) defined a safe product as "any product which, under normal or reasonably foreseeable conditions of use including duration and, where applicable, putting into service, installation and maintenance requirements, does not present any risk or only the minimum risk compatible with the product's use ...". Similarly, the US Air Force (2000) described product safety as "the application of engineering and management principles, criteria, and techniques to achieve acceptable hazard risk, within the constraints of operational effectiveness, time, and cost, throughout all phases of the system life cycle". In creativity measurement, Cropley et al. (2011) and Cropley and Kaufman (2012) have included the measurement of product safety in the creativity assessment of functional creative products.

iii. Human-factors (ergonomics)

Another element that contributes to user satisfaction and efficiency of use is the extent to which the product is comfortable, straightforward, and easy to control (Karwowski, Soares, & Stanton, 2011). In Brooke's System Usability Scale (SUS), product's usability is characterised by statements such as "I think that I would like to use this product frequently", "I felt very confident using the product", and "I found that the various functions in this product were well integrated" (Bangor, Kortum, & Miller, 2008). The process of studying and understanding the interactions between human and different components of a system is referred to as human-factors (Helander, 2006). In creativity literature, researchers have included the evaluation of product ergonomic specifications in creativity assessment. In Besemer and O'Quin's CPSS (1999) the *elaboration and synthesis* factor consist of human-factors indicators such as *meaningfulness*, *understandability*,

simpleness, and *straightforward*. Other researchers who also considered measures of product's ergonomics as an indicative for creativity are (Cropley et al., 2011).

iv. Sustainability

Sustainability contains three main elements including Environment, economy, and society (Gupta, Dangayach, & Singh, 2015). Sustainable products are the ones that are able to offer benefits at all three areas during their life cycle. Environmental sustainability encourages products to consume natural resources at a rate that they can be replaced by themselves. Social sustainability includes adhering to human rights and improving life quality, health and well-being along with reducing the ecological impacts. A sustainable product must also provide economic advantages to consumers by cutting the unnecessary costs arise during different stages of the product's life cycle, from the selection of raw material to consumption and recycling. Recent studies show that customers of companies that produce and promote green products are far more satisfied compared to the other businesses (Yazdanifard & Mercy, 2011). Consumers demands have tended towards an eco-friendlier lifestyle and they tend to be willing to spend their resources for a sustainable product. In creativity measurement, Cropley et al. (2011) and Cropley and Kaufman (2012) have included the evaluation of product's sustainability in the creativity assessment of functional creative products.

v. Reliability

Product reliability is defined as "the time period over which a product meets the standards of quality for the period of expected use" (Levin & Kalal, 2003). Reliability is essential because an unreliable product can significantly reduce customer's interest in using the product. Researchers believe reliability is positively affecting current customers degree of satisfaction (Baziuk, 1995). An adequate level of reliability makes it possible for the consumer to rely on product and hence use it more frequently and perceive it as more effective. In line with this, Harold et al. (2009) emphasised the necessity of maintainability and reliability for enhancing product effectiveness. In a study by the American Society for Quality Control (Quality Progress, vol. 18 (Nov), pp. 12-17, 1986), the result of 1000 interviews revealed similar outcomes. Their investigations inserted product reliability and maintainability in the top five most important factors to the customer. In

the extant literature, Cropley et al. (2011) have considered a form of reliability (i.e., durability) in the creativity measurement of functional products.

vi. Aesthetics (attractiveness and emotion)

Aesthetic is among the crucial elements affecting user experience especially when other domain products lack this quality. Aesthetics include the ability of the product in stimulating the user's sensory perception either visually, tactility, auditory, olfactory, taste, and/or emotion. From the creativity point of view, the attractiveness of the product impacts viewers' perception of creativity. Attractiveness is simply not just around beauty and style; it rather includes the extent to which product is skilful, precise, and proficient in performing the task (elegance). Often researchers measure product creativity with elegance factor (Cropley et al., 2011). Gelernter (1998) characterises *elegance* as the confluence of aesthetics and usability. In other words, above and beyond the importance of performance characteristics, it is essential for the product to have a delicacy in action and this must be perceived by the users. Other researchers defined and measured the *elegant* with three factors including *functionality*, *aesthetics* and *simplicity* (Nikou, 2017). In creativity literature, Horn et al. (Horn & Salvendy, 2006b) examined product creativity considering consumer's perspective. They indicated that attractiveness and emotional impact on the user also contribute to product creativity. Other researchers who have emphasised the need for aesthetic qualities in creative products are (Besemer & O'Quin, 1986, 1999; Christiaans, 2002; Cropley et al., 2011).

The amount of emotional impact that a product can cause on the user is linked to product's performance and aesthetic characteristics. Emotions emerge from experiences that users acquire from owning and/or using the product in different circumstances. In other words, a part of consumer satisfaction is inspired by the degree to which the consumer perceives the product as technically capable and aesthetically efficient attractive and pleasing compared to other domain products. Horn et al. (Horn & Salvendy, 2006b) have reported that almost 10% of the total variance accounted for product creativity is due to the Emotion factor. In creativity research studies, product's attractiveness has been assessed with indicators such as *elegance*, *gracefulness*, *attractive*, and *refined* (Cropley et al., 2011; O'Quin & Besemer, 1989), whereas

product's ability in emotionally stimulating the observers is described with terms including *pleasingness*, *convincingness*, and *delightfulness* (Cropley et al., 2011; Horn & Salvendy, 2006b).

vii. Product importance (Relevance)

Another aspect affecting product creativity is related to the degree to which the product appears as essential and relevant to the observer. For instance, if the judgments of two individuals on technical and aesthetic characteristics of a product seem to be similar while their overall creativity evaluation is different, then it implies that one rater did not recognise the solution as crucial and necessary for the given problem. In creativity research literature, several authors have included measures of importance as one of the indicatives of product creativity. The *resolution* factor in Besemer and O'Quin's CPSS assesses product's ability in successfully performing what it is supposed to do. This factor consists of items such as *relevant*, *appropriate*, and *logical*. Cropley et al. (2011) has denoted that product creativity is dependent on *relevance and effectiveness* factor. Horn et al. (Horn & Salvendy, 2006b, 2009) measured product creativity using a three dimensional instrument consisting of *affect*, *importance*, and *novelty* factors. In creativity literature, researchers often referred to *importance* factor with indicators such as *crucial*, *ideal*, *logical*, *relevant*, *appropriate*, and *making sense* (Besemer & O'Quin, 1986; Horn & Salvendy, 2006b; Takai, Esterman, & Midha, 2015).

As this framework was only a concept, a case was designed to test and determine the reliability of creativity measurement and confirm the validity of the constructs. The following section describes the research methodology and research concerns for conducting an appropriate factor analysis.

Table 5-1 The framework for evaluating creativity of functional consumer products (factor patterns and creativity indicators)

Factors	No.	Negative	Positive	Factors	No.	Negative	Positive	Factors	No.	Negative	Positive
Novelty	1	Conventional	Original	HFs	5	Inoperable	Operable	Attrac.	9	Imbalance	Balance
	2	Usual	Unusual		6	Unergonomic	Ergonomic	Emotion	1	Disgusting	Desirable
	3	Ordinary	Unique		7	Intolerable	Tolerable		2	Dissatisfying	Satisfying
	4	Average	Revolutionary		8	Uncomfortable	Comfortable		3	Unpleasant	Pleasant
	5	Frequent	Infrequent	Reliability	1	Flimsy	Durable		4	Contempt	Admirable
	6	Expected	Surprising		2	Weak	Strong		5	Boring	Exciting
Performance	1	Non-functional	Functional		3	Unsound	Sound	Relevance	1	Irrelevant	Relevant
	2	Futile	Practical		4	Fragile	Solid		2	Unnecessary	Necessary
	3	Inefficient	Efficient		5	Frail	Robust		3	Inessential	Essential
	4	Incapable	Capable		6	Unreliable	Reliable		4	Inappropriate	Appropriate
	5	Ineffective	Effective		7	Questionable	Trustworthy		5	Unreasonable	Reasonable
	6	Unchanging	Adaptable		8	Poorly built	Well-built		6	Worthless	Valuable
	7	Inflexible	Flexible	Sustainability	1	Unsustainable	Sustainable		7	Incompatible	Compatible
	8	Fixed	Modifiable		2	Synthetic	Organic	Total	65		
	9	Immutable	Alterable		3	Non-recyclable	Recyclable				
	10	Static	Changeable		4	Irrecoverable	Renewable				
Safety	1	Does not protect	Protects		5	Environmentally harmful	Environmentally friendly				
	2	Unsafe	Safe		6	Pollutive	Pollution free				
	3	Hazardous	Non-hazardous	Attractiveness	1	Repulsive	Appealing				
	4	Insecure	Secure		2	Repelling	Charming				
	5	Harmful	Harmless		3	Coarse	Elegant				
	6	Painful	Innocuous		4	Crude	Refined				
Human-factors (HFs)	1	Difficult	Easy to use		5	Messy	Neat				
	2	Complicated	Simple		6	Disordered	Arranged				
	3	Ambiguous	Comprehensible		7	Disorganised	Organised				
	4	Complex	Straightforward		8	Poorly proportioned	Well-proportioned				

5.3. Research Methodology

In this thesis, a quantitative research method was designed to test the reliability and validity of creativity measurement using the proposed instrument, novice raters, and a specific category of consumer products as case study (lounge chairs). This section outlines the methodology used for data collection, piloting, and analysing the data in order to establish the research direction in the next chapter. This section only covers the basics and prerequisites require for conducting an appropriate research. All information about participants, material (products), and judging procedure are explained in the next chapter where the main study is reported.

Survey design

To collect novice raters's opinion on 65 items per each product, a web-based survey was considered. The rationale for selecting this type of survey was due to the benefits it offers over other methods (e.g., paper surveys and interviews). Theuri and Turner (Theuri & Turner, 2002) outlined some of the advantages of web-based surveys over other approaches as follows:

Unlike paper surveys, a web-based survey does not involve printing costs and contrary to interview it does not involve labour costs. In addition, web-based surveys may positively affect the number of responses as participants are usually given the freedom to choose the right time, place, and/or mood to answer the questions and hence they are more useful for larger sample sizes which is the case in this research. Web-based surveys also enable the researcher to operate complex surveys while reducing the bias by using integrated services available in most survey software/providers. The data collection process for web-based surveys usually takes a couple of days (or hours), whereas in other methods this may take up to weeks and/or months. The use of web-based surveys removes the geographical restrictions and make it possible to collect information from participants all around the world as long as they access the internet.

For establishing the survey, four main sections are considered, including: consent form, demographic questions, presenting the products, and main survey questions (65 items). A consent form was written at the beginning of the survey to explain the purpose of the research as well as to notify participants around several issues including their right to withdraw at any time, confidentiality of their responses, amount of risk involved in the research, and that they

are free to whether or not to take part in the survey. Demographic questions were included in the survey to gather certain background information about the respondents (Sue & Ritter, 2007). The information attained in this section used to describe the participants characteristics (e.g., age, gender, and ethnicity distribution) as well as to support the researcher in targeting a specific population for the study (e.g., the profile of the respondents on 'job title' question was filtered to only consider students). Prior to rating the products, participants were presented with all products including all specifications and appearance of the products and were instructed to observe and compare the products to one another. This was required for an unbiased data and a fair comparison between the products.

The final part of the survey contains the judgments of novice raters on the 65 creativity indicators proposed in this study. The Likert-type scaling system was used to measure each variable; where all responses are considered as continuous variables to allow application of interval-based techniques for analysis. Likert-type data are by nature ordinal variables (i.e., ordered categorical data, meaning that the distance between two points of measurement may not be equal with the distance between another pair of units). Nevertheless, in creativity research literature, it is widely accepted to treat creativity and related items as continuous variables (Amabile, 1996; Piffer, 2012). In addition, several studies support the use of parametric statistics with Likert type data (Carifio & Perla, 2007; Norman, 2010). In this context, Subedi (Subedi, 2016) argued that non-parametric statistics are useful for analysing Likert items; although parametric statistics can be applied for Likert scales (sum or mean of a number of Likert items).

In this survey, variables were measured on a 7-point scaling system to better indicate consumer's perception of each criteria.

Sample size

In the relevant research literature, there is no simple method to decide on the sample size required for factor analysis. As a result, researchers suggested a number of guidelines for the minimum samples size needed to achieve a stable component patterns in the long term. In general, authors' suggestions on minimum sample sizes range from 100 to 1000 cases or over, with 100 indicating a poor model and 1000 or over showing an excellent model for factor analysis

(AL Comrey & Lee, 1973). Kline (1979) recommended the use of no less than 100 cases, whereas Hutcheson and Sofroniou (1999) recommended a sample size between 150 to 300 cases. Nevertheless, other analysts argued that the process of making decisions regarding sample sizes should not only include a general guideline but instead, researchers must consider the ratio of sample size (*N*) to the number of measured variables (*P*). In this context, Hair, Anderson, Tatham, and Black (1995) recommends a ratio of 20:1 (i.e., at least 20 observations per each variable in the model), whereas in a less controversial approach, another group of researchers advised the ratio of 10:1 (Bryant & Yarnold, 1995) or 5:1 (Everitt, 1975) case for each item.

Meanwhile, another group of authors believe that the use of rule thumbs does not always yield the right data, and hence suggested that each case must be separately assessed before being able to show whether the data is suitable for factor analysis. Some of parameters to be assessed after data collection are communality of variables, factor loadings, and degree of number of factors to number of variables (Preacher & MacCallum, 2002; Velicer & Fava, 1998). For instance, Tabachnick and Fidell (2007) and Hair et al. (2010) suggested that the communality of the variables must be higher than 0.3 and 0.5 respectively. Factor loadings must be large enough and there should be at least 3 variables for each factor.

Another widely reported measure of sample adequacy is the Kaiser-Meyer-Olkin (KMO) test. This statistic ranges from 0 to 1 with 0 indicating a lack of appropriateness and 1 showing an ideal dataset for factor analysis. The KMO shows the extent to which the observed data is suitable for factor analysis by measuring the proportion of variance in the dataset that might arise because of latent factors. Researchers recommend the KMO value to be greater than 0.6 for ideal factor analysis (Hair et al., 2010). The Bartlett's test of sphericity is another method to check the viability of the data for factor analysis. The statistic compares the correlation matrix for the observed data with the identity matrix (a matrix at which items are assumed as uncorrelated) to show whether there is any correlation among the measured variables that can be described using fewer factors. For this statistic to show adequate data, the chi-square value must be significant at ($p < 0.05$) (Hair et al., 2010).

Pretesting the survey

The present research consists of two phases, a small pilot study and a main study using larger population. Considering that this is the first study to be performed using the proposed indicators, it was necessary to assess whether the survey instructions as well as the questions (65 items) are indeed able to collect reliable data for the main study. Therefore, a pilot study was conducted with the following objectives:

1. To measure the internal consistency (inter-item reliability) among the items constructing a latent factor (e.g., to assess the consistency between indicators of novelty factor such as originality, unusualness, and uniqueness).
2. To examine the item wording and to ensure items are easy to understand, unambiguous, and contain no technical word/phrase.
3. To confirm the clarity of instructions, forms and layout of the survey.
4. To assess the efficiency of the administration procedure.

Note: All information regarding participants, material, and task procedure are explained in the next chapter where the main study is reported.

5.4. Quantitative data analysis

In this thesis, a quantitative research method was designed to test the relationships between the hypothesised creativity indicators on a cluster of consumer products selected as case study. The following section explains the quantitative methods considered for data analysing. The IBM-SPSS (version 24) and AMOS software were used to perform all statistical computations.

Reliability

The first step after development of the scale is to assess whether all subscales in the construct are measuring the same factor (Cronbach, 1951). In other words, a high consistency between items in a construct assures that the judgment of an individual on these items would be in the similar range. In creativity research, internal consistency is referred to as inter-item reliability and is mostly calculated by Cronbach's coefficient alpha (Besemer & O'Quin, 1999; Cropley et al., 2011) (see equation 5.1). Alpha was established by Lee Cronbach in 1951, and it is indicated by a

number in the 0-1 range. In general, the higher the ratio of alpha (α) the lower the error variance in the test scores. For instance, if the reliability of a construct is measured at 0.9, this means 0.19 of the test score is due to random error (i.e., 0.9×0.9=0.81; 1−0.81=0.19) (P. Kline, 1994). Researchers usually report an excellent consistency if the (α) coefficient is equal or greater than 0.9 (Bland & Altman, 1997; Nunnally & Bernstein, 1994). The alpha of 0.8 ≤ α < 0.9 illustrates a 'good' level of internal consistency whereas any value between 0.7 ≤ α < 0.8 shows an 'acceptable' inter-item reliability across the items in a test. A coefficient alpha below 0.7 often interpreted as unacceptable consistency or indicates the presence of systematic errors in the scale.

In addition to measuring internal consistency among the item in the test, it is also recommended to assess the degree of agreement among the raters. In creativity research, an adequate ratio of inter-rater (inter-judge) reliability is referred to as the validity of assessment. Many scholars have employed the alpha in analysing the ratios of inter-judge reliability by transposing the items with raters in the data matrix (Cropley et al., 2011; Kaufman, Baer, et al., 2008).

$$\alpha = \frac{N}{N-1}\left(1 - \frac{\sum Var(X_i)}{Var(S)}\right)$$ Equation (5.1) (Cappelleri et al., 2014)

Where N is number of items, $Var(X_i)$ is the variance of *i*th item, and $Var(S)$ is the total variance.

Exploratory Factor Analysis (EFA)

Exploratory factor analysis intends to determine the optimum underlying relationships between variables to which the original observation can be described with a minimum number of latent factors while presenting the same concept together with the maximum portion of the variance (DeCoster, 1998). In the current study, EFA was utilised to determine how raters perceived creativity as well as to identify the number of latent factors and their underlying factor patterns.

In the current study, exploratory factor analysis was conducted on the mean ratings of 65 dimensions of judgment. To extract the factor loadings, Principal Axis Factoring (PAF) was used to numerically estimate the variable's communalities (i.e. the variance of variable accounted by the common factors plus residual) and generate a reduced correlation matrix to compute

loadings that best represent the original observation. Afterwards, loadings were rotated under oblique rotation algorithm to indicate the best possible factor structure. The oblimin rotation system was selected over orthogonal approach for its assumption that correlations exist between the factors. A cut-off point of 0.4 was used as a criterion to accept loadings on to one factor (Stevens, 1992).

Confirmatory Factor Analysis (CFA)

Confirmatory Factor Analysis is a statistical technique to test and confirm the factor structure proposed either from theory or an empirical study (e.g., EFA). CFA provides useful information with which researchers are able to decide whether they should accept or reject the hypothesised factor pattern. Contrary to exploratory factor analysis, CFA is driven by theory implying that every attribute of the observed dataset (proposed model) must be specified before the analysis. Given that the theory exists, CFA then attempts to reproduce (estimate) a population covariance matrix to be used for comparison purposes with the observed covariance matrix (Prudon, 2015). In other words, CFA employs appropriate estimation algorithms (e.g., Maximum Likelihood Estimation) to numerically predict various parameters in the model, including primary factor loadings, factor variances and covariances, as well as error variances. These parameters indeed form a factor structure that is statistically adjusted to the given sample matrix (observed data), maximising high factor loadings while minimising low cross-loadings (Prudon, 2015). Then, CFA compares the implied covariance matrix with the observed covariance matrix and drives the residual covariance matrix from the differences. These distinctions build the basis for measuring the goodness-of-fit of the predicted construct.

In creativity studies, researchers often employed confirmatory factor analysis to investigate the construct validity (including convergent and discriminant validity) of the hypothesised factor patterns (Besemer & O'Quin, 1999; Cropley & Kaufman, 2012). Convergent validity includes testing the relationships between different measures in a single construct and illustrate how well the observed measures (items) are able to explain the latent factor. Fornell and Larcker (1981) stated that a construct has convergent validity when the Average Variance Extracted (AVE) value is greater than 0.5; in other words, the average factor loadings for all items in that construct is

not less than 0.7. In this context, Hair et al. (2010) also suggested the same criterion for AVE, and added that factor loadings for the items in a construct must not be less than 0.5, while composite reliability (CR) ratio should be greater than 0.7. On the other hand, discriminant validity denotes the degree of distinction (unrelatedness) between several constructs in a single instrument (Farrell, 2010). According to the criteria suggested by Fornell and Larcker (1981), to establish discriminant validity between the various constructs in a measurement model, the AVE for every two latent factors must be larger than the square correlation (shared variance) between the two factors.

In addition to these, CFA also utilises several model fit indices to show how well the predicted model corresponds with the observed data (or theory). These measures provide the level of goodness-of-fit for the implied model; in other words, the extent to which the model is able to reproduce the sample data. Next section (model fit) clarifies the model fit indexes used in this thesis while gives information about their cut-off value for a good fit. The fit indices considered in this thesis are Chi-Square statistic (X^2), RMSEA, GFI, CFI, TLI, and NFI.

Model fit

In confirmatory factor analysis, the adequacy of model fit is assessed with several statistical tests. One of the widely reported measures of goodness-of-fit is the chi-square test (X^2). The chi-square test examines the degree of distinction between the distribution found in the observed data and the distribution that is extracted from the implied covariance matrix (Prudon, 2015). The null-hypothesis in this chi-square test, contrary to the conventional chi-square tests, is that the two covariance matrices indeed represent the same distribution (or they are very similar). As a result, researchers expect a non-significant chi-square value for a good model fit (i.e., $p > 0.05$). However, the chi-square test requires the data to come from a normal distribution, which make the use of this test inappropriate for any dataset that is not normally distributed. In addition, X^2 test is sensitive to large sample sizes, meaning that the X^2 value may yield a significant value even for small differences between the distributions (Gatignon, 2010). As a result, scholars suggested the chi-square to the degree of freedom ratio (CMIN/DF) to overcome this problem (Wheaton, Muthen, Alwin, & Summers, 1977). In literature, researchers believe for a good model

fit the ratio of CMIN/DF should not go beyond 3 (Hair et al., 2010). However, again similar to the chi-square test, the CMIN/DF ratio is also sensitive to sample size and hence it may incorrectly suggest a poor fit when the implied model actually fits the data. Therefore, other researchers have proposed the use of different fit indices along with chi-square and CMIN/DF tests to assess the goodness-of-fit. In general, all other fit indices belong to one the two classes of fit: tests of absolute fit or tests of incremental fit (Widaman & Thompson, 2003). The tests of absolute fit include the use of chi-square test to compare the implied covariance matrix with the sample data just similar to chi-square test; however, they often consider several different parameters in their measurement to eliminate the sensitivity to sample size, such as degree of freedom and/or number of indicators in the model. On the other hand, incremental indices do not rely on chi-square test only and include in their measurement a baseline model (null model) to compare it with the X^2 value. In contrast to tests of absolute fit, incremental tests consider a third matrix (null model) in addition to the former implied and observed covariance matrices to allow an assessment that is less sensitive to sample size. Among the fit indices used in this thesis for assessing model fit, Steiger's (1990) root mean square error of approximation (RMSEA) as well as Jöreskog and Sörbom's (1984) goodness of fit index (GFI) are tests of absolute fit while the rest (i.e., NFI, TLI, and CFI) belong to the tests of incremental fit. Overall, a good model fit requires the value of each fit index to be acquired in the acceptable threshold suggested in the literature. Table 5-2 indicates mostly reported model fit indices and their cut-off value for a good model fit.

RMSEA measures the differences between the implied covariance matrix and the observed data but considers the number of estimated parameters (degree of freedom) in the model. For a good model fit the value of the RMSEA index must be obtained below or equal to 0.08 (MacCallum, Browne, & Sugawara, 1996), while some other researchers suggesting a cut-off value below or equal to 0.06 as threshold a good model fit (Hu & Bentler, 1999). It is worth noting that a RMSEA value between 0.08 and 0.10 suggests neither a good nor bad fit. The RMSEA is one the most popular fit indices because of its least dependency to sample size.

The goodness of fit index (GFI) was established to eliminate concerns regarding the use of chi-square test (Jöreskog & Sörbom, 1984). This index has a range between 0 to 1 where zero indicates no model fit and one denotes a perfect fit. Researchers suggest the cut-off point of 0.90

for a good fit and recommended a higher value when factor loadings and sample sizes are low (Miles & Shevlin, 1998). However, just like other indices the GFI has its limitation. This index has a tendency for a low value when the model contains a large number of degree of freedoms compared to sample size (Sharma, Mukherjee, Kumar, & Dillon, 2005). In contrast, the GFI rises when the number of measured variables increases (MacCallum & Hong, 1997). Considering the limitations of the GFI, it is not advised to rely solely on the value of this index for measure of model fit.

Normed fit index (NFI) measures the distinctions between the chi-square value obtained from the implied covariance matrix and the X^2 value from the null model. In this statistic, the null model is indeed the most severe condition at which all measured variables are considered to be uncorrelated (Prudon, 2015). NFI ranges between 0 to 1 where zero implies a poor fit (or no fit) and 1 signifies the perfect fit. Most scientists suggest that any value greater than 0.90 as an indicative of a good fit. Nevertheless, the NFI has bias for sample sizes greater than 200 (Bentler, 1990), and thus assessing the goodness-of-fit solely using NFI index does not reflect a good fit in reality (R. Kline, 2005). Therefore, researchers have suggested the Non-Normed Fit Index (NNFI, also referred to as the Tucker-Lewis index) to address this problem. The NNFI can exceed the conventional 0-1 interval but just similar to other incremental fit indices, NNFI suggests a good fit when the index value is closer to 1 with any value larger than 0.95 indicating a good fit (Hu & Bentler, 1999).

The comparative fit index (CFI) also developed to resolve the NFI's bias over the large sample sizes. Similar to NFI index, the CFI compares the chi-square value of the implied matrix with the X^2 value of the null matrix (a model that assumes no correlation between the measured variables); however, it is developed to neglect the sensitivity to sample size. This statistic index ranges from 0 to 1 with values closer to 1 suggesting a good fit. The cut-off value for the CFI index is a value greater than 0.9 (Maiyaki, 2012); while some authors suggested more controversial threshold of equal or greater than 0.95 (Hu & Bentler, 1999).

Table 5-2 List of model fit indices and their level of acceptance

Name of index	Cut-off value for a good fit
Chi-Square (X^2)	*P*-value > 0.05
CMIN/DF	Chi-Square/df < 3.0
Goodness of Fit Index (GFI)	GFI > 0.90
Root Mean Square of Error Approximation (RMSEA)	RMSEA ≤ 0.06
Comparative Fit Index (CFI)	CFI > 0.90
Non-Normed Fit Index (NNFI) or Tucker-Lewis Index (TLI)	TLI > 0.95
Normed Fit Index (NFI)	NFI > 0.90

Regression analysis

In this research, to understand which of the indicators has the most influence on creativity, raters were also asked to judge the products on creativity item in addition to rating the 65 indicators proposed earlier. Then, multiple linear regression analysis was used to assess the effects of implied indicators as independent variables on creativity judgments as the dependent variable. Implied indicators are those items that left from the total after being subjected to factor analysis.

5.5. Chapter conclusion

Criterion-based techniques are formed to assess the creativity of a wide range of products using a series of constant variables and constructs. Nevertheless, creativity is not an attribute of the product, but instead, it must be sensed and perceived by appropriate observers. As a result, the idea of evaluating creativity using constant indicators may not always reflect the correct consumer's perception of creativity for the given product. In contrast, this study aimed to propose a theoretical framework for assessing creativity of functional consumer products especially those in the shopping-goods category. It is worth noting that this instrument does not intend to suggest a single construct for creativity measurement, but instead attempts to support researchers who want to generate their own unique instrument for a specific product category by presenting the relevant creativity indicators for the consumer products. This framework evaluates products from 9 primary dimensions and 65 indicators including *novelty*, *performance quality*, *safety*, *human-factors*, *sustainability*, *attractiveness*, *emotion*, and *relevance*.

This chapter addressed the relationships between creativity and each of the criteria proposed, while the literature to support the hypothesised model was reviewed. Novelty is the crucial factor but not sufficient. The existence of novelty in creative products is a must, meaning that even the most practical and/or aesthetically efficient products cannot be creative unless they comprise new and original features. However, novelty on its own does not describe creativity. An essential characteristic of the creative artefact is the ability of the product to address the user's demand effectively. Product's (technical) performance quality signifies how well the product is able to accomplish its planned task and thus plays an essential role in recognising creativity. Another element affecting consumers perception of creativity is the safety factor. Product safety is a necessary part of the product development process, and in general is referred to as a mandatory requirement by laws, regulations, and/or standards. Apart from these, consumer want the solution to be easy to control, comfortable, and straightforward, implying the need for human-factors assessment.

A product's usefulness can be enhanced by improving reliability. A reliable solution enables the user to trust the product and its performance capabilities without worrying about system failure, making it more effective than any unreliable product. Another element that can add value to the product is sustainability. Consumers of sustainable products are more satisfied as they know that their usage of the product has less or no effect on the environment and/or society. Aesthetic is among the crucial elements affecting user experience especially when other domain products lack this quality. Aesthetics include the ability of the product in stimulating the user's sensory perception either visually, tactility, auditory, olfactory, taste, and/or emotion. From the creativity point of view, the attractiveness of the product impacts viewers' perception of creativity. Attractiveness is simply not just around beauty and style; it rather includes the extent to which the product is skilful, precise, and proficient in performing the task. On the other hand, emotions emerge from experiences that users acquired from owning and/or using the product in different circumstances. In other words, a part of consumer satisfaction is inspired by the degree to which the consumer perceives the product as technically capable and aesthetically attractive and pleasing compared to other domain products. Finally, the last aspect affecting product creativity is related to the degree to which the product appears as essential and relevant to the observer.

Given that the model proposed here is formed based on theory by synthesising the existing literature around product creativity and criterion-based techniques, it is necessary to verify the assessment reliability as well as the construct validity of the instrument. Therefore, a quantitative research method was designed while considering a group of consumer products as case study.

In summary, this study includes the design of a survey for data gathering, determining an appropriate panel of judges (novice raters), and conducting a small pilot study to ensure that the hypothesised items are internally consistent as well as to assess the quality of survey instructions. This chapter also covered various techniques used for analysing the data including reliability, factor analysis, and regression analysis.

6. Confirming the Validity of the Consumer Product Creativity Evaluation Framework

In the research literature, several methods have been developed to assess the level of product creativity including indirect methods (peer and teacher nomination and awards), global judgments (consensual assessment technique), and criterion-based instruments (creative product semantic scale or creative solution diagnosis scale). In the consensual technique, creativity is measured using domain relevant experts and so far, it has been referred to as "gold standard" of creativity (Baer & McKool, 2009b). CAT assumes no definition for creativity and instead relies on the degree of consistency among the experts, in the sense that if expert judges agree that the product is creative, then it is creative. CAT is a reliable method as revealed by several researchers in the field (Amabile, 1996; B. Hennessey et al., 2011, 2008). However, depending on the situation (i.e. product functionality, judges, time, and budget), CAT might not be as effective as it should be especially when it comes to troubleshooting the shortcomings of a product that is just appeared to have a lower creativity compared to others in an evaluation. The access to experts is limited most of the time and even if the availability is not an issue, the use of experts usually involves a higher cost compared to ordinary people. As described by Rubera et al. (2011), creativity is not an attribute of the product but rather it is a notion that must be sensed by appropriate observers, and in that sense the use of CAT cannot be replaced with anything else. However, criterion-based techniques as an alternative solution for (product) creativity assessment could be beneficial to researchers since they do not need a panel of expert judges and because of that they are less time consuming, and the cost is usually reasonable. In these methods, reliability is achieved by preventing the (novice) raters to judge creativity based on

their own understanding of the concept. Instead, they are instructed to give their opinions indirectly by rating a list of pre-identified indicators where the relationships between these indicators and creativity had been already investigated and established by researchers in the field. Overtime, several criterion-based instruments have emerged and most of them are shown to be reliable approaches when employed together with non-experts (Besemer & O'Quin, 1999; Cropley & Kaufman, 2012; Horn & Salvendy, 2009).

On the other hand, the way these instruments are built could be an issue by itself because most of the time they are designed by exploring what is perceived by the observers. Where these observers usually give their opinions on a particular subject (e.g. a group of chairs, lamps, and mouse traps) without paying attention to the effects of product functionality on the assessment. The function of a product is an important element in creativity assessment as it could influence the way the viewers are looking at creativity. The product's function can affect viewers mindset and based on that they may choose to give a strong opinion on a specific matter and a weak vote to other topics, but this could be altered given what product functionality is in question. Observers depending on product's functionality prioritize their point of view regarding the product. On the other hand, the function also affects the way a product presents itself to the observer. For instance, an oxygen capsule is an essential product and thus any safety feature that is added to it will be easily recognized by the observer, whereas the safety feature might be a less important element for a chair. In other words, products with different functionality need different elements in order to be recognised as useful. As a result, the underlying constructs of creativity might differ from product to product depending on what product functionality is in question. However, in the relevant literature of criterion-based instruments researchers did not pay attention to this important effect. Most of the studies in this domain were conducted (Cropley & Kaufman, 2012; Horn & Salvendy, 2009; O'Quin & Besemer, 1989) with a focus on developing a universal tool capable of assessing a wide range of products for creativity (functional and non-functional), while no attention was given to the effect of product functionality on the list of indicators chosen for the instrument.

Therefore, this research is an attempt to fill the gap in knowledge by proposing a framework for creativity measurement which allows the researchers and practitioners to establish their own

specific instrument using the pre-defined constructs proposed in the previous chapter (see chapter 5, section 5.2). This model introduces a new and quick way of creating unique instruments that are specifically designed to represent the consumers' point of view around a particular product. This model consists of 9 primary dimensions for creativity and 65 indicators that are carefully selected to describe each dimension from various angles (e.g., novelty factor is described with originality, unusualness, uniqueness, and revolutionary). In order to assess creativity more precisely, this theoretical model is focused on functional consumer products (i.e., products that serve some practical function to the user) to allow selection of more accurate and relevant creativity factors and indicators. This chapter aims to test the construct validity as well as the measurement reliability of the proposed theoretical framework by selecting a group of consumer products as case study (a series of lounge chairs).

6.1. Research aim and objectives

The aim of this study is to investigate the development process of a unique criterion-based instrument for a specific category of products (lounge chairs). Also, to inspect whether the extracted model is able to validly assess the creativity for the selected products. Hence, the following objectives have been set to fulfil this aim:

Objective 1: Conduct a web-based survey to collect information about consumers' opinions on 65 indicators proposed in the previous chapter over 5 lounge chairs selected as case study.

Objective 2: Examine the overall inter-rater reliability of the judgments to see whether different judges can give consistent ratings to the indicators.

Objective 3: Construct a unique criterion-based instrument by subjecting the data into an exploratory factor analysis and finding out the underlying structure that is perceived by the observers for the given products.

Objective 4: Analyse the degree of internal consistency among the items that are forming each construct as well as evaluate the convergent and discriminant validity of the obtained model by submitting the data into a confirmatory factor analysis.

Objective 5: Examine the predictive validity of the obtained model to see which of the factors in the model are able to significantly predict the ratings of creativity given by non-experts. Also, to analyse the predictive power of individual items to see which of the items has the most capacity to explain majority of the variations in the dataset.

6.2. Methodology

In this study, a quantitative research method was employed to collect and analyse data as well as to establish the construct validity of the model extracted. A web-based survey medium was selected for its advantages over other methods and suitability in operating large sample sizes (also see chapter 5, section 5.3). In summary, this survey contained 4 parts including consent form, demographic questions, product presentation panel, and the main questions (65 indicators). In order to assess the overall feasibility of the research, survey instructions, viability and readability of the survey questions, a small pilot study was performed. Once the survey accepted, then the main study was conducted using a larger sample size to construct and validate the underlying structure for the indicators proposed. The following sections explain the process in detail.

Main study

i. Material

In the process of selecting stimuli for this study, one criterion was to pick products that are easily recognisable at a glance. For example, observers must be able to effortlessly distinguish various characteristics of the products using only a series of two-dimensional images. Another criterion was to select a spectrum of products that possess the same functionality while offering different features, functions, and/or styles to allow comparison. Moreover, products must be selected in a way to facilitate the collection of data that are normally distributed. Therefore, products must range from low to medium and to high creative to allow raters to choose their choices from all available options. Consequently, a set of 10 lounge chairs were selected from a pool of chairs nominated from various reliable sources including but not limited to Taschen's 1000 chairs (Charlotte & Peter, 2005). The nomination process of the stimuli was performed rather arbitrarily to ensure that the selection of products would consist of a wide range of items so that they can

be graded from low to high creative by the raters. The only requirement was that they need to be a lounge chair (a long reclining chair) while the attention paid to the novelty factor including features such as unusual structure, colour, material and/or usage. Then, a small group of domain-relevant experts were asked to judge all chairs for creativity. The panel of judges consisted of 5 experts with an adequate level of experience in the field. Three faculty members were from the design engineering department, one was from art department, and the last judge was a specialist in design and manufacturing of furniture from industry.

To gather experts' points of views regarding the level of creativity for the selected chairs, a web-based survey was conducted. First, raters read and agreed to a consent form, and then were given guidance for the task. Next, judges were presented with a few two-dimensional images as well as the technical specifications of the products to allow comparison. Experts were advised to work independently and try not to consult about their ratings with their fellow workers. To prevent the occurrence of systematic errors in the study, the order of presenting the products was randomized for each rater. Judges were instructed to observe and compare the products to one another before starting the assessment. Then they were asked to evaluate each product for creativity using a 5-point Likert scale with 1 representing 'Not at all creative' and 5 representing 'Extremely creative'. After recording information, the data were analysed for inter-judge reliability where results revealed a good consistency among the raters ($\alpha = 0.819$). Therefore, from the 10 original lounge chairs, five products were selected for the main study including at least one very low creative (average score value of 1.4), one very high creative (average score value of 4.2), while the other three chairs were carefully selected in a way to illustrate the medium range of creativity from medium-low to medium-high creative (average value scores include 2, 3, 3.6).

ii. Participants

Over 1000 university students from various universities in the United Kingdom were asked to take part in the study. From all records, 288 students have completed the survey. Out of which, 248 students were currently enrolled in an undergraduate program, 32 enrolled in a postgraduate program, and 8 were studying a doctoral degree. The sample contained 143 respondents who defined themselves as female (49.7%), and 145 respondents who defined

themselves as male (50.3%). The most frequent age group was 18-24 years-old (88.5%), followed by 25-29 years-old (6.6%). The ethnic breakdown of the participants was 187 (64.9%) White; 58 (20.1%) Asian or British Asian; 23 (8%) Black/African/Caribbean/Black British; 14 (4.9%) Mixed/Multiple ethnic groups; and 6 (2.1%) who selected other ethnic groups.

iii. Task procedure

The process of data collection was conducted online, where judges read and agreed to consent conditions, and then were given instruction for the task. First, raters were presented with two-dimensional pictures to visually assess products' appearance and become familiar with their technical specifications. The order of presenting the products was randomized for each participant to eliminate the occurrence of systematic errors in the study. Participants were asked to rate each of the lounge chairs using the 65 indicators proposed in the previous chapter (see chapter 5, section 5.1, Table 5-1). The order of questions was also randomized to control for fatigue during the data collection process. Each item (indicator) was assessed on a 7-point Likert-type scale, ranging from 1 to 7 between bipolar adjectives such as *usual-unusual*. In addition to 65 items, judges were asked to rate each product for creativity using a 7-point Likert-type scale with 1 representing 'uncreative' and 7 representing 'creative'. The process of data gathering was administrated by the Qualtrics software company (see https://www.qualtrics.com/uk/).

Pretesting the survey

To assess the viability of the survey instructions as well as to examine the readability of the 65 indicators, a small pilot study was conducted. Major objectives of this study include: 1) to measure the internal consistency among the items forming a construct (e.g., *originality*, *unusualness*, and *uniqueness* for the Novelty factor), 2) to examine the item wording and to assure that items are easy to understand and contain no technical word/phrase, and 3) to confirm the correctness of the survey instructions and the overall administration procedure.

Initially, the survey was distributed to 6 PhD students as well as two faculty members at the Dyson School of Design Engineering (Imperial College London) to reduce errors in the form of participants misunderstanding a question or the instructions. Individuals were asked to inspect the overall format of the survey and to comment on any question that is unclear. At this stage,

no issue was detected by the respondents. Furthermore, the length of the survey as well as the time required for completing it was also assessed and once again no issue was pointed out by the reviewers. Then to assure that whether the items forming the constructs are indeed measuring the same concept, the process of data collection for the main study included a soft launch before gathering the whole data. Therefore, the collection of data was stopped when the number of respondents who completed the survey reached 30.

At this stage, the ratio of internal consistency among the items (indicators) was assessed using the Cronbach's coefficient alpha. The overall inter-item reliability between the 65 indicators was achieved at an excellent range ($\alpha = 0.986$). Also, the reliability of the items that belong to each factor was measured separately. The internal consistency between the indicators forming the novelty factor was obtained at 0.986; similarly, consistency of the items under the performance factor was 0.922, and from there 0.829 for the safety factor, 0.886 for the human factor, 0.942 for the reliability factor, 0.871 for the sustainability factor, 0.938 for the attractiveness factor, 0.836 for the emotion factor, and at last 0.907 for the relevance factor. As the outcomes did not indicate any significant issue, the process of data collection was continued. Next section covers the results of the main study.

Statistical methods

After completing the data collection process, all records were subjected to data analysis including reliability tests, factor-analysis, and linear regression analysis. Cronbach's alpha was used to assess whether all subscales in the construct are measuring the same factor (inter-item reliability). Also, to measure the overall degree of consistency between the raters (inter-judge reliability). Exploratory factor analysis was performed to determine the perception of raters who judged five lounge chairs on the 65 creativity indicators. In addition, to identify the number of latent factors and their corresponding factor structure. Then, the data were exposed to a confirmatory factor analysis to test and verify the factor patterns suggested by the EFA. Finally, multiple linear regression analysis was employed to determine the level of impact that each of the remaining indicators had on creativity (see chapter 5, section 5.4 for further information regarding the data analysis).

6.3. Results

In the present study, 288 novice raters judged five lounge chairs on the 65 items in addition to rating the products for creativity. Initially a data matrix including 1440 (288 × 5 = 1440) × 66 (65 indicators + 1 creativity ratings) was prepared. Prior to analysis, the data was inspected for outliers. The unengaged participants were detected by measuring the standard deviation (SD) for the ratings given by each rater for all 66 questions per product type. All respondents with zero variation (i.e., SD = 0, or respondents who answered the exact same value for every question in the survey) were removed from the data. This resulted in eliminating 147 cases from the original data matrix (1440 - 147 = 1293 cases; 632 cases with male raters and 661 cases with female raters). Then, the data was examined for normality using the skewness and kurtosis tests. According to Field (Field, 2013), data can be considered normal if the value of skewness and kurtosis do not exceed ±1.96. In the current study, all items including creativity itself had skewness and kurtosis in the acceptable threshold. The skewness for the entire dataset ranged between -0.610 and 0.179, while the kurtosis was ranged between -1.235 and -0.212, indicating that the data is normally distributed.

To inspect whether the mean of creativity judgments is different between female and male raters, an independent-sample t-test was conducted. There was homogeneity of variances, as tested by Levene's test for equality of variance (p = 0.129). Results indicated that female raters on average scored the creativity (M = 5.06, SD = 1.789) higher than the male raters (M = 4.79, SD = 1.840), a statistically significant difference, M = 0.271, 95% CI [0.073, 0.469], $t(1291) = 2.683$, $p < 0.01$. The inter-judge reliability for the entire dataset was achieved at an excellent range (α = 0.970). The consistency among the raters was also assessed for each product and results are illustrated in Table 6-1. Overall, a high reliability was obtained between the raters assessing each product which allowed the data to be combined and form a larger sample size, making it more appropriate for factor analysis.

Table 6-1 Inter-judge reliability for each product as well as the entire dataset

	Chair A	Chair B	Chair C	Chair D	Chair E	Total
Cronbach's alpha coefficient	0.934	0.978	0.929	0.965	0.951	0.970

To analyse the underlying factor patterns of the judgments of college students rating five lounge chairs on the 65 indicators, an exploratory factor analysis was conducted. Prior to testing, the degree of factorability of the measurement variables was evaluated. The Kaiser-Meyer-Olkin (KMO) measure of sample adequacy was obtained at 0.989, which is greater than the commonly recommended threshold 0.6 (Hair et al., 2010); whereas, Bartlett's test of sphericity was also significant ($x^2(2080) = 61417.564,\ p < .001$). Initially, the analysis of EFA using the PAF extraction method suggested a solution with six factors (i.e., six factors had eigenvalues equal or greater than 1) which explained 57.4% of the variance. However, further scree test implied a five-factor solution since the slope of eigenvalues is levelled off after the fifth factor. Therefore, a factor analysis using the PAF extraction method was repeated with five factors.

Table 6-2 The five-factor model extracted from the EFA analysis

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Appealing	**0.882**	0.024	-0.032	-0.028	-0.012
Charming	**0.831**	0.016	0.009	-0.055	0.010
Elegant	**0.723**	0.018	0.034	-0.020	0.024
Desirable	**0.863**	0.032	-0.023	-0.010	-0.037
Satisfying	**0.855**	0.067	-0.001	-0.023	-0.047
Pleasant	**0.883**	-0.042	-0.012	-0.004	-0.043
Appropriate	**0.622**	-0.074	0.014	0.117	0.125
Reasonable	**0.728**	-0.101	0.027	0.090	0.057
Original	0.057	**0.664**	0.068	-0.059	0.084
Unusual	-0.118	**0.836**	-0.050	0.031	0.012
Surprise	-0.024	**0.781**	0.052	0.002	-0.028
Unique	0.106	**0.729**	-0.023	0.069	-0.025
Adaptable	0.111	0.053	**0.692**	0.015	-0.008
Modifiable	-0.087	-0.011	**0.798**	0.067	-0.027
Alterable	0.115	-0.023	**0.698**	0.015	0.016
Changeable	-0.045	0.019	**0.857**	-0.072	0.007
Recyclable	-0.050	0.071	0.121	**0.564**	-0.027
Environmentally-Friendly	-0.012	-0.022	0.025	**0.761**	0.057
Pollution free	0.093	0.064	0.074	**0.710**	-0.040
Durable	0.205	0.034	-0.019	-0.034	**0.659**
Solid	0.106	0.016	-0.016	-0.016	**0.734**
Robust	0.117	-0.006	0.020	0.052	**0.655**

For interpreting the underlying structure for the five-factor, an oblique rotation algorithm was employed. All items with loadings smaller than 0.4 were considered as insignificant and hence removed from the model. Similarly, items that cross-loaded at smaller than 0.2 onto any pairs of factors were also eliminated from the study. Table 6-2 shows the structure after the omission of insignificant measured variables. This model accounted for 62.19% of the variation in the dataset and in general, fulfils the common conditions for a meaningful structure, including a minimum of three indicators for each factor. The majority of the scales in the model obtained items with high factor loadings. In this regard, Comrey and Lee (A Comrey & Lee, 1992) reported that any loading greater than 0.7 is "excellent", greater than 0.63 is "very good", >0.55 is "good", and >0.45 is "fair". It appears that the first latent factor with 8 indicators contains a fusion of some of the factors from the original model proposed in chapter 5 (including attractiveness, emotion, and relevance); thus, this factor was labelled Aesthetics and Relevance.

Table 6-3 The degree of internal consistency between the items of each constructs

	Novelty	Versatility	Ecological footprint
Coefficient alpha	0.854	0.862	0.761
	Reliability	Aesthetics and relevance	Overall consistency
Coefficient alpha	0.836	0.938	0.938

The second factor consists of indicators measuring the level of originality of the product and hence classified as Novelty. Further items loaded onto the third factor seem to express the product's ability in adjusting and adapting itself with the working environment; therefore, referred to as Versatility. The fourth factor includes items linked to the degree of impact that the product has on the environment during its life-cycle, thus labelled Ecological footprint. The items that loaded onto the fifth factor were associated with the product's ability to last longer and stay functional during the intense conditions; thus, classified as Reliability. Before analysing the validity of the constructs, the ratio of internal consistency was measured between the items of each scale (see Table 6-3). For Novelty factor (4 indicators) the coefficient alpha was 0.854, for the Versatility factor (4 items) was 0.862; for the Ecological footprint factor (3 items) was 0.761; for the Reliability factor (3 items) was 0.836; and for the Aesthetics and Relevance factor (8 items)

was achieved at 0.938, further indicating that the underlying sub-items of each scale are uniformly describing the same concept (Cronbach, 1951; Nunnally & Bernstein, 1994).

A confirmatory factor analysis was carried out in order to verify the validity of the five-factor solution suggested by the EFA analysis. A model was estimated using the Maximum Likelihood estimation method and results are illustrated in Figure 6-1. In reference to model fit, the value of the chi-square test was significant ($x^2(199) = 520.965, p = 0.000$); however, given the large sample size in this study (>200), the ratio of chi-square to the degree of freedom (CMIN/DF) is a better index to assess the model fit (Wheaton et al., 1977). In this model, the CMIN/DF was attained in the acceptable threshold < 3 (CMIN/DF = 2.618). Furthermore, the goodness-of-fit index (GFI) was 0.965. Likewise, the comparative fit index (CFI) was 0.981, the normed fit index (NFI) was 0.97, the non-normed fit index (NNFI or Tucker-Lewis fit index (TLI)) was .978, indicating a good fit between the implied model and the observed data, i.e., the specified model is able to reproduce the observed data in a way that a high percentage of the variance is explained by the estimated model.

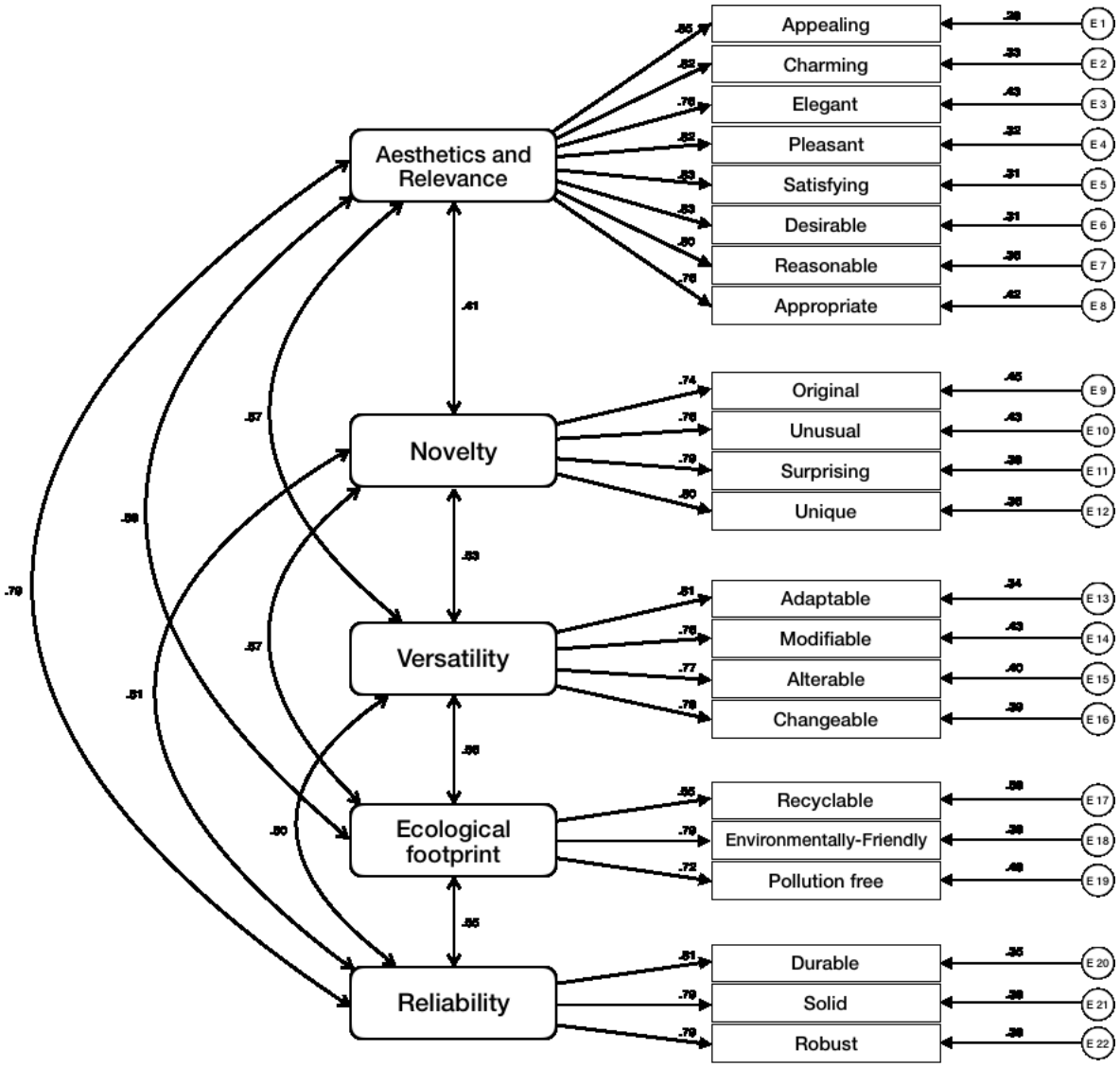


Figure 6-1 Standardized estimates of the five-factor model (x^2=520.923, df=199, p<0.001, RMSEA=0.035, CFI=0.981)

The value of the root mean square error of approximation (RMSEA) was attained at 0.035 which further indicated a good fit between the implied covariance matrix and the observed data. In this model, the composite reliability (CR) and average variance extracted (AVE) were computed for each latent factor, and results were in the acceptable range; i.e., greater than 0.7 for CR and greater than 0.5 for AVE (Hair et al., 2010) (see Table 6-4). Discriminant validity was attained between the constructs in the model (Aesthetics and Relevance, Novelty, Versatility, Ecological footprint, and Reliability). Table 6-4 gives the values of AVE as well as the squared correlations between the constructs.

Table 6-4 The AVE, Composite Reliability (CR), and squared correlations between the constructs

	CR	AVE	Aesthetics and Relevance	Novelty	Versatility	Ecological footprint	Reliability
Aesthetics and Relevance	0.938	0.654					
Novelty	0.855	0.595	0.171				
Versatility	0.862	0.609	0.321	0.396			
Ecological footprint	0.765	0.522	0.339	0.442	0.434		
Reliability	0.836	0.629	0.624	0.263	0.252	0.422	

According to the criteria suggested by Fornell and Larcker (1981), to establish discriminant validity the AVE for each pairs of factors should be larger than the squared correlation (i.e. shared variance) between the factors. In this study, the AVE values for all latent factors were obtained at larger ratios compared to the squared correlations between all pairs of factors. The outcomes indicated that discriminant validity exists within the constructs.

Table 6-5 Factor score predictions of creativity ratings

Predictors of creativity	Partial R^2	R^2	df	F value	*p*-value
Novelty	0.576	0.576	1291	1753.630	0.000
Aesthetics and Relevance	0.045	0.621	1290	154.183	0.000
Versatility	0.003	0.624	1289	9.965	0.002
Ecological footprint	0.001	0.626	1288	4.647	0.031

To test the instrument's ability to predict the judgment of creativity, a stepwise linear regression analysis between latent factors' scores (as independent variables) and creativity ratings (as dependent variable) was conducted. Factor scores were calculated by averaging the scores from the list of items loaded onto each latent factor extracted in the EFA analysis (see Table 6-2). Results indicated that Novelty, Aesthetics and Relevance, Versatility, and Ecological footprint significantly predict the judgment of creativity ($R^2 = 0.626\ F(4, 1288) = 537.815,\ p < 0.001$; Table 6-5). In this model, the Novelty factor explained almost 58% of the variance in the dataset ($R^2 = 0.576,\ p < 0.001$), whereas the rest of factors each explained less than 5% of the total variance.

Table 6-6 Significant predictors of product creativity for the lounge chairs

Significant predictors	Standardized coefficients (Beta)	95% CI	*p*-value
Unique	0.291	0.241 – 0.339	0.000
Original	0.155	0.109 – 0.201	0.000
Surprising	0.159	0.113 – 0.211	0.000
Charming	0.070	0.018 – 0.125	0.009
Unusual	0.154	0.105 – 0.197	0.000
Appealing	0.066	0.012 – 0.115	0.016
Changeable	0.061	0.022 – 0.099	0.002
Robust	0.047	0.005 – 0.095	0.030
Environmentally-Friendly	0.043	0.003 – 0.098	0.037
Desirable	0.053	0.002 – 0.105	0.044

Furthermore, to find out how well each item (indicator) predict the judgment of creativity, an additional stepwise linear regression analysis was carried out between 22 items (all items in the model) as independent variables and creativity ratings as the dependent variable. The results of this regression yielded in a model with 10 significant predictors which accounted for $R^2 = 0.635$ $F(10,1282) = 222.736,\ p < 0.001$ of the variation in the dataset; a large size effect according to Cohen (1988) (see Table 6-6).

6.4. Discussion

The present study showed the applicability of the proposed theoretical framework in developing a unique criterion-based instrument for the measurement of creativity for a specific category of consumer products. Overall, a high degree of inter-rater reliability was achieved, indicating that novice raters are able to understand the meaning of the questions consistently, while they agree with each other on the level of products' performance with respect to a particular product characteristic. The values of inter-rater reliability attained in this study are in the similar range with those of other researchers in the field (Cropley & Kaufman, 2012; Cropley et al., 2011). High levels of agreement between the non-experts imply that no training is required for such raters to assess the product creativity using the instrument obtained. This allows for a much more affordable and quicker assessment compared to the consensual assessment technique which is in general, a time-consuming process as it demands searching for experts in the domain (Cropley et al., 2011).

The factor analysis revealed a five-factor solution for measuring the creativity of lounge chairs, including Aesthetics and Relevance, Novelty, Versatility, Ecological footprint, and Reliability (see Table 6-7). It is worth noting that this model is a unique instrument developed for the creativity assessment of lounge chairs where the factor patterns are likely to represent the way the consumers perceive creativity when they see similar products in the domain. The Aesthetics and Relevance dimension contains items related to the product's ability to stimulate observers by either visually, tactility, and emotion. For these types of products, it appears that the extent to which observers consider the chair as relevant and important to the task also correlates with the product's aesthetics characteristics. Novelty refers to the ability of the product to offer dissimilar, unique, and/or original features where no other domain relevant product can provide. The third factor (Versatility) denotes the capability of the product in adapting and adjusting itself with the working environment in favour of user experience. The Ecological footprint dimension refers to product's impact on the environment throughout all phases of its life-cycle. Reliability relates to the ability of the product to last longer, withstand intense wear and pressure. High ratios of inter-item reliability between the sub-scales of each construct verify the consistency of the instrument over time. Other researchers who established criterion-based instruments also reported comparable ratios of reliability (Cropley et al., 2011; Horn & Salvendy, 2006b; O'Quin & Besemer, 1989). In this model, the majority of the constructs attained "good" to "excellent" reliabilities, which means that an adequate level of association between the items of each scale exists, and thus they are expected to give similar effects if they are subjected to a different group of raters. This to some extent ensures the validity of the instrument over repetitive consumption at different time intervals.

The model obtained in this study is a valid measurement tool. All constructs comprise similar or higher factor loadings compared to those of other researchers in the field such as CPSS (O'Quin & Besemer, 1989), CSDS (Cropley & Cropley, 2005; Cropley & Kaufman, 2012), and PCMI (Horn & Salvendy, 2009). The outcomes of the CFA analysis also confirm the convergent and discriminant validity of the constructs. Overall, the factor structure attained in this is in line with the relevant literature. This study among other investigations (Besemer & O'Quin, 1999; Cropley, 2016; Horn & Salvendy, 2006b) confirms that Novelty is the crucial factor affecting product creativity. This is

also consistent with the general definition of product creativity (Plucker et al., 2004; Runco & Jaeger, 2012).

Table 6-7 The unique criterion-based instrument obtained for creativity measurement of lounge chairs

Aesthetics and Relevance	Novelty	Versatility	Ecological footprint	Reliability
Appealing	Unique	Adaptable	Recyclable	Durable
Charming	Unusual	Modifiable	Environmentally-Friendly	Solid
Elegant	Surprising	Alterable	Pollution free	Robust
Desirable	Original	Changeable		
Satisfying				
Pleasant				
Reasonable				
Appropriate				

The Aesthetics and Relevance factor in this study is to some degree equivalent to the *Affect* factor extracted in Horn and Salvendy's PCMI model (2009). Both factors consist of items associated with product's emotional impact on the user as well as the product's ability to attract the evaluators' attention (attractiveness). However, the Aesthetics and Relevance dimension also includes the associated items with importance factor (i.e., how important/relevant is the product for the task). Possibly, the main reason for this contradiction is laid on the difference between products' functionality. Lounge chairs are fundamentally different products compared to normal chairs. They are designed to provide a relaxing moment for the user or occasionally they are placed in certain areas of the room to draw in observer's attention. They are also utilized to enhance the overall appearance of the space and to convey calmness into the place. As a result, observers may see a lounge chair as relevant to the task only if it is capable of emotionally stimulating the viewers in addition to being able to offer a design that is elegant and appealing.

The rest of the factors (i.e., versatility, ecological footprint, and reliability) are less recognized as distinct dimensions; however, their impacts on product creativity as well as product's usefulness have been already addressed by the researchers in the field. For instance, adaptability is referred to as a technique that can improve the overall creativity of an existing product. In this regard, the SCAMPER tool is relevant which is composed of seven strategies in the form of a checklist to help individuals to increase creativity (Eberle, 1971). Among the items in the list include adaptability

and modifiability. In the present study, the term *versatility* pertains the ability of the chair to make changes of any kind (to its structure) which can convert the product into a more useful and relevant solution for the task, enhancing product's usability. Consumers of adaptable products are able to use the product more frequently as they are able to change the product characteristic into the configuration that they most demand it. Product's ecological footprint influences the product's effectiveness. In this regard, Yazdanifard and Mercy (2011) have examined the effects of green marketing strategies on customer satisfaction and reported that customers of environmentally friendly products are more fulfilled compared to products that lack this quality. Reliability is another factor that can enhance creativity among the lounge chairs. Reliability of the product allows the formation of a bond between consumer and product and plays an important role in product's usability. Recent studies showed that as consumers' perception of reliability increase, Consumers will trust the outcomes of the product and hence use it more frequently and consistently (Page, 2014).

The results of the regression analysis on the five-factor model of product creativity demonstrated that 4 out of 5 latent factors (i.e., Aesthetics and Relevance, Novelty, Versatility, and Ecological footprint) significantly predict the creativity judgment of the raters. This shows that a high percentage of variance for the judgment of creativity can be explained by the model obtained in this study, further indicating the validity of the model in assessing product creativity accurately. In the relevant literature, Horn and Salvendy (2009) also reported comparable predictive validity for the measurement of creativity using the PCMI instrument. Overall, the factor patterns extracted and confirmed in this study signify the importance of product's functionality in the formation of criterion-based instruments. The theoretical framework proposed in this thesis gives the researchers and practitioners the ability to quickly and validly establish new and unique factor patterns for any category of functional products. The results of the lounge chair case study confirmed the inter-item reliability of the indicators as well as the applicability of the original framework in producing unique models for creativity measurement; where only a set of specific factors and indicators are found to be able to correctly measure the creativity.

In the present study, to test the validity of the creativity assessment, the end results of creativity measurement (i.e. creativity grades/rankings) obtained from novice raters (248 university

students) were compared with those of experts. That is the average values for creativity scores given by experts were compared with the average values for the scores given to all relevant indicators of creativity provided by novice raters. According to the results, the order of creativity rankings were identical between the panel of expert judges and novice raters even though the actual numbers were different (average values for creativity ratings by expert judges were as follow: item A = 1.4, item B = 2, item C = 3, item D = 3.6, item E = 4.2; average values for scores given to all relevant indicators of creativity produced by novice raters were as follow: item A = 3.62, item B = 4.34, item C = 4.37, item D = 4.46, item E = 4.62 (for list of items see Appendix)). Considering that the assessment of creativity is a relative concept, then it is normal for the actual numbers to not match with each other, however, the rankings of creativity are the same between the two panels of judges. Here, the main focus is not to find out the correlation between the two panels of judges, instead, this comparison intends to verify that whether the orders of creativity rankings are the same between the new framework and the panel of expert judges. In the research literature, the use of expert judges for assessing product creativity has been regarded as the "gold standard" (Baer & McKool, 2009b) and it is essential to ensure that the new measurement tool is genuinely capable of obtaining results that are as identical as expert judges.

6.5. Research limitation

The five-factor model obtained in this study revealed the capacity of the theoretical framework in generating distinct criterion-based instrument for product creativity. Using this framework researchers are able to establish precise measurement tools to better capture consumers' perception of creativity based on product's functionality. However, there are a number of issues that could influence the quality of the data gathered. The process of data collection using surveys usually involves stress and fatigue especially when participants are asked to answer a large number of questions. In the present study, each student was instructed to respond over 300 questions in a single attempt. Although in the current study the order of questions had been randomized to prevent errors caused by fatigue, still it does not guarantee whether participants gave their answers to the questions with the exact same quality. In other words, it was impossible to completely eliminate the fatigue's impact from the ratings.

The outcomes of this study may be affected (slightly) from the inconsistency between the images shown to the judges. This is among the research limitations as there was no physical access to the products and thus all images were obtained from available illustrations provided by the designer and/or manufacturer. Although this effect may seem insignificant, however, the way the product is presented to the consumer, without doubt, contributes to consumer's feelings and as a result may affect their decisions on whether the product is creative. In each picture, a different scene setup such as distinctive backgrounds, arrangements, people, and conditions was used, and this inconsistency might have caused some changes to the judgements of the raters.

Moreover, the results of this study might have also been affected by culture factors. Even though the respondents are from an extensive range of cultures and ethnicity (such as Asian, Black, and White), yet the majority of the participants have White backgrounds; and thus, outcomes of this study cannot be generalized to all cultures and ethnicity. In fact, the model of product creativity attained in this study is likely to represent the perspectives of young consumers in the United Kingdom environment.

Even though the theoretical framework proposed in this thesis aims to ease the process of generating unique criterion-based instruments that are dedicated to a specific category of products, still it cannot guarantee whether the model would represent the actual perception of consumers in the real-world scenario. The main reason for this issue is the lack of a reference value for creativity to allow a comparison to be made to inspect if this model or any other model are indeed addressing the genuine structure in which consumers perceive creativity. This is an important restriction affecting the whole field. Nevertheless, the model obtained in this study is factorially valid and statistically reliable and significant. This study revealed the ability of the framework in successfully generating a distinct structure for product creativity which is indicative of the applicability of the instrument in future studies.

6.6. Future work

The current study indicated the ability of the framework proposed in generating a unique instrument for product creativity by employing a group of non-experts. However, it is important to further test the capability (validity and reliability) of the theoretical framework in capturing

consumers' perception of creativity using different types of products. This an important step toward the improvement of the basic ideal model as well as to gain information about the practicality and efficiency of the framework.

This study has used non-experts to establish the factor structure; however, experts might have different opinions. An interesting study is to utilise the theoretical framework to compare the factor patterns for a specific category of products once using non-experts and once using domain relevant experts. This can reveal the capability of the framework in extracting the actual (true) factor structure as well as demonstrating the difference between the two groups of raters.

6.7. Chapter conclusion

The ability to assess and control creativity is essential for innovative firms and product designers. In this context, the consensual assessment technique dominates the field of product creativity measurement. The CAT relies on domain relevant experts for the reliable assessment of creativity. However, the process of searching for appropriate judges is typically time-consuming and involves a high cost and energy. Aside from experts, the outcome of the CAT usually does not provide useful information to guide designers and managers on how creativity can be improved during the product development process. Alternatively, criterion-based instruments are able to conduct reliable creativity measurement using a set of constant constructs (and indicators) and by employing non-experts as judges. One drawback of using constant criteria for measuring product creativity is that the outcomes may not accurately reflect observers' perception of product creativity for all products. Creativity is not an attribute of the product but rather it is a concept that must be perceived by appropriate observers. As a result, the attributes that reflect creativity are distinctive based on what product functionality is in question. To overcome this issue, researchers require to establish unique factor patterns for each category of products.

Previously, this thesis proposed a theoretical framework to enable researchers and practitioners to establish their own unique criterion-based instrument using the indicators suggested in the framework. The current study showed a successful application of the framework in generating a valid and reliable instrument for measuring creativity of lounge chairs using 288 college students.

The factor analysis illustrated the process in which the nine primary dimensions reduced into a five-factor model to describe the way viewers perceive creativity for lounge chairs. High inter-rater reliability achieved in this study indicates the correctness and ease of use of the indicators proposed. Also, it verifies that researchers and practitioners can confidently employ these indicators to establish their own product creativity models using novice raters with no formal training. In addition, the high ratios of internal consistency between the subscales of each construct specify the ability of the instrument in providing valid measurements in the long run.

Overall, the nine dimensions and the 65 indicators proposed in this thesis resemble ideal factor patterns in which observers are able to distinguish the difference between all scales; however, in the real-world scenario viewers might not always be able to psychologically separate two or more factors depending on what product functionality is in question. This study, in general, confirms the ability of the theoretical framework in generating unique measurement models for creativity. Compared to other instruments, this framework is relatively efficient in the long run, because once a unique model is produced and established in a domain, then there is no need to repeat the study using a large number of raters and instead all assessments can be performed using a much smaller panel (e.g. 100 participants) making it cost effective. In the meantime, the measurement accuracy is kept at a high level due to evaluating creativity using only relevant indicators that are carefully calculated with respect to the functionality of the product in question. In addition, because of a high degree of internal consistency between the indicators, there is no need to train the raters. Researchers and managers of the innovative firms can employ this tool to find out how consumers of specific category of products perceive creativity in different time intervals at a lower cost.

7. Overall Discussion and Conclusion

Consumer products form an important part of everyday life helping people address essential and desirable activities. These products typically facilitate consumers with a selection of functional solutions to help them to perform tasks more efficiently. Today, with the advancement in technology the demand for consumer products is increasing and, in the meantime, companies need to compete with one another in order to enhance their market share and operate on an economically viable basis. Creativity can be an important element in the product development process since it induces the creation of new and useful ideas which in turn let innovative businesses to identify themselves with other domain competitors. However, in order to be able to systematically implement creativity in the design process, companies need to find out how to recognize creativity while having the ability to understand the factors that positively impact creativity.

This thesis explores the measurement of product creativity aiming at addressing the need for further improvements in two contexts. First, to provide the ability to measure product creativity without the consideration of the process (i.e., individual, process, and press). At present, much of what is assessed by the panel of experts is narrowed to artistic products created in response to a pre-defined assignment using identical equipment and motivation (chapter 3). Under such conditions, the influence of environmental factors on creativity would not be addressed. Measuring product creativity without the consideration of the process, would then allow companies to comprehend insufficiencies in other components of the creative process and facilitate the implementation of creativity into the design process. Empirical studies conducted in this thesis showed a clear demonstration of how to employ expert judges for the creativity assessment of non-parallel functional products (i.e., products produced under dissimilar conditions), and further considerations were addressed to assure the robustness and reliability of the assessment (chapter 4).

The second aspect that requires further improvement, is to understand how to enhance the assessment procedure to have the ability to determine what factors are responsible for creativity based on the product's functionality. Currently, researchers have developed general criterion-based instruments which allow them to employ non-experts to reliably assess creativity with a set of constant variables whose influence on creativity have already been uncovered in the relevant research literature (chapter 3). Nevertheless, considering that creativity changes its shape and form according to product's functionality, the issue that arises is that whether the use of constant variables and factor patterns can genuinely represent the consumer perception of creativity. To address this problem, it would be necessary to generate a unique measurement instrument for each product category thinking about the key performance characteristics of the domain that is in question. A note to keep in mind at this stage is that the people who perceive creativity at present time are different from people who would perceive it tomorrow. In other words, people change their demands over time and creativity is no exception to this matter and it would change its shape and form over time as consumers will change their definition of creativity from time to time. However, it is suspected that the definition of creativity for a category of product would remain constant (changes are likely but not significant) for a period of time and that is what this research is aiming at to identify. This thesis proposed a theoretical framework to assist researchers or interested parties to develop their own unique criterion-based instruments from a pool of items whose relationship with creativity have been addressed in the relevant literature (chapter 5). A case study using a series of lounge chairs was conducted to illustrate the ability of the theoretical framework in extracting unique factor patterns indicating how observers would perceive creativity for these products (chapter 6). This chapter provides a summary of the main findings, guidelines for implementing the measurement approaches, and directions for the future work.

7.1. Summary of the main findings

Chapter 1 presents the concept of creativity and emphasises the value of creativity for innovative businesses. Creativity offers the ability to produce something new and useful which founds the basis for innovation. Creativity helps companies to distinguish themselves in the market and allows them to improve their market share and economic viability. Historically, the majority of

creativity assessments have been based within artistic domains where products have been produced in response to an identical assignment. Thus, the first goal of this research was arranged to investigate the capability of a panel of experts for assessing the creativity of non-parallel functional products. On the other hand, to further improve the measurement quality, a theoretical framework was planned to enable researchers or interested parties such as a company design team, to create their own unique measurement instrument based on the functionality of the product that is in question.

Chapter 2 explores the role of creativity in engineering design. Engineering design is a discipline that facilitates designers (and engineers) with the optimal selection of tools and strategies to enable them to better understand design problems while enhances their productivity through the application of creative thinking techniques. In general, the study of creativity in the engineering domain is founded onto four main areas including *person*, *process*, *press*, and *product*. This model implies a system in which a creative thought emerges as a result of the interaction between the individual(s), the thinking procedure, and the environment. Although several factors contribute to explain the border between creative and uncreative individuals, yet the most influential reason is due to the distinctions between personality traits such as conscientiousness, agreeableness, and openness to experience. For instance, individuals with an open mind have a tendency for divergent thinking which form the basis of the creative thinking. Agreeableness aids people to face problems with a friendly manner and enhances their ability to cooperate and avoid conflict.

The creative process, on the other hand, comprises the thinking procedure which individuals employ to generate new ideas and/or to optimize current ideas to better fit the problem. Several stages of the creative process include preparation (i.e., seeking and organising information), incubation (refers to unconscious cognitive problem-solving), insight (the ‘Aha’ moment), elaboration and evaluation (i.e., assessing the practicality and functionality of the ideas). The workplace is another element that creativity is dependent on it. Several attributes of the working climate can indeed influence the performance of individuals including but not limited to having the freedom to select methods for accomplishing the task, organisational encouragement, Supervisory encouragement, and having access to sufficient resources. Finally, the product is the

result of this interaction and it is recognized creative when it contains features that are both original (novel) and effective to the consumer.

Despite this simple definition; however, the process of measuring creativity is not a straightforward task. Creativity is a complex and dynamic concept since it appears in different forms and shapes depending on environment, and product's functionality. Meanwhile, the factors that contribute to the formation of creative thoughts are themselves distinctive within each step. For instance, the list of factors that have an effect on the person's creativity differs from those which define an appropriate environment that promotes creativity. As a result, the process of recognizing creativity is performed separately for each form of creativity. For the person, creativity is assessed by conducting self-report inventories which include asking individuals to rate themselves on a series of questions about personality characteristics. Clearly, these characteristics have been previously established as determinants of creativity by researchers who studied the background of the most creative individuals in the past. Creative processes are measured using divergent thinking tests which evaluate individuals' responses to a number of divergent thinking tasks based on four batteries including fluency (quantity), flexibility (variations among the solutions), originality, and elaboration (tendency to provide further details in the solution). Creative climates are assessed by scales developed in the form of a questionnaire. The aim of these questionnaires is to assess the quality of the working environment by several attributes such as workload pressure, team working, and having a task that is challenging.

Among the four aspects of creativity, product is an important form of creativity (Plucker & Makel, 2010). They are the result of the creative process which means no matter how efficient each step is, if the final product does not reflect the expected creativity then it simply suggests that one or maybe all of the influential components are lacking enough effectiveness. Consequently, measurement of product creativity is the most significant type of creativity measurement as it not only allows to understand the degree of creativity but also it represents the level of efficiency in other forms of creativity.

Chapter 3 reviews the current product creativity measurement methods. Three main categories of measurement methods were discussed including indirect measurement, global-judgments, and criterion-based techniques. The indirect measurement employs the assessment of product creativity as a means to investigate the background of the most creative individuals. These methods require experts in the domain to nominate the most creative person of their field while in their judgments, they may consider both quantity (number of creative works) and quality factors. Nevertheless, these methods are unable to provide useful information for the application of functional products since they are mainly focused on creative individual rather than creative product. The global-judgments; however, are applicable as they assume no specific definition for creativity but rather identify it by the judgments of experts in the domain. The consensual assessment technique is the most well-known approach in this category and has been regarded as the "gold standard" of creativity measurement (Baer & McKool, 2009b).

In the past, there have been studies where researchers used panel of judges to evaluate the creativity of products mostly produced in artistic domains (e.g. art, poetry, and musical compositions) and under tightly controlled conditions (Baer, 1994; Baer et al., 2009; Hickey, 2001; Kaufman, Baer, et al., 2008). Therefore, chapter 4 investigates the capability of the consensual technique to assess creativity for products produced under non-experimental conditions. In this study, 104 experts (in three groups) were invited to a web-based survey to judge the creativity of three categories of functional consumer products including home-accessories, electric-vehicles, and smoke alarms. The judges were also asked to rate the products on nine dimensions in addition to creativity (i.e., novelty, technical performance, safety, human factors, reliability, sustainability, attractiveness, emotion, and relevance). Initially, the results of inter-judge reliability indicated that CAT is a reliable method for products produced under dissimilar conditions; however, the outcomes of exploratory factor analysis revealed that for complex products even domain relevant experts are unable to separate the border between creativity and other factors. This simply implies that experts did not psychologically separate the judgment of creativity from the judgment of other factors such as product's technical performance and/or aesthetic appeal. As a result, the assessment of product creativity using the consensual technique for complex products should be performed with caution. In this study, the

outcomes of regression analysis revealed that the significant predictors of creativity are different between the products. This further proves the dynamic nature of creativity and signifies that depending on product's functionality, judges will choose different product attributes to make their decisions about creativity.

The last method whose capabilities for assessing functional consumer products have been considered in this research is criterion-based techniques (Chapter 3). These approaches are able to assess creativity by a set of constant variables and use of non-experts, a useful characteristic that allow for more affordable and quicker measurement. However, the dynamic nature of creativity influences the way viewers perceive creativity. It appears that depending on the functionality of the product that is in question, observers recognize creativity by only considering specific set of parameters whereas for other types of products they might choose to deliver their opinion with different factors. To solve this issue, researchers require to narrow their exploration and try to establish instruments that are specifically designed with relevant indicators based on each application.

In chapter 5, a theoretical framework for evaluating the creativity of functional consumer products is proposed (see Table 7-1). The aim of this framework is to assist researchers to develop their own unique criterion-based instrument depending on what product functionality is in question. The framework comprises 65 indicators from nine broad categories including *novelty*, *performance*, *safety*, *human-factors*, *sustainability*, *reliability*, *attractiveness*, *emotion*, and *relevance.*

Table 7-1 The theoretical framework for evaluating functional product creativity

Novelty	Performance	Safety
Conventional | Original	Non-functional | Functional	Does not protect | Protects
Usual | Unusual	Futile | Practical	Unsafe | Safe
Ordinary | Unique	Inefficient | Efficient	Hazardous | Non-hazardous
Average | Revolutionary	Incapable | Capable	Insecure | Secure
Frequent | Infrequent	Ineffective | Effective	Harmful | Harmless
Expected | Surprising	Unchanging | Adaptable	Painful | Innocuous
	Inflexible | Flexible	
	Fixed | Modifiable	
	Immutable | Alterable	
	Static | Changeable	

Human-Factors	Reliability	Sustainability
Difficult | Easy to use	Flimsy | Durable	Synthetic | Organic
Complicated | Simple	Weak | Strong	Unsustainable | Sustainable
Ambiguous | Comprehensible	Unsound | Sound	Non-recyclable | Recyclable
Complex | Straightforward	Fragile | Solid	Irrecoverable | Renewable
Inoperable | Operable	Frail | Robust	Environmentally harmful |
Unergonomic| Ergonomic	Unreliable | Reliable	Environmentally friendly
Intolerable | Tolerable	Questionable | Trustworthy	Pollutive | Pollution free
Uncomfortable | Comfortable	Poorly built | Well=built	

Attractiveness	Emotion	Relevance
Repulsive | Appealing	Disgusting | Desirable	Unnecessary | Necessary
Repelling | Charming	Dissatisfying | Satisfying	Irrelevant | Relevant
Coarse | Elegant	Unpleasant | Pleasant	Inessential | Essential
Crude | Refined	Contempt | Admirable	Inappropriate | Appropriate
Messy | Neat	Boring | Exciting	Unreasonable | Reasonable
Disordered | Arranged		Worthless | Valuable
Disorganised | Organised		Incompatible | Compatible
Poorly proportioned | Well-proportioned		
Imbalance | Balance		

In this chapter, the relationship between creativity and each of the factors is reviewed and a synthesis of the relevant literature on each criterion is presented. Novelty is the crucial factor for creativity which means no product can be considered creative without being detected as novel first. However, novelty on its own does not represent creativity. Indeed, an essential characteristic of the creative product is its ability to effectively solve user's problem. In this context, one of the influential factors is product's technical performance. Customers

continuously demand products that are able to perform tasks efficiently while providing them with possibility to complete their task in a variety of situations (adaptability). Overall, greater performance capabilities can enhance the usefulness and therefore, the current framework has included several items to assess product's performance capacity. Next is product safety which usually is referred to as a mandatory requirement by laws and/or regulations and hence is able to positively impact product's effectiveness. Another aspect that could positively influence the usefulness of the product is associated with how users communicate with the product. It is essential for consumers to be able to effortlessly interact with different parts of the product as this directly influences their overall level of satisfaction and hence can affect their perception of creativity. The current framework has included several items linked to human-factors to address this issue.

Today, manufacturers are under pressure to shift their production toward sustainable products by laws, policy makers, and market. Consumers seek to find products that are compatible with the environment to assure that their consumption does not negatively influence the ecological system. In the current framework, metrics of sustainability were included in order to measure viewers' perception on product's sustainability. Another factor whose existence in the product positively impacts usefulness is the ability of the product to perform its intended task with the same quality over a pre-defined time interval (reliability). This element allows users to trust the outcome of the product and hence use it more frequently without the worry of failure. On the other hand, aesthetic aspects of the product also contribute to creativity. In this thesis, aesthetics is defined as the ability of the product to stimulate observers from both areas of attractiveness and emotion. The amount of emotional impact that a product can have on consumer, in general, contributes on how the consumer perceive creativity. Emotional power of the product is indeed equivalent to sum of attributes that users find them important, pleasant, and satisfying in accordance with product's functionality (e.g., technical performance, safety, and/or reliability). The degree to which the product appears as appealing in the minds of the viewers also contributes to the overall product's effectiveness especially when other domain products are lacking this quality. Lastly, consumers need to consider that the solution is perfectly fitting the

situation before judging it for creativity. Thus, based upon how relevant and appropriate the solution is, observers may comprehend different values for creativity.

In the real-world scenario, however, occasionally certain factors are combined with each other and instead appear in the form of a new and unique attribute dependent on what product functionality is in question. In this context, the theoretical framework proposed in this thesis allows the researchers to identify unique (factor) structures that are more precise in the sense that it can illustrate exactly how viewers recognise creativity in a particular product.

In chapter 6, the ability of the theoretical framework to generate a specific instrument for the creativity assessment of a particular category of consumer products was investigated. In this study, 288 university students judged five lounge chairs using the 65 indicators introduced in the framework. An excellent inter-rater reliability between the raters was achieved, indicating that researchers do not require to perform any formal training before conducting assessment using the framework proposed. Factor analysis revealed a five-factor model in which some constructs had almost identical indicators as similar as they proposed in the original framework, and others had items from multiple factors. All constructs have been examined for internal consistency using the Cronbach's alpha coefficient. 'Good' to 'excellent' inter-item reliabilities were attained indicating that items under each construct are indeed measuring the same concept and that their measurement is expected to be consistent over time. In this study, a stepwise regression analysis was conducted to further inspect the capability of the new instrument in effectively anticipating viewers' viewpoints on creativity. Results revealed that 4 out of 5 factors of the new instrument are able to significantly predict the judgments of creativity. Overall, this study showed a successful application of the theoretical framework in generating a reliable and factorially valid criterion-based instrument for the assessment of lounge chairs. Researchers and practitioners can now utilize this framework to develop their own creativity instruments depending on what product functionality is in question. Once a unique model is generated and its reliability established in the domain there is no need to repeat the study using a large dataset. Instead, a smaller panel of judges (e.g. 100 raters) can be utilized to carry out the assessment and because the framework is designed with internal consistency in mind there would be no need to train the raters. This dramatically reduces the cost, energy, and time consumption and provide researchers

in the relevant fields with a powerful tool to assess creativity in a much more efficient and precise way. In fact, every time that a new model is born, it will be a simpler version of the original model (see chapter 5, section 5.2) as it would only contain relevant criteria and indicators with respect to the area that is in question. The reduction in the number of indicators, as well as the simplification of the underlying structure together, would make it simpler for raters to understand the questions and help to reduce the amount of time required for answering the questions and hence enhance efficiency.

7.2. Guidelines for implementing the measurement approaches

This thesis showed a reliable assessment for functional products produced under non-experimental conditions. In addition, a theoretical framework was proposed allowing researchers to develop their own unique criterion-based instruments considering the functionality of the product that is in question. However, in order to be consistent with the outcomes of this thesis, researchers are recommended to take into consideration the following notes before conducting any assessment.

In general, this thesis supports the use of two types of measurement including consensual assessment technique and criterion-based instruments. However, depending on the level of information that researchers intend to extract from the assessment as well as considering several influential factors, they are advised to choose one of the three approaches presented in this guideline. Figure 7-1 shows the flow diagram for the selection of appropriate measurement approach based on the researcher's intention and other influential criteria. First, researchers need to take a new product and examine its functionality with other similar products in the domain. The product is novel if its functionality does not exist; at this stage, creativity cannot be measured as the product's usefulness (effectiveness) is undefined. Alternatively, researchers require to identify a number of competing artefacts to allow comparison. When selecting the artefacts, an extra attention must be paid to ensure that an adequate level of diversity exist among the items. A lower level of diversity may artificially inflate the inter-judge reliability whereas an intense level of diversity may raise the issue of fatigue and hence the growth of error in the measurement. Unfortunately, there is no clear instruction in the literature showing the

right approach for selecting the stimuli for the assessment. Overall, larger number of products (not extremely large) can improve the probability of capturing normally distributed data which further can strengthen the precision in the measurement. Once products are identified, researchers need to decide on the level of information that they mean to extract from the measurement. If they aim to only define the product's degree of creativity and have access to domain experts, then they are advised to select the consensual assessment technique. Otherwise, if the access to experts is restricted, researchers should choose among one of the existing criterion-based instruments such as CPSS, CSDS, and PCMI (see chapter 3, section 3.4).

To use CAT, researchers are advised to follow Amabile's (1982) recommended procedures for a reliable assessment (also see chapter 3, section 3.3), however, other researchers over time have revised and updated some of its protocols (Baer & McKool, 2009a; Cseh & Jeffries, 2019; Kaufman, Plucker, et al., 2008). For instance, researchers tested different types of raters as judges to see if the need for experts can be eliminated, however, almost all studies yet support the idea of using experts as judges and argue that the validity of CAT highly depends on it (Baer et al., 2009; Byrne et al., 2003; Kaufman, Plucker, et al., 2008). Others investigated the perfect size of the panel of judges and in this respect, Kaufman, Plucker, et al. (2008) advise that for most studies a panel consisting of between 5 to 10 judges is sufficient. To enhance measurement reliability, researchers also reviewed different types of scales and suggested the optimal scale to be consisted of between 5 to 7 points (Preston & Colman, 2000). This study investigated the reliability of the CAT for the assessment of non-parallel functional products and results indicated reliable outcomes, however, a larger panel of judges is recommended as discussed in chapter 4. The results of this thesis indicated that a reliable panel of judges (for non-parallel functional products) should include approximately between 10-15 judges. It is clear that as the degree of complexity in the product increases the greater panel of judges is required for a reliable assessment (complexity means a product consists of too many components with a high degree of diversity. Also, the connectivity between the components is complicated (see page 88 for more information)). For such products, researchers are suggested to break down the product into constituent systems and then carry out the measurement separately for each system using appropriate judges per section. After identifying the judges, all products should be presented to

judges in random order to allow them to observe and compare the products to one another before judging them for creativity. Once the judgment is performed, the data must be analysed for inter-rater reliability as suggested by researchers (Amabile, 1982). If acceptable reliability was achieved, the average scores for creativity can be computed and products can be ranked from a lower to higher creativity. Alternatively, criterion-based instruments can be used if access to domain relevant experts is restricted. To achieve a reliable assessment, a minimum of 100 non-experts is recommended (Kaufman, Baer, et al., 2008). Similarly, all products must be presented to judges in random order before starting the assessment. Judges should then be asked to rate each product on the list of indicators provided within the instrument. Once the process of data collection is completed, the data must be analysed for inter-judge reliability and if acceptable reliability is achieved then creativity can be computed by once calculating the average scores for each indicator and then for the entire items respectively. At last, products should be ranked from a lower to higher creativity.

On the other hand, if researchers prefer to extract certain information from the measurement to better understand creativity and to be able to implement it within the product development process; they are advised to use the theoretical framework proposed in chapter 5. Under this condition, researchers first need to identify an appropriate panel of judges by selecting a sample of targeted consumers. This sample must have adequate size to allow performing reliable and valid factor analysis. Considering the rule of thumbs for sample size selection (incl. a minimum sample size of 150-300 (Hutcheson & Sofroniou, 1999), and/or the minimum ratio of sample size to measured variables of 10:1 (Bryant & Yarnold, 1995)) as well as the number of indicators in the theoretical framework, a minimum of 220 participants is recommended (the procedure for calculating the minimum sample size is as follow: 65 X 10 = 650 respondents. Considering that each judge should assess a minimum of three products then this value (650) should be divided into three; Hence, 650 / 3 = 216.66. As a result, a minimum of 220 participants should be considered for utilising the theoretical framework). In general, the greater the panel of judges the better and more reliable the outcomes of the factor analysis and assessment. Once an appropriate panel of judges is selected, researchers may present all the products to judges in random order to allow them to observe and compare the products to one another. Next, judges

should be asked to rate each product on the 65 indicators proposed in the theoretical framework (see Table 7-1). Then, the data must be analysed for inter-rater reliability and if acceptable reliability is achieved, researchers may combine the data and conduct a factor analysis to further obtain a unique measurement instrument representative of how consumers perceive creativity for the product that is in question. To obtain a reliable model, all criteria for an effective factor analysis must be followed as discussed in chapter 5 (see chapter 5, section 5.4). Once a unique model obtained, creativity can be computed by calculating the average scores for each indicator left in the model and then computing the average score of the entire dataset. Finally, products should be ranked from a lower to higher creativity.

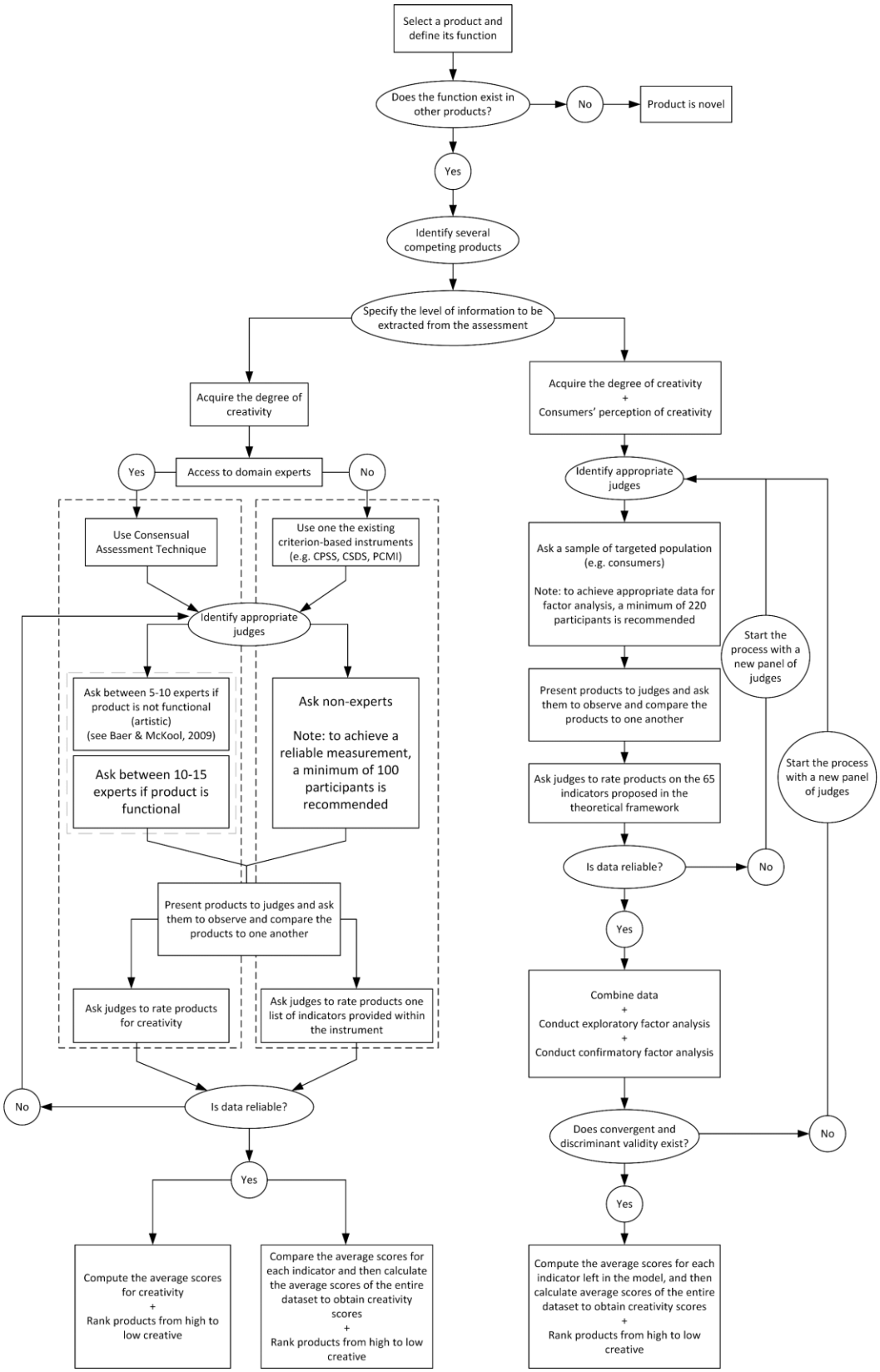


Figure 7-1 Guidelines for the measurement of product creativity

7.3. Directions for future research

The research in this thesis suggests several new directions for future studies. These include further examination of the current application of the existing measurement tools for the assessment of functional (engineering) products. This research showed that CAT is a reliable method for assessing products produced under distinct conditions. This opens new doors for research on creativity in engineering domains by making it possible for researchers to examine the efficiency of each stage of the creative process. For instance, the creativity of products produced under different environmental conditions by the same individual(s) can be assessed with the aim to understand what environmental factors impact the overall creativity positively.

The outcomes of the current study indicated that the usage of CAT with complex products (i.e. products with multiple internal components, complex connectivity between different elements, and complicated management system) should be further investigated. Although, in this study experts were able to reliably assess the creativity of non-experimental products at home-accessories and smoke-alarm groups, however, the assessment of electric vehicles revealed that even expert judges are having trouble to distinguish the border between creativity and other influential factors. Results of this study still confirm the CAT as a powerful tool for creativity assessment of non-parallel products but suggest for more investigations to be done in this area to fully understand its capabilities over functional products produced under non-experimental conditions and especially those that possess complex structures. Although one strategy to resolve the issue is to first decompose the product into its constituent systems, an alternative solution could be to explore whether different protocols for data gathering can be used to enhance measurement quality. For instance, a comparison can be arranged between once gathering data using conventional questionnaires (i.e., using surveys and two-dimensional images to gather experts' opinions about creativity), and once by physically presenting the products to experts and see whether their overall engagement and reliability is influenced.

In the second part of this thesis, a theoretical framework for generating new criterion-based instruments based on products' functionality was presented. Additional empirical studies are needed to further verify the applicability and compatibility concerning this foundational model.

Such study could be to apply the model into as many different applications as possible and compare the results of measurement with those generated by already existing methods. Also, an attempt could be done to optimise the indicators used within this model by either refining them and/or adding in new items. Another study may take a longitudinal approach to explore the differences between factor patterns produced for a specific product application to see whether consumer perception of creativity is changing with time.

7.4. Chapter conclusion

Creativity is a critical concept for innovative firms in that it provides the basis for innovation and smoothen the path for further enhancing growth and profitability. However, in order to implement creativity in the product development process, companies need to know how to recognise creativity from the contrary. Although creativity emerges from the interaction between individuals, their thinking process, and press; yet the product of this interaction is identifiable without the consideration of the process itself. In fact, the description of product creativity does not rely on any of the earlier components, but rather it depends on the degree of novelty and usefulness. Creative solutions must be unique and original while they should be able to effectively fulfil user demands. Nevertheless, an absolute definition of product creativity does not exist. Creativity is a dynamic concept and hence any effort for defining it within a closed boundary is indeed in contrast with the essence of it. For this reason, the majority of assessments in this context are performed by subjective judgments of appropriate evaluators. In the past, the consensual assessment technique has been employed to assess product creativity mostly in artistic domains such as visual art, poetry, and musical compositions where artefacts were produced under tightly controlled conditions (Baer, 1994; Baer et al., 2009; Hickey, 2001; Kaufman, Baer, et al., 2008; K. Tsai, 2016). This thesis investigated the reliability of using the panel of experts for evaluating the creativity of non-parallel functional product (i.e., products produced under non-experimental conditions). The results showed that CAT is a reliable method for these types of products. However, researchers are advised to use a relatively larger panel of experts (10-15) compared to those reported in artistic domains (5 to 10). In addition, results of factor analysis indicated that experts are incapable to distinguish the distinction between creativity and other factors for complex products. A complex product is a product where multiple products (or

systems) are integrated into one product and often it is capable of solving multiple problems at a time. This thesis suggests that the creativity assessment of complex products by means of experts as judges to be carried out with caution.

In the second part of this thesis, a theoretical framework was proposed to enable the researchers and practitioners to develop their own unique criterion-based instruments considering the functionality of the product that is in question. A case study was conducted asking 288 university students to judge a series of lounge chairs making use of 65 indicators proposed in the framework. High inter-judge reliability was achieved indicating that no formal training is required and that novice raters agree with each other on their ratings. Factor analysis revealed a five-factor model including Aesthetics and Relevance, Novelty, Versatility, Ecological footprint, and Reliability. The internal consistency between the items within each construct achieved at a "good" to "excellent" range indicating that the outcomes are valid, and they are expected to return the same outcomes over time. The five-factor model has both convergent and discriminant validity as tested by confirmatory factor analysis. Overall, this model indicated the reliability and applicability of the theoretical framework for generating a unique measurement instrument for the application of lounge chairs. Researchers can now utilise the theoretical framework proposed in this thesis to better understand how certain viewers (consumers) perceive creativity for a particular product. For instance, researchers and practitioners in the marketing sector who intend to study (product) creativity would benefit from this framework as they could employ it to create a unique criterion-based instrument for the area of their interest and use it to assess creativity of whole products in that territory. Once a unique model is generated and its reliability is established there is no need to repeat the study using a large panel and instead a smaller panel of judges (e.g. 100 novice raters) can be formed to evaluate the products without training them. In addition, any model would contain a list of indicators which are highly correlated with creativity with respect to the domain that is in question and thus product designers would be able to get insights and enhance creativity by paying more attention to these criteria during the product development process.

References

Altshuller, G. (1999). *The Innovation Algorithm: TRIZ, Systematic Innovation and Technical Creativity* (L. Shulyak & S. Rodman, Eds.). Worcester, Massachusetts: Technical Innovation Center, Inc.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, *43*(5), 997–1013. https://doi.org/10.1037/0022-3514.43.5.997

Amabile, T. M. (1983). The Social psychology of creativity:A Componential Conceptualization. *Journal of Personality and Social Psychology*, *45*(2), 357–376.

Amabile, T. M. (1988). A Model of Creativity and Innovation in Organizations. *Research in Organizational Behavior*, *10*, 123–167.

Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, Colorado, Colorado: Westview Press.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). Assessing the work environment for creativity. *Academy of Management Journal*, *39*(4), 1154–1184.

American Marketing Association. (1948). Report of the definitions committee. *Journal of the Academy of Marketing Science*, *13*(6), 202–217.

Anderson, J., & Gerbing, D. (1988). Structural Equation Modelling in Practice: A Review and Recommended Two Step Approach. *Psychological Bulletin*, *103*, 411–423.

Anderson, N. R., & West, M. A. (1998). Measuring climate for work group innovation: Development and validation of the team climate inventory. *Journal of Organizational Behavior*, *19*(3), 235–258.

Arthur, B., & VanGundy, J. (1988). *Techniques of Structured Problem Solving* (2end ed.).

Springer.

Baer, J. (1993). *Creativity and divergent thinking: A task-specific approach*. New Jersey: Lawrence Erlbaum Associates.

Baer, J. (1994). Generality of creativity across performance domains: A replication. *Perceptual and Motor Skills*, *79*, 1217–1218.

Baer, J., & Frese, M. (2003). Innovation is not enough: climates for initiative and psychological safety, process innovations, and firm performance. *Journal of Organizational Behavior*, *24*(1), 45–66.

Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the Consensual Assessment Technique to Nonparallel Creative Products. *Creativity Research Journal*, *16*(1), 113–117. https://doi.org/10.1207/s15326934crj1601_11

Baer, J., Kaufman, J. C., & Riggs, M. (2009). Brief Report: Rater-Domain Interactions in the Consensual Assessment Technique. *The International Journal of Creativity & Problem Solving2*, *19*(2), 87–92.

Baer, J., & McKool, S. S. (2009a). Assessing Creativity Using the Consensual Assessment Technique. *Handbook of Research on Assessment Technologies, Methods, and Applications in Higher Education*, 65–77. https://doi.org/10.4018/978-1-60566-667-9.ch004

Baer, J., & McKool, S. S. (2009b). The Gold Standard for Assessing Creativity. *International Journal of Quality Assurance in Engineering and Technology Education*, *3*(1), 81–93.

Bangor, A., Kortum, P., & Miller, J. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594.

Barron, F. (1963). *Creativity and Psychological Health*. Oxford: Van Nostrand.

Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). Latent variable models and factor analysis: A unified approach. In *Book*. Chichester: Wiley.

Baxter, M. (1995). *Product design: a practical guide to systematic methods of new product*

development. New York: Chapman & Hall.

Baziuk, W. (1995). BNR/NORTEL: Path to Improve Product Quality, Reliability, and Customer Satisfaction. *The 6th International Symposium of Software Reliability Engineering*, 256–263. Toulouse, France.

BCG. (2018). The Most Innovative Companies in 2018. *Boston Consulting Group*, 4–28.

Bentler, P. (1990). Comparitive Fit Indexes in Structural Models. *Psychological Bulletin*, *107*(2), 238–246.

Besemer, S. (1984). How Do You Know It's Creative? *Gifted Child Today*, *7*(2), 30–35.

Besemer, S. (1998). Creative Product Analysis Matrix: Testing the model structure and a comparison among products_Three novel chairs. *Creativity Research Journal*, *11*, 333–346.

Besemer, S. (2000). To buy or not to buy: Predicting the willingness to buy from creative product variables. *Korean Journal of Thinking and Problem-Solving*, *10*(2), 5–18.

Besemer, S., & O'Quin, K. (1986). Analyzing creative products: Refinement and test of a judging instrument. *Journal of Creative Behavior*, *20*, 115–126.

Besemer, S., & O'Quin, K. (1987). Creative product analysis: Testing a model by developing a judging instrument. In S. Isaksen (Ed.), *Frontiers of creativity research: Beyond the basis* (pp. 341–357). Buffalo, NY: Bearly Ltd.

Besemer, S., & O'Quin, K. (1999). Confirming the Three-Factor Creative Product Analysis Matrix Model in an American Sample. *Creativity Research Journal*, *12*(4), 287–296.

Besemer, S., & Treffinger, D. J. (1981). Analysis of Creative Products: review and synthesis. *Journal of Creative Behavior*, *15*, 158–178.

Bland, J., & Altman, D. (1997). Statistics notes: Cronbach's alpha. *BMJ*, 314–572.

Bloch, P., Brunel, F., & Arnold, T. (2003). Individual differences in the centrality of visual product aesthetics: concept and measurement. *Journal of Consumer Research*, *29*, 551–565.

Boden, M. (1991). *The Creative Mind: Myths and Mechanisms*. New York: Basic Books.

Boden, M. (2003). *The creative mind: Myths and mechanisms*. London, UK: Routledge.

Brinkman, D. (1999). Problem finding, creativity style and musical compsitions of high school students. *Journal of Creative Behavior*, *33*(1), 62–68.

Bruner, J. (1962). *On Knowing essays for the left hand*. London: Harvard University Press.

Bryant, F., & Yarnold, P. (1995). Principle components analysis and exploratory and confirmatory factor analysis. In L. Grimm & R. Yarnold (Eds.), *Reading and Understanding Multivariate Statistics* (pp. 99–136). Washington, DC: American Psychology Association.

Byrne, C., MacDonald, R., & Carlton, L. (2003). Assessing creativity in musical compositions: flow as an assessment tool. *British Journal of Music Education*, *20*(3), S0265051703005448. https://doi.org/10.1017/S0265051703005448

Cagan, J., & Vogel, C. (2002). *Creating Breakthrough Products: Innovation from Product Planning to Program Approval*. Upper Saddle River, NJ: FT Press.

Cappelleri, J., Zou, K., Bushmakin, A., Alvir, J., Alemayehu, D., & Symonds, T. (2014). *Patient-Reported Outcomes: Measurement, Implementation, and Interpretation*. Boca Raton: Chapman & Hall/CRC.

Carifio, J., & Perla, R. (2007). Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. *Journal of Social Sciences*, *3*(3), 106–116.

Carson, S. *Creativity and mental illness*. , (2006).

Carson, S., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the Creative Achievement Questionnaire. *Creativity Research Journal*, *17*, 37–50.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychomrika*, *12*, 197–220.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. (1970). *Handbook for the Sixteen personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.

Chakrabarti, A., & Khadilkar, P. (2003). A Measure for Assessing Product Novelty. *International Conference on Engineering Design ICED*, 19–21. Stockholm.

Chakrabarti, A., Sarkar, P., Leelavathamma, B., & Nataraju, B. (2005). A functional representation for aiding biomimetic and artifical inspiration of new ideas. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *19*, 113–132.

Charlotte, F., & Peter, F. (2005). *1000 Chairs*. New York: Taschden.

Chen, C., Kasof, J., Himsel, A. J., Greenberger, E., Dong, Q., & Xue, G. (2002). Creativity in drawings of geometric shapes: A cross-cultural examination with the consensual assessment technique. *Journal of Cross-Cultural Psychology*, *33*(2), 171–187. https://doi.org/10.1177/0022022102033002004

Chen, M. H., & Kaufmann, G. (2008). Employee creativity and R&D: A critical review. *Creativity and Innovation Management*, *17*(1), 71–76. https://doi.org/10.1111/j.1467-8691.2008.00471.x

Childs, P., & Fountain, R. (2011). Commercivity. *International Conference on Engineering and Product Design Education*. London, UK: City University.

Childs, P. R. N. (2014). *Mechanical Design Engineering Handbook*. London: Butterworth-Heinemann.

Childs, P. R. N. (2018). *Private communication*. London.

Childs, P. R. N., Hamilton, T., Morris, R., & Johnston, G. (2006). Centre for Technology Enabled Creativity. *4th Engineering and Product Design Education Conference*, 367–372. Salzburg, Austria.

Christiaans, H. (2002). Creative as a Design Criterion. *Creativity Research Journal*, *14*(1), 41–54.

Christiaans, H., & Venselaar, K. (2005). Creativity in design engineering and the role of knowledge: Modelling the expert. *International Journal of Technology and Design Education*, *15*(3), 217–236. https://doi.org/10.1007/s10798-004-1904-4

Clapham, M. M. (2004). The convergent validity of the Torrance Tests of Creative Thinking and creativity interest inventories. *Educational and Psychological Measurement*, *64*(5), 828–841.

Coakes, S. J., & Steed, L. G. (1999). *SPSS: Analysis without Anguish*. Australia: John Wiley and Sons Ltd.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sicences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Colangelo, N., Kerr, B., Hallowell, K., Huesman, R., & Gaeth, J. (1992). The Iowa Inventiveness Inventory: Toward a measure of mechanical inventiveness. *Creativity Research Journal*, *5*, 157–163.

Comrey, A, & Lee, H. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Comrey, AL, & Lee, H. (1973). *A first course in factor analysis*. New York: Academic Press.

Copeland, M. (1924). *Principles of Merchandising*. Chicago: A.W.Shaw Co.

Costa, P., & McCrae, R. (1985). *The NEO personality Inventory Manual*. Odessa: Psychological Assessment Resources.

Costa, P., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI): Professional manual*. Odessa: Psychological Assessment Resources.

Couger, D., Higgins, L. F., & McIntyre, S. C. (1993). (Un)Structured creativity in information systems organizations. *MIS Quarterly*, *17*(4), 375–398.

Cox, G. (2005). *Cox Review of Creativity in Business: Building on the UK's Strengths*. London: HM Treasury, Executive Summary.

Cramond, B., Matthews-Morgan, J., Bandalos, D., & Zuo, L. (2005). A report on the 40-years follow-up of the Torrance Tests of Creative Thinking: Alive and well in the new millennium. *Gifted Child Quarterly*, *49*, 283–291.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrica*, *16*, 297–334.

Cropley. (2000). Defining and measuring creativity: Are creativity tests worth using? *Roeper Review*, *23*(2), 72–79.

Cropley, D. (2016). Creativity in Engineering. In G. E. Corazza & S. Agnoli (Eds.), *Multidisciplinary Contributions to the Science of Creative Thinking* (pp. 155–173). Singapore: Springer.

Cropley, D., & Cropley, A. (2005). Engineering creativity: A systems concept of functional creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity Across Domains: Faces of the Muse* (pp. 169–185). https://doi.org/10.4324/9781410611925

Cropley, D., & Cropley, A. (2008). Elements of a universal aesthetic of creativity. *Psychology of Aesthetics, Creativity and the Arts*, *2*, 155–161.

Cropley, D., & Kaufman, J. C. (2012). Measuring Functional Creativity: Non-Expert Raters and the Creative Solution Diagnosis Scale. *Journal of Creative Behavior*, *46*(2), 119–137.

Cropley, D., Kaufman, J., & Cropley, A. (2011). Measuring Creativity for Innovation Management. *Journal of Technology Management & Innovation*, *6*(3), 13–30.

Cseh, G., & Jeffries, K. (2019). A Scattered CAT: A Critical Evaluation of the Consensual Assessment Technique for Creativity Research. *Psychology of Aesthetics Creativity and the Arts*, *13*(2), 159–166.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the Psychology of Discovery and Invention*. New York: HarperCollins.

Csikszentmihalyi, M., & Sawyer, R. K. (2014). Shifting the Focus from Individual to Organizational Creativity. In *The Systems Model of Creativity* (pp. 67–71). New York: Springer.

Csikszentmihalyi, M., & Wolfe, R. (2000). New conceptions and research approaches to creativity: Implications of a systems perspective for creativity in education. *International Handbook of Giftedness and Talent*, *2*, 81–91.

Daly, S., Seifert, C., Yilmaz, S., & Gonzalez, R. (2016). Comparing ideation techniques for beginning desingers. *Journal of Mechanical Design2*, *138*(10), 101–108.

Danneels, E., & Kleinschmidt, E. (2001). Product innovativeness from the firm's perspecrtive: its dimensions and their relation with project selection and performance. *Journal of Product Innovation Management*, *18*(6), 357–373.

DeCoster, J. (1998). Overview of Factor Analysis. Retrieved October 27, 2018, from http://stat-help.com/factor.pdf

Demirbilek, O., & Sener, B. (2003). Product design, semantics and emotional response. *Ergonomics*, *46*, 1346–1360.

Design Council. (2018). The Design Process: What is the Double Diamond? Retrieved October 22, 2018, from https://www.designcouncil.org.uk/news-opinion/design-process-what-double-diamond

Dodds, R., Ward, T., & Smith, S. (2003). A review of experimental literature on incubation in problem solving and creativity. In M. Runco (Ed.), *Creativity Research Handbook (Vol. 3)*. Cresskill, NJ, NJ: Hampton Press.

Duckworth, A. L., & Quinn, P. (2009). Development and Validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, *91*(2), 166–174.

Dym, C. L. (2006). Engineering design: So much to learn. *International Journal of Engineering Education*, *22*(3), 422–428. https://doi.org/10.3362/0262-8104.1984.013

Eberle, R. (1971). *SCAMPER: Creative games and activities for imagination development*. Buffalo, NY: DOK Publishing.

EC, & EP. (2002). Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety. *Official Journal of the European Communities*, 14. https://doi.org/10.1016/j.jclepro.2010.02.014

Ekong, G., Long, C., & Childs, P. (2013). Application of creativity tools to Gas Turbine Engine Compressor Clearance Control. *Proceeding of ASME International Mechanical Engineering*

Congress and Exposition. California, USA.

Ekvall, G. (1996). Organizational climate for creativity and innovation. *European Journal of Work and Organizational Psychology*, *5*(1), 105–123.

Eric, O. (2003). Manifestation of divergent-convergent thinking in question asking and decision making processes of design teams: A performance dimension. In U. Lindemann (Ed.), *Human behaviour in design: Individuals, teams, tools*. Berlin: Springer.

Everitt, B. (1975). Multivariate analysis: The need for data, and other problems. *British Journal of Psychiatry*, *126*, 237–240.

Farrell, A. (2010). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009). *Journal of Business Research*, *63*, 324–327.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). London: SAGE Publications.

Ford, C. (1996). Theory of individual creative action in multiple social domains. *Academy of Management Review*, *21*(4), 1112–1142.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39–50.

Freeman, C., Son, J., & McRoberts, L. B. (2015). Comparison of noivce and expert evaluations of apparel design illustrations using the Consensual Assessment Technique. *International Journal of Fashion Design, Technology and Education*, *8*(2), 122–130. https://doi.org/10.1080/17543266.2015.1018960

French, M. (1985). *Conceptual design for engineers*. London: The Design Council.

Galenson, D. (2001). *Painting outside the lines*. Cambridge: Harvard University.

Gardner, H. (1989). *To open minds*. New York: Basic Books.

Gatignon, H. (2010). *Confirmatory Factor Analysis*. New York: Springer.

Gelernter, D. (1998). *Machine beauty: Elegance and the heart of Technology* (1st ed.). New

York: Basic Books.

Goldenberg, J., & Mazursky, D. (2002). *Creativity in product innovation*. Cambridge, UK: Cambridge University Press.

Götz, K. O., & Götz, K. (1979). Personality characteristics of successful artists. *Perceptual and Motor Skills*, *49*, 919–924.

Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of Personality and Social Psychology*, *37*, 1398–1405.

Grabner-Kräuter, S. (2018). Consumer Goods. In R. Kolb (Ed.), *The SAGE Encyclopedia of Business Ethics and Society* (pp. 641–642). Thousand Oaks: SAGE Publications, Inc.

Griffin, R. W. (1983). Objective and social sources of information in task redesign: A field experiment. *Administrative Science Quarterly*, *28*, 184–200.

Guastello, S. (2009). Creativity and personality. In T. Richards, M. Runco, & S. Moger (Eds.), *The Routledge Companion to Creativity* (pp. 267–278). New York: Routledge.

Guilford, J. P. (1950). Creativity research: Past, present, and future. *American Psychologist*, *5*, 444–454.

Gupta, S., Dangayach, G., & Singh, A. (2015). Key Determinants of Sustainable Product Design and Manufacturing. *12th Global Conference on Sustainable Manufacturing*, 99–102.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1995). *Multivariate data analysis* (4th ed.). Saddle River, NJ: Prentice Hall Inc.

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis: A global prespetive* (7th ed.). Saddle River, NJ: Pearson Education.

Harold, S., Balaban, N. C., & Pecht, M. (2009). Product Effectiveness and Worth. In P. M (Ed.), *Product Reliability, Maintainability, and Supportability Handbook* (pp. 1–18). London: Taylor & Francis Group.

Hebert, T. P., Cramond, B., Neumeister, K. L. S., Miller, G., & Sivian, A. F. (2002). *E. Paul*

Torrence: His life, accomplishments, and legacy (The nation). Storrs, CT: The University of Connecticut.

Helander, M. (2006). *A guide to Human Factors and Ergonomics* (2nd ed.). London: Taylor & Francis Group.

Hennessey, B. A. (2010). The Creativity-Motivation Connection. In J. C. Kaufman & R. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 342–365). New York: Cambridge University Press.

Hennessey, B., & Amabile, T. M. (1988). Story-telling: A method for assessing children's creativity. *Journal of Creative Behavior*, *22*(4), 235–246.

Hennessey, B., Amabile, T. M., & Mueller, J. (2011). Consensual Assessment. In M. Runco & S. Pritzker (Eds.), *Encyclopedia of Creativity2*. Boston.

Hennessey, B., Gia, K., Guomin, Z., & Weiwei, S. (2008). A multi-cultrual application of the Consensual Assessment Technique. *The International Journal of Creativity & Problem Solving*, *18*(2), 87–100.

Hickey, M. (1995). *Qualitative and quantitative relationships between children's creative musical thinking processes and products.* Northwestern University, Evanston, Illinois.

Hickey, M. (2001). An application of Amabile's consensual assessment technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, *49*(3), 234–244. https://doi.org/10.2307/3345709

Hicks, M. J. (1991). *Problem Solving in Business and Management* (1st ed.). London: Chapman & Hall.

Hocevar, D. (1980). Intelligence, Divergent Thinking, and Creativity. *Intelligence*, *4*, 25–40.

Hocevar, Daniel, & Bachelor, P. (1989). A taxonomy and critique of measurements used in the study of creativity. In J. Glover, R. Ronning, & C. Reynolds (Eds.), *Handbook of creativity* (pp. 53–75). New York: Plenum Press.

Hocevar, Dennis. (1979). The Development of the Creative Behavior Inventory (CBI). *Conference of the Rocky Mountain Psychological Association*. Las Vegas, Nevada.

Horn, D., & Salvendy, G. (2006a). Consumer-Based Asssessment of Product Creativity: A Review and Reappraisal. *Human Factors and Ergonomics in Manufacturing*, *16*(2), 155–175. https://doi.org/10.1002/hfm

Horn, D., & Salvendy, G. (2006b). Product creativity: conceptual model, measurement and characteristics. *Theoretical Issues in Ergonomics Science*, *7*(4), 395–412.

Horn, D., & Salvendy, G. (2009). Measuring Consumer Perception of Product Creativity: Impact on Satisfaction and Purchasability. *Human Factors and Ergonomics in Manufacturing*, *19*(3), 223–240.

Howard, T., Culley, S., & Dekoninck, E. (2008). Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Design Studies*, *29*, 160–180.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Thousand Oaks: SAGE Publications.

IBM. (2010). Capitalizing on Complexity. *IBM Global Services*, *1518*(3), 1–75. https://doi.org/10.2190/HFLG-14N9-KF8L-4FMD

IEEE. (2018). What is an Electrical Engineering. Retrieved October 22, 2018, from https://www.ieee.org/content/dam/ieee-org/ieee/web/org/educ/IEEE_Career_brochure_3.1.07.pdf

IMechE. (2018). Definition of Mechanical Engineering. Retrieved October 22, 2018, from https://www.imeche.org/membership-registration/become-a-member/definition-of-mechanical-engineering

Isaksen, S. G., & Ekvall, G. (2007). *Assessing the Context for Change: A Technical Manual for the*

SOQ - Enhancing Performance of Organizations, Leaders and Team for over 50 Years (2nd ed.). Orchard Park, NY: The Creative Problem Solving Group, Inc.

Isaksen, S. G., & Ekvall, G. (2010). Managing of Innovation: The Two Faces of Tension in Creative Climates. *Creativity and Innovation Management*, *19*(2), 73–88.

Jones, C. J. (1992). *Design methods: Seeds of human futures*. London: John Wiley.

Jöreskog, K., & Sörbom, D. (1984). *Lisrel VI. Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods*. Mooresville, Indiana: Scientific Software.

Kanter, R. (1983). *The change masters*. New York: Simon and Schuster.

Kanter, R. (1988). When a thousand flowers bloom: Structural, collective and social conditions for innovation in organizations. *Research in Organizational Behavior*, *10*, 169–211.

Karwowski, W., Soares, M., & Stanton, N. (2011). *Human Factors and Ergonomics in Consumer Product Design* (2nd ed.). London: Taylor & Francis Group.

Kaufman, J. C. (2012). Counting the Muses: Development of the Kaufman Domain of creativity Scale (K-DOCS). *Psychology of Aesthetics*, *6*(4), 298–308.

Kaufman, J. C., & Baer, J. (2004). Sure, I'm creative-But not in mathematics! *Empirical Studies of the Arts*, *22*, 143–155.

Kaufman, J. C., & Baer, J. (2012). Beyond new and appropriate: Who decides what is creative? *Creativity Research Journal*, *24*(1), 83–91.

Kaufman, J. C., Baer, J., & Cole, J. C. (2009). Expertise, domains, and the consensual assessment technique. *Journal of Creative Behavior*, *43*(4), 223–233. https://doi.org/10.1002/j.2162-6057.2009.tb01316.x

Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the Consensual Assessment Techqnique. *Creativity Research Journal*, *20*(2), 171–178. https://doi.org/10.1080/10400410802059929

Kaufman, J. C., Baer, J., & Gentile, C. A. (2004). Differences in Gender and Ethnicity as Measured by ratings of Three Writing Tasks. *Journal of Creative Behavior*, *38*(1), 56–69.

Kaufman, J. C., Beghetto, R., & Ronald, A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, *13*(1), 1–12.

Kaufman, J. C., Cole, J. C., & Baer, J. (2009). The construct of creativity: A structural model for self-reported creativity ratings. *Journal of Creative Behavior*, *43*, 119–134.

Kaufman, J. C., Lee, J., Baer, J., & Lee, S. (2007). Captions, consistency, creativity, and the consensual assessment technique: New evidence of reliability. *Thinking Skills and Creativity*, *2*(2), 96–106. https://doi.org/10.1016/j.tsc.2007.04.002

Kaufman, J. C., Niu, W., Sexton, J. D., & Cole, J. C. (2010). In the eye of the beholder: Differences across ethnicity and gender in evaluating creative work. *Journal of Applied Social Psychology*, *40*(2), 496–511.

Kaufman, J. C., Plucker, J. A., & Baer, J. (2008). *Essentials of creativity assessment*. New York: Wiley.

Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal*, *18*(1), 3–14.

King, L. A., McKee Walker, L., & Broyles, S. (1996). Creativity and the five-factor model. *Journal of Research in Personlaity*, *30*(2), 189–203.

Kline, P. (1979). *Psychometrics and psychology*. London: Academic Press.

Kline, P. (1994). *An easy guide to factor analysis*. New York: Routledge.

Kline, R. (2005). *Principles and Practice of Structural Equation Modelling* (2nd ed.). New York: The Guilford Press.

Kozbelt, A., Beghetto, R., & Runco, M. (2010). Theories of Creativity. In J. C. Kaufman & R. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 20–47). New York: Cambridge University Press.

Kristensson, P., Gustafsson, A., & Archer, T. (2004). Harnessing the creative potential among users. *Journal of Product Innovation Management*, *21*, 4–14.

Krumm, G., Lemos, V., & Filippetti, V. A. (2014). Factor structure of the Torrance Tests of Crearive Thinking Figural Form B in Spanish-speaking children: Measurement invariance across gender. *Creativity Research Journal*, *26*, 72–81.

Lee, S., Lee, J., & Youn, C. (2005). A variation of CAT for measuring creativity in business products. *The Korean Journal of Thinking & Problem Solving*, *15*, 143–153.

Levin, M., & Kalal, T. (2003). *Improving Product Reliability: Strategies and Implementation*. Chichester: John Wiley and Sons Ltd.

Lin, C. Y. Y., & Liu, F. C. (2012). A cross-level analysis of organizational creativity climate and perceived innovation: The mediating effect of work motivation. *European Journal of Innovation Management*, *15*(1), 55–76. https://doi.org/10.1108/14601061211192834

Lindstrom, J., Sas, D., Lideskog, H., Lofstrand, M., & Karlsson, L. (2015). Defining "Functional Products" through their constituents. *International Journal of Product Development*, *20*(1), 1–24.

MacCallum, R., Browne, M., & Sugawara, H. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods*, *1*(2), 130–149.

MacCallum, R., & Hong, S. (1997). Power Analysis in Covariance Structure Modeling Using GFI and AGFI. *Multivariate Behavioral Research*, *32*, 193–210.

MacKinnon, D. (1962). The nature and nurture of creative talent. *American Psychologist*, *17*(7), 484–495.

MacKinnon, D. (1965). Personality and the realization of creative potential. *American Psychologist1*, *20*, 273–281.

MacKinnon, D. W. (1968). Creativity: Psychological aspects. In D. L. Sills (Ed.), *International Encyclopedia of the Social Sciences*. New York: Macmillan.

Maher, M. (2010). Evaluating creativity in humans, computers, and collectively intelligent systems. *The 1st DESIRE Network Conference on Creativity and Innovation in Design*, 22–28. Lancaster, UK.

Maiyaki, A. (2012). Statistics for social science: structural equation modeling approach. *Elixir International Journal: Social Science*, *49*, 9930–9934.

Martindale, C., & Dailey, A. (1996). Creativity. Primary process cognition and personality. *Personality and Individual Differences*, *20*, 409–414.

Mayer, R. (1999). Fifty years of creativity research. In R Sternberg (Ed.), *Handbook of creativity* (pp. 449–460). Cambridge, UK, UK: Cambridge University Press.

McCrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences*, *36*, 587–596.

McCrae, R. R., Costa, P. T., & Busch, C. M. (1986). Evaluating comprehensiveness in personality systems: The California Q-Set and the five-factor model. *Journal of Personality*, *57*, 17–40.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, *69*, 220–232.

Miles, J., & Shevlin, M. (1998). Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis. *Personality and Individual Differences*, *25*, 85–90.

Miller, B., Vehar, J., Firestein, R., Thurber, S., & Nielsen, D. (2011). *Creativity unbound: An introduction to creative process*. Evanston, US: FourSight.

Mital, A., Desai, A., Subramanian, A., & Mital, A. (2014). *Product Development: A Structured Approach to Consumer Product Development, Design, and Manufacture* (Second). London: Elsevier.

Moghimi, S., & Subramaniam, I. D. (2013). Employee's creative behavior: The role of organizational climate in Malaysian SMEs. *International Journal of Business & Management*, *8*(5), 1–13.

Molloy, O., Tilley, S., & Warman, E. A. (1998). *Design for Manufacturing and Assembly: Concepts, Architechtures and Implementation*. US: Springer.

NASA. (2018). Engineering Design Process. Retrieved October 22, 2018, from https://www.nasa.gov/audience/foreducators/best/edp.html

Netemeyer, R., Bearden, W., & Sharma, S. (2003). *Scaling Procedures: Issues and Applications*. Thousand Oaks: SAGE Publications.

Nikou, T. (2017). *Following the User's Footsteps: A Data-driven Framework for Measuring and Modeling Elegant Design in Engineering*. Clemson Univresity.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Adv Health Sci Educ Theory Pract*, *15*(5), 625–632.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.

O'Quin, K., & Besemer, S. (1989). The development, reliability and validity of the revised Creative Product Semantic Scale. *Creativity Research Journal*, *2*, 268–279.

O'Quin, K., & Besemer, S. (2011). Creative Products. In M. Runco & S. Pritzker (Eds.), *Encyclopedia of Creativity* (pp. 273–281). London: Academic Press.

Oldham, G., & Cummings, A. (1996). Employee Creativity: Personal and Contextual Factors at Work. *Academy of Management Journal*, *39*(3), 607–634.

Oman, S., & Tumer, I. (2009). The potential of creativity metrics for mechanical engineering concept design. *The 17th International Conference on Engineering Design (ICED'09)*, 145–156.

Osborn, A. F. (1963). *Applied imagination: Principles and procedures of creative problem-solving*. New York: Charles Scribner's.

Page, T. (2014). Product attachment and replacement: implications for sustainable design. *International Journal of Sustainable Design*, *22*(3), 256–282.

Pahl, G., & Beitz, W. (1984). *Engineering design*. London: The Design Council.

Pahl, G., & Beitz, W. (1996). *Engineering Design: A systematic Approach* (2nd ed.). London: Springer-Verlag.

Pahl, G., Beitz, W., Feldhusen, J., & Grote, K. H. (2013). *Engineering Design: A Systematic Approach* (3rd ed.; Springer-Verlag, Ed.). London.

Parnes, S. J. (1977). Guiding Creative action. *Gifted Child Quarterly*, *21*, 460–476.

Piffer, D. (2012). Can creativity be measured? An attempt to clarify the notion of creativity and general directions for future research. *Thinking Skills and Creativity*, *7*(3), 258–264. https://doi.org/10.1016/j.tsc.2012.04.009

Plucker, J. A., Beghetto, R., & Dow, G. (2004). Why isn't creativity more important to educational psychologists? Potential, pitfalls, and future directions in creativity. *Educational Psychologist*, *39*, 83–96.

Plucker, J. A., & Makel, M. C. (2010). Assessment of Creativity. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 48–73). New York: Cambridge University Press.

Podolyakina, N. (2017). Estimation of the Relationship between the Products Reliability, Period of Their Warranty Service and the Value of the Enterprise Cost. *Procedia Engineering*, *178*, 558–568.

Pols, A. (2002). *Insight in problem solving*. University of Utrecht, Utrecht, Netherlands.

Preacher, K., & MacCallum, R. (2002). Exploratory Factor Analysis in Behavior Genetics Research: Factor Recovery with Small Sample Size. *Behavior Genetics*, *32*, 153–161.

Preston, C., & Colman, A. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*(1), 1–15.

Prudon, P. (2015). Confirmatory factor analysis as a tool in research using questionnaires: a critique. *Comprehensive Psychology*, *4*(10).

Puccio, G., Mance, M., & Murdock, M. (2011). *Creative leadership: Skills that drive change* (2nd ed.). London: SAGE Publications, Inc.

Pugh, S. (1991). *Total Design: Integrated Methods for Successful Product Engineering*. New Jersey: Addison-Wesley.

Raja, U., & Johns, G. (2010). The joint effects of personality and job scope on in-role performance, citizenship behaviors, and creativity. *Human Relations*, *63*(7), 981–1005.

Rhodes, M. (1961). An Analysis of Creativity. *The Phi Delta Kappan*, *42*(7), 305–310.

Rimm, S., & Davis, G. A. (1979). GIFT: An instrument for the identification creativity. *Journal of Creative Behavior*, *10*, 178–182.

Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality*, *69*, 617–640.

Rodríguez, N., Perez, M., & Gutierrez, J. (2007). Interfunctional trust as a determining factor of a new product performance. *European Journal of Marketing*, *41*(5–6), 678–702.

Rothenberg, R., & Hausman, C. (1976). *The Creative question*. Durham, NC: Duke University Press.

Rubera, G., Ordanini, A., & Griffith, D. (2011). Incorporating cultural values for understanding the influence of perceived product creativity on intention to buy: An examination in Italy and the US. *Journal of International Business Studies*, *42*, 459–476.

Runco, M. (1989). The creativity of children's art. *Child Study Journal*, *19*(3), 177–189.

Runco, M. (2002). Creativity, cognition and their educational implications. In J. Houtz (Ed.), *The educational psychology of creativity* (pp. 25–58). Cresskill, NJ: Hampton Press.

Runco, M. (2003). Idea evaluation, divergent thinking, and creativity. In *Critical creative processes*. Cresskill, NJ, NJ: Hampton Press.

Runco, M. (2007). *Creativity: Theories and themes: Research*. Burlington, MA: Elsevier Academic Press.

Runco, M. (2008). Creativity and education. *New Horizons in Education*, *56*, 107–115.

Runco, M., & Jaeger, G. (2012). The Standard Definition of Creativity. *Creativity Research Journal*, *24*(1), 92–96.

Runco, M., Miller, G., Acar, S., & Cramond, B. (2010). Torrance Tests of Creative Thinking as predictors of personal and public achievement: A fifty-year follow-uo. *Creativity Research Journal*, *22*, 361–368.

Runco, M., Noble, E., Reiter-Palmon, R., Acar, S., & Ritchire, T. (2011). The Genetic Basis of Creativity and Ideational Fluency The Genetic Basis of Creativity and Ideational Fluency. *Psychology Faculty Publications*, *44*.

Runco, M., Plucker, J., & Lim, W. (2001). Development and Psychometric Integrity of a Measure of Ideational Behavior. *Creativity Research Journal*, *13*(3&4), 393–400.

Sarkar, P., & Chakrabarti, A. (2011). Assessing design creativity. *Design Studies*, *32*, 348–383.

Sawyer, R. K. (2006). *Explaining Creativity: The Science of Human Innovation* (1st ed.). USA: Oxford university Press.

Scott, S. G., & A, R. (1994). Determinants of Innovative Behavior: A Path Model of Individual Innovation in the Workplace. *Academy of Management Journal*, *37*(3), 580–607.

Sethi, R., & Sethi, A. (2009). Can quality-oriented firms develop innovative new products? *Journal of Product Innovation Management*, *26*(2), 206–221.

Sethi, R., Smith, D., & Park, C. (2001). Cross-Functional Product Development Teams, Creativity, and Innovativeness of New Consumer Products. *Journal of Marketing Research*, *38*, 73–85.

Shah, J., Vargas-Hernandez, N., & Smith, S. (2003). Metrics for measuring ideation effectiveness. *Des Stud*, *24*(2), 111–134.

Shalley, C. (1995). Effects of Coaction, Expected Evaluation, and Goal Setting on Creativity and Productivity. *Academy of Management Journal*, *38*(2), 483–503.

Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. (2005). A simulation study to investigate the

use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research*, *58*(1), 935–943.

Shneiderman, B. (2000). Creating creativity: user interfaces for supporting innovation. *ACM Transactions of Computer-Human Interaction*, *7*(1), 114–138.

Siegel, S. M., & Kaemmerer, W. F. (1978). Measuring the perceived support for innovation in organizations. *Journal of Applied Psychology*, *63*(5), 553–562.

Simonton, D. (1977). Eminence, Creativity, and Geographic Marginality: A Recursive Structural Equation Model. *Journal of Personality and Social Psychology*, *35*(11), 805–816.

Simonton, D. (1995). Exceptional personal influence: An integrative paradigm. *Creativity Research Journal*, *8*(4), 371–376.

Smith, B. (1993). Interpersonal behaviors that damage the productivity of creative problem solving groups. *Journal of Creative Behavior*, *27*(3), 171–187.

Smith, S., & Dodds, R. (1999). Incubation. In M. Runco & S. Pritzker (Eds.), *Encyclopedia of Creativity*. San Diego, CA, CA: Academic Press.

Sobek, D. K. (2005). Transitions: From Conceptual Ideas to Detail Design. *Proceeding of the American Society for Engineering Education Annual Conference & Exposition*.

Soldz, S., & Vaillant, G. (1999). The Big Five personality traits and the life course: A 45-year longitudinal study. *Journal of Research in Personlaity*, *33*, 208–232.

Steiger, J. (1990). Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Research*, *25*, 173–180.

Stein, M. (1974). *Stimulating creativity* (Vol. 1). New York: Academic Press.

Sternberg, Robert, & Lubart, T. (1995). *Defying the Crowd: Cultivating Creativity in a Culture of Conformity*. New York: Free Press.

Sternberg, Robert, & Lubart, T. (1999). The concept of creativity: Prospects and paradigms. In Robert Sternberg (Ed.), *Handbook of creativity* (pp. 3–15). New York: Cambridge University

Press.

Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Subedi, B. (2016). Using Likert Type Data in Social Science Research: Confusion, Issues and Challenges. *International Journal of Contemporary Applied Science*, *3*(2), 36–49.

Sue, V., & Ritter, L. (2007). *Conducting Online Surveys*. Thousand Oaks: SAGE Publications, Inc.

Svensson, N. (1974). *Introduction to engineering design*. Bath: Pitman.

Systemsinnovation. (n.d.). Complex Systems Features.

Tabachnick, B., & Fidell, L. (2007). *Using Multivariate Statistics* (5th ed.). Boston: Pearson Education.

Takai, S., Esterman, M., & Midha, A. (2015). An Approach to Study Associations Between Design Concepts and Design OUtcomes. *Journal of Mechanical Design*, *137*(4).

Taylor, A. (1975). An emerging view of creative actions. In I. Taylor & J. Getzels (Eds.), *Perspectives in creativity* (pp. 297–325). Chicago: Aldine.

Theuri, P., & Turner, L. (2002). Conducting Survey Research Through an Enhanced Online Web Survey Procedure. *Journal of Internet Commerce*, *1*(4), 37–54.

Thompson, G., & Lordan, M. (1999). A review of creativity principles applied to engineering design. *Proceeding of the Institution of Mechanical Engineers Part E-Journal of Process Mechanical Engineering*, 17–31.

Torrance, E. P. (1974). *The Torrence Tests of Creative Thinking-Norms-Technical manual research edition-verbal tests, forms A and B; figural tests, forms A and B*. Princeton, NJ: Personnel Press.

Torrance, E. P., & Ball, O. E. (1984). *The Torrance Tests of Creative Thinking Streamlined (revised) manual, Figural A and B*. Bensenville, IL: Scholastic Testing Service, Inc.

Tsai, C. T., & Kao, C. F. (2004). The relationships among motivation orientations, climate for

organization innovation, and employee innovative behavior: a test of Amabile's motivational synergy model. *Journal of Management*, *21*(5), 571–572.

Tsai, K. (2016). Quasi-Experts' Assessments of Creative Products: An Evaluation Using a Sample of Design Students. *British Journal of Education, Society & Behavioural Science*, *12*(3), 1–7. https://doi.org/10.9734/BJESBS/2016/21057

Tsai, K. C. (2013). Assessing Creative Products by Experts. *International Journal of Sciences*, *2*(May).

Ulrich, K. T., & Eppinger, S. D. (2004). *Product Design and Development* (Fifth). New York: McGraw-Hill/Irwin.

US Air Force. (2000). *Air Force System Safety Handbook*. Kirtland, NM: Air Force Safety Agency.

Valgeirsdottir, D., & Onarheim, B. (2015). Beyond Creativity Assessment: Comparing Methods and Identifying Consequences of Recognised Creativity. *The Third International Conference on Design Creativity (3rd ICDC)*, 1–8.

Van de Ven, A. H. (1986). Central Problems in the Management of Innovation. *Management Science*, *32*, 590–607.

Velicer, W., & Fava, J. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, *3*, 231–251.

Von Nordenflycht, A. (2007). Is Public Ownership Bad for Professional Service Firms? Ad Agency Ownership, Performance, and Creativity. *Academy of Management Journal*, *50*(2), 429–445.

Walker, A. M., Koestner, R., & Hum, A. (1995). Personality correlates of depressive style in autobiographies of creative achievers. *Journal of Creative Behavior*, *29*, 75–94.

Wallace, D. B., & Gruber, H. E. (1989). *Creative people at work: Twelve cognitive case studies*. New York: Oxford university Press.

Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the*

creativity-intelligence distinction. New York: Rinehart & Winston.

Wallas, G. (1926). *The Art of Thought*. New York: Harcourt Brace.

Ward, T. B., & Kolomyts, Y. (2010). Cognition and Creativity. In J. C. Kaufman & R. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 93–112). New York: Cambridge University Press.

Weerasiri, R., Zhengang, Z., & Perera, T. (2012). Innovative and Creativity of Small and Medium Scale Enterprises (MSEs) in Sri Lanka: A Review. *International Forum of Researchers Students and Academician*, *2*(1).

Weisberg, R. (1999). Creativity and Knowledge: A Challenge to Theories. In Robert Sternberg (Ed.), *Handbook of creativity* (pp. 226–250). New York: Cambridge University Press.

Wheaton, B., Muthen, B., Alwin, D., & Summers, G. (1977). Assessing Reliability and Stability in Panel Models. *Sociological Methodology*, *8*, 84–136.

White, A., Shen, F., & Smith, B. (2002). Judging Advertising Creativity Using the Creative Product Semantic Scale. *Journal of Creative Behavior*, *36*(4), 241–253.

White, A., & Smith, B. (2001). Assessing advertising creativity using the Creative Product Semantic Scale. *Journal of Advertising Research*, *41*(6), 27–34.

Widaman, K., & Thompson, J. (2003). On Specifying the Null Model for Incremental Fit Indices in Structural Equation Modeling. *Psychological Methods*, *8*(1), 16–37.

Wolfradt, U., & Pretz, J. (2001). Individual Differences in Creativity: Personality, Story Writing, and Hobbies. *European Journal of Personality*, *15*, 297–310.

Wong, Y. L., & Siu, K. W. M. (2012). A model of creative design process for fostering creativity of students in design education. *International Journal of Technology and Design Education*, *22*(4), 437–450. https://doi.org/10.1007/s10798-011-9162-8

Yazdanifard, R., & Mercy, I. (2011). The impact of Green Marketing on Customer satisfaction and Environmental safety. *International Conference on Computer Communication and*

Management, 637–641. Singapore.

Yuan, X., & Lee, J.-H. (2014). A quantitative apprach for assessment of creativity in product design. *Advanced Engineering Informatics*, *28*, 528–541.

Appendix: Questionnaire

Welcome

This questionnaire explores the relationships between a series of attributes for five lounge chairs. The outcomes will be used to better understand innovation within consumer products. The information you give us will be treated confidentially. Please feel free to be completely honest, and you will not be judged in any way on the response you give. The survey will take approximately 30 minutes to complete.

Instruction

You are going to observe and compare 5 lounge chairs, then indicate your opinions around a series of attributes for each product using a 7-point scale. You may choose '1' if your opinion toward an attribute is extremely negative, '4' if your opinion is neutral, and '7' if your opinion is extremely positive. Otherwise, select from other options.

You need to compare the products to one another, as your judgments should not be based on an absolute standard that you are considering as ideal. For instance, if you think product (A) is moderately unusual, but it is relatively the most unusual in comparison to other products presented in this survey, then you should give 7. Otherwise, choose the option that is most sensible (see below for an example).

Background

---

1. What is your age?

( ) Less than 18 years old

( ) 18-24

( ) 25-29

( ) 30-34

( ) 35-39

( ) 40-44

( ) 45-49

( ) 50-54

( ) 55 years or older

2. What is your gender?

( ) Male

( ) Female

( ) Don’t wish to state

3. Which race/ethnicity best describes you?

( ) White

( ) Asian/Asian British

( ) Black/African/Caribbean/Black British

( ) Mixed/multiple ethnic groups

( ) Other ethnic group

4. Which statement best describes your current situation? *

( ) Working (paid employee)

( ) Working (self-employed)

( ) Student (full-time)

( ) Not working (unemployed)

( ) Others

5. Please specify what degree are you currently studying:

( ) University undergraduate degree (e.g. BA, BS)

( ) University postgraduate degree (e.g. MA, MS, Meng, MBA)

( ) Doctoral degree (PhD)

---

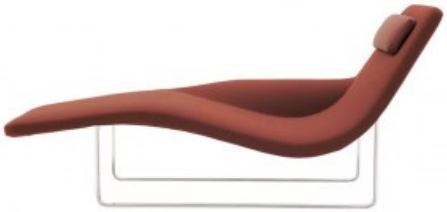
*Note: At this stage, participants are allowed to continue only if they select 'Student (full-time)' to best describe their current situation.

Product Presentation

Here are the 5 chairs that you need to compare and judge. Please spend some time familiarising yourself with the products.

A

B



C

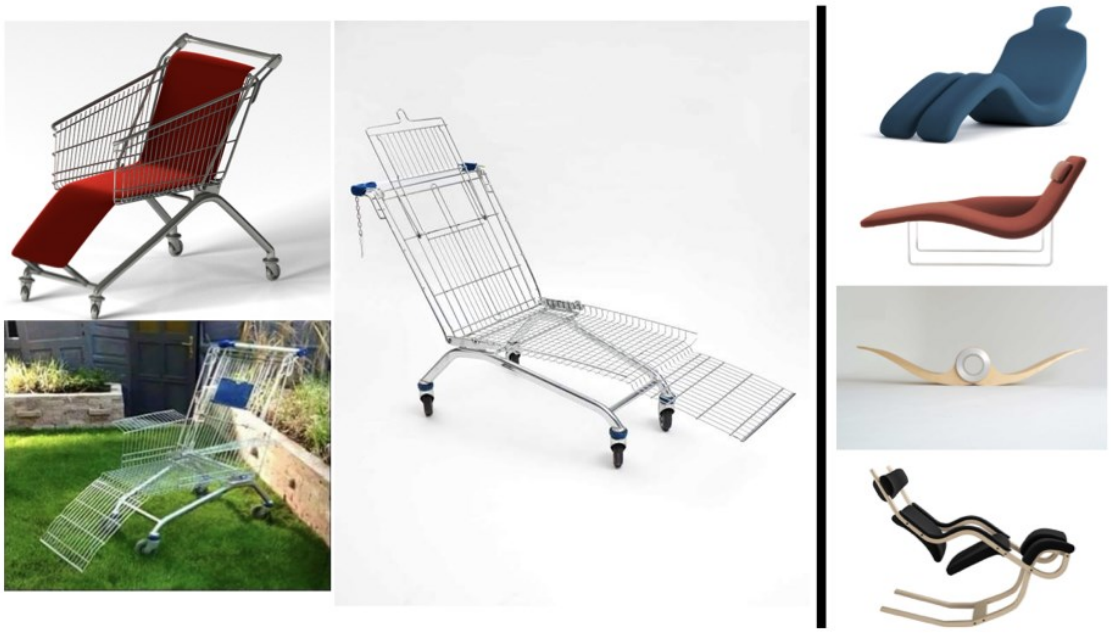
D

E

Please compare the above product (left) to other chairs (right) and select the option that most closely matches your immediate thought about each attribute. Do not spend long time thinking about each attribute.

	Extremely	Moderately	Slightly	Neither	Slightly	Moderately	Extremely	
Conventional	( )	( )	( )	( )	( )	( )	( )	Original
Usual	( )	( )	( )	( )	( )	( )	( )	Unusual
Ordinary	( )	( )	( )	( )	( )	( )	( )	Unique
Average	( )	( )	( )	( )	( )	( )	( )	Revolutionary
Frequent	( )	( )	( )	( )	( )	( )	( )	Infrequent
Expected	( )	( )	( )	( )	( )	( )	( )	Surprising
Non-functional	( )	( )	( )	( )	( )	( )	( )	Functional
Futile	( )	( )	( )	( )	( )	( )	( )	Practical
Inefficient	( )	( )	( )	( )	( )	( )	( )	Efficient
Incapable	( )	( )	( )	( )	( )	( )	( )	Capable
Ineffective	( )	( )	( )	( )	( )	( )	( )	Effective
Unchanging	( )	( )	( )	( )	( )	( )	( )	Adaptable
Inflexible	( )	( )	( )	( )	( )	( )	( )	Flexible
Fixed	( )	( )	( )	( )	( )	( )	( )	Modifiable
Immutable	( )	( )	( )	( )	( )	( )	( )	Alterable
Static	( )	( )	( )	( )	( )	( )	( )	Changeable
Does not protect	( )	( )	( )	( )	( )	( )	( )	Protects
Unsafe	( )	( )	( )	( )	( )	( )	( )	Safe
Hazardous	( )	( )	( )	( )	( )	( )	( )	Non-hazardous
Insecure	( )	( )	( )	( )	( )	( )	( )	Secure
Harmful	( )	( )	( )	( )	( )	( )	( )	Harmless
Painful	( )	( )	( )	( )	( )	( )	( )	Innocuous
Difficult	( )	( )	( )	( )	( )	( )	( )	Easy to use

Complicated	( )	( )	( )	( )	( )	( )	( )	Simple
Ambiguous	( )	( )	( )	( )	( )	( )	( )	Comprehensible
Complex	( )	( )	( )	( )	( )	( )	( )	Straightforward
Inoperable	( )	( )	( )	( )	( )	( )	( )	Operable
Unergonomic	( )	( )	( )	( )	( )	( )	( )	Ergonomic
Intolerable	( )	( )	( )	( )	( )	( )	( )	Tolerable
Uncomfortable	( )	( )	( )	( )	( )	( )	( )	Comfortable
Flimsy	( )	( )	( )	( )	( )	( )	( )	Durable
Weak	( )	( )	( )	( )	( )	( )	( )	Strong
Unsound	( )	( )	( )	( )	( )	( )	( )	Sound
Fragile	( )	( )	( )	( )	( )	( )	( )	Solid
Frail	( )	( )	( )	( )	( )	( )	( )	Robust
Unreliable	( )	( )	( )	( )	( )	( )	( )	Reliable
Questionable	( )	( )	( )	( )	( )	( )	( )	Trustworthy
Poorly built	( )	( )	( )	( )	( )	( )	( )	Well-built
Synthetic	( )	( )	( )	( )	( )	( )	( )	Organic
Unsustainable	( )	( )	( )	( )	( )	( )	( )	Sustainable
Non-recyclable	( )	( )	( )	( )	( )	( )	( )	Recyclable
Irrecoverable	( )	( )	( )	( )	( )	( )	( )	Renewable
Environmentally harmful	( )	( )	( )	( )	( )	( )	( )	Environmentally friendly
Pollutive	( )	( )	( )	( )	( )	( )	( )	Pollution free
Repulsive	( )	( )	( )	( )	( )	( )	( )	Appealing
Repelling	( )	( )	( )	( )	( )	( )	( )	Charming
Coarse	( )	( )	( )	( )	( )	( )	( )	Elegant
Crude	( )	( )	( )	( )	( )	( )	( )	Refined
Messy	( )	( )	( )	( )	( )	( )	( )	Neat
Disordered	( )	( )	( )	( )	( )	( )	( )	Arranged
Disorganised	( )	( )	( )	( )	( )	( )	( )	Organised
Poorly proportioned	( )	( )	( )	( )	( )	( )	( )	Well-proportioned
Imbalance	( )	( )	( )	( )	( )	( )	( )	Balance
Disgusting	( )	( )	( )	( )	( )	( )	( )	Desirable
Dissatisfying	( )	( )	( )	( )	( )	( )	( )	Satisfying
Unpleasant	( )	( )	( )	( )	( )	( )	( )	Pleasant
Contempt	( )	( )	( )	( )	( )	( )	( )	Admirable
Boring	( )	( )	( )	( )	( )	( )	( )	Exciting
Unnecessary	( )	( )	( )	( )	( )	( )	( )	Necessary
Irrelevant	( )	( )	( )	( )	( )	( )	( )	Relevant
Inessential	( )	( )	( )	( )	( )	( )	( )	Essential
Inappropriate	( )	( )	( )	( )	( )	( )	( )	Appropriate
Unreasonable	( )	( )	( )	( )	( )	( )	( )	Reasonable
Worthless	( )	( )	( )	( )	( )	( )	( )	Valuable
Incompatible	( )	( )	( )	( )	( )	( )	( )	Compatible
Uncreative	( )	( )	( )	( )	( )	( )	( )	Creative



Please compare the above product (left) to other chairs (right) and select the option that most closely matches your immediate thought about each attribute. Do not spend long time thinking about each attribute.

	Extremely	Moderately	Slightly	Neither	Slightly	Moderately	Extremely	
Conventional	( )	( )	( )	( )	( )	( )	( )	Original
Usual	( )	( )	( )	( )	( )	( )	( )	Unusual
Ordinary	( )	( )	( )	( )	( )	( )	( )	Unique
Average	( )	( )	( )	( )	( )	( )	( )	Revolutionary
Frequent	( )	( )	( )	( )	( )	( )	( )	Infrequent
Expected	( )	( )	( )	( )	( )	( )	( )	Surprising
Non-functional	( )	( )	( )	( )	( )	( )	( )	Functional
Futile	( )	( )	( )	( )	( )	( )	( )	Practical
Inefficient	( )	( )	( )	( )	( )	( )	( )	Efficient
Incapable	( )	( )	( )	( )	( )	( )	( )	Capable
Ineffective	( )	( )	( )	( )	( )	( )	( )	Effective
Unchanging	( )	( )	( )	( )	( )	( )	( )	Adaptable
Inflexible	( )	( )	( )	( )	( )	( )	( )	Flexible
Fixed	( )	( )	( )	( )	( )	( )	( )	Modifiable
Immutable	( )	( )	( )	( )	( )	( )	( )	Alterable
Static	( )	( )	( )	( )	( )	( )	( )	Changeable
Does not protect	( )	( )	( )	( )	( )	( )	( )	Protects
Unsafe	( )	( )	( )	( )	( )	( )	( )	Safe
Hazardous	( )	( )	( )	( )	( )	( )	( )	Non-hazardous
Insecure	( )	( )	( )	( )	( )	( )	( )	Secure
Harmful	( )	( )	( )	( )	( )	( )	( )	Harmless

Painful	( )	( )	( )	( )	( )	( )	( ) Innocuous
Difficult	( )	( )	( )	( )	( )	( )	( ) Easy to use
Complicated	( )	( )	( )	( )	( )	( )	( ) Simple
Ambiguous	( )	( )	( )	( )	( )	( )	( ) Comprehensible
Complex	( )	( )	( )	( )	( )	( )	( ) Straightforward
Inoperable	( )	( )	( )	( )	( )	( )	( ) Operable
Unergonomic	( )	( )	( )	( )	( )	( )	( ) Ergonomic
Intolerable	( )	( )	( )	( )	( )	( )	( ) Tolerable
Uncomfortable	( )	( )	( )	( )	( )	( )	( ) Comfortable
Flimsy	( )	( )	( )	( )	( )	( )	( ) Durable
Weak	( )	( )	( )	( )	( )	( )	( ) Strong
Unsound	( )	( )	( )	( )	( )	( )	( ) Sound
Fragile	( )	( )	( )	( )	( )	( )	( ) Solid
Frail	( )	( )	( )	( )	( )	( )	( ) Robust
Unreliable	( )	( )	( )	( )	( )	( )	( ) Reliable
Questionable	( )	( )	( )	( )	( )	( )	( ) Trustworthy
Poorly built	( )	( )	( )	( )	( )	( )	( ) Well-built
Synthetic	( )	( )	( )	( )	( )	( )	( ) Organic
Unsustainable	( )	( )	( )	( )	( )	( )	( ) Sustainable
Non-recyclable	( )	( )	( )	( )	( )	( )	( ) Recyclable
Irrecoverable	( )	( )	( )	( )	( )	( )	( ) Renewable
Environmentally harmful	( )	( )	( )	( )	( )	( )	( ) Environmentally friendly
Pollutive	( )	( )	( )	( )	( )	( )	( ) Pollution free
Repulsive	( )	( )	( )	( )	( )	( )	( ) Appealing
Repelling	( )	( )	( )	( )	( )	( )	( ) Charming
Coarse	( )	( )	( )	( )	( )	( )	( ) Elegant
Crude	( )	( )	( )	( )	( )	( )	( ) Refined
Messy	( )	( )	( )	( )	( )	( )	( ) Neat
Disordered	( )	( )	( )	( )	( )	( )	( ) Arranged
Disorganised	( )	( )	( )	( )	( )	( )	( ) Organised
Poorly proportioned	( )	( )	( )	( )	( )	( )	( ) Well-proportioned
Imbalance	( )	( )	( )	( )	( )	( )	( ) Balance
Disgusting	( )	( )	( )	( )	( )	( )	( ) Desirable
Dissatisfying	( )	( )	( )	( )	( )	( )	( ) Satisfying
Unpleasant	( )	( )	( )	( )	( )	( )	( ) Pleasant
Contempt	( )	( )	( )	( )	( )	( )	( ) Admirable
Boring	( )	( )	( )	( )	( )	( )	( ) Exciting
Unnecessary	( )	( )	( )	( )	( )	( )	( ) Necessary
Irrelevant	( )	( )	( )	( )	( )	( )	( ) Relevant
Inessential	( )	( )	( )	( )	( )	( )	( ) Essential
Inappropriate	( )	( )	( )	( )	( )	( )	( ) Appropriate
Unreasonable	( )	( )	( )	( )	( )	( )	( ) Reasonable
Worthless	( )	( )	( )	( )	( )	( )	( ) Valuable
Incompatible	( )	( )	( )	( )	( )	( )	( ) Compatible
Uncreative	( )	( )	( )	( )	( )	( )	( ) Creative

Please compare the above product (left) to other chairs (right) and select the option that most closely matches your immediate thought about each attribute. Do not spend long time thinking about each attribute.

	Extremely	Moderately	Slightly	Neither	Slightly	Moderately	Extremely	
Conventional	( )	( )	( )	( )	( )	( )	( )	Original
Usual	( )	( )	( )	( )	( )	( )	( )	Unusual
Ordinary	( )	( )	( )	( )	( )	( )	( )	Unique
Average	( )	( )	( )	( )	( )	( )	( )	Revolutionary
Frequent	( )	( )	( )	( )	( )	( )	( )	Infrequent
Expected	( )	( )	( )	( )	( )	( )	( )	Surprising
Non-functional	( )	( )	( )	( )	( )	( )	( )	Functional
Futile	( )	( )	( )	( )	( )	( )	( )	Practical
Inefficient	( )	( )	( )	( )	( )	( )	( )	Efficient
Incapable	( )	( )	( )	( )	( )	( )	( )	Capable
Ineffective	( )	( )	( )	( )	( )	( )	( )	Effective
Unchanging	( )	( )	( )	( )	( )	( )	( )	Adaptable
Inflexible	( )	( )	( )	( )	( )	( )	( )	Flexible
Fixed	( )	( )	( )	( )	( )	( )	( )	Modifiable
Immutable	( )	( )	( )	( )	( )	( )	( )	Alterable
Static	( )	( )	( )	( )	( )	( )	( )	Changeable
Does not protect	( )	( )	( )	( )	( )	( )	( )	Protects
Unsafe	( )	( )	( )	( )	( )	( )	( )	Safe
Hazardous	( )	( )	( )	( )	( )	( )	( )	Non-hazardous
Insecure	( )	( )	( )	( )	( )	( )	( )	Secure
Harmful	( )	( )	( )	( )	( )	( )	( )	Harmless
Painful	( )	( )	( )	( )	( )	( )	( )	Innocuous

Difficult	( )	( )	( )	( )	( )	( )	( )	Easy to use
Complicated	( )	( )	( )	( )	( )	( )	( )	Simple
Ambiguous	( )	( )	( )	( )	( )	( )	( )	Comprehensible
Complex	( )	( )	( )	( )	( )	( )	( )	Straightforward
Inoperable	( )	( )	( )	( )	( )	( )	( )	Operable
Unergonomic	( )	( )	( )	( )	( )	( )	( )	Ergonomic
Intolerable	( )	( )	( )	( )	( )	( )	( )	Tolerable
Uncomfortable	( )	( )	( )	( )	( )	( )	( )	Comfortable
Flimsy	( )	( )	( )	( )	( )	( )	( )	Durable
Weak	( )	( )	( )	( )	( )	( )	( )	Strong
Unsound	( )	( )	( )	( )	( )	( )	( )	Sound
Fragile	( )	( )	( )	( )	( )	( )	( )	Solid
Frail	( )	( )	( )	( )	( )	( )	( )	Robust
Unreliable	( )	( )	( )	( )	( )	( )	( )	Reliable
Questionable	( )	( )	( )	( )	( )	( )	( )	Trustworthy
Poorly built	( )	( )	( )	( )	( )	( )	( )	Well-built
Synthetic	( )	( )	( )	( )	( )	( )	( )	Organic
Unsustainable	( )	( )	( )	( )	( )	( )	( )	Sustainable
Non-recyclable	( )	( )	( )	( )	( )	( )	( )	Recyclable
Irrecoverable	( )	( )	( )	( )	( )	( )	( )	Renewable
Environmentally harmful	( )	( )	( )	( )	( )	( )	( )	Environmentally friendly
Pollutive	( )	( )	( )	( )	( )	( )	( )	Pollution free
Repulsive	( )	( )	( )	( )	( )	( )	( )	Appealing
Repelling	( )	( )	( )	( )	( )	( )	( )	Charming
Coarse	( )	( )	( )	( )	( )	( )	( )	Elegant
Crude	( )	( )	( )	( )	( )	( )	( )	Refined
Messy	( )	( )	( )	( )	( )	( )	( )	Neat
Disordered	( )	( )	( )	( )	( )	( )	( )	Arranged
Disorganised	( )	( )	( )	( )	( )	( )	( )	Organised
Poorly proportioned	( )	( )	( )	( )	( )	( )	( )	Well-proportioned
Imbalance	( )	( )	( )	( )	( )	( )	( )	Balance
Disgusting	( )	( )	( )	( )	( )	( )	( )	Desirable
Dissatisfying	( )	( )	( )	( )	( )	( )	( )	Satisfying
Unpleasant	( )	( )	( )	( )	( )	( )	( )	Pleasant
Contempt	( )	( )	( )	( )	( )	( )	( )	Admirable
Boring	( )	( )	( )	( )	( )	( )	( )	Exciting
Unnecessary	( )	( )	( )	( )	( )	( )	( )	Necessary
Irrelevant	( )	( )	( )	( )	( )	( )	( )	Relevant
Inessential	( )	( )	( )	( )	( )	( )	( )	Essential
Inappropriate	( )	( )	( )	( )	( )	( )	( )	Appropriate
Unreasonable	( )	( )	( )	( )	( )	( )	( )	Reasonable
Worthless	( )	( )	( )	( )	( )	( )	( )	Valuable
Incompatible	( )	( )	( )	( )	( )	( )	( )	Compatible
Uncreative	( )	( )	( )	( )	( )	( )	( )	Creative

Please compare the above product (left) to other chairs (right) and select the option that most closely matches your immediate thought about each attribute. Do not spend long time thinking about each attribute.

	Extremely	Moderately	Slightly	Neither	Slightly	Moderately	Extremely	
Conventional	( )	( )	( )	( )	( )	( )	( )	Original
Usual	( )	( )	( )	( )	( )	( )	( )	Unusual
Ordinary	( )	( )	( )	( )	( )	( )	( )	Unique
Average	( )	( )	( )	( )	( )	( )	( )	Revolutionary
Frequent	( )	( )	( )	( )	( )	( )	( )	Infrequent
Expected	( )	( )	( )	( )	( )	( )	( )	Surprising
Non-functional	( )	( )	( )	( )	( )	( )	( )	Functional
Futile	( )	( )	( )	( )	( )	( )	( )	Practical
Inefficient	( )	( )	( )	( )	( )	( )	( )	Efficient
Incapable	( )	( )	( )	( )	( )	( )	( )	Capable
Ineffective	( )	( )	( )	( )	( )	( )	( )	Effective
Unchanging	( )	( )	( )	( )	( )	( )	( )	Adaptable
Inflexible	( )	( )	( )	( )	( )	( )	( )	Flexible
Fixed	( )	( )	( )	( )	( )	( )	( )	Modifiable
Immutable	( )	( )	( )	( )	( )	( )	( )	Alterable
Static	( )	( )	( )	( )	( )	( )	( )	Changeable
Does not protect	( )	( )	( )	( )	( )	( )	( )	Protects
Unsafe	( )	( )	( )	( )	( )	( )	( )	Safe
Hazardous	( )	( )	( )	( )	( )	( )	( )	Non-hazardous
Insecure	( )	( )	( )	( )	( )	( )	( )	Secure

Harmful	( )	( )	( )	( )	( )	( )	( )	Harmless
Painful	( )	( )	( )	( )	( )	( )	( )	Innocuous
Difficult	( )	( )	( )	( )	( )	( )	( )	Easy to use
Complicated	( )	( )	( )	( )	( )	( )	( )	Simple
Ambiguous	( )	( )	( )	( )	( )	( )	( )	Comprehensible
Complex	( )	( )	( )	( )	( )	( )	( )	Straightforward
Inoperable	( )	( )	( )	( )	( )	( )	( )	Operable
Unergonomic	( )	( )	( )	( )	( )	( )	( )	Ergonomic
Intolerable	( )	( )	( )	( )	( )	( )	( )	Tolerable
Uncomfortable	( )	( )	( )	( )	( )	( )	( )	Comfortable
Flimsy	( )	( )	( )	( )	( )	( )	( )	Durable
Weak	( )	( )	( )	( )	( )	( )	( )	Strong
Unsound	( )	( )	( )	( )	( )	( )	( )	Sound
Fragile	( )	( )	( )	( )	( )	( )	( )	Solid
Frail	( )	( )	( )	( )	( )	( )	( )	Robust
Unreliable	( )	( )	( )	( )	( )	( )	( )	Reliable
Questionable	( )	( )	( )	( )	( )	( )	( )	Trustworthy
Poorly built	( )	( )	( )	( )	( )	( )	( )	Well-built
Synthetic	( )	( )	( )	( )	( )	( )	( )	Organic
Unsustainable	( )	( )	( )	( )	( )	( )	( )	Sustainable
Non-recyclable	( )	( )	( )	( )	( )	( )	( )	Recyclable
Irrecoverable	( )	( )	( )	( )	( )	( )	( )	Renewable
Environmentally harmful	( )	( )	( )	( )	( )	( )	( )	Environmentally friendly
Pollutive	( )	( )	( )	( )	( )	( )	( )	Pollution free
Repulsive	( )	( )	( )	( )	( )	( )	( )	Appealing
Repelling	( )	( )	( )	( )	( )	( )	( )	Charming
Coarse	( )	( )	( )	( )	( )	( )	( )	Elegant
Crude	( )	( )	( )	( )	( )	( )	( )	Refined
Messy	( )	( )	( )	( )	( )	( )	( )	Neat
Disordered	( )	( )	( )	( )	( )	( )	( )	Arranged
Disorganised	( )	( )	( )	( )	( )	( )	( )	Organised
Poorly proportioned	( )	( )	( )	( )	( )	( )	( )	Well-proportioned
Imbalance	( )	( )	( )	( )	( )	( )	( )	Balance
Disgusting	( )	( )	( )	( )	( )	( )	( )	Desirable
Dissatisfying	( )	( )	( )	( )	( )	( )	( )	Satisfying
Unpleasant	( )	( )	( )	( )	( )	( )	( )	Pleasant
Contempt	( )	( )	( )	( )	( )	( )	( )	Admirable
Boring	( )	( )	( )	( )	( )	( )	( )	Exciting
Unnecessary	( )	( )	( )	( )	( )	( )	( )	Necessary
Irrelevant	( )	( )	( )	( )	( )	( )	( )	Relevant
Inessential	( )	( )	( )	( )	( )	( )	( )	Essential
Inappropriate	( )	( )	( )	( )	( )	( )	( )	Appropriate
Unreasonable	( )	( )	( )	( )	( )	( )	( )	Reasonable
Worthless	( )	( )	( )	( )	( )	( )	( )	Valuable
Incompatible	( )	( )	( )	( )	( )	( )	( )	Compatible
Uncreative	( )	( )	( )	( )	( )	( )	( )	Creative

Please compare the above product (left) to other chairs (right) and select the option that most closely matches your immediate thought about each attribute. Do not spend long time thinking about each attribute.

	Extremely	Moderately	Slightly	Neither	Slightly	Moderately	Extremely	
Conventional	( )	( )	( )	( )	( )	( )	( )	Original
Usual	( )	( )	( )	( )	( )	( )	( )	Unusual
Ordinary	( )	( )	( )	( )	( )	( )	( )	Unique
Average	( )	( )	( )	( )	( )	( )	( )	Revolutionary
Frequent	( )	( )	( )	( )	( )	( )	( )	Infrequent
Expected	( )	( )	( )	( )	( )	( )	( )	Surprising
Non-functional	( )	( )	( )	( )	( )	( )	( )	Functional
Futile	( )	( )	( )	( )	( )	( )	( )	Practical
Inefficient	( )	( )	( )	( )	( )	( )	( )	Efficient
Incapable	( )	( )	( )	( )	( )	( )	( )	Capable
Ineffective	( )	( )	( )	( )	( )	( )	( )	Effective
Unchanging	( )	( )	( )	( )	( )	( )	( )	Adaptable
Inflexible	( )	( )	( )	( )	( )	( )	( )	Flexible
Fixed	( )	( )	( )	( )	( )	( )	( )	Modifiable
Immutable	( )	( )	( )	( )	( )	( )	( )	Alterable
Static	( )	( )	( )	( )	( )	( )	( )	Changeable
Does not protect	( )	( )	( )	( )	( )	( )	( )	Protects
Unsafe	( )	( )	( )	( )	( )	( )	( )	Safe
Hazardous	( )	( )	( )	( )	( )	( )	( )	Non-hazardous
Insecure	( )	( )	( )	( )	( )	( )	( )	Secure
Harmful	( )	( )	( )	( )	( )	( )	( )	Harmless

Painful	( )	( )	( )	( )	( )	( )	( )	Innocuous
Difficult	( )	( )	( )	( )	( )	( )	( )	Easy to use
Complicated	( )	( )	( )	( )	( )	( )	( )	Simple
Ambiguous	( )	( )	( )	( )	( )	( )	( )	Comprehensible
Complex	( )	( )	( )	( )	( )	( )	( )	Straightforward
Inoperable	( )	( )	( )	( )	( )	( )	( )	Operable
Unergonomic	( )	( )	( )	( )	( )	( )	( )	Ergonomic
Intolerable	( )	( )	( )	( )	( )	( )	( )	Tolerable
Uncomfortable	( )	( )	( )	( )	( )	( )	( )	Comfortable
Flimsy	( )	( )	( )	( )	( )	( )	( )	Durable
Weak	( )	( )	( )	( )	( )	( )	( )	Strong
Unsound	( )	( )	( )	( )	( )	( )	( )	Sound
Fragile	( )	( )	( )	( )	( )	( )	( )	Solid
Frail	( )	( )	( )	( )	( )	( )	( )	Robust
Unreliable	( )	( )	( )	( )	( )	( )	( )	Reliable
Questionable	( )	( )	( )	( )	( )	( )	( )	Trustworthy
Poorly built	( )	( )	( )	( )	( )	( )	( )	Well-built
Synthetic	( )	( )	( )	( )	( )	( )	( )	Organic
Unsustainable	( )	( )	( )	( )	( )	( )	( )	Sustainable
Non-recyclable	( )	( )	( )	( )	( )	( )	( )	Recyclable
Irrecoverable	( )	( )	( )	( )	( )	( )	( )	Renewable
Environmentally harmful	( )	( )	( )	( )	( )	( )	( )	Environmentally friendly
Pollutive	( )	( )	( )	( )	( )	( )	( )	Pollution free
Repulsive	( )	( )	( )	( )	( )	( )	( )	Appealing
Repelling	( )	( )	( )	( )	( )	( )	( )	Charming
Coarse	( )	( )	( )	( )	( )	( )	( )	Elegant
Crude	( )	( )	( )	( )	( )	( )	( )	Refined
Messy	( )	( )	( )	( )	( )	( )	( )	Neat
Disordered	( )	( )	( )	( )	( )	( )	( )	Arranged
Disorganised	( )	( )	( )	( )	( )	( )	( )	Organised
Poorly proportioned	( )	( )	( )	( )	( )	( )	( )	Well-proportioned
Imbalance	( )	( )	( )	( )	( )	( )	( )	Balance
Disgusting	( )	( )	( )	( )	( )	( )	( )	Desirable
Dissatisfying	( )	( )	( )	( )	( )	( )	( )	Satisfying
Unpleasant	( )	( )	( )	( )	( )	( )	( )	Pleasant
Contempt	( )	( )	( )	( )	( )	( )	( )	Admirable
Boring	( )	( )	( )	( )	( )	( )	( )	Exciting
Unnecessary	( )	( )	( )	( )	( )	( )	( )	Necessary
Irrelevant	( )	( )	( )	( )	( )	( )	( )	Relevant
Inessential	( )	( )	( )	( )	( )	( )	( )	Essential
Inappropriate	( )	( )	( )	( )	( )	( )	( )	Appropriate
Unreasonable	( )	( )	( )	( )	( )	( )	( )	Reasonable
Worthless	( )	( )	( )	( )	( )	( )	( )	Valuable
Incompatible	( )	( )	( )	( )	( )	( )	( )	Compatible
Uncreative	( )	( )	( )	( )	( )	( )	( )	Creative

Thank you!

Your contribution has been recorded