POLITECNICO DI TORINO

Corso di Laurea Magistrale

In Engineering and Management

Tesi di Laurea Magistrale

Electroencephalogram: the definition of the assessment methodology for verbal responses and the analysis of brain waves in an Idea Creativity experiment

Relatore:
Prof. Francesca Montagna

Candidato:
Alessandro Candusso

Anno Accademico 2018/2019

Abstract

The Science of Design is a multidisciplinary approach used to study designers and design activities scientifically. Its final aim is to provide techniques and strategies in order to support design processes. In the beginning, these researches were relying on protocol analysis, observational research method, which was lacking information regarding the cognitive subconscious processes of designers. For this reason, in the last decades, neurophysiologic tools have been adopted in order to integrate the investigation with new data.

In particular, the present experiment aims to identify neurophysiologic patterns through the electroencephalogram (EEG) during the creative ideation process in the earliest phases of product development. In order to do so, an Alternative Uses Task experiment has been designed, in which participants were asked to find an alternative or standard solutions to everyday objects in a limited amount of time. The experiment was relying on the model proposed by P.J. Guilford (1967), involving the iteration of two phases, named Divergent Thinking and Convergent Thinking, respectively allowing designers to generate a solution space and to choose the most suitable idea from it to solve a stated problem.

In the creative ideation process field, the neurocognitive approach is still at the beginning, and several inconsistent results have been obtained due to the lack of a standard in the design of the experiment; therefore, the aim of the present research is the prosecution of the design of the experiment developed by Colombo (2019) based on the state of art of the field, extending further data in the literature regarding the ideation design cognitive processes, to test a clear and defined experiment pipeline and to provide the information required in order to replicate the experiment in further studies.

Besides, it cannot be excluded, being this branch just at its beginning, that the output of the present experiment could be adopted in future studies in order to improve the results of the ideation process through the control of environmental variables or the implementation of new strategies for designers.

Table of contents

Introduction ... 1
The Problem ... 1
The Purpose of the Research ... 2
Methodology ... 3
Document structure ... 5
2 Context ... 7
2.1 Product development ... 7
2.2 Conceptual design & idea generation ... 9
2.3 Science of Design: Protocol analysis & main issues ... 12
2.4 Neuroscientific experiments in Design & divergent thinking ... 13
2.5 Creativity: convergent thinking vs divergent thinking ... 15
3 Literature review: assessment methodologies for verbal responses in design creativity ... 19
3.1 AUT history & main criteria ... 20
3.2 Review of idea generation and idea evaluation methods ... 23
3.2.1 Idea evaluation & the lack of a standard ... 25
3.2.2 Jauk et al. experiment (2012): CAT methodology (Amabile, 1982) ... 28
3.2.3 Shah et al. (2000, 2003) ... 31
3.2.4 Dean et al. (2006) ... 37
3.2.5 Other relevant assessment methods ... 42
4 The experiment (Colombo, 2019) ... 47
4.1 The experiment phases (Colombo, 2019) ... 48
4.1.1 The DOE pre-data gathering ... 48
4.1.2 The focus of the present research ... 51
4.2 Design of the assessment methodology for verbal responses in design creativity ... 53
4.2.1 Parameters ... 54
4.2.2 Judges ... 60
4.2.3 Guidelines ... 61
4.2.4 The procedure ... 61
4.3 Brain Signals ... 62
4.3.1 The Brain and the cognitive processes ... 63
4.3.2 Brain waves (alpha waves) ... 66
4.3.3 Independent component analysis (ICA) ... 69
4.4 Design of the experiment ... 71

4.4.1 Inputs: TRP & database ... 72
4.4.2 "3 areas" ... 76
4.4.3 "6 areas" ... 80
4.4.4 "7 areas" ... 85
5 Results ... 91
5.1 Literature comparison ... 91
5.2 Design of experiment chosen ... 93
5.3 "2x2x6" results in detail ... 94
5.4 Between-subjects factor "Background" ... 98
5.5 Between-subjects factor "Gender" ... 104
5.6 Between-subjects factor "Degree level" ... 108
Conclusions ... 111
Appendix ... 115
A.1 Personal data ... 115
A.2 Stimuli ... 117
A.3 Idea assessment literature review ... 118
A.4 Guidelines ... 119
A.5 Other data for the design of configuration (2x2x3, 2x2x6, 2x2x7) ... 132
A.6 Other data for 2x2x6 ... 135
A.7 Designers vs Engineers ... 139
A.8 Male vs Female ... 146
A.9 Bachelor vs Master ... 153
References ... 158
Sitography ... 167

List of tables

Table 1 Main creative process models (Warr & O'Neill, 2005)........ 16
Table 2 Comparison of criteria in the psychology literature........ 23
Table 3 Articles review summary; Idea Evaluation Methods (Dean, Hender, Rodgers, & Santanen, 2006)........ 26
Table 4 Dimensions identified in creativity literature (Dean, Hender, Rodgers, & Santanen, 2006)........ 38
Table 5 Definitions of the Quality dimensions and sub-dimensions (Dean, Hender, Rodgers, & Santanen, 2006)........ 40
Table 6 Derivations of General Constructs (Dean, Hender, Rodgers, & Santanen, 2006)........ 41
Table 7 Inter-judge reliabilities calculated in the experiment of Dean et al. (2006)........ 42
Table 8 Comparing criteria in psychological methodologies........ 45
Table 9 Summary of the designed criteria........ 59
Table 10 Macro database explanation........ 74
Table 11 Detailed database explanation........ 74
Table 12 Tests of Within-Subjects Effects (means 2x2x3)........ 77
Table 13 Tests of Within-Subjects Effects (all population 2x2x3)........ 78
Table 14 Condition * Hemisphere * Area (means 2x2x3)........ 79
Table 15 Condition * Hemisphere * Area (all population 2x2x3)........ 79
Table 16 Tests of Within-Subjects Effects (means 2x2x6)........ 81
Table 17 Tests of Within-Subjects Effects (all population 2x2x6)........ 82
Table 18 Condition_ * Hemisphere * Area (2x2x6 means)........ 83
Table 19 Condition_ * Hemisphere * Area (all population 2x2x6)........ 84
Table 20 Tests of Within-Subjects Effects (means 2x2x7)........ 86
Table 21 Tests of Within-Subjects Effects (all population 2x2x7)........ 87
Table 22 Condition_ * Hemisphere * Area (means 2x2x7)........ 88
Table 23 Condition_ * Hemisphere * Area (all population 2x2x7)........ 89
Table 24 Participants' data........ 116
Table 25 Stimuli presented to participants........ 117
Table 26 further literature review illustration........ 118
Table 27 Example of judges' Excel sheet........ 131
Table 28 Mauchly's Test of Sphericity (2x2x3 means)........ 132
Table 29 Mauchly's Test of Sphericity (2x2x3 all population)........ 133
Table 30 Mauchly's Test of Sphericity (2x2x6 means)........ 133
Table 31 Mauchly's Test of Sphericity (2x2x7 means)........ 134
Table 32 Mauchly's Test of Sphericity (2x2x7 all population)........ 134
Table 33 Tests of Within-Subjects Effects........ 136
Table 34 Mauchly's Test of Sphericity........ 136
Table 35 Estimates marginal means - Area........ 137
Table 36 Pairwise Comparisons........ 138
Table 37 Estimate marginal means - Condition * Area........ 138
Table 38 Estimate marginal means - Hemisphere * Area........ 138
Table 39 Participants selected for the Background comparison........ 139
Table 40 Tests of Within-Subjects Effects (Background)........ 141

Table 41 Estimated marginal means - Background * Area 141
Table 42 Estimated marginal means - Background * Hemisphere 142
Table 43 Estimated marginal means - Background * Hemisphere * Area 142
Table 44 Estimated marginal means - Background * Condition 143
Table 45 Estimated marginal means - Background * Condition *Hemisphere * Area............ 144
Table 46 Levene's Test of Equality of Error Variances 145
Table 47 Participants selected for the Gender comparison 147
Table 48 Tests of Within-Subjects Effects (Gender) 148
Table 49 Estimated marginal means - Gender * Area 149
Table 50 Estimated marginal means - Gender * Hemisphere 149
Table 51 Estimated marginal means - Gender * Hemisphere * Area 150
Table 52 Estimated marginal means - Gender * Condition * Hemisphere * Area 152
Table 53 Levene's Test of Equality of Error Variances 152
Table 54 Participants selected for the Degree level comparison 153
Table 55 Tests of Within-Subjects Effects (Degree’s level) 155
Table 56 Estimated marginal means - Degreelevel * Hemisphere * Area................................ 156
Table 57 Levene's Test of Equality of Error Variances 157

List of figures

Figure 1.1 Generic creative process (Warr & O'Neill, 2005)........ 9
Figure 1.2 Conceptual design process (Cantamessa & Montagna, 2016)........ 10
Figure 1.3 Ideation process: inputs and outputs (Hay, Duffy, & Grealy, 2019)........ 11
Figure 3.1 Experiment phases........ 48
Figure 3.2 Block presentation in Colombo (2019) experiment........ 50
Figure 3.3 Stimulus presentation in Colombo (2019) experiment........ 51
Figure 3.4 brain composition - (Soraghan, 2010)........ 63
Figure 3.5 Cerebrum composition – Source: www. 5inc-elements.com........ 64
Figure 3.6 neuron illustration (Puppo, 2018)........ 65
Figure 3.7 Synaptic transmission – Source www.thestudentscientist.org........ 65
Figure 3.8 illustration of an ideal wave........ 67
Figure 3.9 Brain waves classification (Younan, 2012)........ 67
Figure 3.10 Example of EEG rough output (Glass, Prieur, Molnar, & Hamiwka, 2006)........ 68
Figure 3.11 Example of a response visual inspection........ 70
Figure 3.12 Example of an independent component brain representation........ 70
Figure 3.13 Electrodes configuration........ 73
Figure 3.14 3-Areas configuration........ 76
Figure 3.15 6-Areas configuration........ 80
Figure 3.16 7 areas configuration........ 85
Figure 4.1 6-areas configuration........ 94
Figure 4.2 Estimated marginal means – Area........ 95
Figure 4.3 Estimated marginal means - Condition*Area........ 96
Figure 4.4 Estimated marginal means - Hemisphere*Area........ 97
Figure 4.5 Estimated marginal means - Background*Area........ 99
Figure 4.6 Estimated marginal means - Background*Hemisphere........ 99
Figure 4.7 Estimated marginal means - Background*Hemisphere*Area........ 100
Figure 4.8 Estimated marginal means - Background*Condition........ 101
Figure 4.9 Estimated marginal means - Background*Hemisphere*Condition*Area (left hemisphere)........ 102
Figure 4.10 Estimated marginal means - Background*Hemisphere*Condition*Area (right hemisphere)........ 102
Figure 4.11 Estimated marginal means - Gender*Hemisphere........ 104
Figure 4.12 Estimated marginal means - Gender*Hemisphere*Area........ 105
Figure 4.13 Estimated marginal means - Gender*Hemisphere*Condition*Area (left hemisphere)........ 106
Figure 4.14 Estimated marginal means - Gender*Hemisphere*Condition*Area (right hemisphere)........ 106
Figure 4.15 Estimated marginal means - Degreelevel*Hemisphere*Area (left hemisphere)........ 108
Figure 4.16 Degreelevel*Hemisphere*Area (right hemisphere)........ 109

Introduction

The present research resumes the experiment designed by Samuele Colombo (2019), providing further support to the phases following the data gathering process. The aim is to investigate the design mechanisms adopted in the earliest phases of product development in industrial design, with a neurophysiologic perspective. Various fields have been encountered in order to complete the present research, including Cognitive Neuroscience, Psychology, Management Engineering, Statistics and Design.

The Problem

In the last decades, the design process is getting increasingly important, permitting to create and develop new and disruptive products. For this reason, several types of research have been conducted in order to improve and support this step. The discipline that scientifically studies designers and design processes is called Science of Design and see its roots in the study of Herbert Simon (1969). There are two distinct approaches aimed at gaining information on the mechanisms involved in creative design that often have been coupled in order to obtain more reliable and meaningful results. The first ones are protocol studies, and the second ones are outcome-based researches (Hay, et al., 2017).

Protocol analysis is an empirical, observational research method focused on the analysis of piece of records, protocols, that in design studies are represented by verbalisation, videotapes or sketches (Yen & Jiang, 2009). Output-based studies are focused on the assessment of the results of the generation process in order to evaluate the outcomes of design processes. These approaches have been used coupled together for several decades to study the design process from a behavioural perspective. However, they lack cognitive information, neglecting data regarding subconscious processes taking place in the designer brain. For this reason, in the last years, several tools have been adopted in order to understand the cognitive design processes. The main ones are neuroimaging and neuroelectric techniques allowing researchers to study the brain of designers at a neurophysiologic level in a standard impersonal manner (Bigler, 2013). In ideation design literature, neuroimaging and neuroelectric techniques are still at the beginning, and several inconsistent results have been obtained due to the lack of a standard in the design of the experiment; therefore, the aim of the present research is the prosecution of the design of the experiment developed by Colombo (2019) based on the state of art of the field.

The Purpose of the Research

The objective of the experiment is to recognise neurocognitive and physiological patterns through instrumentations such as the Electroencephalogram (EEG) and the Eye-tracker in order to evaluate the cognitive activations during the process of idea generation in designer's activities. It has been taken as primary reference the model proposed by Guilford (1950) for the analysis of the ideation process, recognised as being the creativity construct most adopted in literature so far (Runco, 2010). This model states that the ideation process is characterised by two main phases, called Divergent Thinking, a non-linear process aiming at the generation of different ideas about a topic, and Convergent Thinking, aiming at structuring, organising and analysing all the ideas generated through divergent thinking. Alternating one to another, they give the possibility to the designer respectively to generate a solution space and to choose the most suitable idea from it to solve the stated problem. In order to assess these two aspects, it has been adopted a revised version of the Alternative Uses Task (AUT) experiment developed by Guilford (1967), in which participants were asked to find several alternative solutions to everyday objects in a limited amount of time.

In the literature, several neuroimaging and neuroelectric experiments were designed based on the AUT model (Benedek et al., 2013; Benedek et al., 2017; Camarda et al., 2018; Fink & Neubauer, 2008; Fink et al., 2007; Fink et al., 2009; Jauk et al. 2012), in particular, the one proposed by Jauk et al. (2012) that has been taken as the primary reference for the design of the present experiment with the proper modifications to try reaching higher reliability. A reference was needed to obtain comparable results due to the lack of a clear definition of the creativity construct and a standard regarding the experimental methodologies.

The present experiment has as objectives to extend further data in the literature regarding the ideation design cognitive processes, to test a clear and defined experiment pipeline and to provide the information required in order to replicate the experiment in further studies. Furthermore, results will be compared to the ones obtained in the experiment designed by Jauk et al. (2012) to verify the reliability of the design of the experiment adopted.

Besides, it cannot be excluded, being this branch just at its beginning, that the output of the present experiment could be adopted in future studies in order to improve the results of the ideation process through the control of environmental variables or the implementation of new strategies for designers.

The present thesis fits into the experiment in the phases following the data-gathering process, aiming at the design of an idea assessment methodology based on the literature review needed to evaluate verbal responses by participants, the design of the statistical analysis regarding the electroencephalogram outputs and a partial analysis of the data obtained. The present paper has not as objective to obtain final results and to fully describe the population due to the massive quantity of data attained from the data-gathering phase and to the various analysis that could be executed on the population. However, the partial results are needed in order to provide a first feedback on the experiment and for directing further studies.

Methodology

The present paper has been firmly based on the literature in order to allow results comparison. The primary source of information was the thesis developed by Colombo (2019), from which the information required for the experiment contextualization has been extracted. Furthermore, it has been used to understand the state of art of the experiment.

The experiment was designed following the Alternative Uses Task paradigm, highly adopted in the assessment of Divergent Thinking, representing the primary model describing the ideation process in literature. In particular, the experiment was a replication of the one proposed by Jauk et al. (2012) with the proper modification, validated by the neuroscience department of the University of Turin, in order to reach high reliability. Researchers selected 40 participants that took part in the experiment. 40 everyday objects (stimuli) were presented to them through slides divided into two distinct blocks, one including the condition common and other uncommon. It was then required to each participant to display a single function for object according to the block condition, respectively a common solution or an alternative and useful one.

In order to gather neurophysiologic information, an Electroencephalogram has been adopted, with the support of an Eye-tracker to capture more data (neglected for the purpose of the present research). The software Matlab has been adopted to compute the pre-processing and post-processing analysis of the data, and the biomedical engineering department of the University of Turin supported the research within these phases.

It has been conducted a firm literature review for the definition of a methodology for the idea evaluation of responses given by participants. This research was necessary because of the lack of a suitable method that could have satisfied the requirements of the experiment, thus ensuring the replicability of results, referring to the earliest stages of idea generation and including in the assessment a quality component of the idea. This research has been necessary in order to allow future analysis to compare the brain signals with the actual Idea Creativity of the responses.

Regarding the design of the post-experiment phase, a literature review has been conducted in order to understand which could have been the best configuration to be adopted to compute the statistical analysis. A strong assumption has been taken in order to compute this phase of the experiment, that is that the task condition (common or uncommon) was a reliable index of Idea Creativity. This simplification was due to the amount of time required for the evaluation of responses by judges. The analysis has been computed to permit to have partial results allowing a first comparison with the literature, including several factors. SPSS has been adopted to compute repeated measures analysis of variance (ANOVA) considering both within-subject and between-subjects factors.

Document structure

The document is divided into four chapters, excluding the introduction and conclusions.

The first chapter refers to the contextualization of the experiment displaying the basic concepts required to understand the experiment, especially from the behavioural point of view, analysing the models in the literature to describe the idea generation process.

In the second chapter, it is present a literature review aimed at analysing the idea assessment methodologies. This section required a multidisciplinary approach, including research on the psychological field for the Alternative Uses Task analysis and a study concerning the methodologies for the evaluation of products and ideas in the design area.

The third chapter concerns the design of the experiment. It examines the state of art of the experiment at the beginning of the present thesis, and the phases encountered within this study. This section includes the definition of the idea assessment methodology, basic neuroscientific information, the preparation of the EEG signals data and the analysis of several configurations to define which one is more representative of the population.

The fourth chapter is focused on the analysis of results, adopting as within-subject factors the Conditions associated with the answer of participants (Common vs Uncommon), the Hemisphere sides (Left vs Right) and the brain Areas (Frontal ventricular, Frontal dorsal, Central ventricular, Central dorsal, Posterior ventricular, Posterior dorsal). Further repeated measures ANOVAs are displayed considering as between-subjects factors the Background (Engineering vs Design), the Gender (Male vs Female) and the Degree level (Bachelor vs Master) of participants.

The conclusion section includes the main topics encountered in the present research. It gathers the main information regarding the experiment, the results obtained by the analysis and some suggestions aimed at the refinement the experiment for further studies.

1 Context

1.1 Product development

In today's management, there is a tendency to carefully implement and standardise processes (process-oriented view). Product development is not always as structured as other phases (e.g. manufacturing processes); when highly innovative changes are involved, this generation process is chaotic, unpredictable, and includes numerous iterations between activities. It is not possible to know ex-ante the final product configuration, and, consequently, how to set a development strategy. As demonstrated in Schön's studies (Schön, 1983, 1995), designers could follow predetermined procedures, but tends to change them and their problem-solving approaches according to the information gathered during the execution of the activities. Hence, it is possible to follow a clear and structured process only if the content of the product development project includes a low degree of innovation. In addition to that, it follows a predetermined structure only when the project has big sizes, making the planning phase meaningful.

The resource-based view (RBV) gives further support about the complexity of this process stating that all firms are different, and each operation carried out is distinct according to routines existing in the company. Firms' environments are various, and enterprises tend to have different resources and competences.

Product development could be classified into different phases. The sequence presented is not always respected in reality because, as explained before, there are often iterations in which designers tend to go back to modify or improve the work already done. Furthermore, it is common to use the concurrent engineering approach in which designers start the following phase before the completion of the former one.

The phases identified by Cantamessa & Montagna (2016) are:

- Product planning – is the first phase of the product development process. This phase is highly inter-functional and has the aim to define the new product characteristic meeting the market and the available technology of the company. At this point, the company could decide the positioning of the product based on the target market of the firm and the competitors and, consequently, customers' needs and requirements are defined.

- Conceptual design – has the aim to define the technical solution that can meet the user requirements, creating a product concept. During this phase, designers decide to use an already existing product concept or to create a new one.
- System-level design – is the phase in which designers convert the product concept into an architecture, also focusing on the primary subsystem and components. Another critical aspect to be analysed in this phase is the definition of the system interacting during the lifecycle. In this phase, designers are asked to decide the customisation and the outsourcing of the components.
- Detailed design – in this phase, decisions regarding materials and components are carried out. The aim of these is to meet the required specifications, eventually to review and modify the system-level choices.
- Prototyping and testing – this phase aims to verify and validate the solution found in the detailed design phase through prototypes or software simulations. In this activity, it is fundamental to pay attention to regulations and laws concerning the product and its ecosystem.
- Process design – the goal of this phase is to define the processes that will permit to produce, distribute and service the final product.
- Product launch and production – is the final phase aimed at calibrating the output, to reach the market. The first part is carried out through pilot lines for testing purposes, the second increasing the pace of the machine until reaching full production.

The present research is mainly focused on the conceptual design phase, in which designers need to be creative to find alternative and useful solutions. As explained, this phase is complex and needs to be carefully evaluated.

1.2 Conceptual design & idea generation

In the product development process, after the identification of the customer needs, designers have the information required to define and outline the problem. At this point, an iterative process starts involving the exploration and generation of new alternatives and the selection of the best concepts proposed. Designers aim to find a technical solution that, in theory, could fulfil the required functions (Ulrich & Eppinger, 1995). Each combination obtained performing the needed tasks is called concept (Pugh, 1981). After a series of iterations, every round discarding the worst solutions obtained and improving the better ones, a few concepts are kept and compared to each other, to find the best feasible solution for the problem. The model presented below (Fig. 1.1) expresses the generic creative process (Warr & O'Neill, 2005) composed by the iteration of 3 micro phases (problem preparation, idea generation, and idea evaluation).

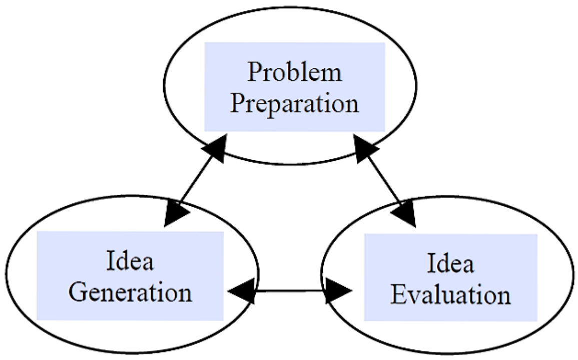

Figure 1.1 Generic creative process (Warr & O'Neill, 2005)

The conceptual design process is composed, as shown in the picture below (Fig 1.2), by an alternation of divergent (idea generation) and convergent (idea evaluation) phases, in which respectively there is a focus on the production of new ideas and the screening for the best one. Designers must work on multiple concepts and decide which are the most suitable ones to be implemented in the product. The divergent phase is more critical in radical innovation products, while in incremental changes, the convergent stage plays a considerable role.

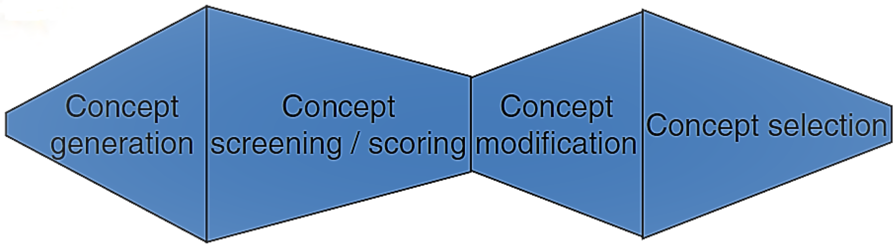

Figure 1.2 Conceptual design process (Cantamessa & Montagna, 2016)

In conceptual design, designers should find the right trade-off between creativity and conservation of the characteristics of the past product. Conceptual design and System-level phases are fundamental in the product development process because the decisions taken in these phases regarding the architecture of the product will have a lock-in effect for its possible future improvements.

Several preliminary approaches have been developed to identify functions that could be implemented in new products (e.g. interviewing lead users or domain experts, reading technical and scientific literature, patent search, bio-inspired design) (Cantamessa & Montagna, 2016). The majority of these approaches rely on the interaction of groups of people, aiming at the generation of new ideas fitting the prerequisites. Several methods called idea generation (ideation) methods have been developed in order to ease the creative process executed by designers and general participants to group reviews, among which the main ones are:

- Brainstorming – a technique developed by Osborn in 1953 (Osborn, 1953) that consists of the free exchange of ideas by participants without any critical comments on the ones presented. A facilitator is usually present to help to have a smoother flow of proposals.
- SCAMPER – a method complementary to brainstorming and, following the same principles, requires participants to proposed ideas associated with predefined categories.
- Wish and wonder & law-breaking – an approach in which designers are asked to proposed ideas relative to an ideal world with different laws of physics. It permits to create a first original solution that usually needs some adaptations for the implementation.
- TRIZ – a more sophisticated methodology that will not be displayed in detail in the present research. It is a technique highly used to simplify human involvement in activities. It relies on the contradiction principle between technical systems and their environment.

Currently, new researches and models have developed to describe the concept creation process. The model developed by Finke et al. (1992) proposed a clear view on what cognitive ideation processes rely; it states “that the ideas created by a person are founded in their knowledge and experiences, and produced through the association, integration, and transformation of internal representation” (Hay et al., 2019). These representations could derive from episodic or semantic memory responses to internal (generated) or external sources. The design ideation process, presented in the picture below (Fig. 1.3), relies on inputs, designer’s knowledge, and outputs, the designer's proposed concepts.

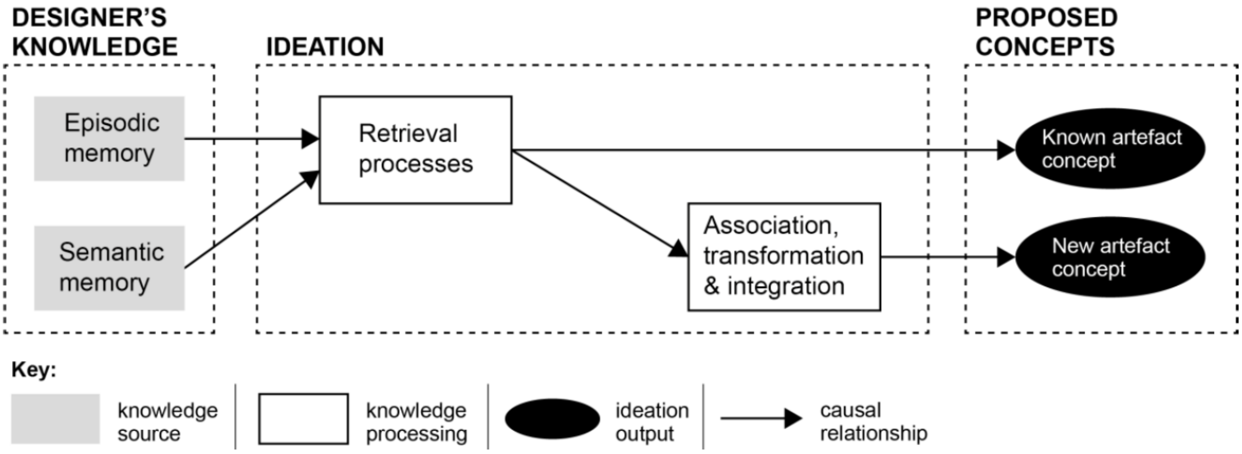

Figure 1.3 Ideation process: inputs and outputs (Hay, Duffy, & Grealy, 2019)

The Alternative Uses Task experiment (AUT), to which this thesis is referring, is an ideational method born to evaluate creativity, more correctly divergent thinking in the psychological field. It was chosen among other design of experiments for allowing more accurate analysis of the idea generation process at a neuroscientific level. It is displayed in detail in the chapters.

One of the goals of this paper is to develop a suitable model for the evaluation of responses given by participants in the AUT experiment (see in detail in Chapter 3.1) proposed by Colombo (2019).

In the next subchapter, the Science of Design, the science studying design processes through a scientific methodology, will be introduced.

1.3 Science of Design: Protocol analysis & main issues

The concept of Science of Design sees its roots in the study conducted by Simon Herbert in 1969, “The Science of the Artificial” (Simon H. A., 1969), in which he suggests to study the design process with a cognitive perspective aiming at discovering and understanding the artificial world. “The design process is the basis and the foundation of all existing products and artefacts” (Simon H. A., 1969).

The terminology Science of Design was coined by Nigel Cross in 2001 (Cross, 2001), who defines it as a “study of principles, practices, and procedures of design” (Simon H. A., 1969). It is a scientific investigation of the design activity, more accurately, the discipline that studies designers and design processes scientifically. It is a multidisciplinary field that involved several sectors (from organisational to psychology) with the final aim to obtain output usable as input for the Design Science. This latter is a discipline focused on the improvement of designers’ methodologies, aiming at providing practical tools that permit designers to be more productive and effective.

Hay (2017) proposed the existence of two distinct approaches aimed at gaining information on the cognitive mechanisms involved in creative design. The first ones are protocol studies, and the second ones are outcome-based methods.

Protocol analysis is an empirical, observational research method that consists of the “direct observation of the designer’s interaction with the problem domain and from models of design reasoning” (Purcell & Gero, 1998). The problem domain, the first aspect highlighted in the paper just mentioned, is composed of two orthogonal dimensions, respectively focused on the artefact at a tangible and an abstracted level. The design reasoning analysis, instead, has the purpose of decomposing the strategy utilised by designers into piece of records, protocols, that in design studies are represented by verbalisation, videotapes or sketches (Yen & Jiang, 2009).

Schön was the first researcher having significant results that, using protocol studies, focused on the analysis of the cognitive process in the design phase. With his studies, Schön discovered that the design process is strongly iterative and based on feedback loops, each one a step towards the final goal. Gero and Kannengiesser (2004), using the protocol approach, discovered that designers tend to use multiple cognitive and logical approaches.

Protocol analysis is a methodology characterised by some drawbacks (Dinar, et al., 2015). They rely mainly on the verbalisation of the thoughts by designers in order to understand the strategies used for problem-solving (Colombo, 2019) and this implies the loss of subconscious concepts that could have been important for the analysis. Furthermore, there is a delay between visualisation and the verbalisation of thoughts. Lastly, the interpretation of responses is a subjective process that lacks objectivity (Dinar, et al., 2015).

The output-based analysis is focused on the assessment of the results of the generation process in order to evaluate the behavioural outcomes of the designer activities (Hay, et al., 2017). They are discussed in detail in the second chapter of the present research.

With the introduction of new technologies, it has been possible to link and integrate output-based approaches with a direct study of the cognitive processes of designers at a neurophysiologic level. This approach permits to gather neuro-cognitive data and to compare them to the behavioural outcome of the generation process. The AUT experiment proposed by (Colombo, 2019) was developed with this purpose and will be analysed in more detail after having cleared other essential concepts needed in order to understand the present research.

1.4 Neuroscientific experiments in Design & divergent thinking

The drawbacks of protocol studies have been overcome with the introduction of different tools in support of the analysis, among which the neurophysiologic ones. They allow researchers to study the changes in response to external stimuli directly from the brain. In the last decades, neuroscience studies took a step forward in the understanding of several cognitive processes, such as attention, memory, language, emotions, etc., but not concerning ideation and creativity processes. This gap is due to several reasons; first of all, due to the lack of a standard regarding the experimental methodologies and the measurement of them (Colombo, 2019; Dietrich & Kanso, 2010) and secondly, because of a lack of results consistency (Saweyer, 2011). Ideation is a difficult characteristic to be assessed in a controlled environment because it is not a cognitive process that could be voluntarily recalled by subjects. The most adopted neurophysiologic techniques are neuroimaging and neuroelectric techniques.

Neuroimaging studies consist of the use of various techniques providing images of the brain's anatomical structure. Some examples of neuroimaging techniques are the functional magnetic resonance imaging (fMRI), the Positron emission tomography (PET), the single-photon emission computed tomography (SPECT), and the near-infrared spectroscopy (NIRS). Neuroelectric researches consist of the study of the brain activation through the analysis of the ERP, voltage variations evoked by an event, in different areas of the brain. The most adopted technique is the electroencephalogram (EEG).

In recent years, several variants of AUT have been implemented, coupled with neuroimaging and neuroelectric measures to assess Divergent Thinking. The experiment of Fink and Neubauer (Fink & Neubauer, 2006) was one of the first to be implemented using this approach. It was aimed at identifying differences in divergent thinking before and after two weeks of creativity training. In 2008, Fink and Neubauer (Fink & Neubauer, 2008) used the same methodology integrated with other tests to analyse the relationship existing between extraversion and originality. In 2009, Fink et al. (Fink et al., 2009) compared four ideation creativity methods, among which the AUT experiment, using neuroimaging techniques.

In 2013, Benedek et al. (Benedek et al., 2013) studied the Alternative Uses Task using the fMRI approach with an expedient: in order to increase the accuracy of the test, after any generation of idea, he asked subjects to explicitly tell whether they recalled memories or they proposed novel ideas. The purpose of it was to analyse creativity from a personal perspective, demonstrating increased activation in brain regions associated with the integration of information in new ways and executively controlled mental stimulation (Hay et al., 2019).

The experiment proposed by Colombo (2019) adopted a similar paradigm to the one by Jauk, Benedek, and Neubauer (Jauk et al., 2012), involving an adapted version of the Alternative Uses Task integrated with EEG measures.

1.5 Creativity: convergent thinking vs divergent thinking

Creativity is a fundamental aspect in designing useful and novel products, and in supporting companies to capture larger market share. It is also highly related to the performance of businesses (Santosh, 2019) and is a quality that could be exploited in several different areas in which decision-making processes are involved; examples could be taken in the field of marketing, strategy, design, management, etc..

It is essential to introduce and study the concept of creativity because the main goal of the AUT experiment is the assessment of divergent thinking, characteristic historically related to creativity in individuals. Furthermore, it is a frequently assessed metric in the literature of idea generation.

Creativity is a concept broadly used that assumes different meanings depending on the context in which it is adopted. Warr (Warr & O'Neill, 2005) distinguished three main types of creativity:

- Creative process – the internal process of an individual through which ideas are generated.
- Creative person – considered as a personality trait of an individual.
- Creative product – as the reflections in the products of signs of creativity.

This section focuses on the explanation of creative processes, while in the second chapter, there is a more in-depth analysis regarding the assessment of creative products and ideas. In literature, there have been developed a lot of different models that describe the creative processes, all including the following phases: an analysis of the problem, a generation of ideas (divergent thinking) and an evaluation of ideas (convergent thinking). In the table below, the chief creative process models are listed and compared to each other.

<table>
<tr><th>Models</th><th colspan="2">Analysis of Problem</th><th colspan="2">Generating Ideas</th><th>Evaluating Ideas</th><th>Donating</th></tr>
<tr><td>Wallas [48]</td><td colspan="2">Preparation</td><td>Incubation</td><td>Illumination</td><td>Verification</td><td>X</td></tr>
<tr><td rowspan="2">Osborn [36]</td><td colspan="4">Idea Generation</td><td rowspan="2">Idea Evaluation</td><td rowspan="2">X</td></tr>
<tr><td colspan="2">Fact-finding</td><td colspan="2">Idea-finding</td></tr>
<tr><td>Amabile [2]</td><td>Problem or task presentation</td><td>Preparation</td><td colspan="2">Response generation</td><td>Response Validation</td><td>X</td></tr>
<tr><td rowspan="2">Shneiderman [44]</td><td colspan="2">Collect</td><td colspan="3">Create</td><td>Donate</td></tr>
<tr><td colspan="6">Relate</td></tr>
</table>

Table 1 Main creative process models (Warr & O'Neill, 2005)

For the present research, the construct proposed by Guildford (Guilford J. P., 1950) is taken as reference for the definition of the design ideation process, characterised by two main phases, called Convergent Thinking and Divergent Thinking. These two processes, alternating one to another, give the possibility to the subject to generate a solution space and to choose the most suitable idea from it to solve the problem. They are both fundamental in the design process to get good results.

Divergent thinking is a non-linear phase, in which the mind of the subject tends to be free-flowing, and connections between ideas are created, according to past experiences of the subject and external stimuli. Divergent thinking tests are mainly based on problems that do not have one single right answer and are focused on the measure of the solution space generated by the subject. It is a valid and reliable indicator of creative production that allows researchers to study creativity in controlled environments. The convergent thinking process has the aim to structure, organise and analyse all the ideas generated through divergent thinking.

Following Shah et al. (2003), creativity is difficult to be assessed analysing the process itself, so the best way to measure it is through an evaluation of its outcome (ideas or products). However in the last decades, as explained in the previous sub-chapters, further methodologies and tools have been created that allows researchers to improve the analysis of the process itself adopting neuroimaging and neuroelectric instruments in order to integrate the outcome of the process with data gathered during it, and to collect data in a standard impersonal manner (Colombo, 2019).

Concretely, this branch of study is still at the beginning, and several conflicting results have been obtained (Dietrich & Kanso, 2010), for this reason, the experiment developed by Colombo (2019) has as objective to extend further data in the literature, linking the neuroscientific and the behavioural approaches.

Therefore, the main aim of the research is to analyse the ideation process through data gathered through EEG instrument and integrating them with the outcomes obtained from the experiment in a behavioural perspective (further information in the third chapter), enabling a broader view on the ideation process. The cognitive analysis of responses by participants will be an indicator of the results of the use of divergent and convergent thinking.

2 Literature review: assessment methodologies for verbal responses in design creativity

This chapter is fully dedicated to the analysis of the literature in order to find a suitable method for the evaluation of the behavioural responses given by participants in the experiment developed by Colombo (2019), allowing the comparison between them and the neuro-signals gathered. The methodology should satisfy precise requirements:

- It should focus on the evaluation of creativity at the idea level.
- It should take a design perspective, therefore covering also the Quality aspect of an idea.
- It should be suitable for the assessment of the earliest phases of product development; thus, it should not be an approach aimed at the evaluation of extremely detailed ideas or final products.
- It should ensure the replicability of results.

A careful analysis of the literature is necessary because most of the methods developed for the assessment of verbal responses lack one or more of the prerequisites just mentioned.

This chapter will be composed of a first review concerning criteria usually adopted in the AUT experiments. Then, a study regarding idea evaluation methods will be conducted. In this part, the most adopted approaches will be analysed in detail: the Consensual Assessment Technique developed by Amabile (1982), the approach created by Shah et al. (2003) and, finally, the approach refined by Dean et al (2006). At the end of each sub-chapter, takeaways and drawbacks of the approaches will be presented.

In the third chapter, the assessment methodology based on the literature review will be defined, disclosing the parameters, the background that the judges should have for the assessment, the guidelines for the raters and, in the end, the actual procedure that judges should follow in order to have a reliable method.

2.1 AUT history & main criteria

In order to understand the AUT proposed by Colombo (2019) and determine how to define the criteria assessment process, it is first essential to analyse the background of the experiment, to understand why it was born, what was the purpose of it, and which were the relevant metrics assessed in it.

Eysenck (1993) made a clear distinction concerning the subject of creativity assessment: the first creativity cluster is at personal level (e.g. Elon Musk, Beethoven), as a personality trait that follows a normal distribution within the population; the second one is associated with the finished artefact (e.g. an auto-mix mug), following a Poisson distribution within a group of ideas. The AUT experiment was born as an experiment to assess the personality traits of an individual, more correctly to his/her divergent thinking inclination.

In the psychology literature, several models have been developed, relying on different definitions and assumptions concerning creativity. Currently, there is not a clear and unambiguous definition that is entirely accepted.

In the earliest researches, creativity was highly associated with intelligence and was assessed using tools like modern IQ tests (Chassell, 1916; Andrews, 1930; McCloy & Meier, 1931). The results of these experiments were unsatisfying, finding that the correlation existing between intelligence and creativity was close to 0.

The turning point was in 1927 with the study of H.L. Hargreaves (Hargreaves, 1927), in which he introduced the concept of fluency as a component of intellectual ability. Fluency is a measure of creativity that focused on the quantitative aspect. The correlation found in his study between IQ and fluency was r=.30, showing that they are related to each other, but not identical.

Guilford (Guilford J. P., 1950, 1967) is recognised as being one of the most significant researchers in the creativity field. Following the studies led by Hargreaves on fluency, he developed a method to assess at the same time both the quality and quantity of an idea. In 1967 he developed a model called Structure of Intellect (Guilford P. J., 1967), fitting both intelligence and creativity and stating that these two personality traits could be considered as orthogonal characteristics. Concretely, it means that a person could be highly creative and have low intelligence and vice-versa.

In 1954 Guilford et al. (1954), in a study involving cadets and student officers, determined the criteria that are still used in many experiments to assess divergent thinking:

- Fluency – regarding the number of ideas proposed by participants.

- Flexibility – regarding the solution space explored and proposed by participants.
- Originality – proposed as a comparative measure regarding the uncommonness of an idea.
- Elaboration – assessing the level of detail in the description.
- Sensitivity to problems – regarding the ability of the subject to deal with the problem proposed (progressively put aside).

After the test proposed by Guilford et al. (1954), many researchers used the same experiment choosing participants with different backgrounds and using different settings. Lowenfeld and Beittel (Beittel & Lowenfeld, 1959) did a research on art and science students using fluency, flexibility, redefinition, sensitivity to problems, originality as metrics and obtaining a correlation of r=.33 with Guilford DT tests; Lauritzen and Wolford (1964) replicated the experiment predicting the originality of the answer with a correlation of r=.48; Barron and Harrington (1981) obtained a correlation of r=.30 between creativity and divergent thinking in students; Barron (1987) replicated the experiment again with a group of adults and obtained the same correlation in originality rating. In 1961 Wallace (1961) demonstrated the existence of a high correlation between problem-solving in customer service and divergent thinking. Elliot in 1964 (Elliot et al., 1964) confirmed 5 of the 8 models proposed by Guilford to assess creativity in public personnel relationships as well. In 1961 two different experiments, one proposed by Mackinnon (1961) on architects and the other by Gough (1961), have shown low correlation using the divergent thinking tests of Guilford.

In 1967, Guilford designed his final version of the Alternative Uses Task (AUT). In this experiment, subjects were asked to generate the highest amount of alternative and uncommon uses to an everyday object, such as a brick or a paperclip, in a limited amount of time. This experiment has been highly used in the following years, and several variants have been developed among which:

- Wallach & Kogan (1965) studied schoolchildren using the AUT test and introducing the concept of uniqueness and productivity (also called “abundance” or “ideational fluency”). In the evaluation of creativity, they used five tests proposed by Guilford and obtaining a correlation of r=.40 in the experiments. In the same research, they tried to correlate IQ test and DT, obtaining unsatisfying results. This experiment was replicated in 1968 by Cropley (1968), reaching the same conclusions.
- Torrance Test of Creative Thinking (TTCT) (Torrance, 1966, 1974, 1990, 1998, 2008) is a test proposed by Torrance in 1966 on personality using sketches, semantic tests, and variants of Alternative Uses Tasks.

Krampen (1993) defined four categories to measure creativity: biographical methods (analysing past data about the subject) self-report measures (relying on reports provided by the subject) peer assessment measures (where subjects evaluate the results of other participants) or psychometric tests (developed to measure individual mental capabilities). The present research focuses only on psychometric tests (Colombo, 2019).

The assessment methodology should be designed considering that the variant of the AUT experiment used for the test, involved mostly the divergent thinking, used by subjects to explore the solution space in the uncommon condition, and convergent thinking, adopted to find the common solution. In the design's perspective, convergent thinking is necessary to find an alternative solution to the problem and the convergent thinking is needed to ensure the feasibility and usefulness of the solution obtained.

In the Experiment of Guilford, as explained before, four main criteria have been adopted: two regarding the assessment of a group of ideas (Fluency and Flexibility) that are not relevant for the experiment here described, and two regarding the assessment of single ideas (Elaboration and Originality). Of the former two, Flexibility could be an attractive criterion to be measured, but with a different scope from the original one, that is to define which stimulus could allow the subject to have a bigger solution space. Elaboration and Originality are, instead, both important to be assessed; Elaboration represents the level of detailed expressed in one idea (Guilford J. P., 1967) and has been measured in different ways in the literature. It could be assessed both subjectively (with rates assigned by judges) or through an objective indicator (e.g. LSA, Forthman, 2018). Originality could be defined as the degree of "uncommonness" of an idea, referred to as "newness" by Sarkar and Chakrabarti (2006). This metric, as Elaboration, could be assessed through objective methods (e.g. using pre-set tables provided for some experiment), or with subjective ratings by judges, that compare ideas proposed by different participants and score them.

An alternative to uncommonness is Uniqueness (Wallach & Kogan, 1965), in which only ideas that are proposed just ones in the whole set are considered Novel. This method presents numerous flaws, such as the dependency of novel ideas on the number of ideas proposed in the set analysed and the reward of ideas that could be considered merely bizarre and not creative at all.

Summing up, AUT experiment was born to assess divergent thinking abilities of a subject and was complementary to other tools used to assess personality traits. The classical AUT experiment relies on two main aspects: the quantitative and qualitative assessment of an idea. For the experiment proposed by Colombo (2019), quantity is not a relevant variable to be assessed.

Criteria	Level assessed	Method	Note
Fluency (Guilford, Torrance)	Set of ideas	Corresponding to the number of alternative uses that a person can think of	The responses should be composed by multiple ideas. NOT APPLICABLE
Originality (Guilford, Torrance)	Single idea	Responses given by 5% of the set are considered unusual (1 point), the ones given by the 1% of the set are uniques (2 points)	2 partially objective methods to assess uncommonness Drawbacks: • uniqueness depends on the size sample • Both consider weird or unfeasible idea as «creative»
Uniqueness (Wallach & Kogan)	Single idea in a set	1 to unique responses and all the other responses receive 0	
Flexibility (Guilford, Torrance	Set of ideas	Concerning the range of ideas in different domains categories proposed by the subject	Could be assessed in a unique criteria called «variety» with the right adjustments
Elaboration (Guilford)	Single idea	corresponding to the level of detail and development of the idea amount of detail of a single response. Could be assessed by the number of words used	Metric important to be assessed. Used as indicator of convergent thinking

Table 2 Comparison of criteria in the psychology literature

2.2 Review of idea generation and idea evaluation methods

To highlight the difference existing between the creativity process in psychological studies, explained before with the Guilford theory of divergent and convergent thinking, and the one in design studies, illustrated in this chapter, it will be taken as reference the terminology coming from Osborn studies (1963), in which he defines the creative process as composed by idea generation and idea evaluation phases. In scientific ideation literature, creativity is defined as the ability to bring something that is both new and useful (Amabile, 1996; Sawyer, 2006; O’Hara, 1999) and not only novel as in the psychological field. Usefulness is needed in idea generation in order to distinguish bizarre ideas from the useful and implementable ones. It is essential to consider this aspect because designers work having constraints and specifications. Therefore, the experiment analysed in this paper has a different focus than the usual AUT experiments.

Cognitive psychologists consider Fluency and Novelty as a primary measurement of creativity, without considering utilities of product and ideas as engineering design should do (Shah & Vargas-Hernandez, 2003). For this reason, in the following subchapter, there will be an analysis concerning the primary methodologies used to assess ideas and, for completeness, products.

Idea generation (ideation) methods are techniques used by designers to generate new ideas or creative solutions to problems (Kudrowitz & Wallace, 2012). The most adopted method of idea generation is "brainstorming" (Osborn, 1953), that relies on the principle of free-flow association permitting participants to create complex and unusual associations.

Kudrowitz and Wallace (2012) divided ideation methodologies following two different classifications:

1) The first distinction regards the structure of the method and how the problem is addressed to subjects. They could be divided into free-form and structured idea generation methods:
    a. In *Free-form* methodologies, there is not an optimal solution, and there are a lot of alternative possibilities to solve the problem (Simon H. , 1973). Some examples of this kind of method are brainstorming, free association, and brainwriting. Free-form methods are also referred to as blue-sky idea generation method.
    b. *Structured problems* usually have an optimal solution and have some practical constraints. The best solution usually is the one that optimally satisfies constraints and specifications. Theory of inventive problem solving (TRIZ) and morphological analysis (Silverstein, 2008) are typical structured problems.

   In 2012, Chulvi et al. (Chulvi et al. 2012) demonstrated that structured methods tend to produce more useful ideas, while blue-sky idea generation methods tend to create more creative ideas.
2) Another distinction developed by Kudrowitz relies on the level of communication allowed to the subject during idea generation processes. They could be divided into traditional and nominal brainstorming methods:
    a. Traditional brainstorming - in which a group of people interacts, sketch and share ideas with the other components of the group. Commonly, there are groups of facilitators that help to ease communication.
    b. The nominal brainstorming – in which participants are not allowed to talk until the end of the brainstorming session. These methods are used in studies in which ideation ability is the primary variable of interest (as the experiment

proposed by Colombo). Nominal brainstorming presents some advantages principally related to the elimination of social pressure presents in within-group interactions. It permits to overcome production blocking, evaluation apprehension, and free-riding.

The AUT experiments can be classified as nominal Free-form problems because subjects do not have strict constraints (unless the condition common/uncommon suggested to them) and take the assessment independently.

In the next chapter, the main approaches to evaluate ideation production will be analysed.

2.2.1 Idea evaluation & the lack of a standard

Idea evaluation is the phase that follows the idea generation and is the process through which the most suitable idea is selected from a set. It could be done explicitly with an objective method (with objective criteria) or implicitly, with a mental screening by the designer before the idea proposal. The aim of this chapter, as anticipated in the previous sections, is to study idea evaluation objective methods, the methodologies adopted for the assessment of the outcomes of idea generation processes, referred to as Idea Creativity by Dean (2006).

During the years, many methods have been developed for the idea evaluation, the majority involving directly or indirectly the assessment of Novelty or Quality metrics. The Novelty aspect has been already highly discussed in the previous chapter, while "in short, a Quality idea is an implementable solution that will solve the problem, regardless of whether or not the idea itself is novel or unusual" (Dean et al., 2006). There is not an actual standard regarding idea evaluation, and this is highly visible from the research conducted by Dean (2006) in which he gathered about 90 studies between 1990 and 2005, finding that several metrics and methodologies have been used.

Method	Abbrev.	Number	%
Ideas were counted but no quality or creativity measures were considered	CINQ	18	20
A single holistic measure of idea quality or creativity or separate holistic measures of both were considered	SHM	21	23
Quality or creativity dimensions were considered but not measured separately	CBNM	6	7
Quality or creativity dimensions were measured separately then combined explicitly	MSCE	11	12
Quality or creativity dimensions were reported separately	DRS	34	38
	Total	90	100

Table 3 Articles review summary; Idea Evaluation Methods (Dean, Hender, Rodgers, & Santanen, 2006)

In the paper proposed by Dean, it is interesting to note that 20 % of the studies analysed were focused entirely on the quantitative aspects of creativity, counting the actual number of ideas generated in one response by the subject. 23 % of the studies considered for the assessment of ideas one single metric related to Creativity or Quality of response. Without proper guidelines given to raters, the inter-judge reliabilities of these studies were low due to some inconsistencies and biases that judges could have experienced regarding criteria definitions.
The remaining 57% of the studies relied on multicriteria methods:

- In 7 % of the studies, judges were asked to rate aggregated criteria, having at disposal the definitions of sub-criteria composing the macro criteria. These methods increase the consistency of results in comparison to the assessment of one single metric without proper guidelines about its composition.
- In 12 % of the studies, judges were asked to assess sub-criteria directly and then results were aggregated into one single metric by the researchers.
- In 38 % of the studies, quality and creativity subdimensions were assessed separately by judges and then not aggregated together, but left apart, as independent measures.

Several metrics have been adopted to assess ideas within these studies. Some examples are: Connolly et al. (1990) assessing Creativity and Workability, Kramer (1997) using Creativity , Feasibility and Effectiveness, Straus and McGrath (1994) adopting Feasibility and Impact, Valacich et al. (1992) assessing Importance, Redmond et al. (1993) analysing Originality and Quality, etc.
From the literature, it is interesting to note that even after the study conducted by Dean the criteria adopted were various, and no standard has emerged yet. Some examples are the research of Sarkar and Chakrabarti (2011), in which they assessed Usefulness and Novelty and the one by Chulvi et al. (2012), in which they assigned to each requirement an importance and then evaluated for each response a degree of satisfaction to find an overall idea rate.

In addition to the variety of metrics adopted, many methods have been developed to assess the same criteria. For this reason, in the next subchapters, the principal methodologies to assess ideas have been analysed. It has been taken as primary reference the research done by Kudrowitz and Wallace (2012) in which they gathered the methods (represented in the table below) adopted to assess product and ideas.

Study	Summary	Products in review	Dimensions of creativity				
Dean *et al.* (2006)	Review of 90 constructs for idea evaluation	Ideas (e.g. increase tourism in Tucson)	Novelty (originality, paradigm relatedness)	Workability (acceptability, implementability)	Relevance (applicability, effectiveness)	Specificity (implicational explicitness, completeness and clarity)	
Besemer and O'Quin (1986, 1999)[a]	Objective metric for creative product evaluation	Artefacts (chairs)	Novelty (surprising, original)	Resolution (logical, useful, valuable, understandable)	Elaboration and synthesis (organic, well-crafted, elegant)		
Shah and Vargas-Hernandez (2000, 2003)	Evaluation of mechanical engineering designs	Artefacts and ideas (mechanical devices)	Novelty (unusual)	Quality (meets specifications)	Variety (explored solution space)	Quantity	
Horn and Salvendy (2006, 2009)[a]	Consumer-based assessment of product creativity	Artefacts (chairs and lamps)	Novelty (frequency, rarity)	Importance (relevance, significance)	Affect (appeal, desire, attraction, delight, stimulation, etc.)		
Amabile (1982)[b]	Subjective assessment method of creativity	Artefacts (artwork and poetry)	Creativity (as determined by appropriate judges)	Creativity Cluster (novel material use, novel idea, effort, detail, etc.)	Technical cluster (technical goodness, organisation, neatness, etc.)	Aesthetic judgment (liking, aesthetic appeal, would you display it?)	
Christiaans (2002)[b]	Creativity as one metric of design review	Artefacts (cabinets and telephone booths)	Creativity (as determined by appropriate judges)	Technical quality	Attractiveness	Interest	Goodness of example

[a]In these studies, each of the dimensions are made up of many sub-dimensions in the form of bipolar adjective scales.
[b]In these studies, the assessment dimensions in addition to 'creativity' do not determine creativity, but were used to find correlations with creativity.

Figure 2.1 Selection of primary studies involving a metric for the evaluation of ideas/artefacts (Kudrowitz & Wallace, 2012)

For the experiment developed by Colombo, the methodologies constructed by Christiians (2002), Horn and Salvendy (2006) and Besemer and O'Quin (1986, 1999), analysed by Kudrowitz and Wallace, are incredibly detailed, for this reason they have been summed up at the end of the chapter, while a more in-depth analysis has been conducted on the methods proposed by Amabile (1982), Shah (Shah et al., 2000, 2003) and Dean (2006). The first methodology that has been analysed, even if it is mainly dedicated to artworks assessment, is the Consensual Assessment Technique by Amabile. It has been dedicated an entire subchapter to it because it was adopted in the experiment developed by Jauk et al. (2012) from which the present experiment is inspired. The following method analysed is the one proposed by Shah et al. (2000, 2003), that goes into depth in idea generation assessment expressing the main characteristics that criteria should have in idea evaluation methodologies. The last method explored is the one proposed by Dean et al. (2006) that refined the work proposed by Shah in order to increase the reliability of the results.

2.2.2 Jauk et al. experiment (2012): CAT methodology (Amabile, 1982)

In the last decades, several AUT experiments aimed at the ideation process analysis have been developed integrating the EEG apparatus; the majority of these approaches were lacking the assessment of the quality aspect of ideas and considering as primary criterion Fluency. Camarda et al. (Carmarda et al., 2018) used as parameters for the creativity evaluation Fluency, Remoteness, and Originality. In 2007, Fink et al. (Fink et al., 2007) assessed Originality and Fluency. In 2009, Fink conducted two different studies (Fink et al., 2009; Fink et al., 2009) both using Fluency as reference dimension of creativity and, concerning the EEG analysis, working on a personality trait level (Creative participants vs non-creative ones). In 2014, Schwab et al. (Schwab et al., 2014) assessed only Originality aspects concerning the behavioural scoring phase. The experiments just mentioned could not be taken as guidelines for the construction of the behavioural assessment method for the experiment of Colombo (2019), because of the lack of focus on the quality aspect of an idea.

The experiment developed by Jauk et al. (2012) was the primary reference concerning the present experiment. In their work, data regarding criteria to evaluate the responses of participants were minimal and share only a few information concerning the method adopted for the responses' evaluation. The main parameter assessed during the AUT was Originality on a 4-points Likert scale ("1" – "very common", "4" – "very uncommon"). For the evaluation were adopted six judges, among which three male and three female students of the University of Graz. The method proposed to assess originality was inspired to the Consensual Assessment Technique (CAT) proposed by Amabile (Amabile, 1982) and the reliability of the results was high according to Cronbach's α (>.85).

The Consensual Assessment Technique (CAT) is a method developed by Amabile (1982) to assess creativity widely used for artworks and related products. This method aims at the assessment of an idea with a single metric. The whole paper is dedicated to the structure of the assessment phase, its preparation, and contains some fundamental assumptions to use the methodology. She provided the following definition of creativity on which the whole approach relies:

"A product or response is creative to the extent that appropriate observers independently agree it is creative. Appropriate observers are those familiar with the domain in which the product was created or the response articulated. Thus, creativity can be regarded as the quality

of products or responses judged to be creative by appropriate observers, and it can also be regarded as the process by which something so judged is produced." (Amabile, 1982).
Besides the definition expressed above, there are two main assumptions regarding creativity that support the method: the first one states that it is possible to obtain reliable judgments of products creativity given an appropriate group of judges; the second states that there are different creativity degrees that could be measured through the CAT.
The method itself is mainly composed of 7 steps, described in the following section:

1. Step – definition and requirements of the tasks
   - Tasks should not depend on specialised skills that only some individuals have undoubtedly developed more than others. (e.g. drawing ability or verbal facility).
   - The tasks should be open-ended enough to permit considerable flexibility in responses (thus permitting to access divergent thinking).
2. Step – judges requirements
   - Judges should have some experience with the domain in question. On the other side, it is not necessary that they have the same level of experience because different backgrounds ensure better results in the assessment process.
   - Judges do not need to have produced themselves highly creative products in the domain field studied.
3. Step - Judges must take their assessment independently, without any training and guidelines about criteria. This choice was taken to enhance different perspectives of judges.
4. Step – During preliminary work, researchers should ask judges to assess other dimensions besides creativity (e.g. technical aspects, aesthetics) to examine the degree of related independence between the criteria assessed.
5. Step – Judges should be instructed to evaluate products comparing one to another on the dimensions specified rather than rating them against an absolute standard. It is essential due to the difficulty encounter in the definition of an undoubted standard regarding product and ideas.
6. Step – Each judge should evaluate products in a different random order to avoid any biases.
7. Step - After judgments are obtained, it is essential to assess inter-judge reliability to ensure the validity of the method constructed. The reliability metric proposed by Amabile is the Cronbach's alpha coefficient

The CAT (Amabile, 1982) was tested under different circumstances in 21 studies concerning art, poetry, and storytelling. It showed high reliability, with Cronbach's alpha coefficient ranging between .72 and .93. In 1996, Amabile (1982) published another research in which reached reliabilities between .70 to .89. Other studies have been reported in the following years, obtaining the same results. (Runco, Johnson, & Bear, 1993; Baer, 1997, 1998; Baer, Kaufman, & Gentile, 2004; Amabile, Conti, & Coon, 1996; Hennessey, 1994; Kaufman, Baer, & Cole, 2009; Runco, Johnson, & Bear, 1993).

The method proposed by Amabile was highly used for decades to assess creativity in several experiment contexts. It has been shown that the evaluation with this method is free from gender, race and ethnicity biases, and it is highly adopted in the field of higher education.

It does not require expert judges, but people that have familiarity with the field assessed.

It presents many advantages, but it is crucial to understand if the method fits the AUT experiment.

Firstly, the CAT method was born for the evaluation of creativity in artistic works such as poetry, thus distant from the assessment of rough ideas. The just mentioned issue represents a huge barrier for the use of CAT, being artistic works and products based on more data than the ones obtained from AUT responses, relying only on single sentences. Furthermore, in the artistic evaluation, there is a lack regarding the quality assessment of an idea; creativity does not take into consideration the feasibility, the applicability and other parameters that are important in the product development phase.

In the CAT, besides creativity, other criteria must be assessed that are redundant and not complementary with it. Besides, being product development a complicated process with a lot of different aspects involved, it is essential to decompose creativity in sub-criteria and to propose to judges clear definitions in order to avoid any biases. The CAT methodology is rarely adopted in the AUT experiments context.

2.2.3 Shah et al. (2000, 2003)

The researches conducted by Shah et al. (Shah, 2000, 2003) were dedicated entirely to the exploration of the idea generation assessment methods, in order to propose a standard methodology, explained through practical examples.

Their methodology was aimed at the assessment of outcomes generated by the ideation processes (and not on the process itself) because a direct analysis of the creative process was considered too complicated and lacking reliability Shah (Shah & Vargas-Hernandez, 2003). In addition to this, Shah et al. stated that "designer's success is judged by how well his/her design meets desired goals and how well he/she has identified the alternative ways of achieving those goals." (Shah & Vargas-Hernandez, 2003).

The approach proposed by Shah could be subdivided into two main phases, the first one regarding preliminary steps to the idea scoring and the second regarding the scoring process and the analysis of data itself.

The preliminary steps proposed by Shah et al. are:

- Decomposition of the design artefact into key functions assigning weights to each of them.
- Analysis of every idea produced defining which functions it fulfils and describe how it fulfils them, at a conceptual or embodiment level.
- Assignation of weights to embodiment and conceptual level.

After having done the preparatory steps, the actual scoring phase starts, and the criteria evaluated. Shah et al. proposed four metrics for the assessment phase, which are Novelty, Quality, Quantity, and Variety. In their opinion, researchers should not try to aggregate them in one single macro-criterion because each of them is created to be independent and represents different aspects of the idea. In the following pages, the methodologies developed by Shah to assess these metrics are displayed.

2.2.3.1 Novelty

Novelty is the most crucial metric regarding creativity assessment. Shah et al. define it through the standard definition "how unusual or unexpected an idea is as compared to other ideas" (Shah, 2003).

The first phase of the novelty assessment is done at the functions level. It can be assessed with two different approaches:

- Defining what function is novel and what is not before analysing any data to avoid biases (done during the preliminary phase).
- Collecting data from the participants and then identify the main attributes (functions) of their ideas. This approach is based on the comparison principle, which relies on the fact that a function belonging to a set, is more novel if the function is less present than others. For this reason, the methodology gives a higher score to the idea with less repetition in a set.

Each description created in the 2nd preliminary step should be evaluated for Novelty with one of the approaches described above. For the experiment designed by Colombo, the first approach is not suitable because it requires an excessive amount of data to be calculated.

At this point, it is possible to assess Novelty at the idea level, as an aggregate of the functions proposed in one single response as a weighted average of functions and stages (embodiment or conceptual). In the formula presented below, M_1 represents the novelty of the idea, f_j represents the weight of each function j, S_{1jk} represents the score of the function j for the k-level (calculated with one of the 2 methods proposed above), p_k represents the weight of at the embodiment and concept levels k.

$$M_1 = \Sigma_j f_j \Sigma_k S_{1jk} * p_k$$

2.2.3.2 Quality

Quality "is a measure of the feasibility of an idea and how close it comes to meet design specifications" (Shah et al., 2000).

The principle proposed by Shah et al. for the assessment of Quality, is similar to the one proposed for the Novelty evaluation, including the weighted average of functions and stages. M_2 represents the Quality of the idea, f_j represents the weight of each function j, S_{2jk} represents the score of the function j for the k-level (calculated with the method chose by the researcher), p_k represents the weight of at the embodiment and concept levels k.

$$M_2 = \Sigma_j f_j \Sigma_k S_{2jk} * p_k / n * \Sigma_j f_j$$

For the calculation of S_{2jk}, Shah et al. do not go into any further details, limiting to cites some methods that could be used such as QFD, Pugh Matrix, Decision Tables and all other popular methods permitting to assess the overall quality scores in general. In the formula presented, the denominator is just aimed at normalising M_2 on a scale 1-10.

The former two metrics described refers to the assessment of single ideas. The following table represents an example of the assessment method proposed by Shah et al. regarding the two just mentioned metrics.

Function		Concept ($p_1 = 0.7$)			Embodiment ($p_2 = 0.3$)		
Name	f_j	**Description**	**Quality** S_{1j1}	**Novelty** S_{2j1}	**Description**	**Quality** S_{1j2}	**Novelty** S_{2j2}
Reach/ Transport	0.5	2 tracks, forward and reverse	10.00	4.00	Caterpillar tracks	8.00	7.00
Abduct/ Deliver	0.5	Gear drives, right angle drives, pushes 1 egg	6.00	8.00	Curtain, up and down / swinging door	8.00	8.00
Total	1.0	Weighted Sum	5.60	4.20	Weighted Sum	2.40	2.25

Total Scores: Quality $M_1 = 8.00$, Novelty $M_2 = 6.45$

Figure 2.2 An example of the assessment of quality and novelty (Shah, Kulkarni, & Vargas-Hernandez, 2000)

2.2.3.3 Variety

Variety is "a measure that explored the solution space generated during the idea generation process" (Shah et al., 2000).

For the assessment of this metric, a genealogy tree is built for each idea proposed by participants, associating at each level of changes a value S_k.

- At the highest level, ideas are distinct from each other by the physical principle used to fulfil each function. $S_1 = 10$.
- At the second level, ideas are differentiated by the working principle, sharing the same physical principle. $S_2 = 6$.
- At the third level, ideas are differentiated by embodiment, sharing the same working principle. $S_3 = 3$.
- At the last level, ideas are differentiated only by the detail level, sharing the same embodiment. $S_4 = 1$.

Below it is represented an example of the genealogy tree proposed by Shah (2003), with respective values of S_k

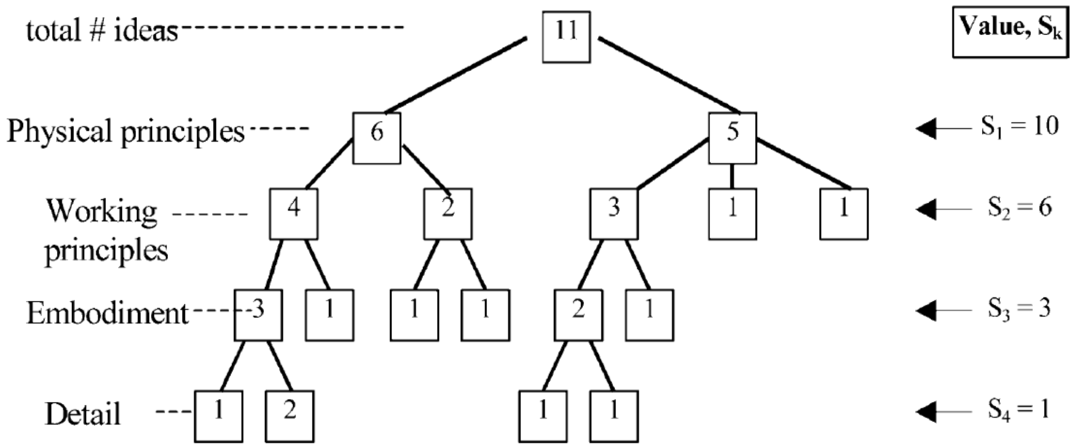

Figure 2.3 An example of a genealogical tree by (Shah & Vargas-Hernandez, 2003)

The overall variety is calculated through the formula:

$$M_3 = \Sigma_j f_j \Sigma_k S_k * b_k / n$$

Where M_3 represents the Variety of the idea, f_j represents the weight of each function j, S_k is associated with the level k (10,6,3,1), b_k represents the number of branches at level k and n is the total number of ideas in one single response.

2.2.3.4 Quantity

Quantity is "the total number of ideas generated in one response" (Shah et al., 2003).
The Quantity assessment is the most straightforward metric proposed by Shah, and it corresponds to the count of the ideas given by participants in one single response.
Quantity and Variety, as explained before, are metrics that assess values concerning a set of ideas, which are not relevant for the experiment conducted by Colombo. These metrics are usually needed to assess creativity at a personal level or in products in a more advanced stage of development. In the method proposed by Shah et al., the number of data available for the judgment of these outcomes is extremely high, not feasible with this research purpose.
In the following table, Johnson et al. (Johnon et al., 2015) tried to highlight the connection existing between psychological creativity criteria and the ones relative to ideation processes. Currently, other studies are trying to link these perspectives together, but there are not still precise correspondence in the results.

Spontaneous Flexibility Test (SFT) Metrics [1]	Engineering Ideation Test (EIT) Metrics [4]
Fluency	Quantity
Flexibility	Variety
Originality	Novelty
Elaboration	Quality

Figure 2.4 Correspondence of ideation metrics in Psychology and Design (Johnson, Caldwell, & Green, 2015)

2.2.3.5 Takeaways

The method proposed by Shah et al. has been developed to assess creativity for idea generation techniques providing a high degree of detail, such as finished products or ideas displayed in full papers with words and sketches (an example is visible in the picture at the end of the sub-chapter). In the experiment designed by Colombo, ideas are simple and usually expressed through a maximum of 15 words. In addition to this, the method of Shah et al. is proposed for structured problems, with constraints to respect and goals to achieve. The present research is aimed at the assessment of an unstructured problem, in which participants can decide which particular function the object can solve, without having specific constraints (apart from the indication given to them for a common or an uncommon function). This fact is typical of the earliest stages of product development processes for highly innovative products. For the experiment, the metrics proposed by Shah et al. need to be simplified to permit the proper evaluation of the responses.

Not all the criteria proposed by Shah are relevant for the present study: Variety and Quantity are metrics that are relative to sets of ideas. In the present research, it is fundamental to define a methodology to assess single ideas and to compare them with the neurological results of the ideation process. Variety could be used to assess the degree of openness provided by the idea generation method, but Quantity is meaningless for the experiment purposes.

The other metrics assessed by Shah et al. are Quality and Novelty. These criteria are essential to be assessed because they represent the evaluation of a single response from the design perspective. The Novelty metric can be frequently seen in the psychological assessment of creativity, as seen in the first part of the chapter, and Quality represents the actual utility of the idea in a new product.

The method proposed by Shah et al. to assess these latter metrics is too complex to be adopted for the experiment of Colombo. Firstly, the experiment is focused on the concept generation and does not take into consideration the embodiment and detail level. Secondly, participants were focused on finding a single function/solution (common or uncommon) per object.

During the literature review, several studies using the approach developed by Shah et al. have been analysed; the main problem incurred by researcher using these metrics, is that there could be misunderstandings between judges about the actual meaning of Quality, thus a low reliability in the results (see the study of Linsey et al., 2011). Therefore, raters need to follow precise instructions, supported by the definition of the criteria to be assessed. In the next subchapter, the method of Dean et al. (2006) will be explained. It relies on the same criteria proposed by Shah et al., but goes deeper into the definition of sub-dimensions that, theoretically, permit to increase the reliability of the results. In addition to it, the latter methodology gives clear guidelines to the raters, providing tables that make the evaluation process more standard and unambiguous.

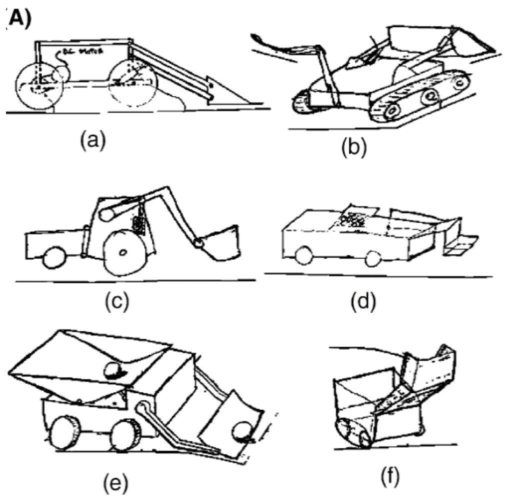

Figure 2.5 An example of sketches of a set of ideas from Shah experiment (Shah & Vargas-Hernandez, 2003)

2.2.4 Dean et al. (2006)

The research conducted by Dean et al. (2006) is a literature review concerning 90 articles published between 1990 and 2005, about empirical studies on creativity and their idea evaluation methods to link them together in a single assessment technique. Several methodologies have been analysed in this paper, and a lot of them presented inconsistencies among each other because of the lack of a standard in the assessment procedure. The first problem was associated with terminology. Many studies were using the same terms but relying on inconsistent definitions. In the same way, a few studies were using the same main concepts (definitions), but referring to them with different terminologies. This lack of consistency led to much confusion in researches, removing the possibility to have unique and readable results (Dean et al., 2006).

For this reason, Dean et al. decided to group all criteria proposed in the previous researches into four main metrics described by MacCrimmon and Wagner (MacCrimmon & Wagner, 1994):

- Novelty – an idea is most novel if nobody has expressed it before.
- Workability – an idea is workable if it does not violate known constraints or if it can be easily implemented.
- Relevance – an idea is relevant if it satisfies goals set by researchers.
- Thoroughness (Specificity) – an idea is thorough if it is worked out in detail.

Definitions proposed by (MacCrimmon & Wagner, 1994)

In addition to the problem highlighted above, there were trouble and inconsistencies even regarding judges, especially concerning their training and scoring phases.

Concerning criteria, Dean et al. proposed a fundamental distinction based on the focus of the assessment method. They could be divided in:

1) Novelty based methods – assessing only novelty metric.
2) Multiple based methods – considering other criteria beyond novelty (e.g. usefulness).

Study	Dimensions				
Part 1. Examples of Novelty-Centric Definitions of Creative Products					
Eisenberger and Selbst, 1994	Rarity				
Redmond et al., 1993	Originality				
Part 2. Examples of Multi-Attribute Definitions of Creative Products					
Woodman et al., 1993	Originality		Usefulness		
Plucker et al., 2004	Novelty		Usefulness		
Bessemer and Treffinger, 1981	Novelty		Resolution	Elaboration and Synthesis	
Wagner, 1996	Originality		Purpose	Implementation	
MacCrimmon and Wagner, 1994	Novelty	Non-obviousness	Relevance	Workability	Thoroughness
Dean et al., 2006	Novelty		Relevance	Workability	Specificity

Table 4 Dimensions identified in creativity literature (Dean, Hender, Rodgers, & Santanen, 2006)

The table above gives some examples of past studies that used Novelty based metrics (Eisenberger & Selbst, 1994; Redmond, Mumford & Teach, 1993) and Multi-Attribute metrics approaches. In the latter ones, criteria assessed were identified and classified following the distinction suggested by MacCrimmon and Wagner.

After various tests, Dean et al. stated that the dimensions proposed were way too general, obtaining low inter-judge reliabilities. They suggested that the reason why inconsistencies were usually high, was mainly due to the lack of proper guidelines for the judges. For this motive, they decided to decompose the dimensions proposed in further sub-dimensions.

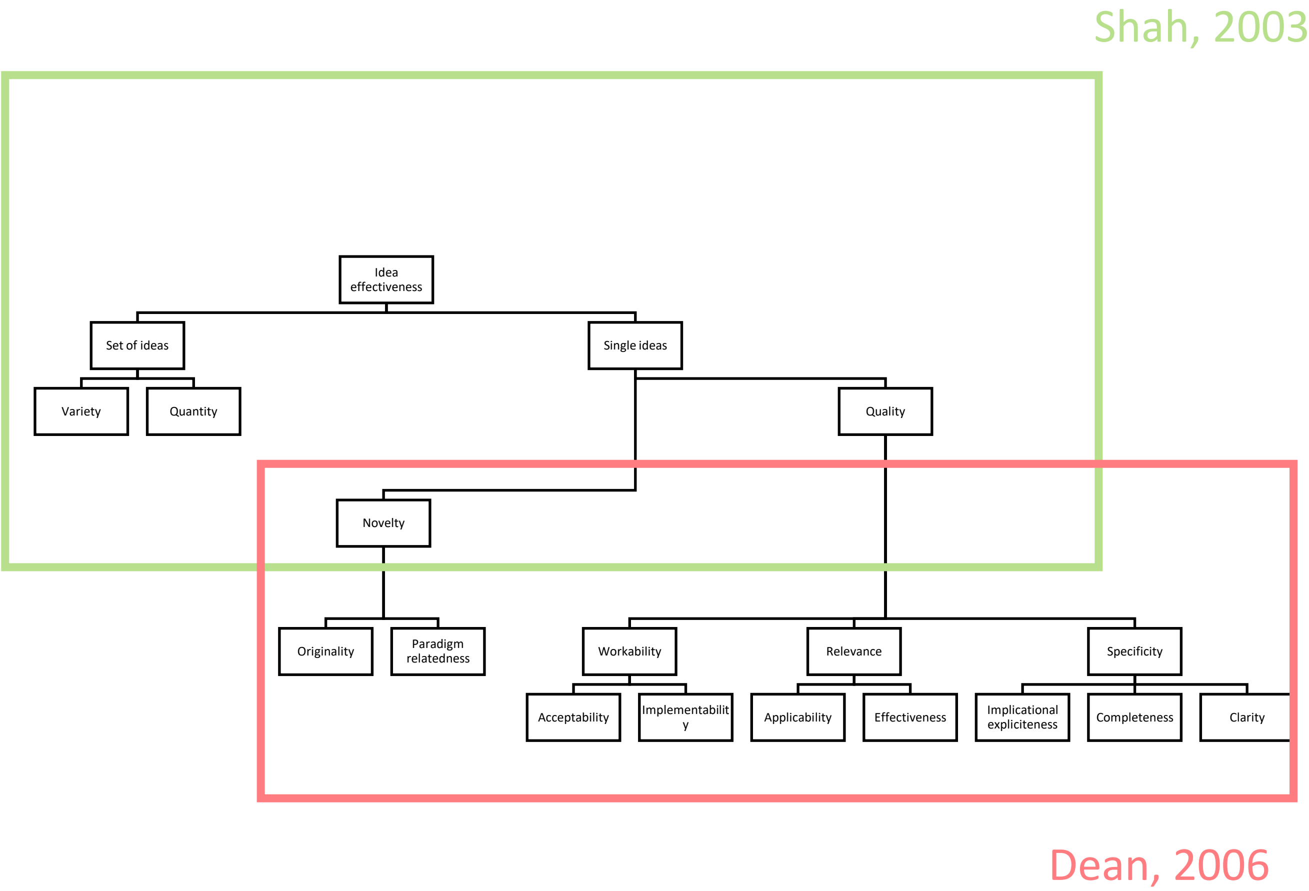

Figure 2.6 Integration of Shah et al. (2003) and Dean et al. (2006) studies (Verhaegen, Verhaegen, Peeters, & Duflou, 2013)

The method proposed by Dean et al. could be considered as an integration of the one proposed by Shah et al. (2000, 2003). The picture above is taken from a research of Verhaegen et al. (Verhaegen et al., 2013), in which they tried to develop a reliable concept of variety suggesting alternative methodologies of assessment. If Shah et al. were trying to evaluate complex concepts (complex ideas or products), Dean et al. focused on the evaluation of more straightforward ideas. For this reason, they overlooked in their research the part of evaluation concerning set of ideas, going into depth in the assessment of single ideas, decomposing the metrics in order to increase the effectiveness of the assessment methodology.

In order to develop a set of metrics that was suitable for the assessment of ideation methods, Dean et al. decomposed the concept of idea creativity in two independent characteristics, idea Novelty and Quality, resuming the concept developed by Shah et al..

The former, Novelty, was divided in Originality and Paradigm Relatedness (Besemer & Treffinger, 1981; Jackson & Messick, 1965; Besemer & O'Quin, 1987). Originality is a metric that includes "Rarity" (Novelty in psychological terms) and originality itself. This expedient was taken to overcome the limitations of psychological studies were ideas were considered creative even if only bizarre. Paradigm relatedness is a criterion apt at the measure of shifting paradigm of the idea proposed by the subject.

Quality was decomposed in further three dimensions: Workability, Relevance, and Specificity. Workability is a parameter that could be defined as the feasibility of an idea at an economical, technical and political level. An idea is workable "if it does not violate known constraints and if it can be easily implemented" (Dean et al., 2006). Relevance measures how much an idea applied to the stated problem and at which degree could the idea solve it. Specificity refers to the degree to which an idea is elaborated, detailed and expressed clearly. In the following table, Dean presented macro and sub-dimensions of his methodology and the related definitions.

#	Dimension	Definition
1	***Novelty****	The degree to which an idea is original and modifies a paradigm.
1.1	Originality	The degree to which the idea is not only rare but is also ingenious, imaginative, or surprising
1.2	Paradigm relatedness	The degree to which an idea is paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational.
2	***Workability (Feasibility)***	An idea is workable (feasible) if it can be easily implemented and does not violate known constraints.
2.1	Acceptability	The degree to which the idea is socially, legally, or politically acceptable.
2.2	Implementability	The degree to which the idea can be easily implemented.
3	***Relevance****	The idea applies to the stated problem and will be effective at solving the problem.
3.1	Applicability	The degree to which the idea clearly applies to the stated problem.
3.2	Effectiveness	The degree to which the idea will solve the problem.
4	***Specificity***	An idea is specific if it is clear (worked out in detail).
4.1	Implicational explicitness	The degree to which there is a clear relationship between the recommended action and the expected outcome.
4.2	Completeness	The number of independent subcomponents into which the idea can be decomposed, and the breadth of coverage with regard to who, what, where, when, why, and how.
4.3	Clarity	The degree to which the idea is clearly communicated with regard to grammar and word usage.

Table 5 Definitions of the Quality dimensions and sub-dimensions (Dean, Hender, Rodgers, & Santanen, 2006)

Precise guidelines were given to judges to rate the dimensions accurately; for each dimension, the definition of the criteria and tables containing marks and instructions (descriptive anchors) were given to judges. All sub-dimensions were ranked on a 1-4 Likert scale, except for the sub-dimensions regarding Specificity that were ranked on a 1-3 scale. Macro-dimensions were then obtained as the sum of the respective scores earned by sub-dimensions, as shown in the table below.

Table 11. Derivations of General Constructs[4]		
Construct	**Range**	**Formula**
Novelty	2-8	Originality + Paradigm relatedness
Workability	2-8	Acceptability + Implementability
Relevance	2-8	Applicability + Effectiveness
Specificity*	2-6	Completeness + Implicational explicitness

Table 6 Derivations of General Constructs (Dean, Hender, Rodgers, & Santanen, 2006)

Dean tested the approach using two unstructured problems. They were both imprinted to the business perspective:

1) The first one was related to an issue met by a restaurant. The problem was stating that the restaurant was targeting students and that in the last period it was losing customers. What could have been a possible solution adopted by the restaurant to retain customers? Ideas proposed were a total of 2,105. After a cleaning phase, the total number of ideas was reduced to 1279.
2) The second problem was proposed at a governmental level. It asked how the city of Tucson could have attracted more tourists. The total number of ideas proposed was 1,019 ideas, and after the screening phase, it was reduced to 692.

Two judges have been used for the idea assessment. The experiment showed good results and the inter-judge reliabilities obtained are reported below.

Inter-judge reliabilities			
Criteria	Sub-dimensions	Experiment 1	Experiment 2
Novelty	Originality	0.766	0.713
	Paradigm relatedness	0.843	0.687
Workability	Acceptability	0.663	0.685
	Implementability	0.713	0.714
Relevance	Applicability	0.658	0.664
	Effectiveness	0.729	0.667
Specificity	Completeness	0.698	0.708
	Implicational explicitness	0.783	0.690
	Clarity	0.618	0.880

Table 7 Inter-judge reliabilities calculated in the experiment of Dean et al. (2006)

The method proposed by Dean is the most suitable one found in the literature with the experiment analysed in the present thesis. For this reason, the takeaways of this method will be directly displayed in chapter 3.2, in which the methodology chosen for the evaluation of verbal responses in the experiment proposed by Colombo is displayed.

2.2.5 Other relevant assessment methods

In this section, other relevant methods proposed for the assessment of products are displayed for completeness. They present the same issues observed in the method proposed by Shah et al. (2000, 2003), in fact, they are focused on the assessment of specific outcome, way too detailed in comparison to ideas generated in the experiment of Colombo (2019).
The first approach analysed is the one proposed by Besemer and O'Quin (Besemer & O'Quin, 1986, 1999). This experiment presented a different perspective on the artefacts assessment; in fact, subjects analysed (N = 185) took judge roles and were asked to evaluate creative objects on a 1-7 Likert scale. The products assessed were three chairs built using unusual materials (food boxes, car parts, and triangular shapes) and one ordinary chair taken as reference as a non-creative product. The scoring method relied on three main criteria, Novelty, Resolution

and Elaboration & Synthesis. These metrics were subdivided into further subdimensions to permit higher reliability. The subjects took the assessment in small-group sessions, in which stimuli were presented in different random order, and the total time required for a group evaluation was about 1 hour. The results showed a normal distribution with small standard deviation, and the inter-groups reliabilities were all above 0.70. As anticipated above, the experiment could not be taken as a reference because was highly time and resource consuming. Regarding the procedure setting, the subdivision of criteria to increase reliabilities and the presentation of stimuli in random order to avoid biases, validate the assessment method proposed in this research.

The second approach proposed is of Horn (Horn & Salvendy, 2006). This method was divided into two main phases; the first one was an initial screening to understand which product in a set could be considered as creative and which not, and the second was a more in-depth analysis concerning those products. The preliminary phase was focused on the assessment of creativity in 4 consumer products for an initial screening through a single metric, on a Likert scale 1-7. The objects analysed were: 3 lamps, 3 toothbrushes, 3 chairs, and 3 toasters. Between those 12 products, 2 lamps and 2 chairs were selected for further studies. For each type of product, one had a high creativity average score (over 4.0), and the other scored a low value (below 4.0). The survey was Web-based, providing a two-dimensional image for each product. The participants selected for the assessment were 24 general consumers. After this preliminary study, to identify the best objects to be studied among the 12, two different assessment methods were adopted: the first one involving 208 participants through e-mail, of which more than 50% were undergraduates, and the second involving a sample of 105 undergraduate and graduate students (Engineering and Mathematics). The surveys were taken independently by subjects, and three macro-criteria were adopted: Overall creativity, Purchase intentions and Satisfaction. As in the experiment analysed before, the macro criteria have been decomposed in sub-dimensions to increase the reliability of the results. The 1st method obtained an overall Cronbach's Alpha of 0.84 and overall inter-judge reliability of 0.81 and the 2nd an overall Cronbach's Alpha of 0.90 and overall inter-judge reliability of 0.88.

The third method analysed is the one developed by Christiaans (2002). The experiment was designed in 2 primary phases. The first aimed at the assessment of design works, precisely of 19 computer cabinets and 25 telephone boots created during a design course. Each object was then proposed to judges through slides. 34 subjects were selected as product judges, among which: 10 designs teachers, 12 senior students and 12, mathematics students. All raters were asked to score all 44 design products on seven main criteria through a 10-point Likert scale. The metrics measured were: Creativity, Technical quality, Attractiveness, Interest, Expressiveness, Integrating capacity and Goodness. The reliabilities were calculated both within and inter groups, obtaining satisfying results (apart from Expressiveness and Integrated capacity). The second phase of the experiment consisted of the replication of the experiment considering different objects and removing Expressiveness and Integrated capacity. Without going in any further detail, even this replication obtained excellent results showing high inter-judge reliabilities. These studies have shown that there are no significant differences in the assessments took by subjects with different levels of experience concerning design.

As anticipated before, these methods are too detailed to be used for an idea generation at the earliest stages of product development, such as the one that has the AUT experiment analysed in the present thesis. They are mainly dedicated to finished products, in which the idea is well structured and presented to judges.

The last method proposed in this section is different from the formers: it is an approach that permits faster ideas evaluation, being low time & resource consuming. NUF (Novel, Useful, Feasible) has been developed by Gray et al. (2010), but it does not have enough references in literature to be considered a reliable one. In this approach, ideas are evaluated on a 1-10 Likert scale, on the three dimensions. Kudrowitz (Kudrowitz & Wallace, 2013) proposed a modified version of the method where he added two further criteria, Clearness (referred to the degree of right communication of the idea) and Product-worthiness (related to the marketable part of the idea).

Method		Creativity field	Criteria	Complexity	Scale	Note
CAT (1982)		Art	Creativity	Low	1-6	This approach is characterized by a lack of guidelines provided to judges in order to highlight different perspectives
Shah (2000,2003)		Design (advance phases)	Quality	High	1-10	Approach dedicated to the assessment of detailed ideas or finished product. It permits a detailed analysis, but it is characterized by an high degree of complexity due to data settings and analysis phases
			Quantity	Low	/	
			Variety	High	1-10	
			Novelty	High	1-10	
Dean (2006)		Design (earliest phases	Novelty	Mid	1-8 (1-4 + 1-4)	Approach dedicated to the analysis of the earliest phase of the product development. It require attention in the preparation of the guidelines but make it easier the assessment for judges
			Feasibility	Mid	1-8 (1-4 + 1-4)	
			Relevance	Mid	1-8 (1-4 + 1-4)	
			Specificity	Mid	1-9 (1-3 + 1-3 + 1-3)	
Other assessment methods	Besemer & O'Quin (1986)	Final-products	Novelty	Mid	1-7	Method involving the division of criteria in further sub-dimensions for increasing reliability
			Resolution	Mid	1-7	
			Elaboration & Synthesis	Mid	1-7	
	Horn & Salvendy (2009)		Creativity	Low	1-7	A study based on a survey presented through online forms
	Christiaans (2002)		Creativity	Low	1-10	Method exploiting the division of creativity in further sub-dimensions
			Technical quality	Low	1-10	
			Attractiveness	Low	1-10	
			Interest	Low	1-10	
			Expressiveness	Low	1-10	
			Integrating capacity	Low	1-10	
			Goodness	Low	1-10	
	NUF		Novelty	Mid	1-10	Practical method in order to analyse fast creativity of responses
			Utility	Mid	1-10	
			Feasibility	Mid	1-10	

Table 8 Comparing criteria in psychological methodologies

3 The experiment (Colombo, 2019)

Within this chapter, the main phases regarding the design of the experiment of Colombo (2019) and the statistical analysis regarding data obtained are displayed.

A subdivision of the experiment phases is proposed in order to highlight the phases computed in the present research and the one conducted before the beginning of it.

The experiment derives from a classic of the ideation literature, the Alternative Uses Task (AUT), in which researchers ask to subject to find a common or uncommon use to everyday objects (further detail in chapter 3.1). Furthermore, data obtained were integrated with the recorded EEG signals.

The aim of this chapter is the representation of all the processes encountered in order to obtained results and to allow the introduction of further factors in the analysis (Gender, Background, Degree level). Furthermore, some basic neuroscientific concepts are provided to the reader in order to permit a better comprehension of the cognitive processes.

It is essential to disclose that the following chapter contains several references to the master thesis developed by Colombo (2019). This is since the present research is the continuation of his work, relying on the experiment designed by him.

3.1 The experiment phases (Colombo, 2019)

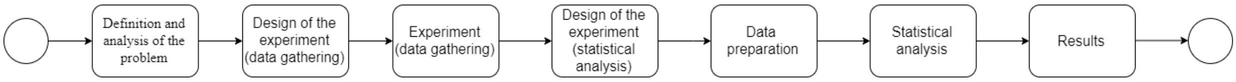

Figure 3.1 Experiment phases

The experiment proposed by Colombo (2019) is composed of different phases, as displayed in the flow chart represented above. This flow chart is a simplification of the actual processes that occurred in the project, in order to give the reader a simpler and more understandable view of the processes. The aim of the following subchapter is to distinguish and describe the processes executed before the start of the present thesis from the one encountered during the development of the present research.

Colombo focused his research on the first three phases of the experiment, defining and analysing the problem, designing the data gathering phase of the experiment and executing it. In addition to these steps, other phases have been outsourced from the present research for the lack of competences required.

The present thesis has been focused on the design of the experiment concerning the analysis of data (post-experiment design), the data preparation, the statistical analysis and the interpretation of the results. The present paper has not as objective to obtain final results and to fully describe the population due to the massive quantity of data attained from the data-gathering phase and to the various analysis that could be executed on the population. However, the partial results are needed in order to provide a first feedback on the experiment and for directing further studies.

3.1.1 The DOE pre-data gathering

As anticipated in the first chapter of the present research, the branch studying neuroimaging and neuroelectric techniques coupled with ideational methods is still at the beginning, and several conflicting results have been obtained (Dietrich & Kanso, 2010). Furthermore, there is still a lack of standards concerning the experimental methodology (Colombo, 2019; Dietrich

& Kanso, 2010) and results consistency (Sawyer, 2011). For these reasons the experiment developed by Colombo (2019) has the objective to extend further data in the literature, allowing the replication of the procedure in further studies.

Among neuroimaging and neuroelectric techniques, the electroencephalogram (EEG) was adopted in the experiment of Colombo (2019) for several reasons:

- *Usability*, being EEG less invasive than the most of neuroimaging techniques.
- *Feasibility* for spatial resolution problems, such as the researched domain.
- *Popularity* in literature researches. It was used in the experiment proposed by Jauk et al. (2012), taken as the primary reference for the study of Colombo.
- *Availability* of the instrumentation for the research team.
- *Knowledge* of the research group taken part in the project that was more confident with the EEG instrumentation.

Electroencephalography is a non-invasive imaging technique mainly dedicated to medical purposes in order to read electrical activity generated by brain structures. The electroencephalogram (EEG) is defined as an alternating type of electrical activity recorded from the scalp surface and measured using metal electrodes and conductive media (Teplan, 2002).

The neuro signals assessment was done adopting the device BrainVision ActiCHamp (developed by BrainProducts GmbH, Germany) including 32 electrodes gathered through a splitter box connected to the ActiCap. Thirty-one electrodes were positioned on the scalp according to the International 10-20 system, and one was connected to the tip of the nose of participants, corresponding to the reference electrode. The sampling process was executed at a 500Hz frequency (Colombo, 2019).

Beyond EEG, other tools have been adopted:

- A Tobii X2-30 Eye Tracker Compact Edition (Tobii) for ocular data acquisition, in order to allow further analysis (not object of the present research).
- Two Logitech c920 cameras by Logitech, in order to record subjects' face and expressions, and eventual excessive body movement, that could have possibly influenced sensors' records during the data gathering process.

The software adopted for data recording was iMotions 7.2, allowing the synchronisation of the EEG, Eye-tracker and cameras signals and the management of the stimuli presented to participants.

Forty participants, 11 females and 29 males, were involved in the experiment in order to have a reliable population. All of them were university students from different backgrounds. Further data concerning participants are presented in the Appendix (A.1).

The task proposed to participants was a variation of the Alternative Uses Task experiment (AUT) proposed by Guilford (1967), aimed at the assessment divergent thinking. This decision was taken since divergent thinking is a valid and reliable indicator of creative production, allowing the researcher to study the ideation process in a controlled environment (Colombo, 2019).

Forty stimuli, common object words, were presented through slides to each subject.

For each participant, the test was divided into two main blocks, one for the "common" condition and one for "uncommon" condition. Each block was composed of 20 stimuli, for a total of 40 trials (Colombo, 2019). The blocks and the stimuli were provided to participants in a different random order enabling the avoidance of any biases. It was then required to each participant to display a single function for object according to the block condition, respectively an ordinary solution in the common block or an alternative and useful one in the uncommon, within 30 seconds.

The main differences with the experiment proposed by Jauk et al. (2012) were: the number of items presented to participants (40 vs 20), the expedient of the subdivision in two distinct blocks according to the condition of the experiment (not present in the experiment by Jauk et al.) and the participants personal data (university students vs male teenagers).

The blocks were introduced with the following procedure:

- The first slide was presenting the block that the participants were going to address ("Block 1", "Block 2"), for 5 seconds.
- The second was presenting the instructions to the subject in order to start the experiment. When the participant was ready to begin, he had to press the space bar.
- The third slide was presenting the "common" or "uncommon" instruction.

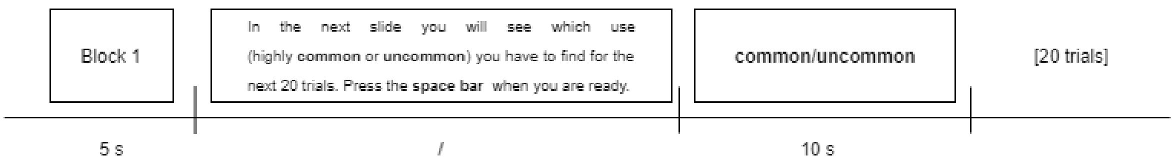

Figure 3.2 Block presentation in Colombo (2019) experiment

At the end of the first block, a two-minute pause was taken by participants. Within each block, the 20 stimuli were presented with the following procedure:

- A blank screen was showed to participants for 5 seconds.
- A static cross was presented in the middle of the screen for 5 seconds in order to have a reference period for the EEG data.

- The stimulus was presented through a word in the middle of the screen appearing for 500 ms in order to avoid biases with eye movements.
- The participant had a maximum of 30 seconds (or the press of the space bar) to think about the solution while watching the cross in the middle of the slide presented.
- The subject verbalised the idea.

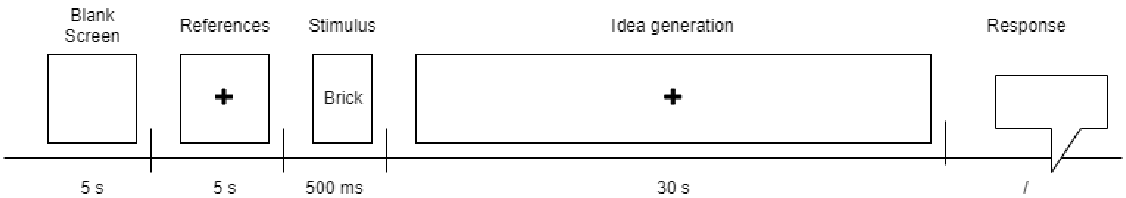

Figure 3.3 Stimulus presentation in Colombo (2019) experiment

At the end of the experiment, a quick questionnaire was administrated to participants in order to collect more data. In the Appendix (A.2), it is possible to see the whole list of the stimuli presented to participants. In the present research, common responses were taken as reference for the convergent thinking, while uncommon ones for divergent thinking.

After the data gathering process, verbal responses have been reported from audio to text on an Excel file, and neuro signals have been cleaned through the help of the videos recorded by cameras. The total number of verbal responses obtained at the end of this process was 1508.

3.1.2 The focus of the present research

The present research has been focused on the analysis of data gathered during the experiment. Firstly, in order to allow the comparison between outcomes (verbal responses) and cognitive data, it has been necessary to convert qualitative data in quantitative data.

Concerning verbal responses, the design of the assessment methodology was done based on the literature review (chapter 2). This phase was critical because the aim of the present experiment is different from the standard AUT test. The former is focused on the assessment of the ideation process and its outcomes, and not on creative personalities as the latter. Furthermore, ideas should be evaluated under an ideational perspective and not a purely psychological one. Therefore, the assessment methodology has been defined carefully, considering all the differences existing in comparison to the original experiment.

Concerning cognitive data, after the first cleaning phase through recorded videos, all the responses have been evaluated using Letswave (an Open-source Matlab EEG signal processing toolbox), allowing the execution of an independent component analysis (ICA) in order to remove all the components of the signals that were registered by the EEG instrumentation, but did not fit the brain signal model, representing noise.

Then, it has been conducted a frequency analysis to observe frequency changes to the external stimuli in the subjects, described using alpha TRP (Task Related Power) values as an indicator of event-related desynchronisation (ERD) and event-related synchronisation (ERS).

The total number of responses at a neuro signals level after this phase was 1157, 464 regarding the common condition and 693 concerning the uncommon.

Another cleaning phase has been conducted in order to eliminate possible outliers from the database, excluding values exceeding 2,5 sigma distance from the mean of all alpha TRP values.

After having gathered all data in a database with the proper settings, several attempts have been conducted in order to understand which could have been the appropriate configuration of the analysis, comparing results for different groups of electrodes (areas) and levels (subject means vs the whole population).

It is crucial to explicit that for the data analysis, it has been assumed that the condition suggested to the participants during the experiment, reflected the creativity and usefulness of the idea proposed. This assumption was necessary to have a first approximate view of the results of the experiment, being the judges' assessment procedure highly time-consuming, requiring to each judge approximatively a total of 6000 scorings.

Finally, a repeated-measures ANOVA (analysis of variance) has been executed on the configuration selected, considering as within factors Area, Hemisphere and Condition, and as between factors Background, Gender, and Degree-level of the participants to the experiment.

3.2 Design of the assessment methodology for verbal responses in design creativity

As seen in the second chapter, it has not been identified an assessment methodology that could have fit completely the experiment proposed by Colombo (2019). The AUT classical experiment and its variants were born in the psychological field, for the assessment of divergent thinking on the personality trait perspective. Therefore, the aim was to assess sets of ideas and not single ones. Furthermore, these methodologies did not take into consideration the Quality of an idea, fundamental characteristic in the design perspective.

More recent studies have tried to link neurophysiologic analysis with AUT tests, but they were either relying on simplified measures (Fink & Neubauer, 2008; Fink, Graif, & Neubauer, 2009; Schwab et al. 2014; Fink et al. 2007; Camarda et al. 2018) or presenting assessment inconsistencies in the perspective of the ideation processes (Jauk, 2012). Therefore, the design of the assessment methodology has been mainly imprinted on the analysis of ideation methods and their evaluation methodologies (Chapter 2.2). Among them, no approach that thoroughly covers the experiment requirements has been identified; some approaches (Shah et al. 2000, 2003; Besemer & O'Quin, 1986, 1999; Horn & Salvendly, 2009; Christiaans, 2002) were based on the evaluation of final products or extremely detailed ideas. Others were used for the assessment of creativity in artistic works, therefore neglecting the design perspective. Several types of research have been conducted on the ideation phase and, in the same way, several inconsistencies have been encountered.

Among the papers analysed, the approach proposed by Dean et al. (Dean et al., 2006) has been identified as the most suitable one for the experiment developed by Colombo (2019):

- Considering two main aspects in the assessment of Idea Creativity, Quality and Novelty.
- Developed in an ideation process perspective.
- Usable in the first phases of product development.
- Being replicable, following proper guidelines.

For this reason, it has been taken as a primary reference for the design of the assessment methodology.

In the next pages, the following topics are displayed: the parameters chosen through the previous literature review, the judges' selection, the construction of the guidelines and the procedure to be followed during the assessment phase with further suggestions to increase the reliability of the method. The judges' selection and the design of the guidelines are integrated with further articles that have been summed up in tables presented in the Appendix (A.3).

3.2.1 Parameters

From the literature review, it has been seen that it is essential to decompose Idea Creativity in sub-dimensions in order to increase the reliability of the methodology, giving to all judges the same shared background and necessary information needed for the evaluation (Horn & Salvendy, 2009; Runco & Jaeger, 2012; Lubart, 2001; Dumas & Dunbar, 2014; Christiaans, 2002). This expedient is visible also in psychological studies, in which researchers tend to decompose creativity in further dimensions (Guilford J. P., 1967; Torrance, 1974; Wallach & Kogan, 1965; Wilson et al., 1954; Hocevar & Michael, 1979). Two main macro aspects have been taken in consideration in order to develop the subdimensions: Novelty, considering the creativity itself of the idea proposed, and Quality, regarding the Implementability, Utility, and Elaboration of the concept (Shah & Vargas-Hernandez; Dean et al., 2006; Chulvi et al.).

3.2.1.1 Novelty

Concerning Novelty, Shah et al. (Shah & Vargas-Hernandez, 2003) proposed an analysis of the parameter at a high level through an approach excessively complicated for the present research. For this reason, Dean et al. (Dean, Hender, Rodgers, & Santanen, 2006) proposed a further subdivision of Novelty into sub-dimensions that have been adopted for the scoring phase of the experiment proposed by Colombo (Colombo, 2019): Originality and Paradigm Relatedness. Nagasundaram & Bostrom (Nagasundaram & Bostrom, 1994) sustained that those are orthogonal measures needed to be considered for the evaluation of creativity. Both criteria will be scored during the judges' assessment phase.

Originality

"The degree to which the idea is not only rare, but is also ingenious, imaginative or surprising" (Dean, Hender, Rodgers, & Santanen, 2006).

This definition could be considered as an aggregation of two historical metrics:

- “Uncommonness” criterion - proposed in psychological studies, it is usually calculated through partially-objective methods. It relies on the number of time that the same idea is proposed by different participants or on pre-set tables, explicating the possible functions (Guilford J. P., 1967; Torrance, 1974; Wallach & Kogan, 1965).
- “Originality” itself – used both in design and psychological studies. It relies on the subjective assessment of originality, highlighting how an idea is perceived by judges.

The definition proposed by Dean is the suitable compromise between the approaches just mentioned because it permits to overcome the flaws of the partially-objective methods, such as the consideration of merely bizarre ideas as creative, and allows to link these 2 criteria in a single score. This approach is supported by the studies conducted by Silvia (Silvia & Wilse, 2008; Silvia et al., 2009; Nusbaum & Silvia 2011) in which he developed and tested a single metric providing to judges the definition of Uncommonness, Cleverness, and Remoteness for the assessment of creativity, obtaining high reliabilities. Furthermore, other studies demonstrated how the correlation existing between originality and uncommonness is high (r=0.91), showing that they could be considered as a single aggregate score (Snelders & Hekkert, 1999).

This metric will be assessed by judges with a 1-4 Likert scale using a table displaying instruction taken directly from the research conducted by Dean (anchor table presented in the appendix A.4).

Paradigm Relatedness

“The degree to which an idea is paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational” (Dean et al., 2006).

This metric is the result of the decomposition of novelty into sub-dimension reflecting the transformational power of the idea. In the literature, there are other examples of the adoption of this metric (Besemer & Treffinger, 1981; Jackson & Messick, 1965; Besemer & O'Quin, 1987; Gryskiewicz, 1980).

The method proposed by Dean to assess Paradigm Relatedness was the only one in his methodology designed ad hoc for the problem; therefore, it has not been possible to implement it directly on the experiment. Create an assessment table for each stimulus would have been excessively time consuming, and it would have lack reliability not having the proper knowledge required to develop it. For this reason, a general anchor table based on the model proposed by Nagasundaram et al. (1994), on which the anchor tables of Dean (2006) relied, will be provided to judges. Besides, the assessment will rely on the definition proposed by Dean and on the practical concept of paradigm from an architectural perspective.

Concerning architecture, the following actions could be taken in order to fulfil new functions, as proposed in the model by Nagasundaram et al. (1994):

- Extension – with the introduction of new components to the object.
- Redesign – with the alteration of the relationships between the elements composing the object.
- Transformational idea – in which both Extension and Redesign could be applied.

Nagasundaram proposed a hierarchical order for the just mentioned changes based on the degree to which the paradigm is shifted from the original object. The assortment starts with Extension changes, considered as the most basic possible changes, followed by Redesign and, lastly, by Transformational ones. This model is highly feasible with the view proposed by Henderson and Clarks (1990). In this view, changes could involve the reference technology or the relationship existing between components (architectural changes). In this optic, it is essential to notice that Henderson and Clarks (1990) refer to architecture as "defined by the main components that make up the product and their mutual relationships". In this perspective, the addition or the removal of new elements that do not change existing relationships between product components are not considered as architectural changes. Besides, for this experiment, changes to the underlying technology should be interpreted as the modification of the object with the addition or removal of new components without taking into consideration the relationship already existing between elements. With the intersection of these two axes (the degree of architectural changes and the degree of reference technology changes), four main kinds of innovations could be identified:

- Incremental innovation, in which "neither the underlying technology nor product architecture will change" (Cantamessa & Montagna, 2016). This kind of product changes has been taken as the "1" reference for the anchor table. It has been associated with the scenario in which no actual changes are done to the original object.
- Modular innovation, in which the "underlying technology does change in one or more functional elements, but product architecture does not" (Cantamessa & Montagna, 2016). These innovations could be associated with the Extension (or Removal) principle proposed by Nagasundaram (1994) and have been associated with the "2" of the anchor table.
- Architectural innovation, in which the "underlying technology does not change, but inter-component relationships do" (Cantamessa & Montagna, 2016). These changes could be viewed as Redesign ones and have been associated with the "3" of the anchor table.

- Radical innovation, in which "both the underlying technology and product architecture change" (Cantamessa & Montagna, 2016). Transformational changes have been associated with this kind of change, and the "4" value will be associated with them.

The former two could be considered as Paradigm Preserving (PP) innovations associated usually with modular architectures, in which changes could be implemented quickly, while the latter as Paradigm Modifying (PM), in which changes require more effort to be implemented.
To have more practical guidelines and follow the structure proposed by Dean, the terminologies proposed by Nagasundaram (Nagasundaram & Bostrom, 1994) have been adopted for the construction of the anchor table.
The assessment technique proposed, permit to use a 1-4 Likert scale, allowing judges not to change the scale used for other criteria (anchor table presented in the appendix A.4).

3.2.1.2 Quality

The Quality aspect of Idea Creativity, as the Novelty criterion, has been decomposed in further sub-dimensions as supported by the experiment of Linsey (2011), in which he obtained low inter-judge reliability assessing the macro criterion directly at the highest level. Concerning this metric, the lack of a standard in ideation processes is a huge limitation. Shah et al. (2003) suggested using standard design tools, such as QFD, Pugh Matrix or decision tables. For the experiment, anchor tables will be adopted to support judges' decisions. The assessed criteria have been developed based mainly on the ones proposed by Dean (2006). The subdimensions assessed are Effectiveness, Feasibility, and Elaboration.

Effectiveness

"The degree to which the idea will solve the problem" (Dean et al., 2006).
This metric has been highly adopted in past researches (Eisenberg & Rhoades, 2001; Kramer & Kuo, 1997; Barki & Pinsonneault, 2001; Fauer, 2004; Valachich & Shearer, 1995), with many inconsistencies regarding the terminology used to name it.
In the methodology developed by Dean et al., this metric was considered as a sub-dimension of Relevance, a dimension assessing how much an idea was pertinent to the stated problem and how much effective the idea could have been to solve it. The stated problem in the AUT design is simple and basic; therefore, there is no need for judges to evaluate the pertinence

(applicability) of the response by participants. For this reason, Effectiveness will be the only part of this criterion that will be assessed. This metric has the purpose to evaluate the extent to which the function proposed by the participant is solving the problem identified by himself.
This criterion will be assessed by judges using a 1-4 Likert scale using a table displaying instruction taken directly from the research conducted by Dean (anchor table presented in the appendix A.4).

Feasibility

"An idea is feasible if it can be easily implemented and does not violate known constraints" (Dean et al., 2006).
In the literature, there are several studies focused on the assessment of this criterion (Briggs et al., 1997; Diehl et al., 1987; Gallupe et al., 1992; Potter & Balthazard, 2004).
Dean (2006) considered a further subdivision concerning this metric, dividing it into Acceptability and Implementability, that, for the present experiment, are excessively detailed. For this reason, both the criteria could be assessed in a single metric at a higher level. Regarding this sub-dimension, it has been necessary to develop a new table constructed merging the one concerning the acceptability and the one relative to the Implementability proposed by Dean.
This metric will be assessed by judges with a 1-4 Likert scale using a table displaying instruction derived from the research conducted by Dean (anchor table presented in the appendix A.4).

Elaboration

"Elaboration is a facility for adding a variety of details to information that has already been produced" (Guilford, 1966).
This metric has been highly used in psychological researches and, in these studies, has been considered as the main criterion regarding the convergent thinking assessment. It has been adopted as a substitute to Specificity proposed by Dean, being the only sub-dimension in his methodology that presented some issues in the assessment phase. Furthermore, it is too detailed for this experiment.
Several methodologies have been developed for the Elaboration assessment, both subjective and objectives. From a historical perspective, in the AUT experiments, the principal methodologies adopted are objective and consists in the count of the number of words utilized in a single response (Forthmann et al., 2018; Harbison & Haarmann, 2014; Forster & Dunbar, 2009; Kudrowitz & Wallace, 2013; Dippo & Kudrowitz, 2015). As said, other methodologies have been developed, such as, in the last decades, the LSA (Latent Semantic Analysis)

(Forthmann et al., 2018), extremely time consuming for this experiment. For this reason, the historical AUT method, word count, will be used for the analysis of the experiment.

Variety

"Variety is a measure of the explored solution space during the idea generation process" (Shah et al., 2003).

This metric is not fundamental for the present experiment; in fact, it usually refers to a set of ideas (Kudrowitz & Wallace, 2012). It is present both in ideation and psychological researches and, in the latter, is known as Flexibility (Guilford, 1967).

It will be used for experimental purposes to assess which stimuli give more significant possibilities to participants to generate a broader solution space.

The solution space has been considered with an architectural perspective and, in order to calculate it, the average of the Paradigm Relatedness results will be adopted for every single stimulus. It will be thus possible to compare the results obtained by every single stimulus.

Methodology	Criteria	Sub-dimension	Definition
Judges score	Novelty	Originality (1-4)	The degree to which the idea is not only rare, but is also ingenious, imaginative or surprising
Judges score		Paradigm Relatedness (1-4)	The degree to which an idea is paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational
Judges score	Quality	Effectiveness (1-4)	The degree to which the idea will solve the problem
Judges score		Feasibility (1-4)	An idea is feasible if it can be easily implemented and does not violate known constraints
Words count		Elaboration	An idea is feasible if it can be easily implemented and does not violate known constraints
Post-assessment		Variety (1-4)	Variety is a measure of the explored solution space during the idea generation process

Table 9 Summary of the designed criteria

3.2.2 Judges

Metrics assessed by judges

Judges should focus on the scoring of 4 main criteria: Originality, Paradigm Relatedness, Effectiveness, and Feasibility. The other criteria, Elaboration, and Variety, will be assessed respectively through the words count calculated directly from the Excel Sheet and as derivative of the Paradigm Relatedness results. According to Dean, all the metrics will be assessed with a 1-4 Likert scale to ease the assessment process to judges (supported by Jauk 2012).

Judges' background

Benedek (Benedek et al., 2013) sustained that there is a crucial difference between designers and other people in the ideation assessment, "that for the former, evaluation is based on the same source of knowledge from which the concept was produced" (Benedek et al., 2013). According to this statement, it is essential that judges have clear ideas concerning the concept of the idea generation process. The level of expertise in the related field, from the literature review, seems not playing a huge role; therefore, students and researchers could have the role of judges. From the literature, high inter-judge correlations were obtained both with teachers (Blohm et al., 2010), students (Mouchiroud & Lubart, 2001), researchers (Johnson, et al., 2014) and with different level of expertise (Christiaans, 2002). Furthermore, Amabile (1982) encouraged judges' diversity in order to obtain better results in the assessment procedure.

Number of judges required

Analysing the literature, it is easy to notice that assessment methods with non-binary scale relying on a low number of judges could show high inconsistencies. Evidence could be seen in the experiments proposed by Guilford (1957), Evans (1983) and Srivathsaval et al (2010), in which through the use of 2 judges, medium-low reliability was found. Differently, optimal results were obtained through the use of 3 or more judges (Fink & Neubauer, 2006; Dean et al., 2006; Radel et al., 2015).

3.2.3 Guidelines

Amabile (1982) thought that guidelines should not have been provided to judges. Nevertheless, this choice had the aim to enhance the different perspectives of judges concerning creativity in artistic works. For the present experiment, being imprinted in a design perspective, one of the main requirements is that the evaluation phase should be replicable, therefore trying to follow and adopt a standard. The best way to achieve this aim is to follow the method proposed by Dean et al. (2006), providing proper guidelines and adopting anchor tables. These charts have the aim to support the scoring process, describing explicitly the characteristics that the responses should meet in order to reach a specific value for a given criterion. These tables will be adopted for all the criteria assessed by judges, Originality, Paradigm Relatedness, Feasibility, and Effectiveness.

Besides, it will be suggested to judges to read all the responses to related to a specific stimulus before to proceed with the evaluation of them in order to have a clearer idea regarding the concepts that could be proposed by participants. This suggestion comes from the fact that there is no reference (absolute standard) for every single object. This expedient was recommended by Shah et al. (2003), who refers to it, talking about the comparison principle. Furthermore, Amabile supported this decision in her research of 1982.

The choice to provide judges with guidelines was based on the literature review, in which assessment methods including the training of judges (Milgram & Arad, 1981; Zarnegar, Hocevar, & William, 1988, Besemer & O'Quin, 1999) or following precise guidelines (Dean et al., 2006; Sarkar & Chakrabarti, 2011) obtained optimal results concerning the inter-judge reliability.

3.2.4 The procedure

The evaluation procedure is divided into two parts; a preliminary phase, in which raters are asked to read the guidelines provided, and the assessment phase, in which judges will score the ideas proposed by participants.

The first step is important in order to give to all raters the same background regarding the topic discussed and to allow them to have a clear view regarding the task required. The guidelines document is composed by three main sections: the first one related to the criteria adopted in literature for the measure of ideational output, the second one presenting the dimensions and sub-dimensions used in the present experiment and the third one presenting the procedure itself and the anchor tables. In the last section, some important additional rules are displayed in order to support the method's validity:

- Judges should take the assessment independently enabling an impartial evaluation of the method through the measurement of the inter-judge reliability. This approach is supported by Amabile (1982) and Dean et al. (Dean et al., 2006) researches.
- Following the suggestions provided by Dean, being the assessment phase extremely time-consuming, judges should take the assessment in several sessions in order to minimise fatigue and to have the lucidity required in order to score the responses correctly.

The scoring phase will be done by judges using Excel files. Each file (one for each judge) is divided into 40 sheets, each one displaying all the functions proposed by participants concerning a single stimulus in order to ease the scoring phase and enabling judges to read all of them before starting the assessment phase of a stimulus.

Following the suggestions proposed by Amabile, responses included inside the Excel sheets are proposed in a different order enabling the avoidance of any biases related to the succession proposed. This methodology is shared by Besemer and O'Quin (Besemer & O'Quin, 1986, 1999).

During the assessment phase, judges should be able to access to a printed version of the guidelines in order to quickly consult the anchor tables provided.

The guidelines provided to participants and an example of an excel sheet are present in the Appendix (A.4).

3.3 Brain Signals

After the completion of the design of the methodology required for the assessment of verbal responses, the analysis regarding the brain signals has been conducted. In the following pages, general information regarding the brain and neuro signals are displayed to allow the reader to understand the experiment exhaustively. Besides, the last subchapter is dedicated to the data preparation process executed in order to allow a reliable analysis.

3.3.1 The Brain and the cognitive processes

The brain is an organ presents in all animals, vertebrate and invertebrate, serving the nervous system. Its main goal is the organisation of the movement. It is an extremely complex organ that has been studied for centuries without having a complete understanding of it yet.
The brain summary in the next subchapters has been divided in the structure of the brain, referred to as brain anatomy, and in its working system, referred to as brain physiology.

3.3.1.1 Anatomy

The human brain is composed of 4 main parts: cerebrum, cerebellum, brain stem, and diencephalon.

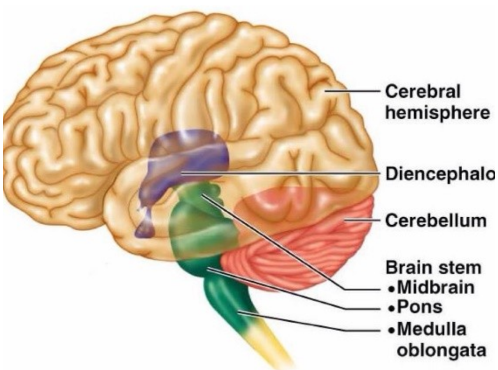

Figure 3.4 brain composition – (Soraghan, 2010)

- The Cerebrum is the largest and uppermost part of the brain, and it contains the cerebral cortex and several subcortical structures, among which basal ganglia, hippocampus and olfactory bulb. It is composed of white matter (myelinated nerve fibre) and the grey matter (outer cortex) and is dedicated to the control of higher-level functions in the human body, such as speech, reasoning, emotions, elaborating, coordinating and learning.
- The Cerebellum is located at the back of the brain, immediately below to the occipital area of the Cerebrum. It is composed of the cortex and white matter, several nuclei, and a fluid-filled ventricle in the middle. It receives information from the sensory system, spinal cord and other parts of the brain in order to coordinate motor movements.
- The Brainstem is located in the posterior part of the brain and represents the continuation of the spinal cord. It is divided into three main sections: midbrain, pons

and medulla oblongata. It is the centre of control of many vital body functions, such as breathing, swallowing and vasomotor control. It is crucial even for sympathetic and parasympathetic autonomic functions.

- The Diencephalon is located deep in the brain, underneath the cerebrum. It is composed of the thalamus and the hypothalamus, and it represents the link between the nervous system and the endocrine system. Its primary function is to relay sensory information between different parts of the central nervous system and to help the brain to interpret signals.

The focus of this research is on the Cerebrum, as explained before, dedicated to the most complex cognitive processes of human thinking. Its wrinkled surface is composed of peaks, called gyri, and depressions called sulci. It is formed by two cerebral hemispheres, that, at first sight, could seem similar and execute some shared vital functions together, but perform even one-side functions.

Both hemispheres could be divided into four lobes, according to the cranial bones protecting them: occipital lobe, parietal lobe, temporal lobe and frontal lobe.

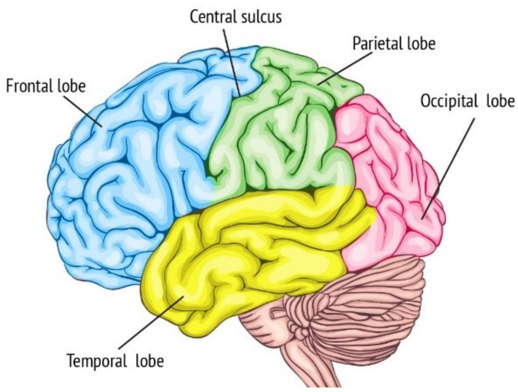

Figure 3.5 Cerebrum composition – Source: www. 5inc-elements.com

- The occipital lobe is located in the back of the brain. It is the seat of the brain's visual cortex, responsible for mapping the visual world with colours, shapes, distances, sizes and depth. It helps both spatial reasoning and visual memory.
- The parietal lobe is located on the middle part of the top of the cerebral cortex, between the occipital and the frontal lobe. It has the responsibility to process sensory information in a short amount of time. It is vital for the perception and integration of information related to senses (hearing, touch, sight, smell and taste). It has an essential role in the self-awareness of the location of a person's body parts in the space.
- The temporal lobe is located in the bottom part of the cerebrum, under the parietal and frontal lobes. Its functions are strictly related to auditory perception, and it is the seat of language comprehension and use. It is even associated with visual perception and facial recognition. It assumes a vital role regarding mnemonic processes, especially concerning long-term memory.
- The frontal lobe is located in the anterior part of the cerebral cortex, ahead of the parietal lobe. It is the seat of the most critical cognitive skills in humans, such as

decision-making processes, emotional expression, problem-solving, judgment and sexual behaviours. It controls the attention of the individual, filtrating information and stimuli that are considered less important than the ongoing process. It controls the language, and it is responsible for motivational control.

3.3.1.2 Physiology

The brain activity is possible because of neurons and the interconnections existing between them. They are composed of a cell body (soma), an axon and dendrites. The cell body contains the nucleus and cytoplasm. The axon connects the cell body to smaller branches called axon terminals. Dendrites are linked to the cell body and receive information from other neurons.

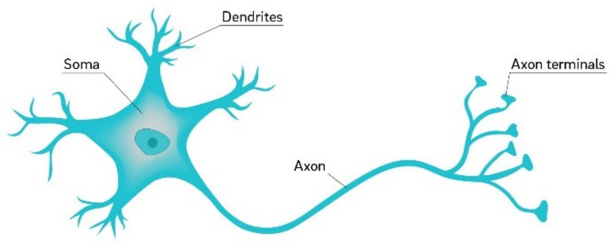

Figure 3.6 neuron illustration (Puppo, 2018)

The connections existing between neurons are called synapses and control their excitatory and inhibitory activities.

The process through which signals are exchanged from one neuron to another is called neurotransmission and, at a chemical level, it consists of the release of neurotransmitters by the vesicles present near to the surface of an axon terminal through its transporters. The neurotransmitters react with receptors present on the dendrites, causing depolarization of the receiving neuron terminations and thus allowing cations to flow across the membrane. The result of this process, at an electrical level, is the presence of an area extramembrane with a negative charge. The resulting electrical field is known as postsynaptic potential (PSP), lasting hundreds of milliseconds, allowing the transmission of a massive quantity of data in a short amount of time. “When the postsynaptic potential involves groups of neurons, the electrical field becomes stronger and therefore detectable by specific

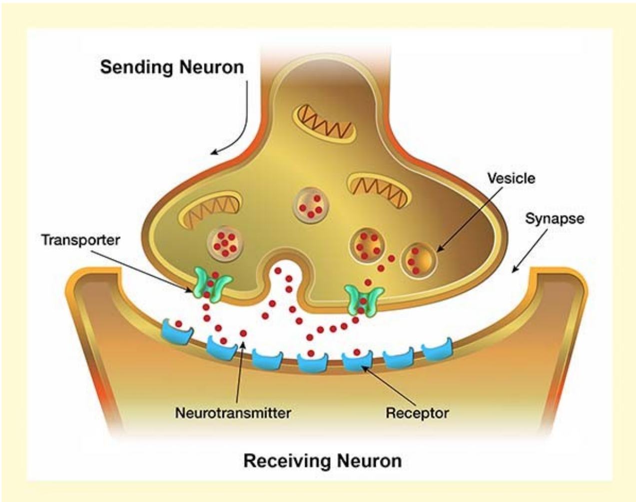

Figure 3.7 Synaptic transmission - Source www.thestudentscientist.org

instruments such as the electroencephalography (EEG)" (Colombo, 2019). The measurement of the electrical field is possible thanks to the orientation of neurons (pyramidal neurons) that are structured perpendicularly to the scalp surface, allowing the propagation of the electric field beyond its outer layer. The orientation of pyramidal neurons is the reason why the recorded signals are perceived as positive or negative by instruments such as EEG; the former are recorded when the excited dendrites of the neuron are farther from the electrode, while the opposite happens in situations in which the excited dendrites of the neuron are closer to the electrode (Colombo, 2019).

This mechanism permits the transmission of information from one neuron to another, creating a complex network that allows every cognitive act. It is essential to say that in the neuronal network there are preferred routes followed by signals, created by the fortification of synapses that are repeatedly utilised. The interconnection could take place within the same cortical area or involving different parts of the brain. The connections could follow both the direction defining a bottom-up process, in which the sensorial stimuli are elaborate by the brain, or top-down processes, in which the signal starts from a higher hierarchy of the thought (therefore, involving the subject attention), then reaching lower stages of elaboration.

3.3.2 Brain waves (alpha waves)

The EEG signals are composed of 5 components (Colombo, 2019):

- Frequency – representing the number of cycles per second and measured in Heartz (Hz).
- Amplitude – measuring the distances between the horizontal axe and the peak of the wave. Differently from the mechanic and electromagnetic, it is measured in microvolt (V).
- Phase – representing the time-position in a cycle and it is measured in degrees or radians.
- Latency – representing the lapse in which the wave is recorded. It is measured in milliseconds (mS).
- Topography – representing the physical location where the wave is recorded. It is usually associated with the electrode position, according to the American Electroencephalographic Society standard (Sharbrough, et al., 1991).

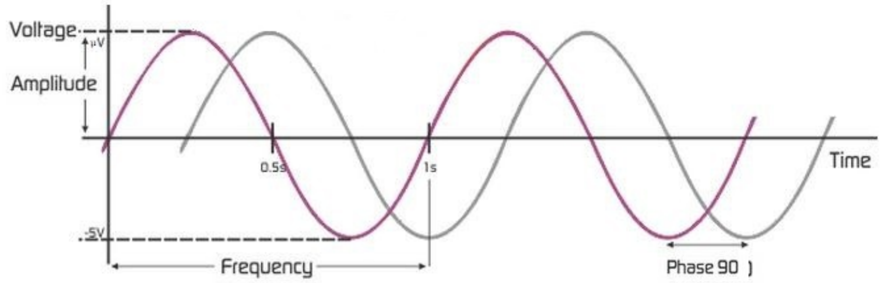

Figure 3.8 illustration of an ideal wave

The most utilised classification concerning human brain waves is the one based on the recorded signals frequencies and includes (Sawant & Zahra, 2010):

- Delta waves – including all the waves in the EEG recording with a frequency below 3.5 Hz. They are typical of the deep sleep phase, in childhood and people presenting organic severe brain diseases.
- Theta waves – ranging between 4 and 7 Hz. They could be observed in childhood, but also emotional stress phases of adults.
- Alpha waves – ranging between 8 and 13 Hz. They could be observed when people are awake in a calm and resting state.
- Beta waves – ranging between 13 and 30 Hz. They occur during mental activity states.
- Gamma waves – including all the waves in the EEG recording with a frequency above 30 Hz. They could be observed during sensory and motor processes, in which high focus is required.

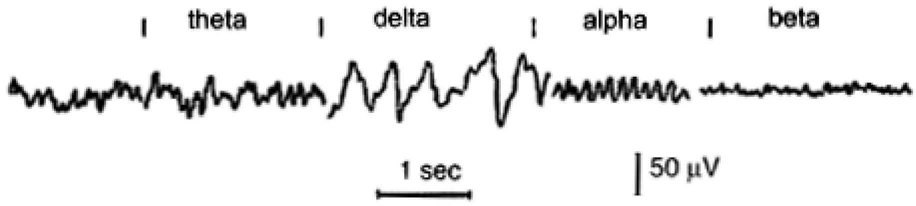

Figure 3.9 Brain waves classification (Younan, 2012)

It is interesting to note that there is an inverse correlation between the frequency and the amplitude of the human brain signals; in fact, high frequencies are usually associated with low amplitudes and vice versa.

The picture (Fig 3.10) below shows an example of an EEG output, in which the axes represent respectively the time (horizontal) and the amplitude (vertical) of the signals recorded by electrodes.

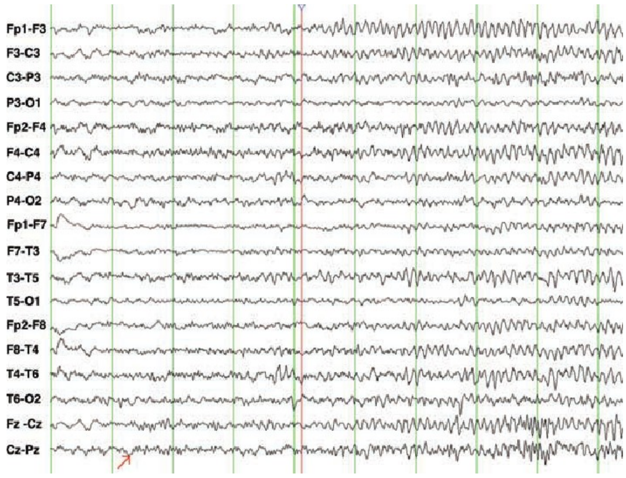

Figure 3.10 Example of EEG rough output (Glass, Prieur, Molnar, & Hamiwka, 2006)

Alpha waves cover an essential role in the study of brain activity. Pfurtscheller et al. (1997) theorised the so-called idling hypothesis, stating that alpha activity is observed in areas of resting neural population. Therefore, desynchronization of alpha frequencies (ERD) was considered as an indicator of cortical motor activation, while alpha synchronization was instead associated with areas not involved in the task (Colombo, 2019). This view was then integrated by another theory stating that alpha waves could be observed even during active processes, the so-called top-down decisions. Other studies confirmed this theory in the following years (Fink & Benedek, 2014; Klimesch, 2012; Kmnesch et al., 2007), therefore the correlation existing between the increase of alpha waves and cognitive processes, especially memory and attention (Colombo, 2019).

Martindale and Hines (1975) were the first ones to theories that alpha waves could allow brain areas to enter in an internally directed state, filtering external stimuli and permitting the use of the area to focus on internal tasks. They sustained that free-associative primary processes, critical in creative ideation processes, happen during relaxed stated in areas with low cortical arousal (Colombo, 2019).

In 2017 Benedek et al. (Benedek et al., 2017) stated that the attention could be directed both within the subject or towards the external environment, respectively through mental representations and physical perceptions. The tendency towards one or the other is variable in each person.

Summing up, the alpha frequency permits people to direct the attention towards internal cognitive processes, partly eliminating the external stimuli and allowing subjects to create elaborated images within their mind. These thoughts are usually directed towards memories or future planning (Colombo, 2019) and allow the subject to have better results in the creative process. Therefore, alpha waves could be considered as an indicator of the attention direct by the person to the creation of creative thoughts.

3.3.3 Independent component analysis (ICA)

The EEG signals obtained from the scalp are the results of a massive amount of neurons potentials. The outputs usually contain, beyond the useful information concerning cerebral activities, redundant or noise information because of the extreme sensibility of this tool, recording also:

- Muscular activity.
- Cardiac activity.
- Eye activity.
- External distortions (e.g. electrical field generated by devices).

In order to have more transparent and precise data, it would be necessary to take the measures directly from different centres of the brain, which would mean surgery (Unguranu, 2004). This would transform the assessment to an invasive measurement, losing the advantage that makes it one of the most adopted techniques in the neuroimaging field.

Therefore, it is required a careful cleaning phase in order to ensure the reliability of data. Several studies have been conducted on the data cleaning phase of the EEG outputs, the majority of which relying on algorithms involving the independent component analysis (ICA) or the principal components analysis (PCA) (Tangkeaingkij et al.; 2009; Subasi & Gursoy, 2011; Jung et al., 1995).

Regarding the experiment, all the signals have been analysed with Letswave, an Open source Matlab EEG signal processing toolbox. It allowed the researchers to compute a visual inspection on the independent components identified by the software in order to understand which could have been associated with brain signals and, instead, which were related to noise. The waves have been filtered in the pre-processing phase through the Fast Fourier Transform (FFT), the most adopted algorithm in frequency analysis in many scientific fields (Freeman & Quian, 2013), ranging from 0,1 to 30Hz. The output of the FFT is a periodogram, also called Power Spectral Density (PSD). It is characterised by the distribution of the power (μV^2) as a function of the frequency spectrum of the signal (Cohen, 2018; Colombo, 2018).

In the picture (Fig. 3.11) below it is represented a screenshot of the ICA spatial filter in which the graph above represents an independent component signal identified by the software, the graph below represents the resulting of the independent components still not removed and the list on the right represents the independent component to be removed by the software. In the red graph, it is easily visible the noise caused by the blinks of the subject, corresponding to the peaks recorded.

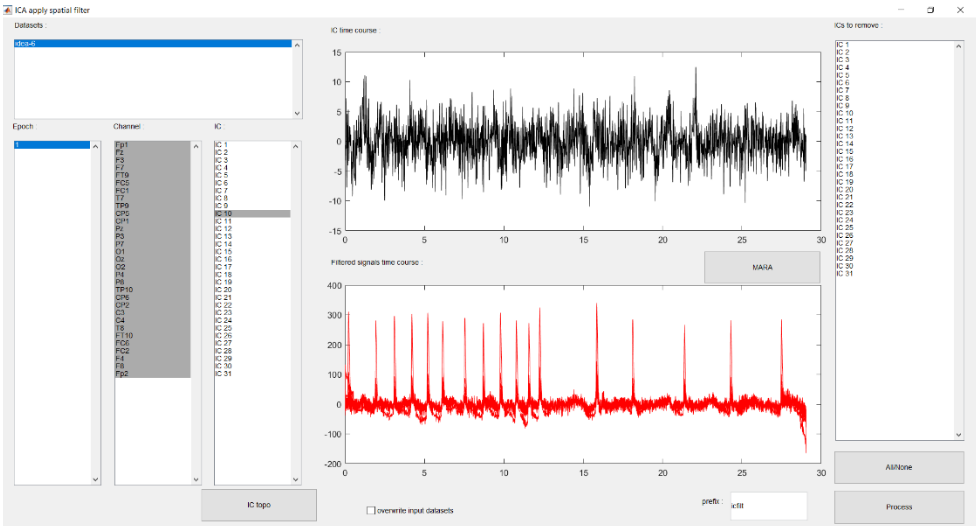

Figure 3.11 Example of a response visual inspection

Letswave also permits to visualise independent components through a brain representation in order to give researchers a better view of the signal analysed. Below an example is displayed.

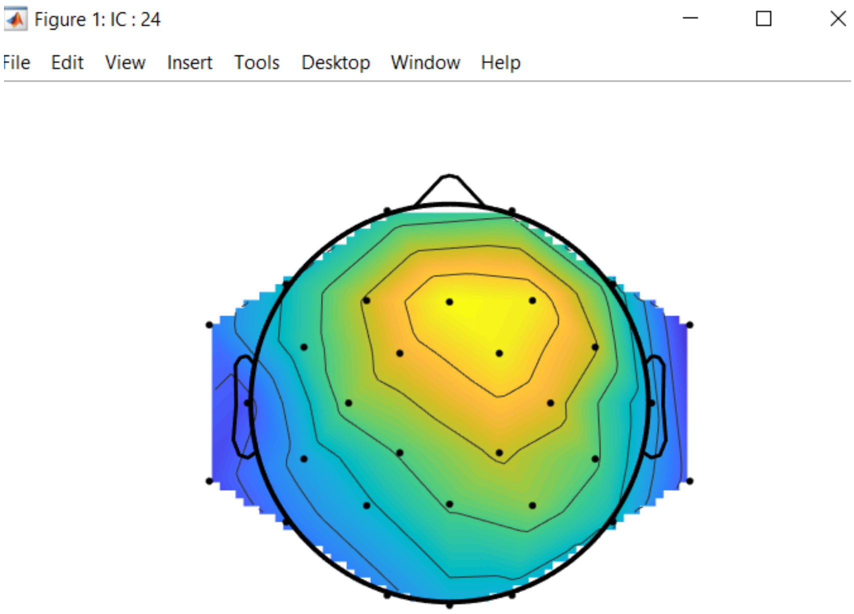

Figure 3.12 Example of an independent component brain representation

The visual inspection process was extremely time-consuming for researchers, requiring the analysis of 31 independent components per idea and per reference (gathered in two blocks of 20 each plus three big references) for each participant. The total number of independent components analysed for this process amounted approximately to 40.000.

3.4 Design of the experiment

In order to compute the statistical analysis, therefore, comparing cortical activity with the outcome of the ideation process, a strong assumption has been taken, that is that participants responded creatively in the uncommon condition of the experiment, while in a non-creative way in the common condition. This simplification was needed in order to compute a first statistical analysis of the data and will be released only after the completion of the assessment by judges. The just mentioned phase will not be part of the present thesis being extremely time-consuming; therefore, it will allow further studies to compare the neurophysiologic results with the actual outcome data.

According to the literature, the analysis should be computed considering three major within-subject factors in order to understand and classify the cortical activities during divergent thinking tasks: Area, Condition and Hemisphere. Therefore, a clusterisation of the electrodes in macro areas was required to start the analysis. In order to understand which configuration could have better described the position of the electrodes, three main subdivisions of the hemispheres have been studied:

- 3 macro areas – this configuration has been considered for its simplicity, dividing the hemispheres into Frontal, Central and Posterior regions.
- 6 macro areas – this configuration has been studied based on the position of the electrodes in the EEG helmet adopted. It divides the hemispheres into Frontal ventral, Frontal dorsal, Central ventral, Central dorsal, Posterior ventral and Posterior dorsal regions.
- 7 macro areas – this configuration has been studied for the high frequency it is adopted in the literature (Jauk et al., 2012; Fink & Neubauer, 2006; Fink et al., 2009;

Fink & Benedek, 2014). However, the experiments just mentioned present some main differences in the configuration of the EEG helmet in comparison to the experiment here studied, and it would require some substantial modifications in the clusters.

Besides, further analysis has been conducted in order to identify any substantial differences between the analysis on the aggregation of all values referred to an electrode for each subject (the mean) and on the whole population. In the former case, the reduction of the number of data analysed could bring a loss of reliability (36 lines vs 354), while in the latter case, some assumptions were needed because of the lack of a univocal relationship between responses common and uncommon of each participant. In Chapter 4, the results of the configuration chosen are displayed in greater details.

3.4.1 Inputs: TRP & database

Using the output of the visual inspection, Power Spectral Density (PSD), it has been possible to calculate the power related to the different wavebands for each electrode of each response. This process has been executed externally to the present research because of the lack of the proper knowledge required for such a task.

The related power has also been obtained for the reference signal in order to highlight the cortical activation differences between the task execution phase and the resting one. Alpha power was then calculated for each electrode employing the Task Related Power (TRP) (Pfurtscheller & Lopes da Silva, 1999), according to the formula (Colombo, 2019):

$$TRP = \log(power_{activation}) - \log\,(power_{reference})$$

Positive of TRPs are associated with ERS, event-related synchronization, thus an increase of the power measured during the task execution in comparison to the reference period, while negative TRPs are associated with ERD, event-related desynchronization.

In order to allow the reader to understand the analysis conducted, it is essential to display the electrodes configuration adopted for the experiment. The picture below (Fig 3.13) represents the position of electrodes in the helmet utilised for the experiment.

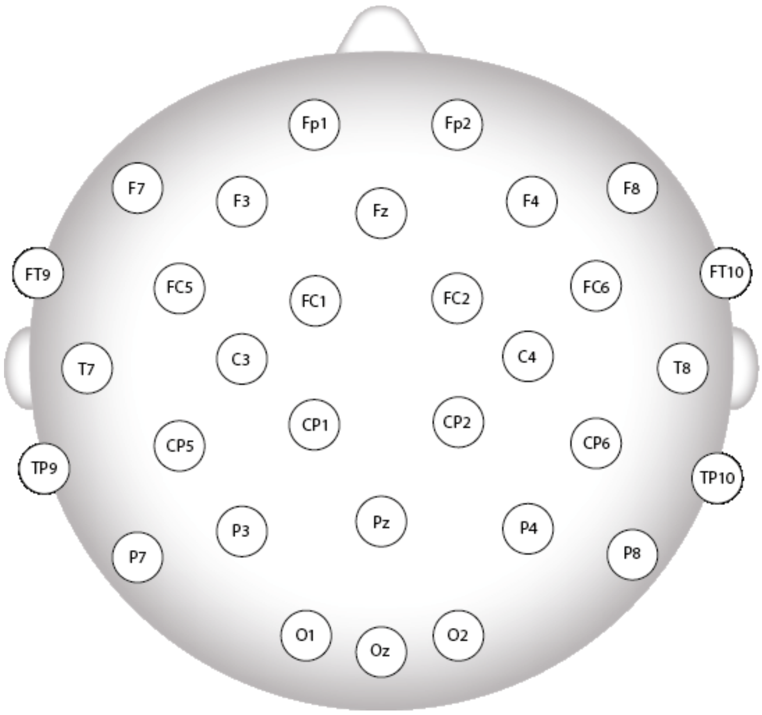

Figure 3.13 Electrodes configuration

As visible from the reported image, each electrode is characterised by letters and a number:

- The letters are associated with the anatomic position of the electrode: F represents the Frontal area, C represents the Central area, T represents the Temporal Area, P represents the Parietal area and O represents the Occipital area. The electrodes to which is associated more than one letter are in proximity of 2 areas (e.g. FC represents the frontocentral region) or are in correspondence of a specific location of the lobe (e.g. Fp corresponding to the Frontopolar region).
- The number identifies the position on the horizontal axe and the related hemisphere: even numbers for the right hemisphere, odd numbers for the left one and the letter z for electrodes allocated on sagittal axe of the brain.

For the analysis, electrodes present on the sagittal axe of the brain (Fz, Pz, Oz) have been neglected in order to permit the comparison between the left and right hemisphere. After a further cleaning phase in which the responses containing values differing for more than $2,5*\sigma$ from the value of the mean of all alpha TRPs were removed, the remaining alpha TRPs of all subjects have been organised in an Excel sheet in order to compute the analysis.
The Excel sheet has been set considering:

- Rows as discriminants of subject and stimulus.
- Columns as discriminants of condition, hemisphere and electrode.

The tables proposed below should be interpreted by the reader as explicative tools, needed for the massive amount of space required to represent the whole configuration of the database in a physical paper.

	Common	Common	Uncommon	Uncommon
	Left	Right	Left	Right
Subject 1	Values: Subject1, Common, Left	Values: Subject1, Common, Right	Values: Subject1, Uncommon, Left	Values: Subject1, Uncommon, Right
Subject 2	Values: Subject2, Common, Left	Values: Subject2, Common, Right	Values: Subject2, Uncommon, Left	Values: Subject2, Uncommon, Right

Table 10 Macro database explanation

In order to understand the sheet configuration, the Values' cells in the picture above should be considered by the reader as matrices, containing on rows stimuli and columns electrodes. Below it is possible to see an example of how data were displayed for the cell "Values: Subject1, Common, Left".

		Common													
		Left													
		Fp1	**F3**	**F7**	**FT9**	**FC5**	**FC1**	**T7**	**C3**	**TP9**	**CP5**	**CP1**	**P3**	**P7**	**O1**
Subject 1	**Stim1**	-0,73	-0,58	-0,35	-0,42	-0,12	-0,27	-0,29	-0,38	-0,21	-0,51	-0,53	-0,26	-0,22	-0,3
	Stim2	0,241	0,227	0,354	0,195	0,175	0,144	-0,02	0,259	-0	0,401	0,464	0,523	0,107	0,234
	Stim3	-0,12	-0,08	-0,27	-0,07	-0,38	-0,33	-0,31	-0,26	0,053	-0,41	-0,13	-0,34	-0,29	-0,31
	Stim4	0,188	0,028	0,167	0,077	-0,16	-0,01	0,401	0,094	-0,07	0,272	0,162	-0,03	0,183	0,008
	Stim5	-0,4	-0,08	0,275	0,727	-0,07	-0,45	-0,06	-0,06	0,197	0,017	-0,14	-0,16	0,23	0,438
	Stim6	1,191	1,028	1,087	0,692	0,416	0,107	0,161	-0,1	0,078	-0,16	-0,52	-0,77	-0,39	-0,4
	Stim7	-0,55	-0,09	-0,78	-0,54	0,106	0,74	0,496	0,759	0,102	0,318	0,664	0,468	0,482	-0,03
	Stim8	-0,2	-0,09	-0,1	0,249	-0,21	-0,3	-0,29	-0,33	-0,51	-0,3	-0,22	-0,22	-0,43	-0,5
	Stim9	-0,38	-0,02	0,172	0,768	0,338	0,204	0,504	0,518	0,777	0,666	0,427	0,641	0,748	0,888
	Stim10	-0,06	-0,86	-0,55	-0,3	-1,55	-1,69	-0,92	-1,07	-0,62	-0,46	-0,65	-0,66	-0,72	-0,76
	Stim11	0,32	0,382	0,392	0,265	-0	0,133	0,173	0,21	0,713	0,199	0,16	0,065	0,265	0,643
	Stim12	0,517	0,649	0,52	0,128	0,781	0,898	0,883	0,575	0,9	0,561	0,38	0,436	0,523	0,542
	Stim13	-0,5	0,702	0,681	-0,1	0,559	0,639	0,341	0,486	0,144	0,256	0,372	0,228	0,081	0,169

Table 11 Detailed database explanation

In order to compute a factorial analysis considering as within-subject factors Condition, Areas and Hemisphere, it has been assumed a univocal relationship between the responses common and uncommon by participants, therefore considering a "subject" as composed of two responses (one common and one uncommon) given by the same participant. This assumption was needed in order to analyse the whole data population.

With the just mentioned settings, the Area configurations have been compared and listed in the next subchapters. Further test (e.g. Mauchly) could be seen in the appendix (A.5). Results have been compared to the ones obtained considering participants means, calculated aggregating the stimuli in the same dataset configuration.

3.4.2 "3 areas"

The 3-area configuration has been the first one to be studied. The subdivision adopted for the left hemisphere, thus applicable symmetrically to the right one, was the following:

- Frontal – including the electrodes Fp1, F7 and F3.
- Central – including the electrodes FT9, FC5, FC1, T7, C3, TP9, CP5 and CP1.
- Posterior – including the electrodes P7, P3 and O1.

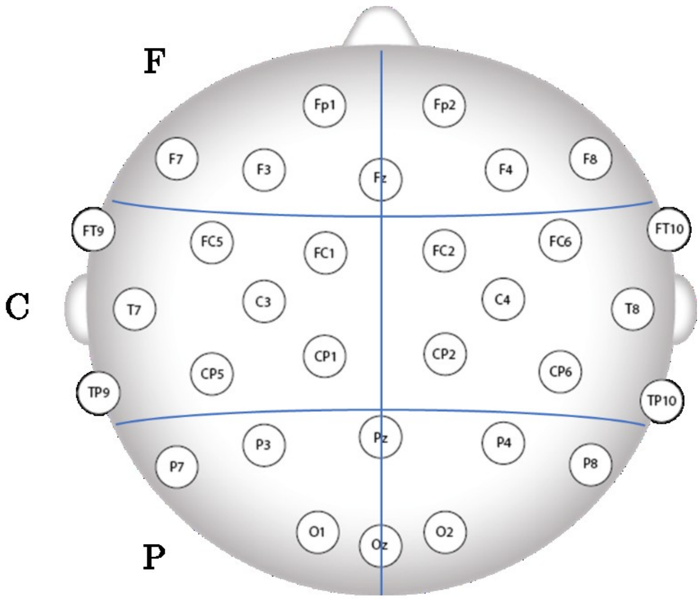

Figure 3.14 3-Areas configuration

In the analysis concerning the means of the different responses given by subjects, levels of significance have been reached only for the factor Area ($F[2,70]= 24.973$, $p< 0.001$, $\eta^2_{partial}= 0.416$) and the interaction Area*Condition ($F[2,70]= 4.557$, $p< 0.05$, $\eta^2_{partial}= 0.115$).

Tests of Within-Subjects Effects (means 2x2x3)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition	Sphericity Assumed	,139	1	,139	1,227	,276	,034
	Greenhouse-Geisser	,139	1,000	,139	1,227	,276	,034
Hemisphere	Sphericity Assumed	,000	1	,000	,024	,878	,001
	Greenhouse-Geisser	,000	1,000	,000	,024	,878	,001
Area	Sphericity Assumed	3,900	2	1,950	24,973	,000	,416
	Greenhouse-Geisser	3,900	1,487	2,624	24,973	,000	,416
Condition * Hemisphere	Sphericity Assumed	,004	1	,004	,216	,645	,006
	Greenhouse-Geisser	,004	1,000	,004	,216	,645	,006
Condition * Area	Sphericity Assumed	,262	2	,131	4,557	,014	,115
	Greenhouse-Geisser	,262	1,308	,200	4,557	,029	,115
Hemisphere * Area	Sphericity Assumed	,052	2	,026	2,472	,092	,066
	Greenhouse-Geisser	,052	1,476	,035	2,472	,109	,066
Condition * Hemisphere * Area	Sphericity Assumed	,014	2	,007	1,060	,352	,029
	Greenhouse-Geisser	,014	1,544	,009	1,060	,338	,029

Table 12 Tests of Within-Subjects Effects (means 2x2x3)

The analysis on the whole population have been showed levels of significance for the factor Area ($F[2,706]= 115,993$, $p< 0.001$, $\eta^2_{partial}= 0.247$) the interactions Area*Condition ($F[2,706]= 26.885$, $p< 0.001$, $\eta^2_{partial}= 0.071$) and Area*Hemisphere ($F[2,706]= 3.887$, $p< 0.05$, $\eta^2_{partial}= 0.011$).

Tests of Within-Subjects Effects (all population 2x2x3)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition	Sphericity Assumed	,310	1	,310	,707	,401	,002
	Greenhouse-Geisser	,310	1,000	,310	,707	,401	,002
Hemisphere	Sphericity Assumed	,057	1	,057	1,342	,248	,004
	Greenhouse-Geisser	,057	1,000	,057	1,342	,248	,004
Area	Sphericity Assumed	21,324	2	10,662	115,993	,000	,247
	Greenhouse-Geisser	21,324	1,395	15,281	115,993	,000	,247
Condition * Hemisphere	Sphericity Assumed	,066	1	,066	1,467	,227	,004
	Greenhouse-Geisser	,066	1,000	,066	1,467	,227	,004
Condition * Area	Sphericity Assumed	2,886	2	1,443	26,885	,000	,071
	Greenhouse-Geisser	2,886	1,248	2,312	26,885	,000	,071
Hemisphere * Area	Sphericity Assumed	,141	2	,071	3,887	,021	,011
	Greenhouse-Geisser	,141	1,494	,095	3,887	,032	,011
Condition * Hemisphere * Area	Sphericity Assumed	,025	2	,013	1,098	,334	,003
	Greenhouse-Geisser	,025	1,529	,017	1,098	,321	,003

Table 13 Tests of Within-Subjects Effects (all population 2x2x3)

As expected, analysing the variance in the interaction Area*Condition*Hemisphere, the standard deviation associated with the means is much higher than the one calculated for the whole population. The respective means of the standard deviations are equal to 0,40 and 0,18.

Condition * Hemisphere * Area (means 2x2x3)

Condition			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Common	Left	Frontal	,131	,041	,047	,215
		Central	-,008	,033	-,075	,058
		Posterior	-,146	,049	-,245	-,048
	Right	Frontal	,154	,042	,068	,240
		Central	-,050	,035	-,121	,021
		Posterior	-,151	,048	-,248	-,054
Uncommon	Left	Frontal	,101	,034	,031	,170
		Central	,007	,034	-,062	,077
		Posterior	-,042	,041	-,126	,042
	Right	Frontal	,136	,040	,054	,217
		Central	,006	,035	-,064	,076
		Posterior	-,062	,044	-,152	,027

*Table 14 Condition * Hemisphere * Area (means 2x2x3)*

Condition * Hemisphere * Area (all population 2x2x3)

Condition			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Common	Left	Frontal	,163	,018	,127	,198
		Central	,031	,017	-,002	,064
		Posterior	-,070	,021	-,110	-,029
	Right	Frontal	,159	,019	,121	,197
		Central	-,007	,018	-,042	,028
		Posterior	-,075	,021	-,117	-,033
Uncommon	Left	Frontal	,114	,016	,082	,146
		Central	,033	,016	,002	,063
		Posterior	,004	,019	-,033	,042
	Right	Frontal	,115	,018	,079	,151
		Central	,023	,015	-,007	,053
		Posterior	,014	,020	-,024	,053

*Table 15 Condition * Hemisphere * Area (all population 2x2x3)*

3.4.3 “6 areas”

The 6-area configuration adopted for the left hemisphere, thus applicable symmetrically to the right one, was the following:

- Frontal dorsal (FD) – including the electrodes Fp1 and F3.
- Frontal ventricular (FV) – including the electrodes of F7 and FT9.
- Central ventricular (CV) – including the electrodes FC5, T7 and CP5.
- Central dorsal (CD) – including the electrodes FC1, C3 and CP1.
- Posterior ventricular (PV) – including the electrodes TP9 and P7.
- Posterior dorsal (PD) – including the electrodes P3 and O1.

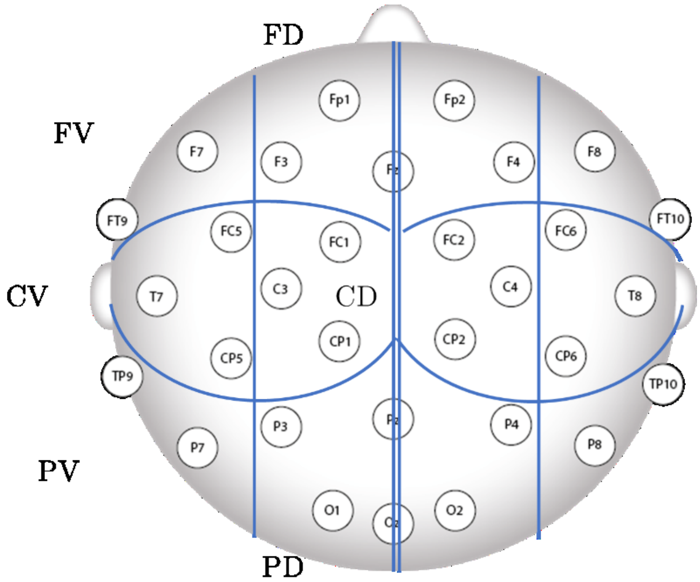

Figure 3.15 6-Areas configuration

In the analysis concerning the means of the different responses given by subjects, levels of significance have been reached only for the factor Area ($F[5,175]= 18.442$, $p< 0.001$, $\eta^2_{partial}= 0.345$) and the interaction Area*Condition ($F[5,175]= 4.298$, $p< 0.005$, $\eta^2_{partial}= 0.109$).

Tests of Within-Subjects Effects (means 2x2x6)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition_	Sphericity Assumed	,494	1	,494	2,097	,156	,057
	Greenhouse-Geisser	,494	1,000	,494	2,097	,156	,057
Hemisphere	Sphericity Assumed	,017	1	,017	,569	,456	,016
	Greenhouse-Geisser	,017	1,000	,017	,569	,456	,016
Area	Sphericity Assumed	5,896	5	1,179	18,442	,000	,345
	Greenhouse-Geisser	5,896	2,664	2,213	18,442	,000	,345
Condition_ * Hemisphere	Sphericity Assumed	,061	1	,061	1,492	,230	,041
	Greenhouse-Geisser	,061	1,000	,061	1,492	,230	,041
Condition_ * Area	Sphericity Assumed	,731	5	,146	4,298	,001	,109
	Greenhouse-Geisser	,731	2,520	,290	4,298	,011	,109
Hemisphere * Area	Sphericity Assumed	,092	5	,018	1,452	,208	,040
	Greenhouse-Geisser	,092	3,096	,030	1,452	,231	,040
Condition_ * Hemisphere * Area	Sphericity Assumed	,047	5	,009	,874	,500	,024
	Greenhouse-Geisser	,047	3,244	,014	,874	,464	,024

Table 16 Tests of Within-Subjects Effects (means 2x2x6)

The analysis on the whole population have been showed levels of significance for the factor Area ($F[5,1765]= 78.452$, $p< 0.001$, $\eta^2_{partial}= 0.182$, the interaction Area*Condition ($F[5,1765]= 19.607$, $p< 0.001$, $\eta^2_{partial}= 0.053$) and the one Area*Hemisphere ($F[5,1765]= 3.771$, $p< 0.01$, $\eta^2_{partial}= 0.011$). Significance level has been almost reached for the factor Hemisphere ($p<0.1$) and the interaction Condition*Hemisphere ($p<0.15$).

Tests of Within-Subjects Effects (all population 2x2x6)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition_	Sphericity Assumed	,736	1	,736	,845	,359	,002
	Greenhouse-Geisser	,736	1,000	,736	,845	,359	,002
Hemisphere	Sphericity Assumed	,256	1	,256	2,728	,099	,008
	Greenhouse-Geisser	,256	1,000	,256	2,728	,099	,008
Area	Sphericity Assumed	35,997	5	7,199	78,452	,000	,182
	Greenhouse-Geisser	35,997	2,820	12,763	78,452	,000	,182
Condition_ * Hemisphere	Sphericity Assumed	,238	1	,238	2,441	,119	,007
	Greenhouse-Geisser	,238	1,000	,238	2,441	,119	,007
Condition_ * Area	Sphericity Assumed	5,205	5	1,041	19,607	,000	,053
	Greenhouse-Geisser	5,205	2,641	1,971	19,607	,000	,053
Hemisphere * Area	Sphericity Assumed	,426	5	,085	3,771	,002	,011
	Greenhouse-Geisser	,426	3,343	,127	3,771	,008	,011
Condition_ * Hemisphere * Area	Sphericity Assumed	,060	5	,012	,688	,633	,002
	Greenhouse-Geisser	,060	3,270	,018	,688	,572	,002

Table 17 Tests of Within-Subjects Effects (all population 2x2x6)

As expected, analysing the variance in the interaction Area*Condition*Hemisphere, the standard deviation associated with the means is much higher than the one calculated for the whole population as in the previous configuration. The respective means of the standard deviations are equal to 0,42 and 0,19.

Condition_ * Hemisphere * Area (2x2x6 means)

Condition_			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Common	Left	FV	,108	,040	,026	,190
		FD	,151	,043	,064	,237
		CV	-,015	,039	-,094	,064
		CD	-,056	,033	-,123	,010
		PV	-,065	,046	-,158	,028
		PD	-,167	,051	-,270	-,063
	Right	FV	,081	,049	-,018	,180
		FD	,160	,042	,074	,245
		CV	-,090	,043	-,178	-,002
		CD	-,081	,036	-,154	-,009
		PV	-,101	,049	-,200	-,002
		PD	-,166	,049	-,265	-,067
Uncommon	Left	FV	,035	,036	-,037	,108
		FD	,119	,038	,043	,195
		CV	,010	,039	-,068	,089
		CD	,024	,041	-,060	,108
		PV	-,005	,044	-,094	,084
		PD	-,041	,039	-,120	,039
	Right	FV	,082	,044	-,007	,171
		FD	,148	,045	,057	,239
		CV	,007	,037	-,067	,082
		CD	,026	,039	-,052	,105
		PV	-,029	,044	-,118	,060
		PD	-,044	,046	-,136	,049

*Table 18 Condition_ * Hemisphere * Area (2x2x6 means)*

Condition_ * Hemisphere * Area (all population 2x2x6)

Condition_			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Common	Left	FV	,136	,020	,098	,174
		FD	,175	,019	,138	,212
		CV	,039	,018	,003	,075
		CD	-,008	,018	-,043	,027
		PV	-,015	,020	-,054	,023
		PD	-,078	,021	-,119	-,037
	Right	FV	,124	,022	,080	,168
		FD	,164	,020	,125	,202
		CV	-,008	,019	-,047	,030
		CD	-,037	,018	-,073	-,001
		PV	-,043	,021	-,084	-,002
		PD	-,080	,022	-,123	-,036
Uncommon	Left	FV	,081	,018	,045	,117
		FD	,123	,016	,091	,156
		CV	,044	,017	,010	,078
		CD	,013	,017	-,020	,046
		PV	,042	,019	,004	,080
		PD	-,006	,019	-,045	,032
	Right	FV	,084	,019	,045	,122
		FD	,127	,019	,090	,164
		CV	,032	,016	,000	,064
		CD	,014	,016	-,017	,046
		PV	,018	,019	-,020	,056
		PD	,019	,020	-,021	,059

*Table 19 Condition_ * Hemisphere * Area (all population 2x2x6)*

3.4.4 “7 areas”

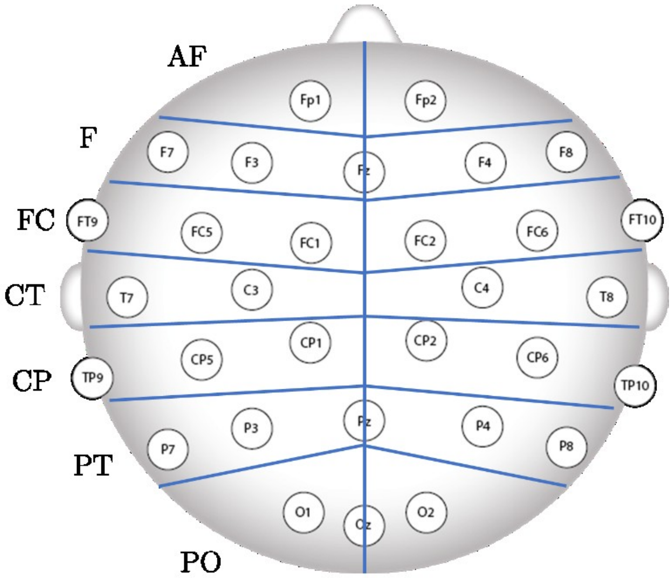

Figure 3.16 7 areas configuration

The 7-area configuration adopted for the left hemisphere, thus applicable symmetrically to the right one, was the following:

- Anteriofrontal (AF) – including the electrode Fp1.
- Frontal (F) – including the electrodes of F3 and F7.
- Frontocentral (FC) – including the electrodes FT9, FC5 and FC1.
- Centrotemporal (CT) – including the electrodes C3 and T7.
- Centroparietal (CP) – including the electrodes TP9, CP5 and CP1.
- Parietotemporal (PT) – including the electrodes P3 and P7.
- Parietooccipital (PO) – including the electrode O1.

In the analysis concerning the means of the different responses given by subjects, levels of significance have been reached only for the factor Area ($F[6,210]= 23.576$, $p< 0.001$, $\eta^2_{partial}= 0.402$) and the interaction Area*Condition ($F[6,210]= 4.931$, $p< 0.001$, $\eta^2_{partial}= 0.123$).

Tests of Within-Subjects Effects (means 2x2x7)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition_	Sphericity Assumed	,516	1	,516	2,025	,164	,055
	Greenhouse-Geisser	,516	1,000	,516	2,025	,164	,055
Hemisphere	Sphericity Assumed	,001	1	,001	,040	,842	,001
	Greenhouse-Geisser	,001	1,000	,001	,040	,842	,001
Area	Sphericity Assumed	11,977	6	1,996	23,576	,000	,402
	Greenhouse-Geisser	11,977	2,130	5,624	23,576	,000	,402
Condition_ * Hemisphere	Sphericity Assumed	,046	1	,046	1,045	,314	,029
	Greenhouse-Geisser	,046	1,000	,046	1,045	,314	,029
Condition_ * Area	Sphericity Assumed	1,076	6	,179	4,931	,000	,123
	Greenhouse-Geisser	1,076	1,892	,569	4,931	,011	,123
Hemisphere * Area	Sphericity Assumed	,094	6	,016	1,190	,313	,033
	Greenhouse-Geisser	,094	3,448	,027	1,190	,318	,033
Condition_ * Hemisphere * Area	Sphericity Assumed	,069	6	,012	1,170	,324	,032
	Greenhouse-Geisser	,069	3,388	,020	1,170	,326	,032

Table 20 Tests of Within-Subjects Effects (means 2x2x7)

The analysis on the whole population have been showed levels of significance for the factor Area (F[6,2118]= 112.795, p< 0.001, $\eta^2_{partial}$= 0.242), the interaction Area*Condition (F[6,2118]= 22.143, p< 0.001, $\eta^2_{partial}$= 0.059) and the one Area*Hemisphere (F[6,2118]= 2.816, p< 0.005, $\eta^2_{partial}$= 0.008). Significance level has been almost reached for the interaction Condition*Hemisphere and the factor Hemisphere (p<0.1) .

Tests of Within-Subjects Effects (all population 2x2x7)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition_	Sphericity Assumed	,721	1	,721	,722	,396	,002
	Greenhouse-Geisser	,721	1,000	,721	,722	,396	,002
Hemisphere	Sphericity Assumed	,267	1	,267	2,780	,096	,008
	Greenhouse-Geisser	,267	1,000	,267	2,780	,096	,008
Area	Sphericity Assumed	73,822	6	12,304	112,795	,000	,242
	Greenhouse-Geisser	73,822	2,179	33,885	112,795	,000	,242
Condition_ * Hemisphere	Sphericity Assumed	,339	1	,339	3,290	,071	,009
	Greenhouse-Geisser	,339	1,000	,339	3,290	,071	,009
Condition_ * Area	Sphericity Assumed	8,431	6	1,405	22,143	,000	,059
	Greenhouse-Geisser	8,431	2,130	3,959	22,143	,000	,059
Hemisphere * Area	Sphericity Assumed	,401	6	,067	2,816	,010	,008
	Greenhouse-Geisser	,401	3,819	,105	2,816	,026	,008
Condition_ * Hemisphere * Area	Sphericity Assumed	,187	6	,031	1,780	,099	,005
	Greenhouse-Geisser	,187	4,207	,044	1,780	,127	,005

Table 21 Tests of Within-Subjects Effects (all population 2x2x7)

As expected, analysing the variance in the interaction Area*Condition*Hemisphere, the standard deviation associated with the means is much higher than the one calculated for the whole population as in the previous configurations analysed. The respective means of the standard deviations are equal to 0,43 and 0,19.

Condition_ * Hemisphere * Area (means 2x2x7)

Condition_			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Common	Left	AF	,264	,061	,140	,388
		F	,069	,035	-,001	,140
		FC	,063	,032	-,001	,128
		CT	-,030	,039	-,109	,049
		CP	-,080	,037	-,156	-,004
		PT	-,145	,049	-,246	-,045
		PO	-,160	,058	-,277	-,043
	Right	AF	,268	,054	,158	,378
		F	,093	,041	,010	,176
		FC	,010	,040	-,072	,091
		CT	-,070	,040	-,151	,011
		CP	-,116	,038	-,194	-,039
		PT	-,144	,047	-,239	-,049
		PO	-,168	,055	-,281	-,055
Uncommon	Left	AF	,181	,044	,091	,270
		F	,055	,034	-,015	,125
		FC	,027	,036	-,045	,099
		CT	,001	,033	-,067	,069
		CP	,010	,040	-,072	,092
		PT	-,012	,043	-,100	,075
		PO	-,057	,047	-,151	,038
	Right	AF	,221	,055	,108	,333
		F	,092	,040	,011	,173
		FC	,044	,037	-,032	,120
		CT	,035	,041	-,048	,118
		CP	-,008	,036	-,080	,065
		PT	-,028	,042	-,112	,057
		PO	-,073	,051	-,176	,031

*Table 22 Condition_ * Hemisphere * Area (means 2x2x7)*

Condition_ * Hemisphere * Area (all population 2x2x7)

Condition_			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Common	Left	AF	,281	,024	,235	,328
		F	,104	,017	,070	,139
		FC	,094	,017	,060	,128
		CT	,027	,018	-,010	,063
		CP	-,029	,018	-,064	,006
		PT	-,077	,021	-,118	-,036
		PO	-,056	,022	-,100	-,012
	Right	AF	,262	,023	,217	,307
		F	,107	,020	,068	,146
		FC	,045	,019	,007	,083
		CT	-,023	,019	-,061	,015
		CP	-,050	,019	-,086	-,013
		PT	-,072	,021	-,114	-,030
		PO	-,079	,024	-,126	-,033
Uncommon	Left	AF	,187	,020	,148	,226
		F	,079	,017	,046	,112
		FC	,050	,016	,017	,082
		CT	,027	,016	-,004	,059
		CP	,028	,018	-,007	,062
		PT	,008	,019	-,031	,046
		PO	,004	,021	-,038	,045
	Right	AF	,195	,022	,151	,239
		F	,079	,019	,043	,115
		FC	,040	,016	,009	,071
		CT	,029	,016	-,002	,061
		CP	,016	,017	-,016	,049
		PT	,023	,019	-,015	,061
		PO	,008	,022	-,035	,052

*Table 23 Condition_ * Hemisphere * Area (all population 2x2x7)*

4 Results

Within the following chapter, results obtained through the statistical analysis have been displayed and commented.
In the first part, a brief literature review has been conducted in order to allow the comparison of the results with the literature.
Subsequently, a more in-depth analysis has been executed on the 6-areas configuration chosen for the analysis, considering as between-subject factors the Background, the Gender and the Degree level of the participants to the experiment.

4.1 Literature comparison

In the creative ideation literature, several contrasting results have been obtained, some of which identify an increase of alpha event-related synchronization (ERS) and others an increase of alpha event-related desynchronization (ERD). Following the literature review conducted by Colombo, most of the studies adopting the AUT approach obtained an increase in alpha event-related synchronization during the divergent thinking task in different areas of right hemisphere of the brain (Colombo, 2019). Some examples of such experiments are presented below.
The study conducted by Fink et al. (Fink & Neubaer, 2006) was aimed at identifying differences in divergent thinking before and after two weeks of creativity training. In both sessions, the experiment showed an alpha synchronization in both the frontal and the posterior region in the divergent thinking task condition. An increase in alpha power has been recorded in the centro-temporo-parietal area in the interaction Area*Hemisphere. After the training, participants showed a stronger alpha ERS in the frontal region for both tasks.

Fink and Neubauer (2008) used the same methodology integrated with other tests to analyse the relationship existing between extraversion and originality. In the experiment, researchers observed higher TRPs from anterior to posterior for both tasks. It was then observed an increase in alpha synchronization in the centro-temporo-parietal area in the right hemisphere for the upper alpha band.

Fink et al. (2009) compared four ideation creativity methods, among which the AUT experiment, using neuroimaging techniques. The researchers observed a general synchronization from anteriofrontal to centrotemporal areas. A slight alpha desynchronization was observed in centroparietal, parietotemporal and parietooccipital regions. The increment of alpha ERS was more evident in the right hemisphere, especially during the AU task. In the interaction Task*Area*Hemisphere, a robust alpha synchronization was recorded in the posterior region of the right hemisphere of the brain.

The experiment allowing a greater degree of comparability with the present research is the one proposed by Jauk et al. (2012), adopted as primary reference for the design of the experiment by Colombo. EEG data have been analysed adopting an ANOVA with within-subject factor Condition, Hemisphere and Area and between-subjects factor Group (creative subjects vs non-creative subjects). The researchers observed a stepwise decrease in TRPs from the anteriofrontal to the posterior Areas. The factor Hemisphere generally showed a decrease in the alpha power in the left hemisphere. According to this, a stronger desynchronization was found in the interaction between Area*Hemisphere in the left hemisphere from the centrotemporal region to the parietooccipital. A significant main effect was observed for the factor Condition, reflecting that a slighter desynchronization has been identified concerning uncommon uses in comparison to the common task. The interaction Condition*Area highlighted that the just mentioned effect differs in the latter factor, in fact no effect has been identified from anteriofrontal to centroparietal areas in the uncommon condition, whereas in the common condition alpha power significantly declined from anterior to posterior regions. The between-subjects factor Group did not show significance. The interaction Condition*Area*Hemisphere*Group revealed significant condition effects, except for the anteriorfrontal region in the lower creative group. The strongest alpha synchronisation has been identified in high creative individuals in the frontal region. On the opposite, lower creative individuals showed the strongest synchronization in right hemisphere posterior area (Jauk, Benedek, & Neubauer, 2012).

4.2 Design of experiment chosen

In order to proceed with the analysis, it had been fundamental to choose a configuration describing the disposition of the electrodes optimally in the EEG helmet.

The analysis computed in the previous chapter did not bring any significant pieces of evidence of the advantages obtainable using a configuration rather than another, apart from the decrease of variances in the results considering the whole populations instead of the means. For this reason, the configuration including 6 areas per hemisphere has been chosen arbitrarily, based on the electrodes disposition of the EEG helmet adopted.

The choice to use the configuration with 7 areas per hemisphere has been excluded because the anterio-frontal and parieto-occipital areas relied on one single electrode each, making the analysis unreliable. The configuration including 7 areas was highly adopted in the literature because of the availability of different EEG caps with different electrodes locations (Jauk et al., 2012; Fink et al.; 2009; Fink et al.; 2009).

At the same way, the 3 areas configuration has been excluded for the imbalance presented between the frontal and posterior regions, measured through 3 electrodes each, and the central regions, measured through 8 electrodes.

4.3 "2x2x6" results in detail

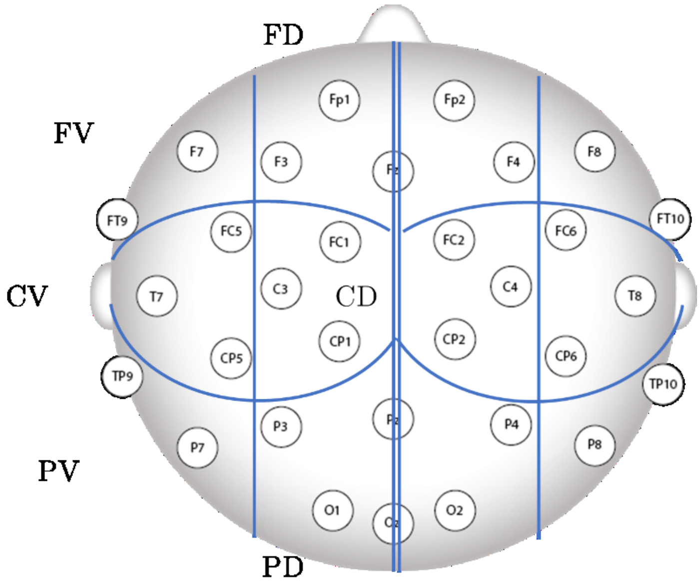

Figure 4.1 6-areas configuration

An analysis of variance (ANOVA) has been computed adopting as within-subject factors Condition (common vs uncommon), Hemisphere side (left vs right) and Area (FV, FD, CV, CD, PV, PD). The results have been fully displayed in the appendix (A.6).

The analysis on the whole population showed main effects for:

- The factor Area ($F[5,1765]= 78.452$, $p< 0.001$, $\eta^2_{partial}= 0.182$), showing a stepwise decrease of alpha synchronization from the anterior region to the posterior one.

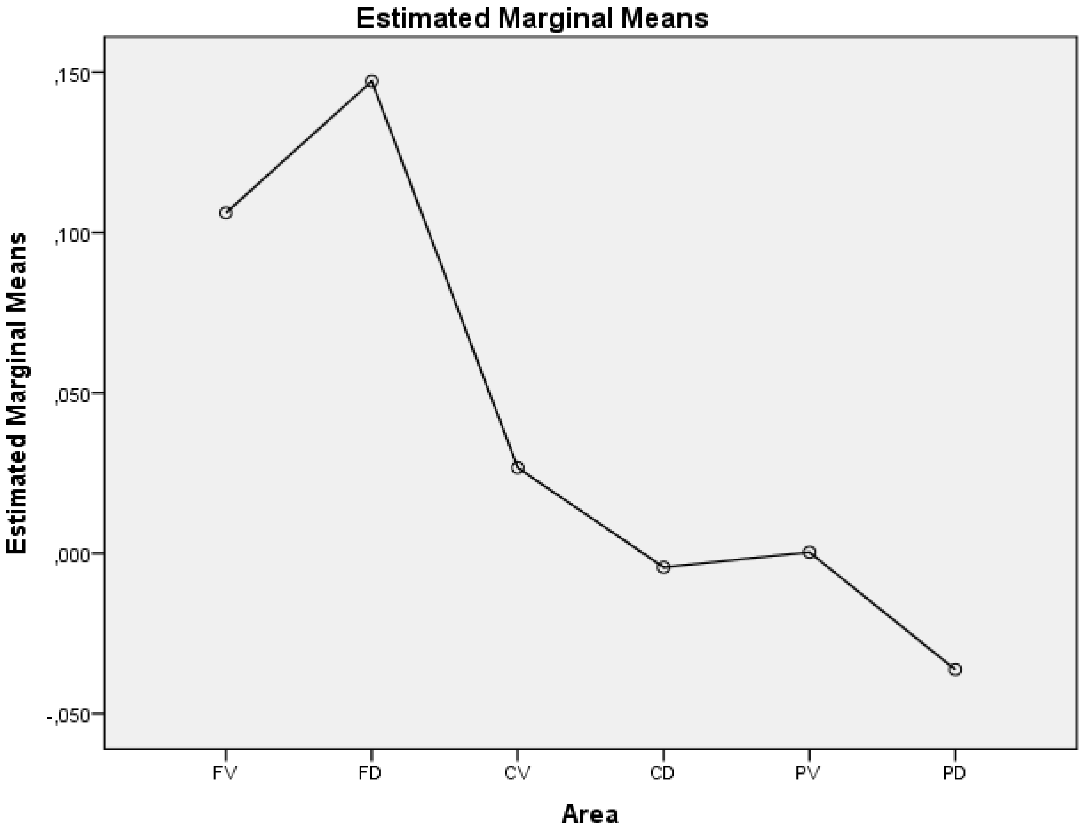

Figure 4.2 Estimated marginal means – Area

The pairwise comparisons showed significance for all the mean differences considered, except for the one calculated between the regions centraldorsal and posterior ventricular, and central dorsal and posterior dorsal.
Frontal dorsal alpha TRPs means reaches almost 0.150 μV, while the area located between the central and posterior regions fluctuate around the null value, reaching its minimum in the posterior dorsal region at -0.36 μV.

- The interaction Area*Condition ($F[5,1765]= 19.588$, $p< 0.001$, $\eta^2_{partial}= 0.053$), highlighting differences between the common and uncommon tasks.

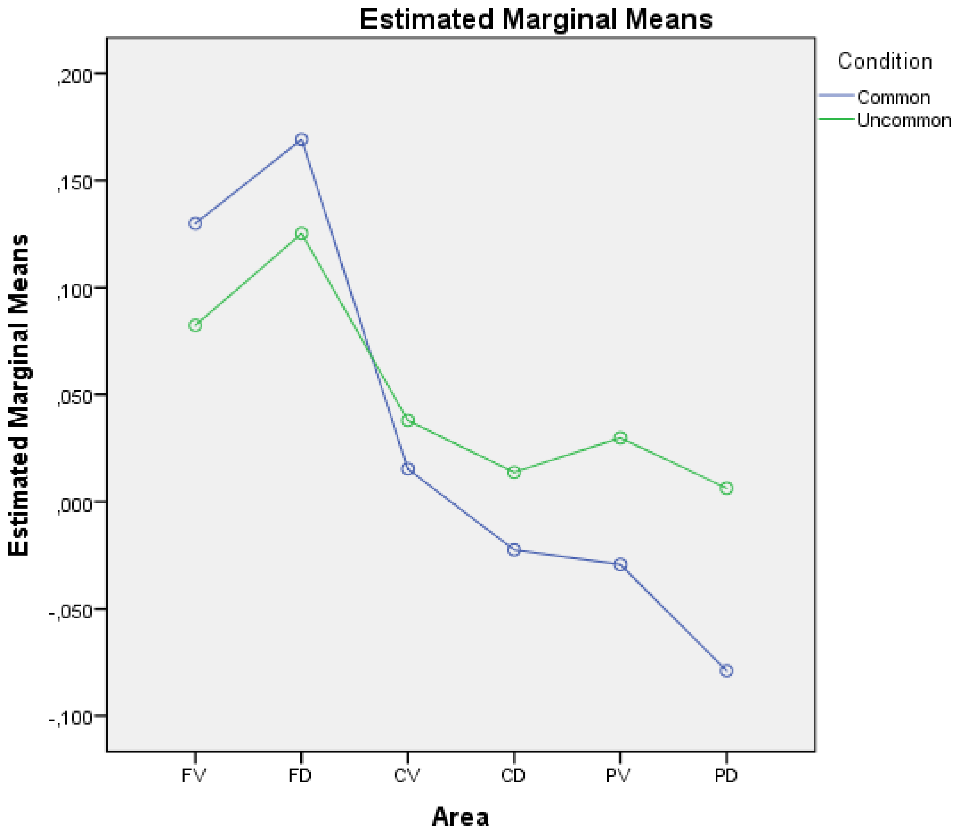

*Figure 4.3 Estimated marginal means - Condition*Area*

The means of alpha TRPs calculated from the common answers have higher values in the frontal regions of the brain (frontal dorsal common 0.169 μV vs uncommon 0.125 μV), while lower in the posterior region of the brain (posterior dorsal common -0.79 μV vs uncommon 0.06 μV) in comparison to the uncommon task results.

Besides, it is interesting to note that in the uncommon condition, it is possible to identify a global synchronization of the brain, always maintaining positive values, while in the common task condition, negative values are obtained in the centraldorsal and posterior regions.

- The interaction Area*Hemisphere ($F[5,1765]= 5.855$, $p< 0.001$, $\eta^2_{partial}= 0.016$).

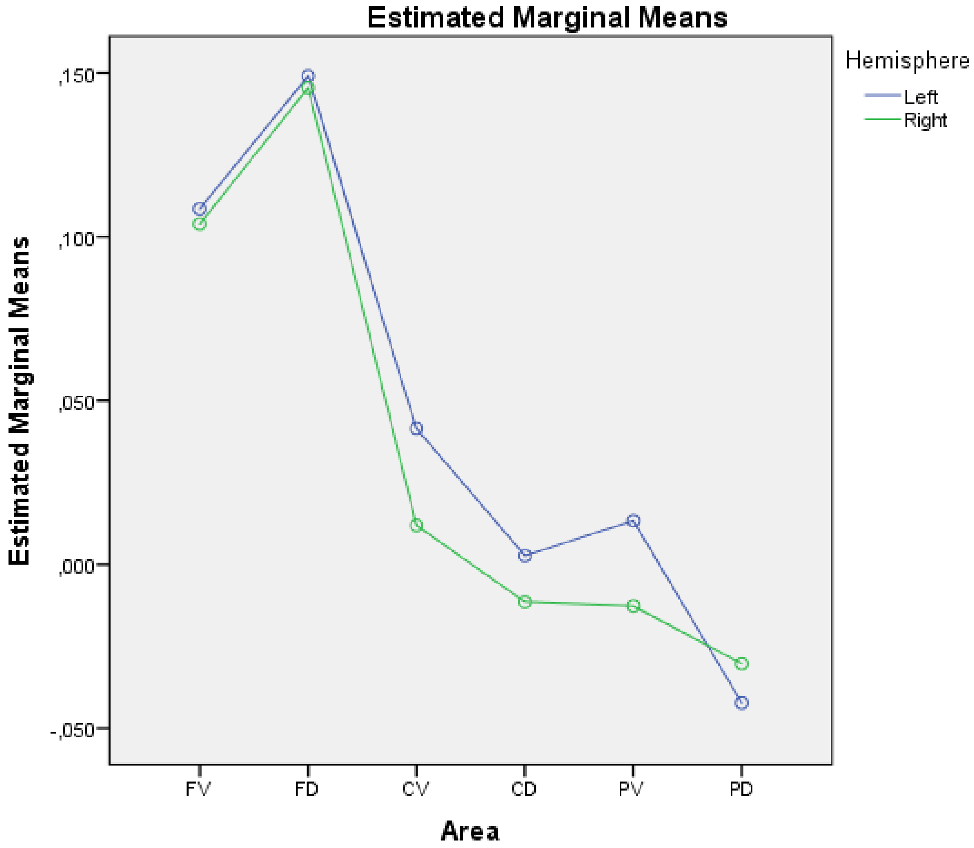

*Figure 4.4 Estimated marginal means - Hemisphere*Area*

Frontal areas had similar activation power in both hemispheres (left 0.149 µV vs right 0.145 µV). In the other regions of the brain, the left hemisphere reaches higher activation, except for the posterior dorsal area in which the right hemisphere reached higher values (left -0.42 µV vs right -0.30 µV).

The results obtained are mainly in line with the literature, showing (i) higher values of alpha TRPs in the Frontal region in comparison to the Central and Posterior ones (Fink & Neubauer, 2008; Fink et al., 2009), (ii) a monotonic decrease of alpha TRPs from the frontal to the posterior areas (Fink et al., 2009), (iii) higher values in the common condition in comparison to the uncommon one in the frontal areas and the opposite in the posterior ones (Fink & Neubauer, 2006), in which was recorded the higher cortical activation differences on conditions. These findings are in line with most of the literature and in particular with the experiment developed by Jauk et al. (2012).

In the present analysis, no main effects could be observed in the interaction Condition*Area*Hemisphere, that have been analysed more in detail in the following sections in the analysis, including between-subjects factors.

4.4 Between-subjects factor "Background"

Some assumptions were required in order to compute the statistical analysis considering as between-subject factor the Background of participants. First of all, subjects presenting left-handedness and the ones associated with poor quality of data have been neglected. Secondly, students with backgrounds differing from Design and Engineering have been removed from the dataset. Two clusters have been created with the remaining 29 participants:

1. Engineering, composed of:
    a. Materials Engineering
    b. Automation Engineering
    c. Advanced Materials Engineering
    d. Physics and Electrical Engineering
    e. Maintenance Engineering
    f. Mechanical Engineering
    g. Industrial economy engineering
    h. Engineering Physics
    i. Material Science and Engineering
    j. Chemistry Engineering
2) Design, composed of:
    a. Industrial Design Engineering
    b. Technical Design
    c. Industrial Engineering

The 2 classifications presented respectively 15 and 14 subjects, allowing a statistical comparison within the population. A repeated-measures ANOVA has been executed considering as within-subject factors Area, Hemisphere side and Condition, and as between-subjects factor Background. The results of the analysis are displayed in detail in the appendix (A.7). Concerning the sample, the analysis showed levels of significance for:

- The interaction Background*Area ($F[5,1470]= 18.776$, $p< 0.001$, $\eta^2_{partial}= 0.060$).

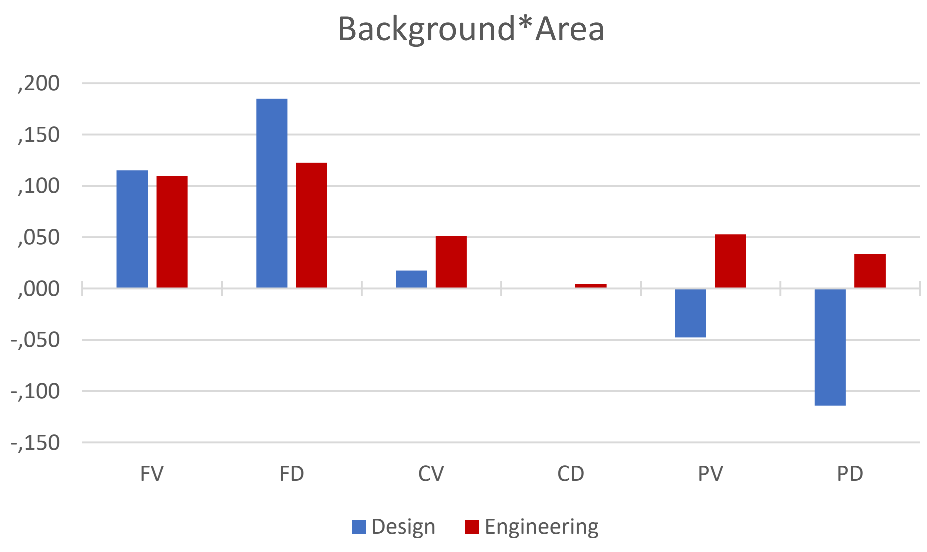

*Figure 4.5 Estimated marginal means - Background*Area*

Results obtained showed higher alpha TRPs values in the frontal regions, while lower in the posterior areas. This difference is more pronounced in the designers. Besides, it is interesting to note that engineers recorded a global synchronization of the whole brain, always maintaining positive values in both tasks, with a minimum in the central dorsal area.

- The interaction Background*Hemisphere ($F[1,294]= 14.956$, $p< 0.001$, $\eta^2_{partial}= 0.048$).

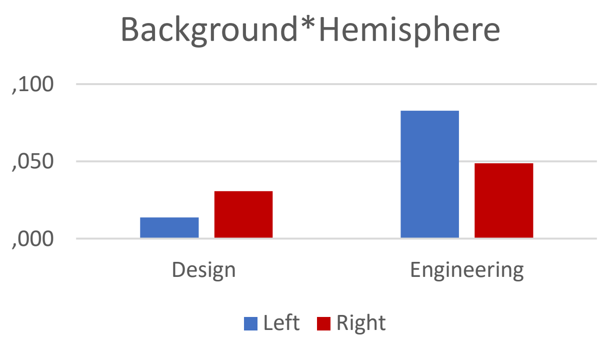

*Figure 4.6 Estimated marginal means - Background*Hemisphere*

In the sample, it is possible to identify a substantial difference in the hemisphere activations between designers and engineers: the former have shown a higher alpha synchronization in the right hemisphere, while the latter a higher desynchronization in the left hemisphere.

- The interaction Background*Hemisphere*Area ($F[5,1470]= 7.266$, $p< 0.001$, $\eta^2_{partial}= 0.024$).

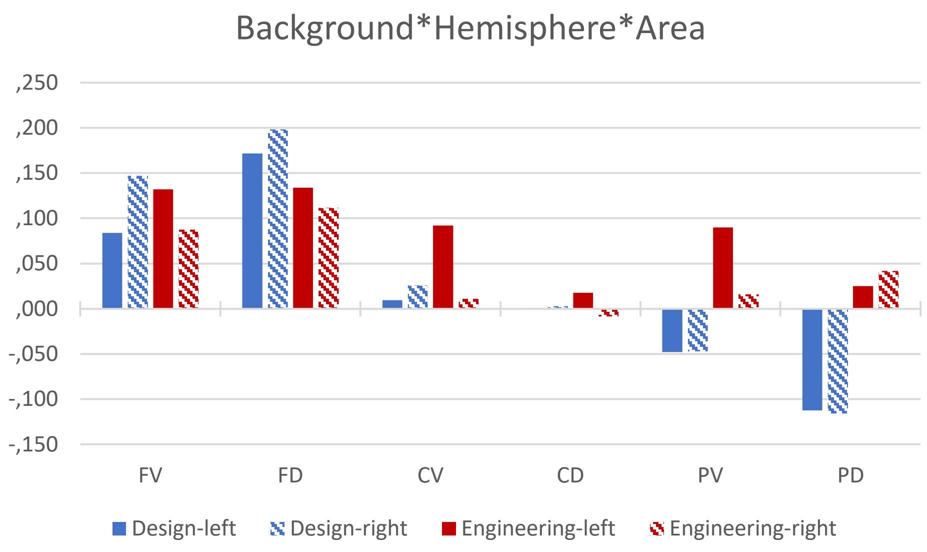

*Figure 4.7 Estimated marginal means - Background*Hemisphere*Area*

Engineers alpha TRPs show a higher synchronization of the frontal and the lateral areas of the left hemisphere, while values close to zero in the dorsal central and posterior regions. On the flip side, Designers recorded higher ERS activation in the right frontal area and a significant deactivation in the occipital region of the brain for both hemispheres.

Regarding the interactions including the factor condition, subject of interest of the present research, main effects have been identified concerning:

- The interaction Background*Condition ($F[1,294]= 11.591$, $p< 0.005$, $\eta^2_{partial}= 0.038$).

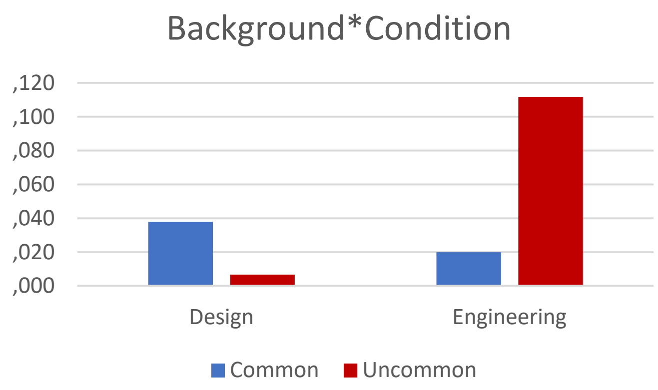

*Figure 4.8 Estimated marginal means - Background*Condition*

The sample shows a substantial difference in the cortical activation pattern of designers and engineers. The former has shown a higher alpha synchronization in the common condition, while the latter a higher ERS in the uncommon task.

- The interaction Background*Condition*Hemisphere*Area ($F[5,1470]= 1.902$, $p< 0.1$, $\eta^2_{partial}= 0.006$) did not show significant results, but are reported for completeness.

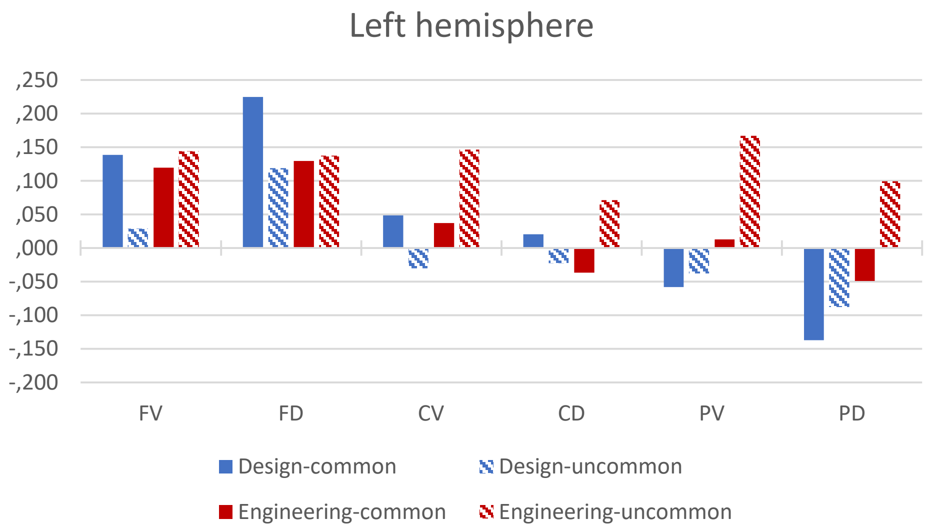

*Figure 4.9 Estimated marginal means - Background*Hemisphere*Condition*Area (left hemisphere)*

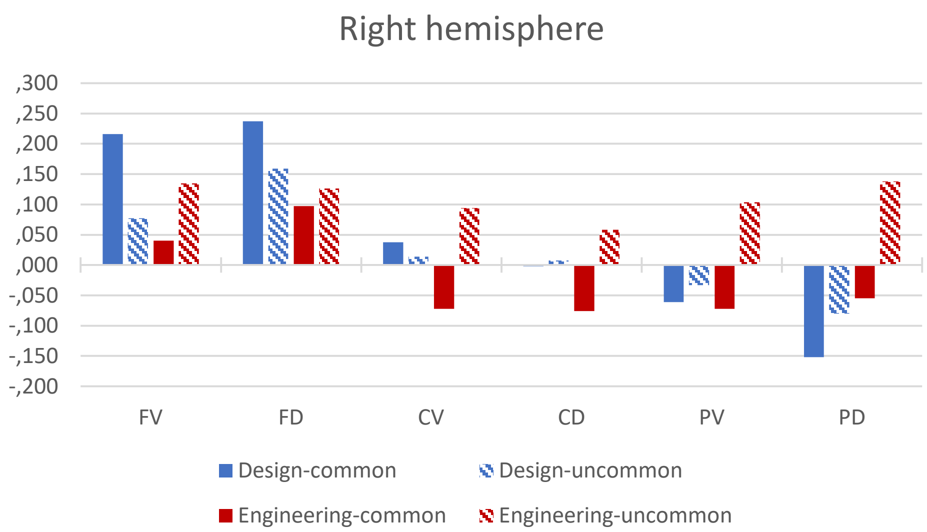

*Figure 4.10 Estimated marginal means - Background*Hemisphere*Condition*Area (right hemisphere)*

It is possible to identify significant differences concerning the cortical activation patterns for engineers and designers. Designers alpha TRPs show a high activation of the frontal area regarding both hemispheres, in particular, right side in the common task. The central and posterior regions underline a lower cortical desynchronization in the uncommon task for the right hemisphere in comparison to the left one. On the other side, engineers alpha TRPs show higher activation of the left hemisphere in all the brain regions in the uncommon task, reaching the peak in the posterior ventricular area. A slight desynchronization is visible in the central and posterior regions in the common condition.

In order to read the following presented results, it is essential to state that they could have been influenced by some biases related to the sample (e.g. all the designers present in the sample are from Sweeden) and they have been based on the strong assumption that the uncommon task produced exclusively creative responses. The latter will be ease in further studies when the judges' assessment is available.

In the literature, it is possible to identify several articles concerning the creative ideation processes of designers from a cognitive perspective, but no comparison has been executed between engineers and designers yet. Nevertheless, following the results, it is possible to recognise typical patterns related to creative people for designers and to logic reasoning for engineers.

The results are mostly in line with the literature showing for designers a slightly higher alpha synchronization in the centro-parieto-occipital area of the right hemisphere in the uncommon task (Fink & Benedek, 2014; Fink et al., 2009; Jauk et al., 2012) and higher alpha TRPs in the frontal area of the brain. On the other side, robust difference between convergent and divergent cognitive processing was observed in posterior rather than in frontal brain regions of engineers. Furthermore, they showed alpha synchronization in centrotemporal to parietooccipital regions in the uncommon production, while a lower synchronization during the common responses, patterns recognised as being typical of low creative individuals (Jauk et al., 2012). From the literature, it is possible to infer that designers registered active involvement of prefrontal structures which has been related to internally focused attention, while engineers have pursued less top-down strategies involving parietal regions that have been associated with loose semantic memory processing (Jauk et al., 2012). Nevertheless, engineers showed a higher synchronization in the left temporal lobe, historically recognised as controlling the comprehension, naming, and verbal memory.

Besides, the interaction Background*Condition highlighted significant differences in the general activation of the brain, in fact, engineers obtained higher TRPs values in the uncommon task in comparison to the common one, while the opposite pattern was identified in designers. It is possible to infer that this gap could be associated to the different education of subjects, permitting to designers to executing divergent tasks easier (lower alpha activation) than the convergent one, while engineers strive more in order to find creative ideas. Furthermore, overall higher alpha TRPs values recorded by engineers are associate in the experiment of Jauk to lower creative individuals.

4.5 Between-subjects factor "Gender"

In order to compute the statistical analysis considering the Gender of participants as a between-subjects factor, some assumptions were required. Similar to the case of the Background, subjects presenting left-handedness and the ones associated with poor quality of data have been neglected. The remaining subjects have been screened in order to obtain a sample composed by the same number of females and males, coupling subjects based on their:

- Background
- Degree level
- Origin and ages

Further details are presented in the appendix (A.8). The resulting sample was composed of 18 subjects, 9 males and 9 females. A repeated-measures ANOVA has been executed considering as within-subject factors Area, Hemisphere side and Condition, and as between-subjects factor Gender. The results of the analysis are displayed in detail in the appendix (A.8). Concerning the sample, the analysis showed levels of significance for:

- The interaction Gender*Hemisphere ($F[1,191]= 4.361$, $p< 0.05$, $\eta^2_{partial}= 0.022$).

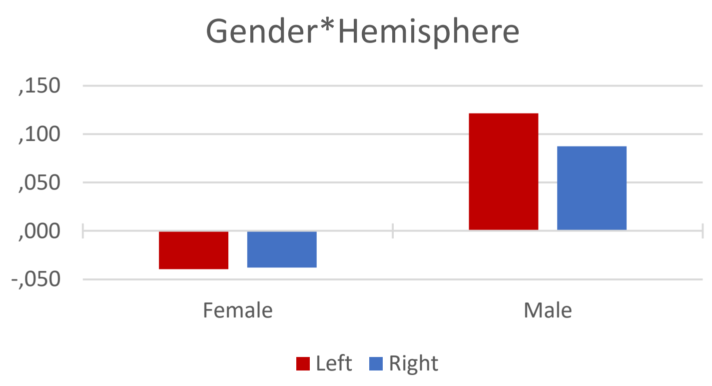

*Figure 4.11 Estimated marginal means - Gender*Hemisphere*

In the sample, it is possible to identify a substantial difference in the hemisphere activation between Females and Males: the former has shown an alpha desynchronization in both hemispheres, while the latter a general synchronization with a peak in the left hemisphere.

- The interaction Gender*Hemisphere*Area ($F[5,955]= 3.666$, $p< 0.01$, $\eta^2_{partial}= 0.019$).

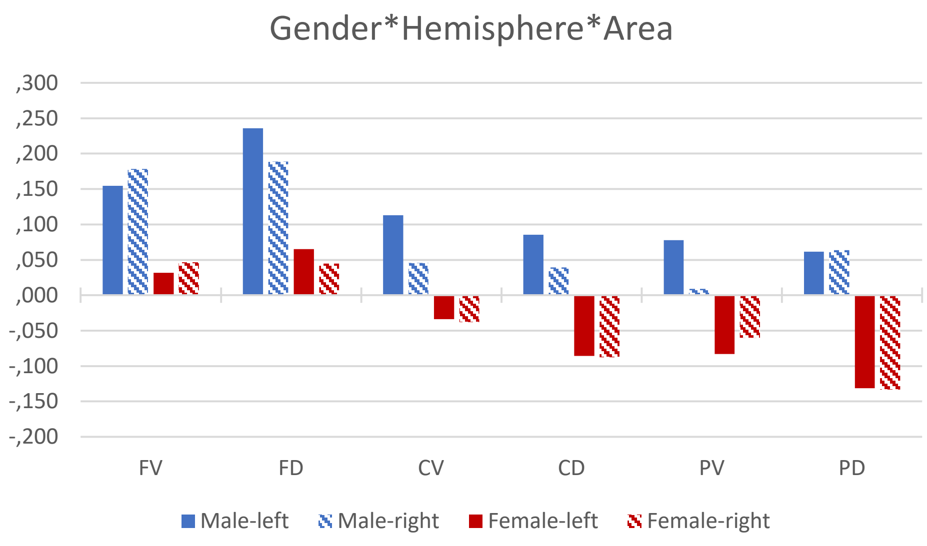

*Figure 4.12 Estimated marginal means - Gender*Hemisphere*Area*

The alpha TRPs values obtained suggested a similar pattern in both genders, characterised by a constant gap within the areas. It is possible to identify a general synchronization in the male brain, while the female one is characterised by synchronization in the frontal areas and by a general desynchronization in the central and posterior regions. Furthermore, a higher synchronization is present in the left hemisphere in comparison to the right one for the male in the frontal dorsal, central and posterior ventricular areas, while no clear differences are visible in the female brain.

Concerning the factor condition, subject of interest of the present research, main effects have been identified concerning the interaction involving all factors, that is Gender*Hemisphere*Condition*Area ($F[5,955]= 9.651$, $p< 0.01$, $\eta^2_{partial}= 0.048$).

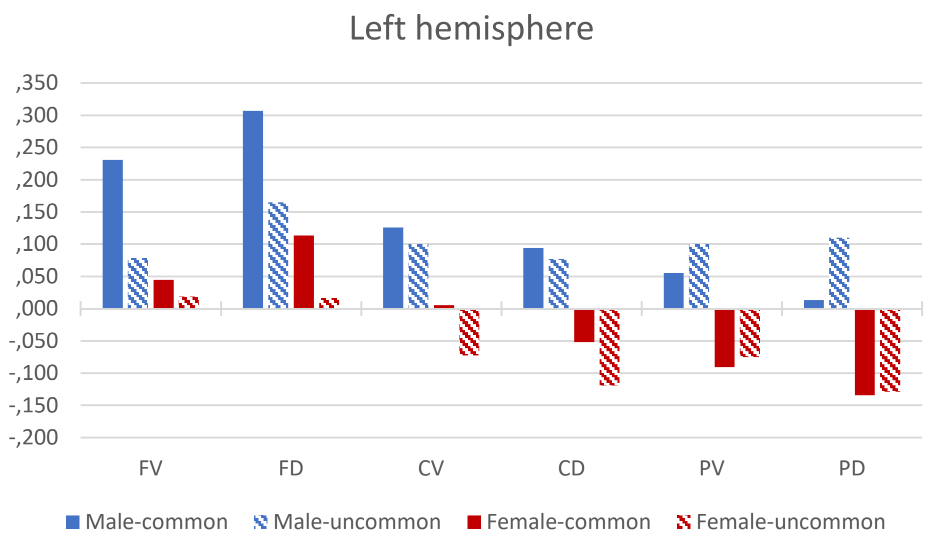

*Figure 4.13 Estimated marginal means - Gender*Hemisphere*Condition*Area (left hemisphere)*

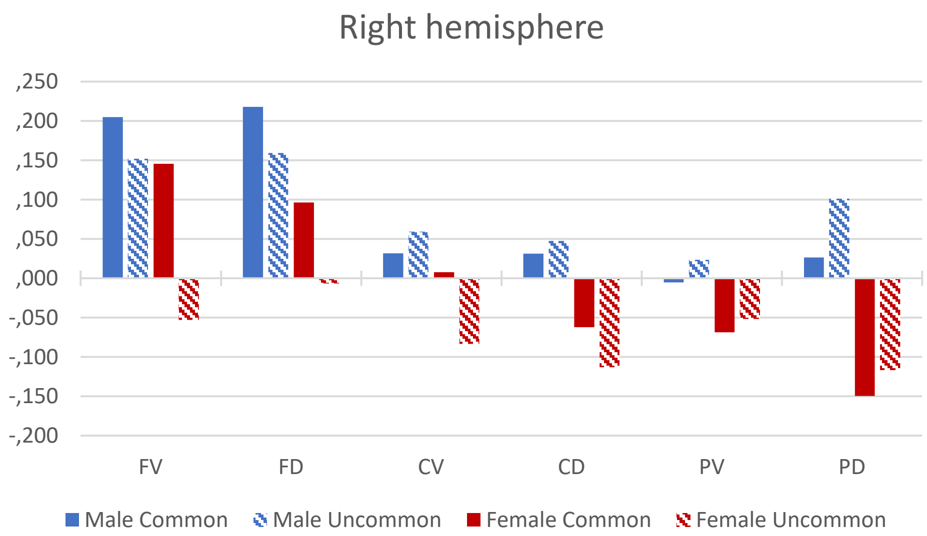

*Figure 4.14 Estimated marginal means - Gender*Hemisphere*Condition*Area (right hemisphere)*

The Highest cortical synchronization has been identified in both genders in the frontal area in the common task: males registered higher alpha TRPs values on the left hemisphere, while females in the frontal areas of the right one. Furthermore, it is possible to identify a constant gender gap for both hemispheres and conditions in all the areas analysed, except for the posterior dorsal one in which the gender cortical activation gap increase.

In the literature, several contrasting results have been obtained concerning the ideation creativity in gender analysis; thus it is not clear which are the causes of the differences identified in the statistical analysis of the present research. Fink and Neubauer (2006) and

Razumikova (2004) studies obtained results in line with the present research, recording higher ERS in the male sample, respectively on the right hemisphere and in the upper alpha band. On the other side, Ramos-Loyo and Guevara (1993) recorded an overall higher alpha synchronization in female, adopting an experiment concerning the differences in general cognitive activity between genders.

From the literature, it is evident that females and males adopt different strategies while working on divergent thinking tasks (Yoruk & Runco, 2014; Janckle, 2018), while there is not a recognizable gap considering the behavioural output perspective (McCarthy et al., 2012). However, it is not possible to identify specifics regions or event-related differences for either gender (Runco & Jager, 2012).

The strategies adopted by participants could thus play a primary role in gender cortical activation differences. After the analysed experiment, brief questionnaires were administrated to subjects in order to capture more data (further details in Appendix A.8). Among the questions, it was required to participants to provide a brief explanation regarding the strategy adopted in order to produce responses for the uncommon task. The sample analysed have shown significative differences among genders; the 9 males of the sample admitted having started their generation process from the common use of the object or to have relied their strategy on their own experience, trying to remember unusual uses of the object that they had seen in the past. On the other side, 7 out of 9 females in the sample stated having relied their generation process on different strategies (e.g. starting from the shape of the object in order to find alternative solutions). The sample does not allow a clear interpretation of the results due to its small size. In other words, it is not clear if the results should be correlated to gender differences, to strategy discrepancies, to both or if it was just a bias due to the small population analysed (9 male vs 9 females). It will be fundamental to integrate the analysis with the evaluation proposed by judges in order to see if the gap associated with the cortical activation could have some consequences on the behavioural outcomes obtained.

4.6 Between-subjects factor “Degree level”

In order to compute the statistical analysis considering the Degree level of participants as a between-subjects factor, some modifications to the dataset have been required. Similar to the previous cases, subjects presenting left-handedness and the ones associated with poor quality of data have been neglected. The remaining subjects have been analysed and screened in order to obtain a sample composed of the same number of females and males, with similar personal data. In order to do so, participants were selected, satisfying the following conditions:

- They were 25 years old or younger.
- They were natives of Italy, Sweeden, Spain or France.

Further details are presented in the appendix (A.9). The resulting sample was composed of 21 subjects, 10 attending the bachelor and 11 the master courses. A repeated-measures ANOVA has been executed considering as within-subject factors Area, Hemisphere side and Condition, and as a between-subjects factor, the Degree level attended. The results of the analysis are displayed in detail in the appendix (A.9).
The statistical analysis including the Degree level brought significant results only concerning the interaction Degreelevel*Hemisphere*Area $F[5,1075]= 10.920$, $p< 0.001$, $\eta^2_{partial}= 0.048$).

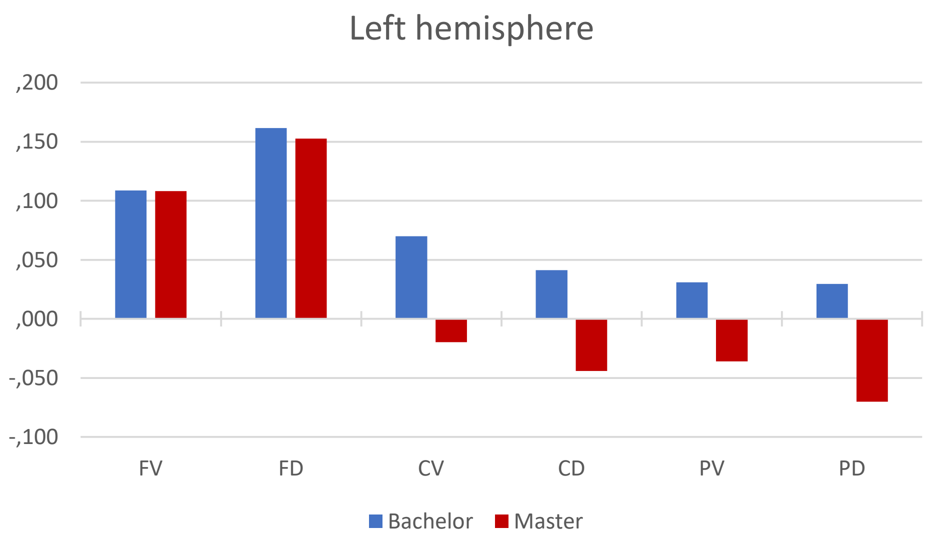

*Figure 4.15 Estimated marginal means - Degreelevel*Hemisphere*Area (left hemisphere)*

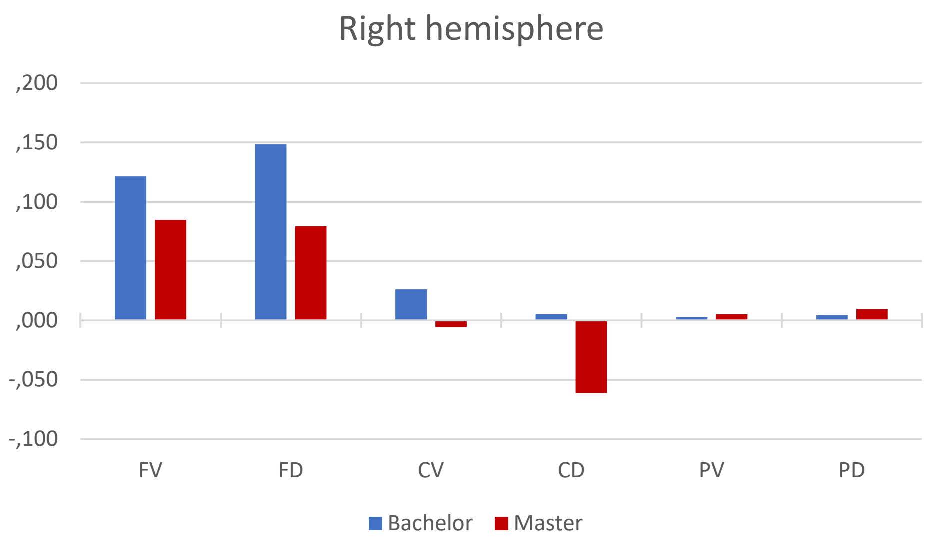

*Figure 4.16 Degreelevel*Hemisphere*Area (right hemisphere)*

Bachelor students alpha TRPs show a high synchronization in the frontal areas for both hemispheres, while a higher synchronization in the right hemisphere considering the central and the posterior regions. On the other side, Master students recorded an unbalance in the frontal region, showing higher alpha TRPs in the left hemisphere. In the central and posterior region, it is possible to identify a desynchronization of the left hemisphere, while values of alpha TRPs fluctuating around zero were obtained in the right hemisphere, apart from a desynchronization recorded in the central ventricular area.

No main effect has been identified considering the factor Condition, the primary focus of the present research. Furthermore, no literature has been identified analysing the attending degree level differences.

Conclusions

From the 1970s, there has been a growing tendency to study design processes to improve their outputs. In the beginning, researches were relying on protocol analysis, allowing recordings of designer's overt behaviours. This approach was lacking cognitive information, neglecting data regarding subconscious processes. In the last decades, new analysis has been possible thanks to new tools such as the Electroencephalogram and the Eye-tracker, allowing neurophysiologic investigations on design processes. This field is still at the beginning and its results highly contrasting due to the lack of a standard concerning the experimental design and the analysis of results. Therefore, the objective of the present experiment is to extend the literature of ideation process in designers through the electroencephalogram (EEG), to adopt and test a clear experiment pipeline and to provide the proper information required in order to replicate the experiment.

Pre data-gathering phases of the experiment have been designed in the Master thesis of Colombo (2019). Following the literature review, an EEG and an Eye-tracker, neglected for the present research, were adopted in order to collect neuro signals data. The experiment was a replication of the one proposed by Jauk et al. (2012) with some slight modifications validated by the neuroscience department of the University of Turin. Forty participants were involved in the experiment, and 40 everyday objects (stimuli) were presented randomly to each of them. The items were presented through slides divided by a two-minutes pause into two distinct blocks, one including the condition common and other uncommon. This expedient was taken in order to allow participants to maintain the same mindset without shifting from a creative paradigm to a non-creative one at each response. It was then required to each participant to display a single function for object according to the block condition, respectively a common solution in the common block or an alternative and useful one in the uncommon, within 30 seconds. The main differences with the experiment proposed by Jauk et al. (2012) were: the number of items presented to participants (40 vs 20), the expedient of the subdivision in two distinct blocks according to the condition of the experiment (not present in the experiment by Jauk et al.) and the participants personal data (university students vs male teenagers). Each stimulus was presented to participants through the adoption of 5 slides: (i) a blank screen for 5 seconds, (ii) a static cross in the middle of the screen for 5 seconds to have a reference period for the EEG recording, (iii) the stimulus word for half a second, (iv) a static cross in the middle of the screen allowing participants to think about the response and (v) a speech balloon representing the lapse of time were the participant were allowed to verbalize the idea. At the end of the experiment, a quick questionnaire was proposed to subjects in order to collect more data. The biomedical engineering department of the Polytechnic of Turin supported and validated the phases of pre-processing and post-processing of neuro-signals.

In order to allow the comparison in further studies between brain waves and verbal responses, in the present research, a methodology to evaluate the Idea Creativity of solutions has been designed, based firmly on the literature. It relies on the evaluation of 5 main criteria related both to the creativity and quality of an idea: (i) Originality, corresponding to the degree to which an idea is not only rare but is also ingenious, imaginative or surprising, (ii) Paradigm Relatedness, concerned to the extent to which an idea is paradigm preserving or paradigm modifying, (iii) Effectiveness, regarding the degree to which an idea will solve a problem, (iv) Feasibility, related to the extent to which an idea can be easily implemented and does not violate known constraints and (v) Elaboration, concerning the facility for adding a variety of details to information that has already been produced. Based on the literature review, it has been decided that the former four metrics will be evaluated by at least three judges through an organised Excel file, while the latter will be evaluated through an objective method relying on words count of responses. Detailed guidelines have been prepared for judges in order to reach high reliability in the evaluation and repeatability of the results. Besides, a further metric has been proposed, referred to as Variety and measuring the explored solution space during the idea generation process by participants. This metric will be calculated as a derivate of the judges' evaluations of Paradigm Relatedness in order to assess the appropriateness of the stimuli for the presented task.

The design of the experiment regarding the statistical analysis of the neurophysiologic data has been chosen based on the literature. The values obtained from the post-processing of the brain waves were the alpha TRPs. Higher alpha TRPs values are associated with a higher alpha activation in the region, theorised being typical of top-down processes. The presence of these waves seems to permit the region to inhibit the perception of external stimuli in favour of internal ones, such as the processes of idea generation involving attention and memory (Benedek, 2018). A strong assumption has been taken in order to compute the statistical analysis phase of the experiment, that is that the task condition (common or uncommon) was a reliable index of Idea Creativity of the solution identified by the participant. This assumption will be released in further studies when the ratings given by judges will be available. In order to compute the analysis, the software SPSS has been adopted. Following the literature, repeated measures ANOVAs have been computed considering as within-subject factors the Conditions associated with the answer of participants (Common vs Uncommon), the Hemisphere sides (Left vs Right) and the brain Areas (Frontal ventricular, Frontal dorsal, Central ventricular, Central dorsal, Posterior ventricular, Posterior dorsal). The Area subdivision just mentioned differed from the one utilised in the literature due to the adoption of 6 regions per hemisphere instead of 7. This choice was due to the different configuration of the electrodes present in the EEG cap in the data-gathering phase that was adopted in the present experiment. Further repeated measures ANOVAs have been computed in order to analyse the between-subjects factors Background (Engineering vs Design), Gender (Male vs Female) and Degree level (Bachelor vs Master) of participants.

The present thesis had not as objective the exhaustive disclosure of the results of the experiment due to the several analysis that could have been implemented on the data, but to provide partial results. The repeated measured ANOVA reported results in line with the literature indicating: (i) greater values of alpha TRPs in the Frontal region in comparison to the Central and Posterior ones, (ii) a monotonic decrease of alpha TRPs from the Frontal to the Posterior areas, and (iii) higher values in the common condition in comparison to the uncommon one in the Frontal areas and the opposite in the posterior ones, in which was recorded the higher cortical activation differences on conditions.

Among the between-subjects factors, the Background obtained more interesting results. Concerning the design ideation processes, the findings are mostly in line with the literature, while the comparison between Engineers and Designers has not been executed in past researches. From the results, it is possible to notice that engineers obtained higher TRPs values in the uncommon Condition in comparison to the common one, while the opposite pattern was identified in designers. It is possible to infer that this gap could be associated to the different education of subjects, permitting to designers to executing divergent tasks easier (lower alpha activation) than the convergent one, while engineers strive more in order to find creative ideas. Furthermore, a slight activation has been identified in the right hemisphere in comparison to the left one in the centro-parieto-occipital region of the designer brain in the uncommon condition, highly mentioned in the literature. Concerning the engineers, a higher activation has been identified in the left central ventricular area (possibly located into the left temporal lobe) that could be associated with prevalent semantic reasoning.

The between-subject factor Degree level did not obtain significant main effects, while the factor Gender did, synthesizable with the higher alpha activation in all the areas of male brains. In order to understand this result, it would have been essential to access the Idea Creativity data rated by judges in order to understand if the just mentioned difference is associable to an actual gap in the outcomes of the genders. In the literature, several contrasting results have been obtained regarding the ideation process in gender comparison (Razumnikova, 2004; Yoruk & Runco, 2014). Abraham (2015) stated that the difference encountered in the neurophysiologic activation is due to different strategies and cognitive styles adopted by male and females. According to this, the questionnaires administrated to subjects highlighted different strategies between males and females; all the males of the population stated that the majority of their uncommon responses were relying on the common use of the object or on their past personal experience, while 7 out of 9 females stated to have adopted different strategies (e.g. starting from the shape of the object in order to find alternative solutions). In conclusion, it is not clear if the results should be correlated to gender differences, to strategy discrepancies, to both or if it was just a bias due to the small population analysed (9 male vs 9 females).

In order to fully disclose the experiment, several further analysis are suggested considering: (i) the Idea Creativity evaluations obtained by judges, (ii) the within-subject factor fatigue, highlighting the performance differences of the participants in the beginning and in the end of the experiment, (iii) the subdivision of the uncommon responses in more intervals in order to underline the alternation of convergent and divergent thinking phases in the responses, (iv) the beta waves activation, especially to highlight gender differences and (v) the subdivision of alpha waves in upper and lower bands.

For further experiments, some expedients could be adopted in order to reach interesting results; following the literature (Hay et al., 2019), it could be interesting to strengthen the personal opinion of participants regarding the creativity of the idea proposed by themselves, being the ideation process mainly based the memory of the subject. It could be assessed through a questionnaire aimed at the auto-assessment of all the uncommon ideas generated during the experiment. In order to obtain reliability in the auto-assessment phase and to not interrupt the generation mindset of the subject, it could be needed to lower the number of items proposed. The just mentioned expedient could bring to the loss of within-subject factor correspondence (common vs uncommon), an issue that could be overcome involving more participants into the experiment.

Appendix

A.1 Personal data

Forty students, 11 females and 29 males, have been selected in order to take part in the experiment with a varied background in Engineering. The mean age was 23.67 years, with a standard deviation of 2.55. Further personal data are presented in the table below.

Subjects	Gender	Age	Nationality	Mother Tongue	Corse of study	Year of Study	Handedness	Poor Eyesight	Neurological Disorder
Subject 1	FEMALE	23	Italian	Italian	Materials Engineering	5th	RIGHT	NO	NO
Subject 2	MALE	25	Italian	Italian	Automation Engineering	5th	RIGHT	NO	NO
Subject 3	MALE	23	Italian	Italian	Material Science	5th	RIGHT	GLASSES	NO
Subject 4	MALE	23	Swedish	Swedish	Civil Technical Design	1st	RIGHT	NO	NO
Subject 5	MALE	26	Syrian	English	Advanced Materials Engineering	4th	RIGHT	GLASSES	NO
Subject 6	MALE	20	Swedish	Swedish	Physics and Electrical Engineering	1st	RIGHT	NO	NO
Subject 7	MALE	23	Portuguese	Portuguese	Maintenance Engineering	4th	RIGHT	GLASSES	NO
Subject 8	FEMALE	21	German	German	Business Engineering	4th	RIGHT	GLASSES	NO
Subject 9	MALE	24	Italian	Italian	Materials Engineering	5th	RIGHT	GLASSES	NO
Subject 10	MALE	24	Swedish	Swedish	Industrial Design Engineering	4th	RIGHT	GLASSES	NO
Subject 11	MALE	27	Swedish	Swedish	Industrial Design Engineering	3rd	RIGHT	LENS	NO
Subject 12	FEMALE	23	Swedish	Swedish	Industrial Design Engineering	3rd	RIGHT	LENS	NO
Subject 13	FEMALE	21	French	French	Mechanical Engineering	4th	RIGHT	GLASSES	NO
Subject 14	MALE	24	Swedish	Swedish	Industrial economy engineering	3rd	RIGHT	NO	NO
Subject 15	MALE	20	Swedish	Swedish	Engineering Physics	2nd	LEFT	NO	NO
Subject 16	MALE	28	Swedish	Swedish	Technical Design	3rd	RIGHT	NO	NO
Subject 17	MALE	26	Brazilian	Portuguese	Material Science and Engineering	4th	RIGHT	NO	NO

Subject 18	FEMALE	26	Spanish	Spanish	Advanced Materials Engineering	4th	RIGHT	NO	NO
Subject 19	MALE	24	Austrian	German	Chemistry Engineering	5th	RIGHT	NO	NO
Subject 20	MALE	24	German	German	Material Engineering	4th	RIGHT	NO	NO
Subject 21	MALE	22	Spanish	Spanish	Mechanical Engineering	4th	RIGHT	LENS	NO
Subject 22	MALE	20	Swedish	Swedish	Computer Science	1st	RIGHT	NO	NO
Subject 23	MALE	25	German	German	Mechanical Engineering	5th	RIGHT	GLASSES	NO
Subject 24	FEMALE	19	Swedish	Swedish	Civil Architecture	1st	RIGHT	LENS	NO
Subject 25	MALE	25	Swedish	Swedish	Industrial Design Engineering	3rd	RIGHT	GLASSES	NO
Subject 26	MALE	23	Swedish	Swedish	Industrial Design Engineering	3rd	RIGHT	NO	NO
Subject 27	MALE	23	Swedish	Swedish	Industrial Design Engineering	3rd	RIGHT	GLASSES	NO
Subject 28	MALE	21	French	French	Computer Science	4th	RIGHT	NO	NO
Subject 29	FEMALE	27	Swedish	Swedish	Industrial Engineering	5th	RIGHT	NO	NO
Subject 30	MALE	22	Swedish	Swedish	Mechanical Engineering	2nd	RIGHT	NO	NO
Subject 31	FEMALE	23	Swedish	Swedish	Technical Design	4th	RIGHT	LENS	NO
Subject 32	FEMALE	21	Swedish	Swedish	Industrial Design Engineering	3rd	RIGHT	NO	NO
Subject 33	MALE	27	Swedish	Swedish	Industrial Design Engineering	4th	RIGHT	GLASSES	NO
Subject 34	MALE	21	Swedish	Swedish	Technical Design	3rd	RIGHT	GLASSES	NO
Subject 35	MALE	32	Swedish	Swedish	Technical Design	4th	RIGHT	NO	NO
Subject 36	FEMALE	31	Swedish	Swedish	Industrial Design Engineering	5th	RIGHT	GLASSES	NO
Subject 37	MALE	22	Swedish	Swedish	Technical Design	3rd	LEFT	GLASSES	NO
Subject 38	FEMALE	19	Romanian	Romanian	Computer Science	2nd	LEFT	GLASSES	NO
Subject 39	MALE	25	Spanish	Spanish	Mechanical Engineering	5th	RIGHT	NO	NO
Subject 40	MALE	22	French	French	Material Science	5th	RIGHT	NO	NO

Table 24 Participants' data

A.2 Stimuli

Below it is represented the list of the stimuli showed to participants during the experiment. They have been translated into the mother tongue of the subjects in order to avoid the language barrier.

Axe	Ball	Basket	Bed	Book	Bra	Bread	Can
Coffin	Coin	Colander	Comb	Fork	Guitar	Gun	Hammer
Hanger	Helmet	Lamp	Magnifier	Mirror	Needle	Net	Paperclip
Pillow	Pot	Rag	Ring	Scissors	Shoe	Sock	Sponge
Stick	Tent	Toothpaste	Trousers	Tyre	Umbrella	Vase	Window

Table 25 Stimuli presented to participants

A.3 Idea assessment literature review

Below, a summary of the literature review containing further articles is presented.

Experiment	Parameter assessed	Number of raters	Scale	Level of direction in the creativity judgment	Kind of expetiment (AUT, artistic work)	Reliability (internal consistency>.70)	Citations
Christensen, P. R. et al. (1957)	Cleverness	2 judges	1-5	-----	Artistic work	Correlation between the 2 judges: .50	234
Harrington et al. (1975)	Creativity	3 advanced graduate in psychology	Scale 1-5	---	AUT	estimated internal reliability were .64 and .82	251
Evans & Forbach (1983)	Quality (worthwihile, infrequency)	2 judges	1-4 0-1	---	AUT	Intercorrelation judges: worthwhile .62 infrequency .80	16
Milgram, R. M. & Arad, R. (1981)	Unusual Quality (fit, cleverness)	3 judges	0-1 0-1 0-1	Trained	Creativity battery (comprending AUT)	Judges were in agreement on 94% of the responses	55
Zarnegar et al. (1988)	Originality	------------------	0-3	Trained on originality concept (clever and unusual)	Derived of AUT	Interjudges reliability 0,9	26
K.J. Gilhooly et al. (2010)	Novelty	1 judge 1 check judge (only 50 responses analysed)	1-7	------	Alternative Uses	Interjudge reliability .85	359
DM Harrington et al. (1983)	Repetitions, too vague, functionless, standard, non-intended uses, non-repeated uses	-----------------	Responses assigned to one of the categories	Teachers	AUT	Results concerning AUT and teacher as judges was low correlated	126
C Mouchiroud, T. Lubart (2001)	Originality	3	1-7 following a normal distribution	Students that have experience wit CAT	AUT	Cronbach's a > .80	96
Hocevar (1979)	Originality	4	0-1	Raters looked for uniques and clever responses	AUT	Interjudge reliability >.80	193
R. Radel et al. (2015)	Fluency Originality Flexibility Elaboration	2	How many responses 0-1 0-1-2 0-1-2	-------	AUT	Krippendorff's alphas: 1 for fluency 1 for originality 0.812 for flexibility 0.834 for elaboration	73
Ramesh Srivathsavai et al. (2010)	Originality Technical feasibility	2	Originality: 1-3, 1-4, 1-10 Technical feasibility: 1-3, 1-4	No definitions given Students of mechanical engineering course	Sketches and word evaluations	Low level of agreement between judge for the scale of 1-10 for originality, medium-high for 1-4 for both criteria and high for 1-3 scale for both the criteria	43
Daniel G. Johnson et al. (2014)	Originality Technical feasibility Innovation (table in backup)	2	1-5 (as proposed by Guilford) 1-4 Number of innovative charact eristics (Saunders, 2011)	2 of the researchers (with a solid knowledge abou t creativity)	Sketches and words evaluation	Weighted kappa 0,9 for originality 0,89 for feasibility 0,8 for innovation	32
I. Blohm et al (2010)	Novelty Relevance Elaboration Feasibility	7	1-7 for al criteria	Experts in the field (professors)	Idea competition	Cronbach's alpha Novelty = 0.956, Relevance = 0.841, Elaboration = 0.887, Feasibility = 0.769.	27

Table 26 further literature review illustration

A.4 Guidelines

Judges' guidelines

This document aims to define clear and precise guidelines in order to allow raters to evaluate, through different criteria, the functions proposed by participants during the experiment. The document is composed of the following sections:

- Measures of ideational Output in the literature (taken directly from the research by Dean 2006).
- Clarification of dimension and sub-dimensions (developed ad hoc for the specific criteria chosen for the assessment phase of this experiment).
- Procedure and anchor tables (the practical guidelines for the scoring phase).

The former 2 are preliminary sections, developed in order to permit judges to have common knowledge about the concept of ideation and the relative criteria. They should be read before the beginning of the evaluation phase. Anchor tables are practical tools that will guide judges through the actual scoring phase.

Measures of Ideational Output in the literature

Early idea-generation research used quantity as a measure of quality, assuming that if a sufficient number of ideas were produced, the resulting idea pool would be more likely to contain high-quality ideas (Osborn, 1953). This positive correlation has been confirmed in some studies (Diehl and Stroebe, 1987; Gallupe et al., 1992; Valacich et al., 1993), but other research has found that the correlation between quantity and quality is tenuous (MacCrimmon and Wagner, 1994), and still others have posited that there is, in fact, a negative correlation between quantity and quality (Graham, 1977; Connolly et al., 1990; Gryskiewicz, 1980). Studies that go beyond merely enumerating ideas require researchers to select a definition of one or more of the three constructs that are typically operationalized as the dependent variable(s): 1) idea quality, 2) idea novelty, which is sometimes referred to as rarity or unusualness, and 3) idea creativity. However, as will be shown later in this paper, the definitions of these three constructs vary considerably among researchers. Therefore, we will now provide a succinct definition of each of these three constructs before examining each one in more detail.

Definition of Idea Quality

We define a quality idea as one that contains the following three characteristics. First, the idea should apply to the problem at hand (Aiken et al., 1996). Second, it should be an effective solution (Valacich, et al., 1995; Kramer and Kuo, 1997). Third, it should be implementable (Diehle and Stroebe, 1987). Each of these attributes is examined in detail later in this paper. According to this definition of idea quality, an idea can be termed a quality idea without it being novel or unusual, which is consistent with conventional definitions of a quality idea. In short, a quality idea is an implementable solution that will solve the problem, regardless of whether or not the idea itself is novel or unusual.

Definition of Idea Novelty

We define a novel idea as one that is rare, unusual, or uncommon (Connolly, Routhieaux, and Schneider, 1993). The most novel idea, then, is an idea that is totally unique; conversely, the least novel idea is the most common one (MacCrimmon and Wagner, 1994). In application, the novelty of any idea must be judged in relation to how uncommon it is in the mind of the idea rater or how uncommon it is in the overall population of ideas.

Definition of Idea Creativity

To define idea creativity, it is helpful to first examine the concept of creativity itself and to differentiate it from idea creativity. Creativity is typically viewed as a characteristic of an environment, a process, a person, or a product (Rhodes, 1961). In terms of idea generation, environments, processes, persons, and groups that generate more novel ideas, or ideas that are not only novel but that also have other desirable attributes, are sometimes considered more creative than sources that produce fewer ideas with these qualities. Creativity can also be measured in terms of the characteristics of a product,

such as an idea. Measures that apply to ideas, the product view, are the focus of this paper.
We define a creative idea as a quality idea that is also novel. That is, it applies to the problem, is an effective and implementable solution, and is also novel (MacCrimmon and Wagner, 1994).

Dean (2006)

Clarification of Dimensions and Sub-Dimensions

"This section describes the approach used to create and refine the evaluation method. Two raters initially attempted to score a training sample of ideas based on MacCrimmon and Wagner's (1994) definitions of novelty, workability, relevance, and specificity but were unable to achieve good inter-rater reliability using these definitions. It became evident that raters held different assumptions about what each dimension meant and that the raters were considering different aspects of each dimension" (Dean, 2006).
Therefore, these criteria were decomposed by Dean (2006) in further subdimensions that, for the purpose of this specific experiment, were extremely detailed. In addition, he provided some practical tools, anchor tables, that helped raters during the assessment phase and permit to increase inter-judge reliability.
Starting from the literature review proposed by Dean, we developed a methodology relying on 2 macro aspects regarding idea creativity discussed before: Novelty, considering the creativity itself of the idea proposed, and Quality, regarding the implementability, utility, and elaboration of the concept (Shah 2000, 2003, Dean 2006, Chulvi 2012).

Novelty

Novelty is a key construct for measuring the creativity of ideas. We now examine two different novelty-related constructs: originality and paradigm relatedness.

Originality

"The degree to which the idea is not only rare, but is also ingenious, imaginative or surprising" (Dean, 2006).

This definition could be considered as an aggregation of 2 historical metrics:

- "Uncommonness" criterion - proposed in psychological studies, usually calculated through partially-objective methods and relying on the counting of the same idea proposed by different participants or on pre-set tables, explicating the functions ex-ante (Guilford 1967, Torrance 1974, Wallach 1965).
- "Originality" itself - the subjective assessment of originality, highlighting how an idea is perceived by judges.

The method proposed by Dean is the suitable compromise between the 2 views because it permits to overcome the flaws of the partially-objective methods, such as the consideration of merely bizarre ideas as creative, and allows to link these 2 criteria in a single metric.
This metric will be assessed by judges with a 1-4 Likert scale using a table displaying instruction taken directly from the research conducted by Dean.

Paradigm relatedness

"The degree to which an idea is paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational" (Dean, 2006).

This metric is the result of the decomposition of novelty into subdimension reflecting the transformational power of the idea. In literature, there are other examples of the adoption of this metric in order to assess the actual changing potential of an idea (Besemer 1981, Jackson, 1965 Besemer, 1987, Nagasundaram 1994, Gryskiewicz 1980).
The assessment will rely on the definition proposed by Dean and on the practical concept of paradigm from an architectural perspective. Concerning architecture, the following actions could be taken in order to fulfill new functions, as proposed in the model by Nagasundaram et al. (1994):

- Extension (Removal) – with the introduction (elimination) of new (old) components to the object.
- Redesign – with the alteration of the relationships between the elements composing the object.
- Transformational idea – in which both Extension and Redesign could be applied.

Nagasundaram proposed a hierarchical order for these changes based on the extent to which paradigm is shifted from the original object. The assortment starts with Extension changes, followed by Redesign and Transformational ones. This model is highly feasible with the view proposed by Henderson and Clarks (1990). In their view, changes could involve the reference technology and/or the relationship between components. In this optic, it is important to notice that Henderson and Clarks refer to architecture as "defined by the main components that make up the product and their mutual relationships". In this perspective, the additions (removals) of new (old) components that do not change existing relationships between product components are not considered as architectural changes. With the intersection of these 2 axes, 4 main kinds of innovations could be identified:

- Incremental innovation, in which "neither the underlying technology, nor product architecture will change" (Cantamessa and Montagna 2016). This kind of product changes will be taken as the "1" reference for the anchor table. It could be associated with the scenario in which no actual changes are involved in the original object architecture.
- Modular innovation, in which the "underlying technology does change in one or more functional elements, but product architecture does not" (Cantamessa and Montagna 2016). These innovations could be associated with the Extension (or Removal) principle proposed by Nagasundaram (1994) and will be associated with the "2" of the anchor table.
- Architectural innovation, in which the "underlying technology does not change, but intercomponent relationships do" (Cantamessa and Montagna 2016). These changes could be viewed as Redesign ones and will be associated with the "3" of the anchor table.
- Radical innovation, in which "both the underlying technology and product architecture change" (Cantamessa and Montagna 2016). Transformational changes will be associated with this kind of changes, and the "4" value will be associated with them.

The former 2 could be considered as Paradigm Preserving (PP) innovations associated usually with modular architectures, in which changes could be implemented easily,

while the latter ones as Paradigm Modifying (PM), in which changes require more effort to be implemented.
For the purpose of having more practical guidelines and of following the structure proposed by Dean, the terminologies proposed by Nagasundaram (1994) will be followed for the construction of the anchor table.
The assessment technique proposed, permit to use a 1-4 Likert scale, allowing judges to not change the scale used for other criteria.

Quality

“in short, a quality idea is an implementable solution that will solve the problem, regardless of whether or not the idea itself is novel or unusual” (Dean, 2006). We now examine two different quality-related constructs: Effectiveness and Feasibility. Specificity is a further subdimension proposed by Dean that will not be taken into consideration by judges, but will be assessed through a different procedure by researchers. Therefore, the analysis of this criterion will be omitted in the following section in order to avoid confusion.

Effectiveness

“The degree to which the idea will solve the problem” (Dean,2006).

This metric has been highly utilized in past researches (Eisenberger 2001, Kramer 1997, Barki 2001, Faure 2004, Valacich 1995), with a lot of inconsistencies regarding the terminology used to name it.
In the methodology developed by Dean, this metric was considered as a subdimension of Relevance, a dimension assessing the extent to which an idea was pertinent to the stated problem and how much effective the idea could have been to solve it. The simplification from the metric proposed by Dean was due to the fact that the only actual constraint that participants had during the experiment was to relate the object, proposed as a stimulus, with the response. For this reason, Effectiveness will be the only part of this criterion that will be assessed. It means that participants should evaluate only the extent to which the object could be effective to fulfill the function proposed by participants.

This criterion will be assessed by judges using a 1-4 Likert scale using a table displaying instruction taken directly from the research conducted by Dean.

Feasibility

"An idea is feasible if it can be easily implemented and does not violate known constraints" (Dean, 2006).

In literature, there are several studies focused on the assessment of this criterion (Briggs 1997, Diehele 1987, Gallupe 1992, Potter, 2004).
Dean considered a further subdivision concerning this metric, dividing it into acceptability and implementability, that, for the purpose of this experiment, is extremely detailed. For this reason, both the criteria could be assessed in a single metric at a higher level. The anchor table will be proposed as the union of the tables relative to the 2 subdimensions proposed by Dean.
This metric will be assessed by judges with a 1-4 Likert scale using a table displaying instruction derived from the research conducted by Dean.

Macro-criterion	Dimension	Definition
Novelty	Originality (1-4)	The degree to which the idea is not only rare, but is also ingenious, imaginative or surprising
	Paradigm Relatedness (1-4)	The degree to which an idea is paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational
Quality	Effectiveness (1-4)	The degree to which the idea will solve the problem without regard for feasibility
	Feasibility (1-4)	An idea is feasible if it can be easily implemented and does not violate known constraints

Procedure and Anchors tables

At the end of this chapter, judges are allowed to start the assessment phase through the use of the Excel file provided. Each sheet includes a single stimulus, and the ideas listed are the ones relative to the stimulus of the sheet. It is important that raters take into consideration the rules provided in this section of the document.
If judges encounter any doubts during the scoring phase concerning criteria, they are invited to read the former pages.

Regarding the Excel file:

All criteria will be rated on a 1-4 Likert scale to ease the process in the judges' perspective. For each row (concerning idea proposed), judges should put only one minuscule "x" for each metrics proposed (one for Originality, one for Paradigm Relatedness, one for Effectiveness and one for Feasibility) in the cell corresponding to the value that they consider the most appropriate one.

Below, the reader can find an example stating that the rater, concerning the 1st idea proposed, evaluated:

- Originality – 1
- Paradigm Relatedness – 1
- Feasibility – 3
- Effectiveness – 2

OBJECT	Object															
Answer	Originality				Paradigm Relatedness				Feasibility				Effectiveness			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
idea 1	x				x						x			x		
idea 2			x					x		x						x

Some rules

- Judges should read at least once the whole document before starting the assessment phase
- It is important that judges score the idea proposed by participants INDEPENDENTLY
- Judges should spread out the assessment phase in several sessions (even more than once per day) in order to have the lucidity needed and to minimize fatigue
- Before the beginning of the assessment of responses regarding each stimulus, judges should look at all the ideas proposed by participants concerning the same stimulus in order to know which will be the concepts that they are going to score

Anchor tables

The following anchor tables are intended as practical tools for the raters to be used during the scoring phase.

Originality

The degree to which the idea is not only rare, but is also ingenious, imaginative or surprising

Score	Level description
4	Not expressed before (rare, unusual) and ingenious, imaginative or surprising; may be humorous
3	Unusual, interesting; shows some imagination
2	Interesting
1	Common, mundane, boring

Paradigm relatedness

The degree to which an idea is paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational

Score	Level description
4	The object present both the addition of new components and the alteration of the relationships bewteen its already existing elements
3	The object present alteration of the relationships between its already existing elements (e.g. recombination of components)
2	The object architecture present some differences attributable to the addition or removal of components
1	The architecture is proposed without significant changes

Effectiveness

The degree to which the idea will solve the problem without regard for feasibility

Score	Level description
4	Reasonable and will solve the problem without regard for feasibility (If you could do it, it would solve the main problem)
3	Reasonable and will contribute to the solution of the problem (It helps, but it is only a partial solution)
2	Unreasonable or unlikely to solve the problem (It probably will not work)
1	Not useful for the problem resolution (It would not work, even if you could do it)

Feasibility

An idea is feasible if it can be easily implemented and does not violate known constraints

Score	Level description
4	Easy to implement common strategies, at low cost, that violate no norms or sensibilities
3	Some changes strategies, somewhat uncommon or unusual, that don't offend sensibilities
2	Significant change or expensive strategy, but not totally impossible to implement. Offends sensibilities somewhat but is not totally unacceptable
1	Totally infeasible to implement or extremely costly solutions, that radically violates laws or sensibilities or Totally unacceptable business practice.

Thank you for your help

OBJECT	window															
Answer	Originality				Paradigm Relatedness				Feasibility				Effectiveness			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
to ventilate the room																
to use it as a mirror to watch yourself																
Telefonare																
if it is raining, you can go by the window and make yourself a video to look dramatic																
to getting light into the house																
as a blackboard																
to grow plants (with the light)																
as a surfboard																
to look out or in																
use it as a door																
use the glass as a magnifier for the sun and start a fire																
look through																
to scrape off ice from another window																
to look through																
to look outside																
make paintings on the glass																
open it to let some fresh air in the house																
shake it and use it as a fan																
to have light inside your house																
to look outside																
make the apartment lighter																
break it and use it as a weapon																
get light into the house																
to look through																
see outside of the house																
watch outside your house and get the best temperature																
to hide behind it																
to shave your hair																
to see what is going on outside																
get light in your room																
see the sun																
look out from your room																
creating tools																
as a cool transparent surfboard																
put it in place of the floor to be able to see below																
to visit a girl without her parents knowing it																
have more light in the room																
Use its handle as a hanger for clothes, when it is open																

Table 27 Example of judges' Excel sheet

A.5 Other data for the design of configuration (2x2x3, 2x2x6, 2x2x7)

Below, further data concerning the analysis of the configuration 2x2x3, 2x2x6 and 2x2x7 are presented. In particular, the Mauchly's Tests are presented in order to verify the sphericity of the samples adopted. This analysis tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix. It is designed considering as within Subjects Design: Condition + Hemisphere + Area + Condition * Hemisphere + Condition * Area + Hemisphere * Area + Condition * Hemisphere * Area.

Mauchly's Test of Sphericity (2x2x3 means)

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Condition	1,000	0,000	0		1,000	1,000	1,000
Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Area	,655	14,406	2	,001	,743	,769	,500
Condition * Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Condition * Area	,471	25,597	2	,000	,654	,669	,500
Hemisphere * Area	,645	14,916	2	,001	,738	,763	,500
Condition * Hemisphere * Area	,705	11,908	2	,003	,772	,801	,500

Table 28 Mauchly's Test of Sphericity (2x2x3 means)

Mauchly's Test of Sphericity (2x2x3 all population)

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Condition	1,000	0,000	0		1,000	1,000	1,000
Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Area	,567	199,859	2	,000	,698	,700	,500
Condition * Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Condition * Area	,397	324,902	2	,000	,624	,625	,500
Hemisphere * Area	,661	145,535	2	,000	,747	,749	,500
Condition * Hemisphere * Area	,692	129,551	2	,000	,765	,767	,500

Table 29 Mauchly's Test of Sphericity (2x2x3 all population)

Mauchly's Test of Sphericity (2x2x6 means)

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Condition_	1,000	0,000	0		1,000	1,000	1,000
Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Area	,040	106,815	14	,000	,533	,581	,200
Condition_ * Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Condition_ * Area	,065	90,251	14	,000	,504	,546	,200
Hemisphere * Area	,118	70,654	14	,000	,619	,686	,200
Condition_ * Hemisphere * Area	,216	50,655	14	,000	,649	,723	,200

Table 30 Mauchly's Test of Sphericity (2x2x6 means)

Mauchly's Test of Sphericity (2x2x7 means)

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Condition_	1,000	0,000	0		1,000	1,000	1,000
Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Area	,000	266,521	20	,000	,355	,379	,167
Condition_ * Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Condition_ * Area	,000	263,126	20	,000	,315	,333	,167
Hemisphere * Area	,099	75,819	20	,000	,575	,645	,167
Condition_ * Hemisphere * Area	,036	108,555	20	,000	,565	,633	,167

Table 31 Mauchly's Test of Sphericity (2x2x7 means)

Mauchly's Test of Sphericity (2x2x7 all population)

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Condition_	1,000	0,000	0		1,000	1,000	1,000
Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Area	,001	2288,151	20	0,000	,363	,365	,167
Condition_ * Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Condition_ * Area	,007	1721,224	20	0,000	,355	,357	,167
Hemisphere * Area	,210	547,964	20	,000	,636	,644	,167
Condition_ * Hemisphere * Area	,260	472,866	20	,000	,701	,711	,167

Table 32 Mauchly's Test of Sphericity (2x2x7 all population)

A.6 Other data for 2x2x6

Below, further data concerning the repeated measures ANOVA executed with within-subject factors Hemisphere side (LEFT vs RIGHT), Condition (COMMON vs UNCOMMON), Area (FRONTAL VENTRICULAR, FRONTAL DORSAL, CENTRAL VENTRICULAR, CENTRAL DORSAL, POSTERIOR VENTRICULAR, POSTERIOR DORSAL) are reported.

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition_	Sphericity Assumed	,736	1	,736	,845	,359	,002
	Greenhouse-Geisser	,736	1,000	,736	,845	,359	,002
Error(Condition_)	Sphericity Assumed	307,623	353	,871			
	Greenhouse-Geisser	307,623	353,000	,871			
Hemisphere	Sphericity Assumed	,256	1	,256	2 ,728	,099	,008
	Greenhouse-Geisser	,256	1,000	,256	2,728	,099	,008
Error(Hemisphere)	Sphericity Assumed	33,176	353	,094			
	Greenhouse-Geisser	33,176	353,000	,094			
Area	Sphericity Assumed	35,997	5	7,199	78,452	,000	,182
	Greenhouse-Geisser	35,997	2,820	12,763	78,452	,000	,182
Error(Area)	Sphericity Assumed	161,969	1765	,092			
	Greenhouse-Geisser	161,969	995,562	,163			
Condition_ * Hemisphere	Sphericity Assumed	,238	1	,238	2,441	,119	,007
	Greenhouse-Geisser	,238	1,000	,238	2,441	,119	,007
Error(Condition_*Hemisphere)	Sphericity Assumed	34,447	353	,098			
	Greenhouse-Geisser	34,447	353,000	,098			
Condition_ * Area	Sphericity Assumed	5,205	5	1,041	19,607	,000	,053
	Greenhouse-Geisser	5,205	2,641	1,971	19,607	,000	,053
Error(Condition_*Area)	Sphericity Assumed	93,711	1765	,053			

	Greenhouse-Geisser	93,711	932,196	,101			
Hemisphere * Area	Sphericity Assumed	,426	5	,085	3,771	,002	,011
	Greenhouse-Geisser	,426	3,343	,127	3,771	,008	,011
Error(Hemisphere*Area)	Sphericity Assumed	39,855	1765	,023			
	Greenhouse-Geisser	39,855	1180,218	,034			
Condition_ * Hemisphere * Area	Sphericity Assumed	,060	5	,012	,688	,633	,002
	Greenhouse-Geisser	,060	3,270	,018	,688	,572	,002
Error(Condition_*Hemisphere*Area)	Sphericity Assumed	30,977	1765	,018			
	Greenhouse-Geisser	30,977	1154,153	,027			

Table 33 Tests of Within-Subjects Effects

Mauchly's Test of Sphericity

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Condition_	1,000	0,000	0		1,000	1,000	1,000
Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Area	,079	889,551	14	,000	,564	,569	,200
Condition_ * Hemisphere	1,000	0,000	0		1,000	1,000	1,000
Condition_ * Area	,094	828,848	14	,000	,528	,533	,200
Hemisphere * Area	,239	502,225	14	,000	,669	,676	,200
Condition_ * Hemisphere * Area	,218	534,130	14	,000	,654	,661	,200

Table 34 Mauchly's Test of Sphericity

Estimates (Area)

Area	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
FV	,106	,014	,079	,133
FD	,147	,014	,120	,175
CV	,027	,012	,003	,051
CD	-,004	,013	-,029	,020
PV	,000	,013	-,026	,026
PD	-,036	,015	-,066	-,007

Table 35 Estimates marginal means - Area

Pairwise Comparisons

(I) Area		Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
FV	FD	-,041*	,010	,001	-,071	-,011
	CV	,079*	,008	,000	,057	,102
	CD	,111*	,012	,000	,075	,146
	PV	,106*	,010	,000	,075	,137
	PD	,142*	,015	,000	,097	,188
FD	FV	,041*	,010	,001	,011	,071
	CV	,121*	,010	,000	,090	,151
	CD	,152*	,011	,000	,120	,183
	PV	,147*	,014	,000	,106	,188
	PD	,184*	,016	,000	,136	,232
CV	FD	-,079*	,008	,000	-,102	-,057
	FV	-,121*	,010	,000	-,151	-,090
	CD	,031*	,007	,000	,011	,052
	PV	,026*	,007	,007	,004	,048
	PD	,063*	,012	,000	,029	,097
CD	FD	-,111*	,012	,000	-,146	-,075
	FV	-,152*	,011	,000	-,183	-,120
	CV	-,031*	,007	,000	-,052	-,011
	PV	-,005	,012	1,000	-,039	,029
	PD	,032	,012	,128	-,004	,068
PV	FD	-,106*	,010	,000	-,137	-,075
	FV	-,147*	,014	,000	-,188	-,106
	CV	-,026*	,007	,007	-,048	-,004
	CD	,005	,012	1,000	-,029	,039
	PD	,037*	,009	,002	,009	,064
PD	FD	-,142*	,015	,000	-,188	-,097
	FV	-,184*	,016	,000	-,232	-,136
	CV	-,063*	,012	,000	-,097	-,029

	CD	-,032	,012	,128	-,068	,004
	PV	-,037*	,009	,002	-,064	-,009

Table 36 Pairwise Comparisons

Condition * Area

Condition		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Common	FV	,130	,019	,093	,167
	FD	,169	,018	,133	,205
	CV	,015	,017	-,019	,050
	CD	-,022	,017	-,057	,012
	PV	-,029	,019	-,066	,008
	PD	-,079	,021	-,120	-,038
Uncommon	FV	,082	,017	,049	,116
	FD	,125	,017	,092	,159
	CV	,038	,015	,008	,068
	CD	,014	,016	-,017	,045
	PV	,030	,017	-,004	,064
	PD	,006	,019	-,031	,044

*Table 37 Estimate marginal means - Condition * Area*

Hemisphere * Area

Hemisphere		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Left	FV	,108	,015	,080	,137
	FD	,149	,013	,123	,175
	CV	,041	,013	,015	,068
	CD	,003	,013	-,023	,028
	PV	,013	,014	-,015	,042
	PD	-,042	,015	-,072	-,012
Right	FV	,104	,016	,073	,135
	FD	,145	,016	,115	,176
	CV	,012	,013	-,014	,038
	CD	-,011	,013	-,038	,015
	PV	-,013	,015	-,042	,016
	PD	-,030	,016	-,062	,001

*Table 38 Estimate marginal means - Hemisphere * Area*

A.7 Designers vs Engineers

Below, data concerning the repeated measures ANOVA executed with within-subject factors Hemisphere side (LEFT vs RIGHT), Condition (COMMON vs UNCOMMON), Area (FRONTAL VENTRICULAR, FRONTAL DORSAL, CENTRAL VENTRICULAR, CENTRAL DORSAL, POSTERIOR VENTRICULAR, POSTERIOR DORSAL) and between-subjects factor Background are reported. Levene's analysis tests the null hypothesis that the error variance of the dependent variable is equal across groups.

Subjects	Corse of study	Background
Subject 1	Materials Engineering	Engineering
Subject 2	Automation Engineering	Engineering
Subject 5	Advanced Materials Engineering	Engineering
Subject 6	Physics and Electrical Engineering	Engineering
Subject 7	Maintenance Engineering	Engineering
Subject 9	Materials Engineering	Engineering
Subject 10	Industrial Design Engineering	Design
Subject 11	Industrial Design Engineering	Design
Subject 12	Industrial Design Engineering	Design
Subject 13	Mechanical Engineering	Engineering
Subject 14	Industrial economy engineering	Engineering
Subject 16	Technical Design	Design
Subject 17	Material Science and Engineering	Engineering
Subject 18	Advanced Materials Engineering	Engineering
Subject 19	Chemistry Engineering	Engineering
Subject 20	Material Engineering	Engineering
Subject 21	Mechanical Engineering	Engineering
Subject 25	Industrial Design Engineering	Design
Subject 26	Industrial Design Engineering	Design
Subject 27	Industrial Design Engineering	Design
Subject 29	Industrial Engineering	Design
Subject 30	Mechanical Engineering	Engineering
Subject 31	Technical Design	Design
Subject 32	Industrial Design Engineering	Design
Subject 33	Industrial Design Engineering	Design
Subject 34	Technical Design	Design
Subject 35	Technical Design	Design
Subject 36	Industrial Design Engineering	Design
Subject 39	Mechanical Engineering	Engineering

Table 39 Participants selected for the Background comparison

Tests of Within-Subjects Effects (Background)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition	Sphericity Assumed	2,769	1	2,769	3,420	,065	,011
	Greenhouse-Geisser	2,769	1,000	2,769	3,420	,065	,011
Condition * Background	Sphericity Assumed	9,383	1	9,383	11,591	,001	,038
	Greenhouse-Geisser	9,383	1,000	9,383	11,591	,001	,038
Error(Condition)	Sphericity Assumed	238,005	294	,810			
	Greenhouse-Geisser	238,005	294,000	,810			
Hemisphere	Sphericity Assumed	,192	1	,192	2,038	,154	,007
	Greenhouse-Geisser	,192	1,000	,192	2,038	,154	,007
Hemisphere * Background	Sphericity Assumed	1,409	1	1,409	14,956	,000	,048
	Greenhouse-Geisser	1,409	1,000	1,409	14,956	,000	,048
Error(Hemisphere)	Sphericity Assumed	27,705	294	,094			
	Greenhouse-Geisser	27,705	294,000	,094			
Area	Sphericity Assumed	32,421	5	6,484	70,847	,000	,194
	Greenhouse-Geisser	32,421	2,861	11,333	70,847	,000	,194
Area * Background	Sphericity Assumed	8,592	5	1,718	18,776	,000	,060
	Greenhouse-Geisser	8,592	2,861	3,003	18,776	,000	,060
Error(Area)	Sphericity Assumed	134,540	1470	,092			
	Greenhouse-Geisser	134,540	841,102	,160			
Condition * Hemisphere	Sphericity Assumed	,430	1	,430	4,432	,036	,015
	Greenhouse-Geisser	,430	1,000	,430	4,432	,036	,015
Condition * Hemisphere * Background	Sphericity Assumed	,034	1	,034	,351	,554	,001
	Greenhouse-Geisser	,034	1,000	,034	,351	,554	,001
Error(Condition*Hemisphere)	Sphericity Assumed	28,551	294	,097			
	Greenhouse-Geisser	28,551	294,000	,097			
Condition * Area	Sphericity Assumed	5,907	5	1,181	23,437	,000	,074
	Greenhouse-Geisser	5,907	2,648	2,231	23,437	,000	,074
Condition * Area * Background	Sphericity Assumed	,436	5	,087	1,731	,124	,006

	Greenhouse-Geisser	,436	2,648	,165	1,731	,166	,006
Error(Condition*Area)	Sphericity Assumed	74,100	1470	,050			
	Greenhouse-Geisser	74,100	778,483	,095			
Hemisphere * Area	Sphericity Assumed	,588	5	,118	5,129	,000	,017
	Greenhouse-Geisser	,588	3,251	,181	5,129	,001	,017
Hemisphere * Area * Background	Sphericity Assumed	,833	5	,167	7,266	,000	,024
	Greenhouse-Geisser	,833	3,251	,256	7,266	,000	,024
Error(Hemisphere*Area)	Sphericity Assumed	33,712	1470	,023			
	Greenhouse-Geisser	33,712	955,682	,035			
Condition * Hemisphere * Area	Sphericity Assumed	,085	5	,017	,925	,464	,003
	Greenhouse-Geisser	,085	3,148	,027	,925	,432	,003
Condition * Hemisphere * Area * Background	Sphericity Assumed	,175	5	,035	1,902	,091	,006
	Greenhouse-Geisser	,175	3,148	,056	1,902	,125	,006
Error(Condition*Hemisphere*Area)	Sphericity Assumed	27,088	1470	,018			
	Greenhouse-Geisser	27,088	925,371	,029			

Table 40 Tests of Within-Subjects Effects (Background)

Background * Area

Background		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Design	FV	,115	,021	,074	,157
	FD	,185	,021	,144	,227
	CV	,018	,019	-,020	,055
	CD	,001	,020	-,038	,040
	PV	-,048	,020	-,087	-,008
	PD	-,114	,023	-,159	-,069
Engineering	FV	,110	,021	,068	,152
	FD	,123	,021	,081	,165
	CV	,051	,019	,014	,089
	CD	,004	,020	-,035	,044
	PV	,053	,021	,012	,093
	PD	,033	,023	-,013	,079

*Table 41 Estimated marginal means - Background * Area*

Background * Hemisphere

Background		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Design	Left	,017	,018	-,017	,052
	Right	,035	,019	-,002	,072
Engineering	Left	,082	,018	,047	,117
	Right	,043	,019	,006	,080

*Table 42 Estimated marginal means - Background * Hemisphere*

Background * Hemisphere * Area

Background			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Design	Left	FV	,084	,022	,041	,127
		FD	,172	,020	,132	,211
		CV	,009	,020	-,031	,049
		CD	-,001	,020	-,040	,038
		PV	-,048	,022	-,091	-,005
		PD	-,112	,023	-,158	-,066
	Right	FV	,147	,024	,100	,194
		FD	,198	,023	,152	,244
		CV	,026	,021	-,015	,066
		CD	,003	,021	-,039	,044
		PV	-,047	,023	-,093	-,002
		PD	-,116	,025	-,164	-,068
Engineering	Left	FV	,132	,022	,089	,175
		FD	,134	,020	,094	,174
		CV	,092	,021	,051	,132
		CD	,017	,020	-,022	,057
		PV	,090	,022	,046	,134
		PD	,025	,024	-,022	,072
	Right	FV	,087	,024	,040	,135
		FD	,112	,024	,065	,158
		CV	,011	,021	-,030	,052
		CD	-,009	,021	-,050	,033
		PV	,016	,023	-,030	,062
		PD	,042	,025	-,007	,091

*Table 43 Estimated marginal means - Background * Hemisphere * Area*

Background * Condition

Background		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Design	Common	,043	,025	-,006	,092
	Uncommon	,010	,021	-,031	,050
Engineering	Common	,006	,025	-,044	,056
	Uncommon	,118	,021	,077	,160

*Table 44 Estimated marginal means - Background * Condition*

Background * Condition * Hemisphere * Area

Background				Mean	Std. Error	95% Confidence Interval	
						Lower Bound	Upper Bound
Design	Common	Left	FV	,139	,030	,079	,198
			FD	,225	,028	,169	,280
			CV	,049	,029	-,007	,105
			CD	,020	,028	-,034	,075
			PV	-,058	,031	-,119	,002
			PD	-,137	,032	-,201	-,073
		Right	FV	,216	,034	,149	,284
			FD	,237	,029	,180	,295
			CV	,037	,030	-,022	,097
			CD	-,002	,029	-,059	,055
			PV	-,061	,033	-,125	,003
			PD	-,152	,034	-,219	-,084
	Uncommon	Left	FV	,029	,027	-,024	,082
			FD	,119	,024	,071	,167
			CV	-,030	,025	-,079	,019
			CD	-,022	,024	-,070	,025
			PV	-,038	,029	-,094	,019
			PD	-,087	,029	-,144	-,030
		Right	FV	,077	,029	,019	,135
			FD	,159	,028	,103	,215
			CV	,014	,024	-,033	,061
			CD	,007	,024	-,039	,054
			PV	-,033	,029	-,090	,023
			PD	-,080	,030	-,139	-,022
Engineering	Common	Left	FV	,120	,030	,060	,179
			FD	,130	,029	,073	,186
			CV	,037	,029	-,020	,094
			CD	-,037	,028	-,092	,018
			PV	,013	,031	-,049	,074
			PD	-,049	,033	-,114	,016

		Right	FV	,040	,035	-,028	,109
			FD	,097	,030	,039	,155
			CV	-,072	,031	-,133	-,012
			CD	-,076	,029	-,134	-,018
			PV	-,072	,033	-,138	-,007
			PD	-,055	,035	-,123	,014
	Uncommon	Left	FV	,144	,027	,091	,198
			FD	,138	,025	,089	,187
			CV	,147	,025	,097	,197
			CD	,072	,025	,023	,120
			PV	,167	,029	,110	,224
			PD	,099	,029	,042	,157
		Right	FV	,135	,030	,076	,193
			FD	,126	,029	,070	,183
			CV	,094	,024	,047	,141
			CD	,058	,024	,011	,105
			PV	,104	,029	,046	,161
			PD	,138	,030	,079	,197

*Table 45 Estimated marginal means - Background * Condition *Hemisphere * Area*

Levene's Test of Equality of Error Variances

	F	df1	df2	Sig.
FV_L_C	,125	1	294	,724
FD_L_C	,001	1	294	,977
CV_L_C	1,982	1	294	,160
CD_L_C	,747	1	294	,388
PV_L_C	,004	1	294	,950
PD_L_C	,034	1	294	,854
FV_R_C	,625	1	294	,430
FD_R_C	,094	1	294	,760
CV_R_C	,098	1	294	,755
CD_R_C	1,682	1	294	,196
PV_R_C	,000	1	294	,999
PD_R_C	1,505	1	294	,221
FV_L_U	,006	1	294	,938
FD_L_U	,087	1	294	,768
CV_L_U	,487	1	294	,486
CD_L_U	,338	1	294	,561
PV_L_U	1,433	1	294	,232
PD_L_U	5,919	1	294	,016
FV_R_U	,277	1	294	,599
FD_R_U	3,636	1	294	,058
CV_R_U	,155	1	294	,694
CD_R_U	1,071	1	294	,302
PV_R_U	1,562	1	294	,212
PD_R_U	3,224	1	294	,074

Table 46 Levene's Test of Equality of Error Variances

A.8 Male vs Female

Below, further data concerning the repeated measures ANOVA executed with within-subject factors Hemisphere side (LEFT vs RIGHT), Condition (COMMON vs UNCOMMON), Area (FRONTAL VENTRICULAR, FRONTAL DORSAL, CENTRAL VENTRICULAR, CENTRAL DORSAL, POSTERIOR VENTRICULAR, POSTERIOR DORSAL) and between-subjects factor Gender are reported. Levene's analysis tests the null hypothesis that the error variance of the dependent variable is equal across groups.

Subjects	Gender	Nationality	Corse of study	Year of Study	Strategy adopted
Subject 1	FEMALE	Italian	Materials Engineering	5th	I was thinking about the object, its dimensions, its shape and then its common use. Then I tried to find another use for it, based on its shape. I thought about "what I use it for", or "how I used it in an uncommon way before"
Subject 2	MALE	Italian	Automation Engineering	5th	The uncommon answers came up in my mind either because I had already experienced them personally or because I saw those kinds of uses somewhere. I think that 60% of my answers did not reflect actual creativity, but was just memory retrieval of something.
Subject 5	MALE	Syrian	Advanced Materials Engineering	4th	I tried to focus on a completly different use from the common, leaving my mind looking around. Someone were really hard. But first, thinking in a common solution, and then change
Subject 8	FEMALE	German	Business Engineering	4th	thinking about the most uncommon way I have used this object in my life
Subject 9	MALE	Italian	Materials Engineering	5th	First I thought about the normal use fo the objects, then I tried to go out of reality. I just tried to go off the grid
Subject 12	FEMALE	Swedish	Industrial Design Engineering	3rd	no particular strategy
Subject 13	FEMALE	French	Mechanical Engineering	4th	I just tiried to figure out a use starting from a mental image of the object and then to see if it could work
Subject 14	MALE	Swedish	Industrial economy engineering	3rd	I know other people use objects for other uses, mainly in design. I tried to keep my mind blank, and avoid thinking about the common use
Subject 18	FEMALE	Spanish	Advanced Materials Engineering	4th	I just pictured stuff in my mind, its shape. Then i thought about the common use and try to get something more uncommon
Subject 21	MALE	Spanish	Mechanical Engineering	4th	I thought about things I did before or to things that I saw. Not a real strategy
Subject 26	MALE	Swedish	Industrial Design Engineering	3rd	I just followed my own experiences with the objects, or some memory form videos
Subject 27	MALE	Swedish	Industrial Design Engineering	3rd	First i thought about the common use of the object, and then I spent time trying to find other multiple uses, compare them and try not to stuck on the first one that i thought

Subject 29	FEMALE	Swedish	Industrial Engineering	5th	It was about getting inspired by the word. I was thinking about the material and what happens when i put it upside-down, or placed it in another environment
Subject 31	FEMALE	Swedish	Technical Design	4th	I tried to think as less as possible of the common use. I tried to see it just as a pure object and try to put it in some context by creating a mental picture of it
Subject 32	FEMALE	Swedish	Industrial Design Engineering	3rd	I thought to have the object in my hand and figure out what could I do with it. Also thought about the common use, and then try to think the opposite
Subject 33	MALE	Swedish	Industrial Design Engineering	4th	I first thought of the most common use, then something similar or figure out something crazy based on the shape of the object
Subject 34	MALE	Swedish	Technical Design	3rd	For some of them I thought about the common use before getting the uncommon use
Subject 36	FEMALE	Swedish	Industrial Design Engineering	5th	I tried to make a mental picture of the objects and see them in other situations

Table 47 Participants selected for the Gender comparison

Tests of Within-Subjects Effects (Gender)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition	Sphericity Assumed	1,290	1	1,290	1,391	,240	,007
	Greenhouse-Geisser	1,290	1,000	1,290	1,391	,240	,007
Condition * Gender	Sphericity Assumed	,468	1	,468	,505	,478	,003
	Greenhouse-Geisser	,468	1,000	,468	,505	,478	,003
Error(Condition)	Sphericity Assumed	177,200	191	,928			
	Greenhouse-Geisser	177,200	191,000	,928			
Hemisphere	Sphericity Assumed	,302	1	,302	3,598	,059	,018
	Greenhouse-Geisser	,302	1,000	,302	3,598	,059	,018
Hemisphere * Gender	Sphericity Assumed	,366	1	,366	4,361	,038	,022
	Greenhouse-Geisser	,366	1,000	,366	4,361	,038	,022
Error(Hemisphere)	Sphericity Assumed	16,013	191	,084			
	Greenhouse-Geisser	16,013	191,000	,084			
Area	Sphericity Assumed	18,502	5	3,700	40,983	,000	,177
	Greenhouse-Geisser	18,502	2,476	7,473	40,983	,000	,177
Area * Gender	Sphericity Assumed	,914	5	,183	2,025	,073	,010
	Greenhouse-Geisser	,914	2,476	,369	2,025	,122	,010

Error(Area)	Sphericity Assumed	86,228	955	,090			
	Greenhouse-Geisser	86,228	472,869	,182			
Condition * Hemisphere	Sphericity Assumed	,015	1	,015	,160	,689	,001
	Greenhouse-Geisser	,015	1,000	,015	,160	,689	,001
Condition * Hemisphere * Gender	Sphericity Assumed	,285	1	,285	3,002	,085	,015
	Greenhouse-Geisser	,285	1,000	,285	3,002	,085	,015
Error(Condition*Hemisphere)	Sphericity Assumed	18,115	191	,095			
	Greenhouse-Geisser	18,115	191,000	,095			
Condition * Area	Sphericity Assumed	4,060	5	,812	16,898	,000	,081
	Greenhouse-Geisser	4,060	2,690	1,510	16,898	,000	,081
Condition * Area * Gender	Sphericity Assumed	,294	5	,059	1,224	,296	,006
	Greenhouse-Geisser	,294	2,690	,109	1,224	,300	,006
Error(Condition*Area)	Sphericity Assumed	45,894	955	,048			
	Greenhouse-Geisser	45,894	513,730	,089			
Hemisphere * Area	Sphericity Assumed	,450	5	,090	4,437	,001	,023
	Greenhouse-Geisser	,450	3,181	,142	4,437	,004	,023
Hemisphere * Area * Gender	Sphericity Assumed	,372	5	,074	3,666	,003	,019
	Greenhouse-Geisser	,372	3,181	,117	3,666	,011	,019
Error(Hemisphere*Area)	Sphericity Assumed	19,381	955	,020			
	Greenhouse-Geisser	19,381	607,518	,032			
Condition * Hemisphere * Area	Sphericity Assumed	,173	5	,035	2,101	,063	,011
	Greenhouse-Geisser	,173	3,171	,054	2,101	,095	,011
Condition * Hemisphere * Area * Gender	Sphericity Assumed	,793	5	,159	9,651	,000	,048
	Greenhouse-Geisser	,793	3,171	,250	9,651	,000	,048
Error(Condition*Hemisphere*Area)	Sphericity Assumed	15,693	955	,016			
	Greenhouse-Geisser	15,693	605,730	,026			

Table 48 Tests of Within-Subjects Effects (Gender)

Gender * Area

Gender		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
FEMALE	FV	,039	,028	-,016	,094
	FD	,055	,028	-9,983E-05	,110
	CV	-,036	,024	-,083	,011
	CD	-,087	,024	-,135	-,039
	PV	-,072	,025	-,120	-,023
	PD	-,132	,029	-,189	-,076
MALE	FV	,167	,028	,112	,221
	FD	,212	,028	,158	,267
	CV	,079	,024	,033	,126
	CD	,063	,024	,015	,110
	PV	,044	,024	-,004	,092
	PD	,063	,028	,007	,118

*Table 49 Estimated marginal means - Gender * Area*

Gender * Hemisphere

Gender		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
FEMALE	Left	-,040	,022	-,083	,004
	Right	-,038	,024	-,086	,010
MALE	Left	,121	,021	,079	,164
	Right	,088	,024	,041	,135

*Table 50 Estimated marginal means - Gender * Hemisphere*

Gender * Hemisphere * Area

Gender			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
FEMALE	Left	FV	,032	,028	-,023	,086
		FD	,065	,027	,012	,118
		CV	-,034	,024	-,081	,014
		CD	-,086	,025	-,135	-,036
		PV	-,083	,026	-,134	-,032
		PD	-,131	,030	-,190	-,073
	Right	FV	,046	,032	-,017	,110
		FD	,045	,030	-,015	,105
		CV	-,038	,027	-,091	,016
		CD	-,088	,025	-,137	-,038
		PV	-,060	,029	-,116	-,004
		PD	-,133	,030	-,192	-,074
MALE	Left	FV	,154	,027	,101	,208
		FD	,236	,026	,184	,288
		CV	,113	,024	,066	,160
		CD	,086	,025	,037	,134
		PV	,078	,025	,028	,128
		PD	,062	,029	,004	,119
	Right	FV	,179	,032	,116	,241
		FD	,189	,030	,130	,248
		CV	,046	,027	-,007	,098
		CD	,039	,025	-,009	,088
		PV	,009	,028	-,046	,065
		PD	,064	,030	,005	,122

*Table 51 Estimated marginal means - Gender * Hemisphere * Area*

Gender * Condition * Hemisphere * Area

Gender				Mean	Std. Error	95% Confidence Interval	
						Lower Bound	Upper Bound
FEMALE	Common	Left	FV	,045	,038	-,030	,120
			FD	,114	,038	,040	,188
			CV	,005	,036	-,065	,076
			CD	-,052	,036	-,123	,019
			PV	-,091	,036	-,163	-,019
			PD	-,134	,040	-,214	-,055
		Right	FV	,146	,047	,054	,237
			FD	,096	,039	,019	,174
			CV	,008	,040	-,071	,087
			CD	-,062	,036	-,134	,009
			PV	-,069	,041	-,149	,012
			PD	-,150	,042	-,233	-,066
	Uncommon	Left	FV	,019	,034	-,048	,085
			FD	,017	,032	-,046	,080
			CV	-,073	,031	-,134	-,012
			CD	-,120	,031	-,181	-,058
			PV	-,075	,037	-,148	-,003
			PD	-,129	,036	-,201	-,057
		Right	FV	-,053	,040	-,132	,026
			FD	-,007	,037	-,080	,066
			CV	-,083	,031	-,145	-,022
			CD	-,113	,029	-,169	-,056
			PV	-,052	,037	-,124	,021
			PD	-,117	,037	-,190	-,044
MALE	Common	Left	FV	,231	,037	,157	,304
			FD	,307	,037	,234	,380
			CV	,126	,035	,056	,196
			CD	,094	,035	,024	,164
			PV	,056	,036	-,015	,126
			PD	,013	,040	-,065	,091
		Right	FV	,205	,046	,115	,295
			FD	,218	,039	,142	,295
			CV	,032	,039	-,046	,110
			CD	,031	,036	-,039	,102
			PV	-,005	,040	-,085	,074
			PD	,026	,042	-,056	,109
	Uncommon	Left	FV	,078	,033	,013	,143
			FD	,165	,031	,103	,227
			CV	,100	,030	,040	,160
			CD	,077	,031	,017	,138
			PV	,100	,036	,029	,172
			PD	,110	,036	,039	,181

		Right	FV	,152	,040	,074	,230
			FD	,159	,036	,087	,231
			CV	,059	,031	-,001	,120
			CD	,047	,028	-,008	,103
			PV	,024	,036	-,048	,095
			PD	,101	,037	,029	,173

*Table 52 Estimated marginal means - Gender * Condition * Hemisphere * Area*

Levene's Test of Equality of Error Variances

	F	df1	df2	Sig.
FV_L_C	,120	1	191	,730
FD_L_C	3,030	1	191	,083
CV_L_C	2,599	1	191	,109
CD_L_C	1,624	1	191	,204
PV_L_C	,093	1	191	,761
PD_L_C	1,162	1	191	,282
FV_R_C	1,597	1	191	,208
FD_R_C	4,309	1	191	,039
CV_R_C	5,878	1	191	,016
CD_R_C	3,379	1	191	,068
PV_R_C	,050	1	191	,824
PD_R_C	,031	1	191	,860
FV_L_U	,497	1	191	,482
FD_L_U	,046	1	191	,830
CV_L_U	,485	1	191	,487
CD_L_U	,082	1	191	,775
PV_L_U	,686	1	191	,409
PD_L_U	5,887	1	191	,016
FV_R_U	1,289	1	191	,258
FD_R_U	2,769	1	191	,098
CV_R_U	1,160	1	191	,283
CD_R_U	,010	1	191	,922
PV_R_U	2,711	1	191	,101
PD_R_U	1,257	1	191	,264

Table 53 Levene's Test of Equality of Error Variances

A.9 Bachelor vs Master

Below, further data concerning the repeated measures ANOVA executed with within-subject factors Hemisphere side (LEFT vs RIGHT), Condition (COMMON vs UNCOMMON), Area (FRONTAL VENTRICULAR, FRONTAL DORSAL, CENTRAL VENTRICULAR, CENTRAL DORSAL, POSTERIOR VENTRICULAR, POSTERIOR DORSAL) and between-subjects factor Degreelevel are reported. Levene's analysis tests the null hypothesis that the error variance of the dependent variable is equal across groups.

Subjects	Age	Mother Tongue	Corse of study	Year of Study	Degree
Subject 1	23	Italian	Materials Engineering	5th	Master
Subject 2	25	Italian	Automation Engineering	5th	Master
Subject 6	20	Swedish	Physics and Electrical Engineering	1st	Bachelor
Subject 9	24	Italian	Materials Engineering	5th	Master
Subject 10	24	Swedish	Industrial Design Engineering	4th	Master
Subject 12	23	Swedish	Industrial Design Engineering	3rd	Bachelor
Subject 13	21	French	Mechanical Engineering	4th	Master
Subject 14	24	Swedish	Industrial economy engineering	3rd	Bachelor
Subject 19	24	German	Chemistry Engineering	5th	Master
Subject 20	24	German	Material Engineering	4th	Master
Subject 21	22	Spanish	Mechanical Engineering	4th	Master
Subject 22	20	Swedish	Computer Science	1st	Bachelor
Subject 25	25	Swedish	Industrial Design Engineering	3rd	Bachelor
Subject 26	23	Swedish	Industrial Design Engineering	3rd	Bachelor
Subject 27	23	Swedish	Industrial Design Engineering	3rd	Bachelor
Subject 28	21	French	Computer Science	4th	Master
Subject 30	22	Swedish	Mechanical Engineering	2nd	Bachelor
Subject 31	23	Swedish	Technical Design	4th	Master
Subject 32	21	Swedish	Industrial Design Engineering	3rd	Bachelor
Subject 34	21	Swedish	Technical Design	3rd	Bachelor
Subject 39	25	Spanish	Mechanical Engineering	5th	Master

Table 54 Participants selected for the Degree level comparison

Tests of Within-Subjects Effects (Degree's level)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Condition	Sphericity Assumed	1,170	1	1,170	1,450	,230	,007
	Greenhouse-Geisser	1,170	1,000	1,170	1,450	,230	,007
Condition * Degreelevel	Sphericity Assumed	,986	1	,986	1,222	,270	,006
	Greenhouse-Geisser	,986	1,000	,986	1,222	,270	,006
Error(Condition)	Sphericity Assumed	173,432	215	,807			
	Greenhouse-Geisser	173,432	215,000	,807			
Hemisphere	Sphericity Assumed	,113	1	,113	1,111	,293	,005
	Greenhouse-Geisser	,113	1,000	,113	1,111	,293	,005
Hemisphere * Degreelevel	Sphericity Assumed	,215	1	,215	2,114	,147	,010
	Greenhouse-Geisser	,215	1,000	,215	2,114	,147	,010
Error(Hemisphere)	Sphericity Assumed	21,838	215	,102			
	Greenhouse-Geisser	21,838	215,000	,102			
Area	Sphericity Assumed	17,879	5	3,576	42,761	,000	,166
	Greenhouse-Geisser	17,879	2,728	6,554	42,761	,000	,166
Area * Degreelevel	Sphericity Assumed	,456	5	,091	1,090	,364	,005
	Greenhouse-Geisser	,456	2,728	,167	1,090	,350	,005
Error(Area)	Sphericity Assumed	89,892	1075	,084			
	Greenhouse-Geisser	89,892	586,484	,153			
Condition * Hemisphere	Sphericity Assumed	,102	1	,102	1,051	,306	,005
	Greenhouse-Geisser	,102	1,000	,102	1,051	,306	,005
Condition * Hemisphere * Degreelevel	Sphericity Assumed	,038	1	,038	,396	,530	,002
	Greenhouse-Geisser	,038	1,000	,038	,396	,530	,002
Error(Condition*Hemisphere)	Sphericity Assumed	20,827	215	,097			
	Greenhouse-Geisser	20,827	215,000	,097			
Condition * Area	Sphericity Assumed	3,608	5	,722	15,850	,000	,069
	Greenhouse-Geisser	3,608	2,859	1,262	15,850	,000	,069
Condition * Area * Degreelevel	Sphericity Assumed	,382	5	,076	1,679	,137	,008

	Greenhouse-Geisser	,382	2,859	,134	1,679	,173	,008
Error(Condition*Area)	Sphericity Assumed	48,946	1075	,046			
	Greenhouse-Geisser	48,946	614,634	,080			
Hemisphere * Area	Sphericity Assumed	,666	5	,133	6,598	,000	,030
	Greenhouse-Geisser	,666	3,563	,187	6,598	,000	,030
Hemisphere * Area * Degreelevel	Sphericity Assumed	1,102	5	,220	10,920	,000	,048
	Greenhouse-Geisser	1,102	3,563	,309	10,920	,000	,048
Error(Hemisphere*Area)	Sphericity Assumed	21,694	1075	,020			
	Greenhouse-Geisser	21,694	766,038	,028			
Condition * Hemisphere * Area	Sphericity Assumed	,178	5	,036	1,882	,095	,009
	Greenhouse-Geisser	,178	3,057	,058	1,882	,130	,009
Condition * Hemisphere * Area * Degreelevel	Sphericity Assumed	,097	5	,019	1,028	,400	,005
	Greenhouse-Geisser	,097	3,057	,032	1,028	,381	,005
Error(Condition*Hemisphere*Area)	Sphericity Assumed	20,366	1075	,019			
	Greenhouse-Geisser	20,366	657,356	,031			

Table 55 Tests of Within-Subjects Effects (Degree's level)

Degree-level * Hemisphere * Area

Degreelevel			Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Bachelor	Left	FV	,109	,026	,057	,160
		FD	,162	,023	,117	,207
		CV	,070	,023	,024	,115
		CD	,041	,023	-,003	,086
		PV	,031	,025	-,018	,080
		PD	,029	,028	-,026	,085
	Right	FV	,121	,028	,066	,177
		FD	,148	,026	,098	,199
		CV	,026	,024	-,021	,073
		CD	,005	,022	-,038	,048
		PV	,003	,026	-,048	,053
		PD	,004	,028	-,052	,060
Master	Left	FV	,108	,025	,059	,157
		FD	,153	,022	,110	,195
		CV	-,020	,022	-,063	,024
		CD	-,044	,021	-,086	-,002
		PV	-,036	,024	-,083	,011
		PD	-,070	,027	-,123	-,017
	Right	FV	,085	,027	,032	,137
		FD	,079	,024	,031	,127
		CV	-,006	,023	-,050	,039
		CD	-,061	,021	-,102	-,020
		PV	,005	,024	-,043	,053
		PD	,010	,027	-,044	,063

*Table 56 Estimated marginal means - Degreelevel * Hemisphere * Area*

Levene's Test of Equality of Error Variances

	F	df1	df2	Sig.
FV_L_C	,192	1	215	,661
FD_L_C	1,053	1	215	,306
CV_L_C	,584	1	215	,446
CD_L_C	4,241	1	215	,041
PV_L_C	1,352	1	215	,246
PD_L_C	1,179	1	215	,279
FV_R_C	1,566	1	215	,212
FD_R_C	,323	1	215	,571
CV_R_C	,179	1	215	,673
CD_R_C	1,035	1	215	,310
PV_R_C	,262	1	215	,609
PD_R_C	,269	1	215	,604
FV_L_U	,253	1	215	,616
FD_L_U	,292	1	215	,589
CV_L_U	1,343	1	215	,248
CD_L_U	8,470	1	215	,004
PV_L_U	5,902	1	215	,016
PD_L_U	9,767	1	215	,002
FV_R_U	1,188	1	215	,277
FD_R_U	5,018	1	215	,026
CV_R_U	,001	1	215	,977
CD_R_U	2,889	1	215	,091
PV_R_U	,576	1	215	,449
PD_R_U	,555	1	215	,457

Table 57 Levene's Test of Equality of Error Variances

References

Abraham, A. (2015). Gender and creativity: an overview of psychological and neuroscientific literature.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique.

Amabile, T. M. (1996). *Creativity and innovation in organizations.*

Amabile, T. M., Conti, R., & Coon, H. (1996). Evidence to support the componential model of creativity: Secondary analyses of three studies.

Andrews, E. G. (1930). The development of imagination in the preschool child.

Baer, J. (1997). Gender differences in the effects of anticipated evaluation on creativity.

Baer, J. (1998). The case for domain specificity of creativity.

Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the consensual assessment technique to nonparallel creative products.

Barki, H., & Pinsonneault, A. (2001). Small Group Brainstorming: Is Electronic Brainstorming the Most Effective Approach?

Barron, F. (1987). Influence of missing attributes on selecting a best multiattributed alternative.

Barron, F., & Harrington, W. (1981). Creativity, intelligence, and personality.

Batey, M., & Furnham, A. (2006). *Creativity, Intelligence, and Personality: A Critical Review of the Scattered Literature.*

Beittel, K., & Lowenfeld, V. (1959). Interdisciplinary criteria of creativity in the arts and sciences.

Benedek, M. (2018). Internally directed attention in creative cognition.

Benedek, M., Muhlmann, C., Jauk, E., & Neubauer, A. C. (2013). Assessment of Divergent Thinking by means of the Subjective Top-Scoring Method: Effects of the Number of Top-Ideas and Time-on-Task on Reliability and Validity.

Benedek, M., Stoiser, R., Walcher, S., & Korner, C. (2017). Neural basis of functional fixdness during creative idea generation: an EEG study.

Besemer, S. P., & O'Quin, K. (1987). Creative product analysis: Testing a model by developing a judging instrument.

Besemer, S. P., & O'Quin, K. (1999). Confirming the three-factor creative product analysis matrix model in an American sample.

Besemer, S. P., & Treffinger, D. J. (1981). Analysis of creative products: Review and synthesis.

Besemer, S., & O'Quin, K. (1986). Analyzing creative products: Refinement and test of a judging instrument.

Bigler, E. D. (2013). *Neuroimaging Biomarkers in Mild Traumatic Brain Injury (mTBI).*

Blohm, L., Bretschneider, U., Leimeister, J. M., & Krcmar, H. (2010). Does collaboration among participants lead to better ideas in IT-based idea competitions? An empirical investigation.

Briggs, R. B., Reinig, M., Shepherd, J., Yen, & et. at. (1997). Quality as a Function of Quantity in Electronic Brainstorming.

Camarda , A., Salvia, E., & Vidal, J. (2018). Neural basis of functiona fixedness during creative idea generation: An EEG study.

Cantamessa, M., & Montagna, F. (2016). *Management of Innovation and Product Development - Integrating Business and Technological Perspective.*

Chassell, L. M. (1916). Tests for originality.

Christiaans, H. H. (2002). *Creativity as a design criterion.*

Chulvi, V., Mulet, E., Chakrabarti, A., Lopez-Mesa, B., & Gonzalez-Cruz, C. (2012). Comparison of the degree of creativity in the design outcomes using different design methods.

Cohen, M. (2018). An overview of neural time series analysis.

Colombo, S. (2019). *Master Thesis: "Brain Activities in Creative Thinking Task".*

Connolly, T., Jessup, L. M., & Valacich, J. S. (1990). Effects of anonymity and evaluative tone on idea generation in computer-mediated groups.

Cropley, A. J. (1968). A note on the Wallach-Kogan tests of creativity.

Cross, N. (2001). Designerly ways of knowing: design discipline versus design science.

Dean, D. L., Hender, J. M., Rodgers, T. L., & Santanen, E. L. (2006). Identifying Quality, Novel, and Creative Ideas: Constructs and Scales for Idea Evaluation.

Diehl, M., & Stroebe, W. (1987). Productivity Loss in Brainstorming Groups: Toward the Solution of a Riddle.

Dietrich, A., & Kanso, R. (2010). A Review of EEG, ERP, and Neuroimaging Studies of Creativity and Insight.

Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., & Vargas-Hernanez, N. (2015). Empirical Studies of Designer Thinking: Past, Present, and Future.

Dippo, C., & Kudrowitz, B. (2015). The effects of elaboration in creativity tests as it pertains to overall scores and how it might prevent a person from thinking of creative ideas during the early stages of Brainstorming and Idea Generation.

Dumas, D., & Dunbar, K. N. (2014). *Understanding fluency and originality: A latent variable perspective.*

Eisenberger, R., & Rhoades, L. (2001). “Incremental Effects of Reward on Creativity,”.

Eisenberger, R., & Selbst, M. (1994). “Does Reward Increase or Decrease Creativity?”.

Elliot, M., Bossom, E., Dupui, M. E., & Masouredis, S. P. (1964). Effect of Ionic Strength on the Serologic Behavior of Red Cell Isoantibodies.

Evans, R. G., & Forbach, G. B. (1983). Facilitation of performance on a divergent measure of creativity: A closer look at instructions to" be creative".

Eysenck, H. J. (1993). Creativity and Personality: Suggestions for a Theory.

Faure, C. (2004). Beyond Brainstorming: The Effects of Different Group Procedures on Selection of Ideas and Satisfaction with the Process.

Fink, A., & Benedek, M. (2014). EEG alpha power and creative ideation.

Fink, A., & Neubauer, A. C. (2006). EEG apha oscillations during the performance of verbal creativity tasks: Differential effects of sex and verbal intelligence.

Fink, A., & Neubauer, A. C. (2008). EEG alpha power and creative ideation.

Fink, A., Benedek, M., Grabner, R. H., Staudt, B., & Neubauer, A. C. (2007). Creativity meets neuroscience: experimental tasks for the neuroscientific study of creative thinking.

Fink, A., Benedek, M., Grabner, R., & Reishofer, G. (2009). The Creative Brain: Investigation of Brain Activity During Creative Problem Solving by Means of EEG and fMRI.

Fink, A., Grabner, R. H., Neubauer, A. C., & Benedek, M. (2009). The Creative Brain: Investigation of Brain Activity During Creative Problem Solving by Means of EEG and fMRI.

Fink, A., Graif, B., & Neubauer, A. C. (2009). Brain correlates underlying creative thinking: EEG alpha activity in professiona vs. novice dancers.

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition.*

Forster, E. A., & Dunbar, K. N. (2009). Creativity evaluation through latent semantic analysis.

Forthmann, B., Oyebade, O., Gunther, F., & Holling, H. (2018). Application of Latent Semantic Analysis to Divergent Thinking is Biased by Elaboration.

Freeman, W., & Quian, R. (2013). Imaging brain function with EEG.

Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., & et al. (1992). Electronic Brainstorming and Group Size.

Gero, J. S., & Kannengiesser, U. (2004). The situated function-behaviour-structure framework.

Glass, H. C., Prieur, B., Molnar, C., & Hamiwka, L. (2006). MICTURITION-INDUCED REFLEX EPILEPSY: A PEDIATRIC CASE AND REVIEW OF THE LITERATURE.

Gough, H. G. (1961). *Techniques for identifying the creative research scientist.*

Gray, D., & Macanufo, J. (2010). *Gamestorming: a playbook for innovators, rulebreakers, and changemakers.*

Gryskiewicz, S. (1980). *A study of creative problem solving techniques in group settings.*

Guilford, J. P. (1950). Creativity.

Guilford, J. P. (1957). Creative abilities in the arts.

Guilford, J. P. (1966). Measurement and creativity.

Guilford, J. P. (1967). Creativity: Yesteray, Today, Tomorrow.

Guilford, P. J. (1967). The Nature of Human Intelligence.

Harbison, J. L., & Haarmann, H. (2014). Automated scoring of originality using semantic representations.

Hargreaves. (1927). The "Faculty" of imagination: An Enquiry concerning the Existence of General "Faculty," or Group Factor of Imagination.

Hay, L., Duffy, A., & Grealy, M. (2019). THE NOVELTY PERSPECTIVES FRAMEWORK: A NEW CONCEPTUALISATION OF NOVELTY FOR COGNITIVE DESIGN STUDIES.

Hay, L., Duffy, A., McTeague, C., Pidgeon, L., Vuletic, T., & Grealy, M. (2017). A systematic review of protocol studies on conceptual design cognition: Design as search and exploration.

Henderson, R. M., & Clarks, K. B. (1990). Architectural innovation: The reconfiguration of existing.

Hennessey, B. A. (1994). The consensual assessment technique: An examination of the relationship between ratings of product and process creativity.

Hocevar, D., & Michael, W. B. (1979). *The effects of scoring formulas on the discriminant validity of tests of divergent thinking.*

Horn, D., & Salvendy, G. (2009). Measuring Consumer Perception of Product Creativity: Impact on Satisfaction and Purchasability.

Jackson, P. W., & Messick, S. (1965). The person, the product, and the response: conceptual problems in the assessment of creativity.

Janckle, L. (2018). Sex/gender differences in cognition, neurophysiology, and neuroanatomy.

Jauk, E., Benedek, M., & Neubauer, A. C. (2012). Tackling creativity at its roots: Evidence for different patterns of EEG alpha activity related to convergent and divergent modes of task processing.

Johnson, D. G., Genco, N., Saunders, M. N., Williams, P., Seepersad, C. C., & Holtta-Otto, K. (2014). An experimental investigation of the effectiveness of empathic experience design for innovative concept generation.

Johnson, T. A., Caldwell, B. W., & Green, M. G. (2015). INVESTIGATING SPONTANEOUS FLEXIBILITY IN CONCEPT GENERATION.

Jung, T. P., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (1995). Removing electroencephalographic artifacts: comparison between ICA and PCA.

Kaufman, J. C., Baer, J., & Cole, J. C. (2009). Expertise, domains, and the consensual assessment technique.

Klimesch, W. (2012). α-band oscillations, attention, and controlled access to stored information.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis.

Kramer, M. W., & Kuo, C. L. (1997). The Impact of Brainstorming Techniques on Subsequent Group Processes.

Krampen, O. (1993). *Effects of Job Application Briefing and Extraction of Subscales in the Assessment with Questionnaires.*

Kudrowitz, B. M., & Wallace, D. (2012). *Assessing the quality of ideas from prolific, early-stage.*

Kudrowitz, B. M., & Wallace, D. (2013). *Assessing the quality of ideas from prolific, early-stage product ideation.*

Lauritzen, A. G., & Wolford, L. W. (1964). Occlusal relationships: The split-cast method for articulator techniques.

Linsey, J. S., Clauss, E. F., Kurtoglu, T., Murphy, J. T., Wood, K. L., & Markman, A. B. (2011). An Experimental Study of Group Idea Generation Techniques: Understanding the Roles of Idea Representation and Viewing Methods.

Lubart, T. I. (2001). *Models of the creative process: Past, present and future.*

MacCrimmon, K. R., & Wagner, C. (1994). Stimulating ideas through creative software.

MacKinnon, D. W. (1961). *Creativity in architects.*

Martindale, C., & Hines, D. (1975). Creativity and cortical activation during creative, intellectual and EEG feedback tasks.

McCarthy, M. M., Arnold, A. P., Ball, G. F., Blaustein, J. D., & De Vries, G. J. (2012). Sex Differences in the Brain: The Not So Inconvenient Truth.

McCloy, W., & Meier, N. C. (1931). Re-creative imagination.

Milgram, R. M., & Arad, R. (1981). Ideational fluency as a predictor of original problem solving.

Mouchiroud, C., & Lubart, T. (2001). Children's original thinking: An empirical examination of alternative measures derived from divergent thinking tasks.

Nagasundaram, M., & Bostrom, R. P. (1994). *The structuring of creative processes using GSS: a framework for research.*

Nusbaum, E. C., & Silvia, P. J. (2011). *Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking.*

O'Hara, L. A. (1999). *13 Creativity and Intelligence.*

Osborn, A. F. (1953). *Applied Imagination: principles and procedures of creative thinking.*

Osborn, A. F. (1963). *Applied imagination: Principles and procedures of creative thinking.*

Pfurtscheller, G., & Lopes da Silva, F. (1999). Event-related EEG / MEG synchronization and desynchronization: basic principles.

Pfurtscheller, G., Nuper, C., Flotzinger, D., & Pregenzer, M. (1997). EEG-based discrimination between imagination of right and left hand movement.

Potter, R. E., & Balthazard, P. (2004). The Role of Individual Memory and Attention Processes During Electronic Brainstorming.

Pugh, S. (1981). *Concept selection: a method that works.*

Puppo, F. (2018). An Optimized Structure-Function Design Principle Underlies Efficient Signaling Dynamics in Neurons, Scientific Reports.

Purcell, A. T., & Gero, J. S. (1998). Drawings and the design process: A review of protocol studies in design and other disciplines and related research in cognitive psychology.

Radel, R., Davranche, K., Fourier, M., & Dietrich, A. (2015). The role of (dis) inhibition in creativity: Decreased inhibition improves idea generation.

Ramos-Loyo, J., & Guevara, M. (1993). *Gender Differences in the EEG During Cognitive Activity.*

Razumnikova, O. (2004). Gender differences in hemispheric organization during divergent thinking: An EEG investigation in human subjects.

Redmond, M. R., Mumford, M. D., & Teach, R. (1993). Putting Creativity to Work: Effects of Leader Behavior on Subordinate Creativity.

Runco, M. A. (2010). *Divergent Thinking, Creativity, and Ideation.*

Runco, M. A., & Jaeger, G. J. (2012). *The standard definition of creativity.*

Runco, M. A., Johnson, D. J., & Bear, P. K. (1993). Parents' and teachers' implicit theories of children's creativity.

Santosh, J. (2019). Design creativity: refined method for novelty assessment.

Sarkar, P., & Chakrabarti, A. (2006). *Assessing degree of novelty of products to ascertain innovative products.*

Sarkar, P., & Chakrabarti, A. (2011). Assessing design creativity.

Sawant, H., & Zahra, J. (2010). Detection and classification of EEG waves.

Sawyer, K. (2006). *Educating for innovation.*

Sawyer, K. (2011). The cognitive neuroscience of creativity: a critical review.

Schön, D. A. (1983). *Reflective practitioner—how professionals think in action.*

Schön, D. A. (1995). *Knowing in action: the new scholarship requires a new epistemology.*

Schwab, D., Benedek, M., Papousek, I., Weiss, E. W., & Fink, A. (2014). *The time-course of EEG alpha power changes in creative ideation.*

Shah, J. J., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness.

Shah, J. J., Kulkarni, S. V., & Vargas-Hernandez, N. (2000). Evaluation of Idea Generation Methods for Conceptual Design: Effectiveness Metrics and Design of Experiments.

Sharbrough, F., Chatrian, G. E., Lesser, R., Luders, H., Nuwer, M., & Picton, T. W. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature.

Silverstein, D. (2008). The innovator’s toolkit: 50+techniques for predictable and sustainable.

Silvia, P. J., & Willse, J. T. (2008). Assessing Creativity With Divergent Thinking Tasks: Exploring the Reliability and Validity of New Subjective Scoring Methods.

Silvia, P. J., Martin, C., & Nusbaum, E. (2009). *A snapshot of creativity: Evaluating a quick and simple method for assessing divergent thinking.*

Simon, H. (1973). *A. structure of ill structured problems.*

Simon, H. A. (1969). The science of artificial.

Snelders, D., & Hekkert, P. (1999). *Association measures as predictors of product originality.*

Soraghan, C. (2010). Development of A Versatile Multichannel CWNIRS Instrument for Optical Brain-Computer Interface Applications.

Srivathsaval, R., Genco, N., Holtta-Otto, K., & Seepersad, C. C. (2010). Study of existing metrics used in measurement of ideation effectiveness.

Straus, S. G., & McGrath, J. E. (1994). Does Medium Matter? The Interaction of Task Type and Technology on Group Performance and Member Reactions.

Subasi, A., & Gursoy, M. (2011). EEG signal classification using PCA, ICA, LDA and support vector machines.

Tangkeaingkij, P., Lursinsap, C., Sanguansintukul, S., & Desudchit, T. (2009). Selecting Relevant EEG Signal Locations for Personal Identification Problem Using ICA and Neural Network.

Teplan, M. (2002). FUNDAMENTALS OF EEG MEASUREMENT.

Torrance, E. P. (1974). Norms technical manual: Torrance Tests of creative thinking.

Torrance, E. P. (2008). *Torrance Tests of Creative Thinking: Figural Forms A und B.*

Ulrich, K. T., & Eppinger, E. D. (1995). *Product design and development.*

Unguranu, M., Bigan, C., Strungaru, R., & Lazarescu, V. (2004). Independent Component Analysis Applied in Biomedical Signal Processing.

Valacich, J. S., & Shearer, T. (1995). Devil's Advocacy and Dialectical Inquiry Effects on Face-to-Face and Computer-Mediated Group Decision Making.

Valacich, J. S., Dennis, A. R., & Nunamaker, J. F. (1992). Group Size and Anonymity Effects on Computer-Mediated Idea Generation.

Verhaegen, P. A., Verhaegen, D., Peeters, J., & Duflou, J. R. (2013). Refinements to the variety metric for idea evaluation.

Wallace, H. R. (1961). Creative thinking: a factor in sales productivity.

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

Warr, A., & O'Neill, E. (2005). *Understanding Design as a Social Creative Process.*

Wilson, R. C., Guilford, J. P., Christensen, P. R., & Lewis, D. J. (1954). A factor-analytic study of creative-thinking abilities.

Yen, C. C., & Jiang, H. (2009). Protocol Analysis in Design Research: a review.

Yoruk, S., & Runco, A. (2014). THE NEUROSCIENCE OF DIVERGENT THINKING.

Yoruk, S., & Runco, M. A. (2014). THE NEUROSCIENCE OF DIVERGENT THINKING.

Younan, P. (2012). Master thesis "Kompression von EEG-Daten mittels räumlich harmonischer Zerlegung".

Zarnegar, Z., Hocevar, D., & William, B. M. (1988). Components of original thinking in gifted children.

Sitography

<<Sistema NeuralScan>>, NeuralScan, http://5inc-elements.com/new_5inc/neuralscan.php

<<Introduction to Neuroscience: Neurotransmitters>>, THE STUDENT SCIENTIST, https://thestudentscientist.org/2018/07/06/introduction-to-neuroscience-neurotransmitters/