INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED15
27-30 JULY 2015, POLITECNICO DI MILANO, ITALY

REQUIREMENTS CHECKLISTS: BENCHMARKING THE COMPREHENSIVENESS OF THE DESIGN SPECIFICATION

Becattini, Niccolo (1); Cascini, Gaetano (1); Rotini, Federico (2)
1: Politecnico di Milano, Italy; 2: Università degli Studi di Firenze, Italy

Abstract

It is commonly recognized that the definition of product requirements is an essential step of any design process. Many techniques have been proposed for building a suitable design specification, i.e. for defining a set of requirements characterized by validity, completeness, operationality, non-redundancy, conciseness and practicability.

Among them, several methods and tools primarily aim at populating the design specification: some of them focus on very specific objectives but are applicable in many different domains (e.g., Design for X). Others are domain specific, but try to cover the entire scope of the specification (e.g., checklists and standards).

This paper describes an abstract-level checklist for requirements definition, suitable for any field of application, aiming at producing exhaustive lists of requirements.

A previous experimental application with Mechanical Engineering students clearly showed that the proposed multi-purpose checklist allows populating design specifications more complete than those defined without any support. This paper follows up demonstrating the capability of the novel checklist against the checklist for conceptual design by Pahl and Beitz.

Keywords: Requirements, Design methods, Design education, Design specification, Requirements checklists

Contact:
Dr.-Ing. Niccolo Becattini
Politecnico di Milano
Dipartimento di Meccanica
Italy
niccolo.becattini@polimi.it

Please cite this paper as:
Surnames, Initials: *Title of paper*. In: Proceedings of the 20th International Conference on Engineering Design (ICED15), Vol. nn: Title of Volume, Milan, Italy, 27.-30.07.2015

1 INTRODUCTION

A proper definition of requirements is the key for an effective and efficient innovation and development process. Such a process requires the designer to search solutions to problems in a space of alternatives that, in turn, has to be limited by the goals to be achieved and ruled by the laws governing the situation at hand (Simon, 1973). In this reference, both problem state and goal state can be characterized by means of requirements: in the former, they get unsatisfactory values, while in the latter they change so as to make the problem solved or not relevant anymore.

In other words, they represent the target conditions to be satisfied and therefore they both support the development of design proposals as well as the evaluation of solutions, for which they can be considered as benchmarking criteria for the related decision-making (e.g. Cross, 2008).

Roozenburg and Eekels (1995) explained the relationship that requirements hold between system characteristics and user's needs and expectations, through an explicit hierarchy that ranges from overall goals, to more specific objectives till the definition of requirements. To this purpose, they also defined the characteristics that a design specification (i.e. the set of requirements) should have in order to be really useful for the development processes:

- **Validity** – requirements should define the criteria for assessing the satisfaction of the design objectives;
- **Completeness** – requirements should cover all the potentially relevant objectives;
- **Operationality** – requirements should be measurable with reference to the objectives;
- **Non-redundancy** – requirements shouldn't be duplicated, if they have the same meaning;
- **Conciseness –** requirements should take into account all the relevant aspects, overlooking the ones that are not relevant;
- **Practicability** - requirements satisfaction should be testable with available information.

Other authors have concentrated their efforts on the development of appropriate design methods and tools for requirement-related purposes. Quality Function Deployment (QFD), for instance, is a largely diffused approach to support the translation of customer requirements into technical requirements, e.g. (Akao, 2004). It is usually presented as part of the House of Quality (HoQ), a graphical and textual model that integrates the analysis of those requirements so that priorities and potential issues can be formulated in a more structured way. Different contributions, moreover, have also aimed at defining appropriate classes in order to characterize requirements and treat them in a more efficient or user-oriented perspective, e.g. the classification of customer requirements by Kano (1984). Other studies, in turn, try to combine the contributions about requirements classification with the existing tools to manage them like QFD, as, e.g., proposed by Matzler et al. (1998).

Among the different methods and tools for dealing with requirements, a relevant role is played by those that primarily aim at populating the design specification. Roughly, they can be organized into two main categories. Those applicable in a great variety of situations and domains of technique but focusing on very specific objectives (e.g. Design for X methods) and domain-specific checklists (sometimes also formalized into standards), which address a wide range of requirements but that present some limitations in the applicability outside the domain they have been developed for. With the overall goal to contribute to the overcoming of this dichotomy, the authors have proposed an abstract-level checklist for requirements definition, suitable for any field of application, but also to produce exhaustive lists of requirements. So far, it has been proficiently used with different purposes in industrial contexts, to start checking its effectiveness, e.g., in order to support benchmarking activities along a project focusing on the substitution of manufacturing technologies (Becattini, 2011); in order to properly define accurate technology descriptions for manufacturing machines using the fluid-bed principle (Becattini, 2013); in order to assess the opportunities of integration with a computer-aided innovation prototype for the prioritization of design problems on the basis of network analysis and conflicting requirements (Becattini and Cascini, 2013).

Moreover, the authors have also recently started a testing activity to statistically evaluate the performance of their proposal and retrieve experimental data so as to also plan further improvements. The experiments carried out so far aimed at checking the performances in populating a design specification with the novel multi-purpose checklist. These tests have been carried out with Mechanical Engineering students and the results have been compared with those of a control group having the same background but with no support by any checklist or other instruments. Two products having different complexity (an iron and a laser printer) have been used as subjects for the definition of requirements along the test. The results of that investigation (Becattini and Cascini, 2014) were presented at the last DESIGN 2014 conference in Dubrovnik and showed that the novel multi-purpose checklist allows people to populate design specifications that are more complete than the ones defined without any support.

This paper follows up that study and aims at checking the capability of the novel checklist against the checklist for conceptual design developed by Pahl and Beitz (2004).

The second Section summarizes the most acknowledged checklists available in literature for defining system requirements. It analyses into further details the characteristics of the two checklists to be benchmarked along this research. Section 3 details the testing methodology and the educational context in which it has been carried out, in order to also facilitate analogous studies in this field. The results of the benchmark are presented and discussed in the fourth Section. The concluding Section summarizes the main outcomes of the research and delineates further investigations to enrich this research.

2 CHECKLISTS FOR POPULATING A DESIGN SPECIFICATION

As mentioned in the introduction, several methods and tools exist of different nature and kind to cope with requirements along the design process. In order to keep the investigation homogenous in terms of tools to be compared, i.e. tools having the same purpose, this Section will exclusively focus the attention on checklists. Pahl and Beitz (latest edition, 2004) originally developed two checklists to guide the identification of requirements for supporting the design process in the conceptual and the embodiment design phases of their systematic approach. Hales and Gooch (2004), several years after updated those checklists and introduced a new one specifically tailored for the detail design stage of their systematic approach. Eder and Hubka (1988), as well as Pugh (1990), have also proposed their own checklist to identify requirements in design processes. An excerpt of Pugh's checklist is available in Roozenburg and Eekels (1995). This said, the following two subsections respectively introduce the requirements checklist for conceptual design by Pahl and Beitz and the abstract-level multi-purpose requirements checklist that has been developed by the authors.

2.1 Pahl and Beitz's checklist for conceptual design

Pahl and Beitz's checklist (2004) have been chosen as a reference for the benchmark for two main reasons. First, the testers participating the study had just started attending a course essentially structured according to the Pahl & Beitz design methodology, thus they were already familiar with its overall logic (not yet with requirements checklists).

Main headings	Examples
Geometry	Size, height, breadth, length, diameter, space requirement, number, arrangement, connection, extension
Kinematics	Type of motion, direction of motion, velocity, acceleration
Forces	Direction of force, magnitude of force, frequency, weight, load, deformation, stiffness, elasticity, inertia forces, resonance
Energy	Output, efficiency, loss, friction, ventilation, state, pressure, temperature, heating, cooling, supply, storage, capacity, conversion.
Material	Flow and transport of materials. Physical and chemical properties of the initial and final product, auxiliary materials, prescribed materials (food regulations etc)
Signals	Inputs and outputs, form, display, control equipment.
Safety	Direct safety systems, operational and environmental safety.
Ergonomics	Man-machine relationship, type of operation, operating height, clarity of layout, sitting comfort, lighting, shape compatibility.
Production	Factory limitations, maximum possible dimensions, preferred production methods, means of production, achievable quality and tolerances, wastage.
Quality control	Possibilities of testing and measuring, application of special regulations and standards.
Assembly	Special regulations, installation, siting, foundations.
Transport	Limitations due to lifting gear, clearance, means of transport (height and weight), nature and conditions of despatch.
Operation	Quietness, wear, special uses, marketing area, destination (for example, sulphurous atmosphere, tropical conditions).
Maintenance	Servicing intervals (if any), inspection, exchange and repair, painting, cleaning.
Recycling	Reuse, reprocessing, waste disposal, storage
Costs	Maximum permissible manufacturing costs, cost of tooling, investment and depreciation.
Schedules	End date of development, project planning and control, delivery date

Figure 1. Pahl and Beitz's requirements checklist

Second, none of the testers had already got in touch with requirements' checklists before the test (this enables the participants to be fresh to both approaches). The checklist has to be used before starting the conceptual design phase when clarifying the task, so as to drive the exploration of solutions and the generation of ideas and concepts accordingly. Actually, Pahl and Beitz, about their checklist just say: "*The checklist […] is a generic one based on ideas described in [a previous] Section. The items in this list are checked against the existing task and its requirements in order to obtain further requirements*" (p. 151, 3rd ed.). The checklist (Figure 1) is globally organized in 17 classes ranging different structural, behavioural and functional features along the lifecycle phases of the product to be developed, with an overall set of 92 items to be checked.

2.2 An abstract-level multi-purpose checklist

As briefly mentioned in the introduction, this checklist has been developed at an abstract level so as to overcome the dichotomy between approaches that are capable of mapping a wide set of requirements in a very specific field of investigation and those focused on a specific target but can be applied in any field.

Furthermore, the multi-purpose checklist is tailored to more effectively support conceptual design tasks. Indeed, it starts defining the detail level of the analysis by identifying the functional outcome that is expected (the product of the function) and the initial conditions of the product before the system to be developed carries out its function. Both the initial and the final condition of the product of the function are represented by means of the Energy-Material-Signal model (Pahl et al., 2004).

Three main axes of investigation are considered along the checklist. They have been defined in accordance with the main drivers defined by the 4th Law of Engineering Systems Evolution, from TRIZ theory. The law says that systems evolve towards the increase of their ideality, where ideality is conceived as the sum of useful functions (UF) the system delivers divided by the sum of harmful functions (HF) it produces or undergoes and the costs (C) to make it work (Equation 1). In other words, technical systems evolve towards the improvement of their performances, the reduction of undesired effects they trigger or suffer and the reduction of any kind of expenses (Altshuller, 1984).

$$Ideality = \frac{UF}{HF + C} \tag{1}$$

This checklist is also conceived as an instrument to enhance thinking instead of purely "checking" the requirements. It aims at stimulating reflections that can support the generation of non-obvious requirements, since trivial ones do not need any methodological tool for their identification. The three axes concerning performances (related to useful functions), side effects (related to harmful functions) and resources consumptions (related to costs) are, therefore, organized so that the designer needs to contextualize the abstract criteria to the specific situation she/he is dealing with, potentially enlightening unexplored alternatives.

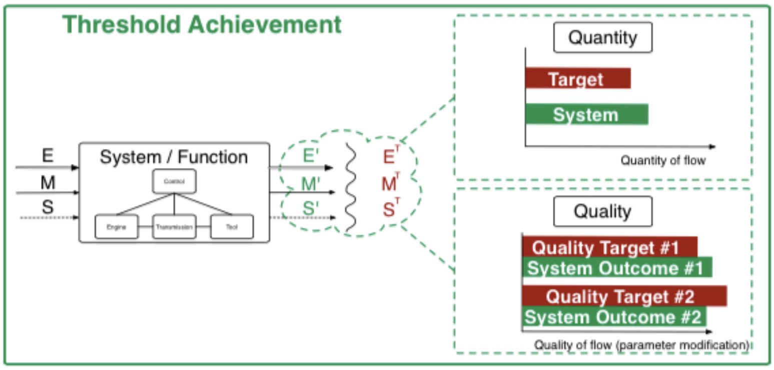
The dimension of performances as a result of useful functions is explored in order to check both the quantitative and the qualitative aspects of the function, according to the following sub-dimensions:

- Threshold achievement: requirements describing the capability to impact the object of the function with the expected extent (see Figure 2);
- Versatility: requirements characterizing the capability to adapt the behaviour of the technical system according to different operating conditions;
- Robustness: requirements accounting the capability of the technical system to obtain the same (stable in values) desired outcome under varying inputs;
- Sensitivity to external conditions: requirements concerning the capability of the technical system in carrying out its function regardless of the conditions of the environment in which it is immersed;
- Controllability: requirements about the capability to set system characteristics and parameters so as to obtain a desired result according to user's will.

The side effects generated by harmful functions are considered from three different viewpoints:

- Impact on the object of the Main Useful Function (e.g. an undesired side effect caused by the same mechanism adopted to deliver the function or as its consequence);
- Impact on the system and subsystems integrity (e.g. an undesired side effect on the technical system as a whole or on its parts);

- Impact on the external environment (e.g. an undesired side effect that compromises some environmental conditions or damages some of the elements that pertains to the world in which the technology/technical system is immersed in).



Product quality: each function changes or stabilize one of the object characteristics to obtain a specific outcome. This criterion aims at defining requirements to measure the system capabilities to achieve the results defined as target (e.g, in a washing machine.: amount of residual stains on clothes, amount of residual odors,...)

Product quantity: Parameters to measure the functional performances in terms of quantity of processable functional outcomes. (e.g., in a washing machine: amount of clothes in a single washing cycle).

Figure 2. An example of abstract criteria for the identification of system requirements (Threshold values for the main function of the system). It is provided as a visual representation, a textual description and an example.

Finally, the costs due to the consumption of resources are further sub-classified into five categories: Resources of Time, Information/Knowledge, Material, Energy and Space
Overall, the abstract-level multi-purpose checklist is organized into three main classes, which are further organized into 13 subclasses for which a set of 53 examples are provided so that one can be facilitated in the contextualization and more effectively inspired for requirements definition. The complete description of the checklist cannot be presented here for space reasons, but it is completely available in Becattini (2013).

The next Section presents the approach followed to benchmark the two checklists against each other in order to evaluate their performance in producing more complete design specifications.

3 TESTING METHOD

The testing methodology has been developed consistently with the objective of comparing the performances of two checklists to populate a complete design specification. In order to properly design the experiment, the authors started identifying the method for data analysis. Indeed, it has to be chosen consistently with the objectives of the research in order to meaningfully identify the test dynamic and the profile of participants. The approach to carry out and analyse the results of the test allows further benchmarking studies to be carried out on a similar basis, so as to also compare results coming from different testers with uniform criteria. This test exclusively focuses on the amount of requirements generated by means of the competing checklists, as a first attempt to evaluate the completeness of a design specification. The authors are planning to extend and refine this test in order to cover a wider range of facets concerning the design specification (e.g.: presence of duplicates, capability to formulate requirements so as to make them measurable and operational, etc.). However, the test dynamic presented in Section 3.2 allows experimenters to extend its applicability to other characteristics of the design specification (e.g. benchmarking non-redundancy, conciseness, etc.) after a small adaptation.

3.1 Test data management and related approach for the analysis.

Once the different contributions have been gathered, the data contained therein has to be extracted and organized to make them suitable to determine if and to which extent one of the alternative checklists overtakes the performances of its competitors. In this perspective, one-to-one comparisons can be conveniently carried out through statistical hypothesis testing. Beyond the specific checklists to be compared in this study, the two alternative hypotheses can be formulated in the following way:

• *H_0: the difference in the average values depends on chance and the <checklist X> do not provide any statistically significant contribution to the definition of requirements if compared to the <checklist Y> (H0: $\mu_x = \mu_y$)*
• *H_a: the difference in the average values depends on the better performances obtained with the <checklist X> against the <checklist Y>. (Ha: $\mu_x > \mu_y$).*

For the present study, the author's checklist is meant as <checklist X>. The checklist by Pahl and Beitz is <checklist Y> and, in general, there should be an appropriate amount of testers per each of the checklist to be benchmarked. In this reference, grouping by randomization is necessary before testing and each group should work with a checklist at a time. Data samples can be, therefore, organized according to the specific checklist the testers have used.
Moreover, it is necessary to carry out the hypothesis test by using the most suitable test statistics that is compliant with the sample composition. In this case the authors suggest using the t-statistics by Student, because of the poor knowledge about the population (e.g.: how much is it dispersed?) even if the individuals in it share a common background. The value of the test statistics can be calculated according to Equation (2) and (3), since they are commonly adopted in cases where hypothesis testing has to be carried out between two samples pooled together, with a cumulative number of participants close or higher than 30 and under the assumption that the two standard deviations of the related populations are equal (as it can be assumed, considering the homogeneous distribution of students in the same classes and a process of grouping by randomization). Specifically, the t-statistics is:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p \sqrt{1/n_1 + 1/n_2}} \tag{2}$$

where x are the observed average values for the distributions, d_0 is the hypothesized difference between the average values of requirements obtained with the two approaches to be benchmarked, as for the null hypothesis ($d_0 = \mu_1 - \mu_2 = 0$) and s_p^2 (pooled variance) is calculated as follows:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \tag{3}$$

In case the data allow an association of each of the results to a specific participant, paired t-tests (not among groups, but within the same group of participants) can be carried out by considering that, on the one hand, they will allow a more precise evaluation of differences by considering the individual response to the treatments under investigation. On the other hand, these results present some limitation in validity if not carried out in different testing round, focusing on the same system for which requirements have to be defined. This last approach is not part of this analysis and will be considered for further studies.

3.2 Test Dynamic

The following procedure allows the test to be carried out following the same steps used by the authors, so as to guarantee internal rigour of the analysis and interoperability of data with tests run in different contexts and with different participants. In brackets, details are reported about the specific test carried out for the benchmark whose results are presented in Section 4.

1. Randomly divide the group of participants into as many groups as the number of alternative requirements checklist to be benchmarked.
   (The 29 students randomly formed two groups. A: 15 people; B: 14, so as).
2. Define a number of testing rounds that is equivalent to the number of systems for which the

participants are required to define a design specification. Account one hour per each round.
(2 rounds were established for this study: during the first one, the system to be considered was an iron; during the second round it was a laser printer).

3. During each round, the groups focus the attention on the same technical system as the subject for defining requirements. Each of the groups work with a requirements checklist that is different from the others.
(The group A first works with the Pahl and Beitz checklist, while B with the checklist proposed by the authors. For the second round, the two groups switch the checklist to be used).
4. Considering the perspective of a benchmarking analysis, provide the testers with the checklist as the original reference sources in literature propose to use them. The media, e.g. paper, text file, spreadsheet,etc, on which the contributions are produced have to be the same in the different rounds, in order to avoid biases due to usability issues depending on the media.
(The test was done simply with pen and paper for both the checklists. Pahl and Beitz's checklist have been proposed as a single printed sheet with an additional sheet where students could write down requirements and, not compulsorily, the category they had been identified from. The checklist proposed by the authors appears as a more consistent document of 3 pages printed on both sides. The criteria were presented with textual and visual explanation for each class of requirements and specific subclasses and sufficient room to write down one or more requirement, upon their identification. Examples in the field of washing machines aim at clarifying the meaning of the proposed criteria.)
5. Gather the outputs from all the participants before switching to the next round, if any, or conclude the testing session.
(The second round has started immediately after the first one).
6. Count the requirements per each of the gathered outputs, without introducing any adjustment or element of evaluation in order to avoid biases due the perspective of the investigation. Each output specification gets a score that corresponds to the amount of requirements that the tester has defined.
(The results have been manually counted; the related data have been organized in spreadsheets and double-checked to verify the correct transcriptions of numbers).
7. Organize the results dividing the contribution according to the specific theme on which to focus the attention, i.e. by round of the testing session, and by the specific checklist used to identify system requirements, i.e. by the specific group that participated the investigation.
(Results are presented as aggregated data in Section 4).
8. Run the statistical analysis described in Section 3.1, so as to verify the hypothesis on an appropriate couple of samples at a time. The two samples should come from different groups, but collected during the same round of investigation.
(Results are presented in Section 4)
9. Define conclusions out of the results of the investigation, considering the degree of statistical significance.
(The results are discussed in Section 4 and summarized in Section 5).

3.3 Participants

The test has been carried out in an educational environment so as to gather a sufficient number of participants and carry out a significant test from which to draw more robust fact-based evidences.
The mechanical engineering MS (mechanical design major) students, 29 volunteers that participated the test, attended a course on "Product Development and Engineering" (Italian name: "Sviluppo e Ingegnerizzazione del prodotto") held at the University of Florence - Italy. Their background is largely homogenous considering that they all hold a BS in mechanical engineering. Along the course, the students get introduced to a systematic approach to design, which is fundamentally grounded on the work by Pahl and Beitz (2004). Different tools and techniques for the different phases of design (Product Planning, Conceptual and Embodiment Design) are taught along the course, in order to facilitate the students in learning both the overall structure of a systematic approach and the tools to deal with the different issues popping up as the design deliverable get defined and refined. This sample is a convenient and relevant one, because it is composed by people that already have higher education knowledge in mechanical design, even if to be further improved and applied on the field, thus representing right candidates for using these instruments in their design activities. Moreover, at

the moment the test was carried out, students were just exposed to the overall structure of Pahl and Beitz's design methodology and the test worked as a way to introduce them to the application of design specification. Being a candidate user of the checklist should represent the discriminating factor for determining the suitability of a person to participate the study.

4 RESULTS OF THE TEST

In order to avoid language related biases, the test has been submitted to students with checklist translated from English to Italian without altering the content of the original sources.
The results of the two rounds of testing have been collected, the requirements counted and the related aggregated data as descriptive statistics are collected in Table 1 (Pahl and Beitz's checklist) and Table 2 (authors' checklist). It is also worth recalling that the test aims at understanding if the author's checklist is capable to overcome the performances of Pahl and Beitz's checklist to identify design requirements.

Table 1. Summary of the results obtained with the Pahl and Beitz checklist

Theme	MAX	min	Average	St. Dev.	Sample size	Round
Iron	41	13	23,53	8,92	14	1
Laser Printer	51	21	30,43	9,84	15	2
Total	51	13	26,86	9,85	29	-

Table 2. Summary of the results obtained with the author's checklist

Theme	MAX	min	Average	St. Dev.	Sample size	Round
Iron	64	32	50,21	8,63	15	1
Laser Printer	63	17	41,13	14,17	14	2
Total	64	17	45,52	12,50	29	-

To this purpose, the following analysis just considers the results from a quantitative perspective, taking into account a more complete design specification, which is populated by a larger number of requirements.
A quick analysis of this descriptive statistics shows that, in general, the authors' checklist has performed better in both the rounds, regardless of the specific theme of its application. The average value for authors' checklist generated requirements (50,21) doubles the ones obtained with Pahl and Beitz's checklist (23,53 requirements) when the tester considered an iron as the focus of the investigation. The test with the laser printer, in turn, shows a similar but less marked behaviour (41,13 requirements and 30,43 respectively with the authors' and Pahl and Beitz's checklist). Moreover, also in terms of variability the two cases present some differences. While the standard deviations get analogous values for the results obtained with the iron, the situation with the laser printer radically changes. The variability of the sample using the authors' checklist is almost 50% bigger than the variability obtained with its own competitor.
In order to verify if and to which extent the current results are significant from a statistical point of view, the hypothesis test becomes necessary. It has been carried out as detailed in Section 3.1.

Table 3. Summary of the test of hypothesis: t-test statics and related p-values.

<table>
<tr><th>Theme</th><th>Checklist</th><th>Sample size</th><th>Degrees of freedom</th><th>S^2_p</th><th>μ_1-μ_2</th><th>t-test value</th><th>p-value</th></tr>
<tr><td rowspan="2">Laser Printer</td><td>Authors'</td><td>14</td><td rowspan="2">27</td><td rowspan="2">76,93</td><td rowspan="2">26,68</td><td rowspan="2">8,19</td><td rowspan="2">0,000 (***)</td></tr>
<tr><td>Pahl and Beitz's</td><td>15</td></tr>
<tr><td rowspan="2">Iron</td><td>Authors'</td><td>15</td><td rowspan="2">27</td><td rowspan="2">146,93</td><td rowspan="2">11,70</td><td rowspan="2">2,38</td><td rowspan="2">0,012 (**)</td></tr>
<tr><td>Pahl and Beitz's</td><td>14</td></tr>
<tr><td rowspan="2">Total</td><td>Author's</td><td>29</td><td rowspan="2">56</td><td rowspan="2">126,69</td><td rowspan="2">18,66</td><td rowspan="2">6,31</td><td rowspan="2">0,000 (***)</td></tr>
<tr><td>Pahl and Beitz's</td><td>29</td></tr>
</table>

According to the results of the hypothesis test (last two columns of Table 3) for the laser printer, it clearly emerges that what was qualitatively assessed with the previous intuitive analysis is also confirmed with statistical significance (p-value <0,005➔***; p<0,01➔**; p<0,05➔*). The null hypothesis has to be rejected in favour of the alternative one with almost no chances to be wrong: the author's checklist overtakes the performances of Pahl and Beitz's checklist.
For what concerns the iron, the p-value states that it is also possible to reject the null hypothesis. Nevertheless, due to the smaller differences between average values and a larger dispersion of data, the probability of being wrong in rejecting the null hypothesis is not null, even if close to 1%.
On the basis of the two results for the two specific rounds of testing, the authors also approached the investigation from a more generic perspective. It means that, consistently with the last row of Table 3, a statistical analysis also for the total of the answers retrieved after both the round of testing. It allows to get further insights about the effectiveness of the checklist, given the (small) differences of results between the two case studies. Moreover, this kind of analysis provide results which are much more independent from the domain of application and the complexity of the system for which requirements are to be defined. The overall analysis not surprisingly shows a behaviour that falls in between the two product-oriented ones. Nevertheless, the overall distribution of data between the two couples of elementary samples in terms of difference of average values and dispersion suggests rejecting the null hypothesis with a probability to be wrong that is practically null. The test demonstrates with statistical significance that the authors' abstract-level multi-purpose checklist overtakes the performances of the one by Pahl and Beitz in populating a more complete design specification.

5 CONCLUSIONS

The paper aims at demonstrating the better performances of a recently developed multi-purpose requirement checklist in populating a design specification. The benchmark was carried out against the Pahl and Beitz's checklist to support conceptual design. This test follows previous studies where the authors' checklist were compared to the intuitive and unsupported definition of requirements. The test has been carried out on 29 students of the mechanical design major of the MS course in mechanical engineering at the University of Florence. The subjects are considered relevant for the analysis, being them potential users of that tool in their working practice. A method to carry out the investigation is defined as a sequence of steps in order to make it reusable for analogous studies. The retrieved data are quantitatively analysed in order to carry out t-pooled hypothesis tests and verify the objective with statistical significance. Two rounds of test have been carried out on technical systems of common use and different complexity (an iron and a laser printer). The results definitely shows that the author's multi-purpose checklist provides better results in terms of completeness of the design specification, measured as the number of requirements it collects.
It is also worth mentioning that, in terms of efforts, the authors' checklist requires higher efforts that during the testing session have been measured in a bigger amount of time to complete the test by the participants (about 30% more, but still within an hour). It might probably depend on the different nature of the checklists: Pahl and Beitz's checklist collects items that can be relevant or not. The authors' checklist, in turn, requires longer time to be read and the relevant concepts are not presented as a list of terms, but with textual, graphical description and examples. The time required for thinking and generating requirements out of the checklist cannot be separated so easily from the time for reading. Tailored talk-aloud protocol analysis tests with individuals might allow this issue to be solved in the future, even if these tests cannot be carried out in classes and require a more accurate setting (e.g. recorders, etc.)
Considering the issue of retrieving testers to carry out statistically significant analysis in many design researches, the proposed testing approach should be considered a strength point of the research. Even if focused on a very specific objective, it can be reused as a common way for carrying out tests, thus allowing for the interoperability of the data among different researchers operating in the field.
For what concerns the further development of this research, the authors also aim at deepening the analysis of the retrieved data with a more fine-grained approach. Data will be cleaned and redundant requirements will be counted and removed for each contribution, so as to also check this specific facet among the characteristics defined by Roozenburg and Eekels about the appropriateness of the design specification.

Moreover, t-paired test will be carried out in order to see if significant shift appears considering the same tester using two different approaches for the same technical system, in order to detect a potential superposition of the two approaches and the related effects. This investigation might be also enriched with benchmark against different requirements checklists as the one mentioned in the second chapter.

REFERENCES

Akao Y.; (2004); Quality Function Deployment: Integrating Customer Requirements Into Product Design; Productivity Press.

Altshuller, G.S.; (1984); “Creativity as an Exact Science: The Theory of the Solution of Inventive Problems”. Gordon and Breach Science Publishers, ISBN 0-677-21230-5, (original publication in Russian - 1979).

Becattini N., Cascini G., Petrali P., Pucciarini A.; (2011); “Production processes modeling for identifying technology substitution opportunities”- Published on the proceedings of the 11th TRIZ Future Conference 2011 – Dublin– Ireland (2nd-4th Nov).

Becattini, N., & Cascini, G. (2013). Mapping Causal Relationships and Conflicts among Design Parameters and System Requirements. Computer-Aided Design and Applications, 10(4).

Becattini N. (2013) Requirements Identification and Characterization in Innovation Process. PhD Dissertation, Politecnico di Milano - Italy

Becattini N., Cascini G., (2014) General-purpose requirements checklist for improving the completeness of a design specification. DESIGN 2014 Conference - Cavtat- Dubrovnik - Croatia May 19-22nd.

Cross, N. Engineering Design Methods: Strategies for Product Design. Fourth edition. Chichester: John Wiley and Sons Ltd.; 2008

Gero, J.S.; (1990); ‘Design prototypes: a knowledge representation schema for design’ AI Magazine Vol 11 No 4 p.26–36

Hales C., Gooch S., (2004) Managing Engineering Design, Springer Science & Business Media

Hubka, V. and Eder, W.E. (1988) Theory of Technical Systems: A Total Concept Theory for Engineering Design, Berlin: Springer.

Kano N.; Seraku N., Takahashi F., Tsuji S.; (1984); "Attractive quality and must-be quality" (in Japanese). Journal of the Japanese Society for Quality Control 14 (2): 39–48.

Matzler K.; Hintertuber H.H.; (1998); How to make product development projects more successful by integrating Kano’s model of customer satisfaction into quality function deployment; Technovation, 18(1) 25–38; Elsevier Science Ltd

Pahl G., Beitz W., Feldhusen J., Grote K.H; (2007) Engineering Design: a systematic Approach, 3rd Edition; Translation from German by Wallace K., Blessing L. (eds.) ; Springer-Verlag, London.

Pugh, S. (1990) Total Design: Integrated Methods for Successful Product Engineering, Wokingham: Addison Wesley.

Roozenburg, N.F.M. & Eekels, J.; (1995); Product Design: Fundamentals and Methods; Wiley; Chicester, New York; ISBN:0471943517.

Simon, H. A.; (1973); The structure of ill structured problems; Artificial Intelligence; Elsevier; Volume 4, Issues 3-4, Winter, Pages 181-201.