

## Research Article

# Estimating Designers' Performance considering Personal Characteristics and External Factors Together

Stefano Filippi 

*DPIA Department, University of Udine, Udine 33100, Italy*

Correspondence should be addressed to Stefano Filippi; [filippi@uniud.it](mailto:filippi@uniud.it)

Received 27 November 2019; Revised 20 July 2020; Accepted 30 July 2020; Published 1 September 2020

Academic Editor: Thomas Mandl

Copyright © 2020 Stefano Filippi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Design team performance evaluation can occur in different ways, all of them requiring considerations on interactions among team members; in turn, these considerations should count on as many pieces of information as possible about individuals. The literature already explains how personal characteristics and/or external factors influence designers' performance; nevertheless, a way to evaluate performance considering several personal characteristics and external factors together is missing. This research tries to fill the gap by developing the Designer's Performance Estimator (DPE), a ready-to-use tool for researchers and practitioners who need to make information about team members as richer as possible.

## 1. Introduction

Due to advances in technology, higher complexity in product development processes, shortages in time and resources, etc., companies must base their design activities more and more on teams rather than individuals [1,2]. A team consists of two or more people who interact together to achieve a common and shared goal or mission [3]. Team performance is the extent to which the team accomplishes that goal or mission [4]. Therefore, team performance evaluation becomes more and more important in modern design contexts both to tune up existing teams and to select the most suitable designers to generate new ones. This evaluation requires effective considerations on interactions among team members; in turn, these considerations should count on as many pieces of information as possible about the performance of the individuals belonging to the team or being candidate for it [5]. The literature already offers methods and tools for job performance evaluation focusing on individuals. These methods and tools range from empirical studies to literature meta-analyses and formal methods. They take one or more personal characteristics (e.g., personality traits, skill, and knowledge) and/or external factors (e.g., contexts, types of design activities and representations of products, users, and environments) into

consideration. Nevertheless, they barely address how mixes of personal characteristics and external factors together influence designers' performance.

The research described in this paper develops the Designer's Performance Estimator (DPE), a tool to quantify designers' performance in terms of how varied, novel, and usable the design results are expected to be, strictly considering personal characteristics and external factors together. This tool delivers immediate information to researchers and practitioners whose domains are already known to it; at the same time, the DPE knowledge base can be improved considering other design activities and this will make its application coverage wider. Specifically, this paper describes both the delivery of immediate information by considering shape-based design activities as already known to the DPE and the general procedure to improve the DPE knowledge base.

There are several possible exploitations of the DPE, depending on design goals and resource availability. For example, the DPE can be used to build design teams very focused on novelty (design goal) in a design context where many designers are at disposal but just one representation (external factor) is present (resource availability) or the DPE can be used to select the best representations (external factors) to maximize the usefulness (design goal) of the

outcome of the design effort of a small team whose composition cannot be changed.

The paper runs as follows. Section 2 reports the background of the research, ranging from the fundamentals of individuals' job performance evaluation to some considerations about possible influences on performance. Section 3, describing the research activities, starts by clarifying the DPE role in team performance evaluations and carries on by reporting the DPE definition and an example of its adoption. Section 4 describes the early validation of the DPE by comparing foreseen designers' performance to actual one in the specific case of the shape-based design activities. Section 5 details possible exploitations of the results of the DPE adoption. Section 6 highlights and discusses the main results of the research and Section 7 closes the paper by summarizing the research and suggesting some perspectives. Finally, Appendix A contains the questionnaire used to collect data during the early DPE validation.

## 2. Background

*2.1. Job Performance Evaluation.* In general, job performance is a multidimensional concept that indicates how well employees perform their tasks, the initiatives they take, the extent to which they complete their tasks, the way they use the resources available, and the time and energy they spend on their tasks alone or in teams [4, 6]. Evaluating job performance consists in judging the employees respect to several dimensions like quality, quantity, planning, and timeliness of work. [7, 8]. The research described in this paper refers specifically to design activities and addresses individuals rather than teams; thus, the focus here is on single designers' performance. Designers' performance evaluation could exploit empirical studies, literature meta-analyses, and formal methods. Bakker et al. [9] conducted an empirical study to assess performance in terms of increments in structural and social job resources, energy, time, and dedication spent for the job, etc. Peeters et al. [7] did a meta-analysis of refereed journals to measure designers' performance in terms of results effectiveness. Salgado [10] developed a formal method to analyze job performance assessments where personality-related five-factor model (FFM-) based inventories and non-FFM-based inventories were applied in different contexts. This analysis highlighted that FFM-based inventories are more reliable in assessing job performance compared to non-FFM-based inventories, especially when focusing on conscientiousness and neuroticism. Azadeh et al. [11] developed a tool to evaluate job performance focusing on stress, health, safety, environment, and ergonomics in petrochemical plants affected by noise and uncertainty. This tool considers seventeen well-ordered steps, from determining the reliability of the data collection procedure (questionnaire) and the definition of input and output to the achievement of the information needed to apply the algorithm that computes the efficiency of each operator. The results allow evaluators to implement corrective actions on low scorers. Lee et al. [12] proposed an approach for evaluating job performance of IT departments of manufacturing industries in Taiwan based on fuzzy

analytic hierarchy processes (FAHP) and balanced scorecards (BSC). This approach has a well-ordered, rigorous structure and functioning; it starts by using questionnaires to define performance indices from the financial, customer, internal business, and learning and growth points of view; the measurement of these indices highlights strengths and weaknesses to focus enhancing/corrective actions on.

*2.2. Possible Influences on Performance.* The many variables that could influence designers' performance can be classified as internal or external, referring to personal characteristics rather than external factors. The internal variables considered in this research are skill, knowledge, and personality; the external variables are design activities and representations. This choice comes from the large literature highlighting that these variables influence job performance much more than others [4, 13–15].

Regarding the internal variables (skill, knowledge, and personality), there is much literature about the definition of skill and knowledge and about their influences on design [16–20]. Personality deserves more attention because this research refers directly to its components, the traits. Personality traits are characteristics of a person that account for consistent behavioral patterns over situations and time [21]. The well-known taxonomy of the big five [6] identifies the following personality traits: extroversion or surgency (from now on, Personal Trait 1–PT1), agreeableness (PT2), conscientiousness (PT3), neuroticism (PT4), and openness to experience/culture (PT5). Much research reports the influences of personality on designers' performance by exploiting the big five because these are considered as good predictors of job task and contextual performance [6, 10, 22–29].

Regarding the external variables (design activities and representations), the research of Sim and Duffy [30], Filippi and Barattin [31], and Gero and Kannengiesser [32] addresses how design activities can be defined and classified. Design activities are sets of actions performed by different actors (e.g., designers and final users) in different contexts (development of electronic devices or mechanical CNC machines, furniture, clothes, etc.), starting from different sources (functions, user needs, shapes, etc.) and ending with the generation of concepts, prototypes, or products. Salas et al. [3] state that the type of design activities and their complexity play a crucial role in influencing designers' performance exactly as individual characteristics like personality traits, skills, task knowledge, motivation, and attitudes do. Sonnentag et al. [18] consider several studies about how different types of tasks can influence job performance and conclude that this influence exists but it is not as heavy as that of other variables like, e.g., cognitive abilities, past experiences, and personality traits. Representations deserve a deeper consideration because this research needs a univocal classification of them based on combinations of precise elements. In the classification of Filippi and Barattin [33], these elements are orthogonal, they cover both classic and more recent product development processes, and they are clearly stated thanks to discrete values assignable to

them. These elements are the environment—real if it corresponds to the physical one, virtual otherwise; the product—real or virtual, with the same meanings as the previous ones; the interaction between environment and product—aware if the product recognizes the environment and behaves accordingly, unaware otherwise; the user—real if the user is a human being, simulated otherwise; the interaction between product and user—direct if this interaction occurs in a natural way, exactly as the user expects it, indirect otherwise. Representations consist of combinations of the values these five elements can assume. The representations considered in the classification are virtual reality (VR)—virtual products and environments; augmented virtuality (AV)—virtual environments and real products; augmented reality (AR)—virtual products and real environments; pure reality (PR)—environments and products are both real; mixed reality (MR)—the combination of AR and AV where at least the environment or the product assumes both its values at the same time. Much research demonstrates the influence of representations on job performance [34–38].

The literature analysis created a solid background for the research described in this paper. On the one hand, it complies with the order and rigorousness suggested by the research approach of Lee et al. [12] and bases data collection on questionnaires as suggested by Azadeh et al. [11]; on the other hand, addressing and classifying what can influence designers' performance allowed defining precisely the set of variables (internal and external factors) used here. The discussion section of this paper will report in detail the relationships between the DPE and the literature described here in terms of affinities and differences.

### 3. Activities

This section opens by clarifying the role of the DPE in team performance evaluations. The descriptions of the DPE definition and an example of its adoption take place afterward.

*3.1. The DPE Role in Team Performance Evaluations.* Figure 1 proposes an example of team performance evaluation exploiting the DPE (represented inside the dotted envelope). The DPE exploits DA (design activities) tables (A); these tables contain the relationships between internal and external variables referred to specific types of design activities (B). Each type of design activities leads to a different DA table. The collection of DA tables makes up the DPE knowledge base (C) (the generation of the DA tables can happen in different times and spaces and, currently, it occurs through user testing conducted by experts/practitioners (D)). Team performance evaluators (E) needing to measure a specific team (F) use the DPE to get as much information as possible about the designers belonging to the team. The DPE generates the pieces of information (G) in the record describing each designer by considering his/her personal characteristics (H), the specific type of design activities that the team will be called to perform (I), and the representations available (J). From that moment on, the

evaluators can apply their knowledge, methods, and tools to compute (K) the performance of the team as a whole, considering all relationships and influences among the designers belonging to it.

This is just one type of DPE exploitation. The DPE can be of help as well when evaluators do not deal with an existing team but they need to select the best designers to build up a new one, when evaluators should have the team ready and the need is to select the best representations to work with, etc. Section 5 lists possible DPE involvements in different situations.

*3.2. DPE Definition.* The definition of the DPE occurs by determining its main data structure named DA table, the procedure to fill it, the metrics to quantify the design results, the output data structure named designer record, and the procedure to fill it.

*3.2.1. The DA Table.* The DA table (please refer to Table 1, ignoring the values in it for the moment because they refer to a specific kind of design activities) puts into relationship internal and external variables using precise metrics. The rows (except for the last two) refer to the internal variables skill, knowledge, and personality traits; the columns correspond to the external variables; the columns correspond to the external variables representations (VR, AV, AR, PR, MR, all of them, or just some, due to the design activities considered), characterized in terms of quantity (Q), variety (V), novelty (N), and usefulness (U), the metrics used in this research, and described later. The whole table refers again to an external variable, a specific type of design activities. Each internal variable can assume five levels as suggested by Likert's scale. The skill (S), defined here as the ability in applying design methods and tools and problem solving techniques, goes from level 1 (no skill)—describing a designer unable to use design methods, tools, and problem solving techniques at all—to level 5 (high skill)—indicating a designer who uses design methods, tools, and problem solving techniques effectively, efficiently, and autonomously. The knowledge (K), defined as the quantity of information owned by designers about design theories, techniques, and processes, uses the same classification: level 1 (no knowledge) describes a designer without any knowledge about design theories, techniques, and processes although he/she basically knows the context; level 5 (high knowledge) indicates a designer who has deep and precise knowledge about design theories, techniques, and processes and about the context. Each personality trait develops through five levels as well, ranging from the opposite of the trait (level 1) to the trait itself (level 5). For example, level 1 of PT1 represents an introvert designer; an extrovert designer corresponds to level 5. The last two rows contain the performance of the best and worst designers for each representation and metric. They are computed automatically and will be used as terms for comparison.

*3.2.2. The Procedure to Fill a DA Table.* If the DPE knowledge base does not contain the DA table related to the specific type of design activities the DPE involvement is focusing on, the procedure to fill it is as follows. Some

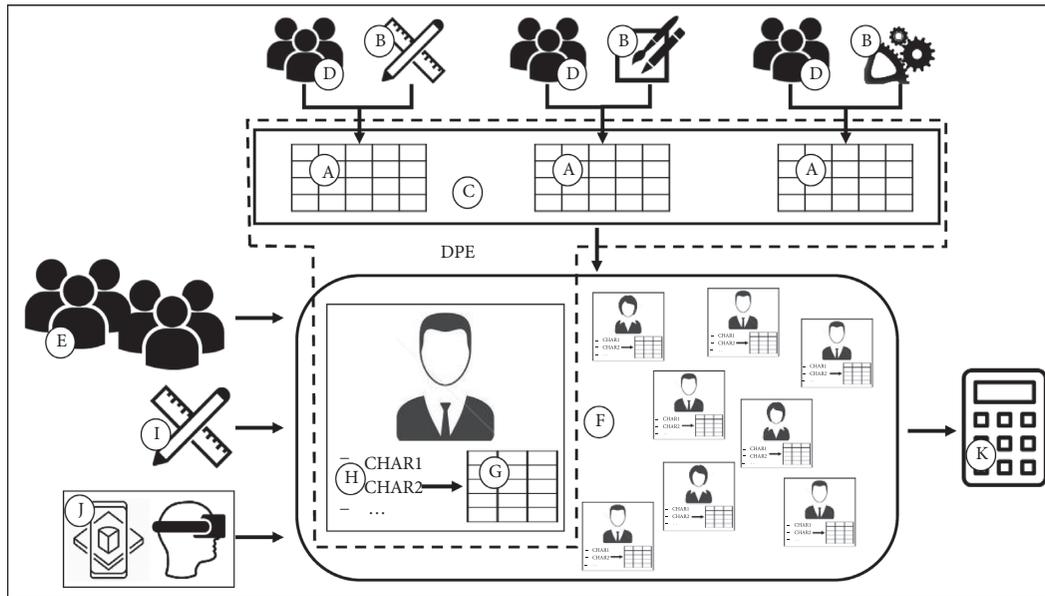


FIGURE 1: Role of the DPE in a team performance evaluation.

designers are selected with respect to their levels of internal variables (rows). The evaluators use a questionnaire to assess the characteristics of designers and classify them against the internal variables. The structure of the questionnaire is reported in Appendix A. It consists of three questions (Q1 to Q3) containing items that designers mark using values between 1 (strongly disagree) to 5 (strongly agree). The first two questions (Q1 and Q2) focus on skill and knowledge. They consist of ten base items each, referring to design methods and tools, equipment to generate prototypes and produce objects, software packages for design, manufacturing, and/or demonstrations, etc., as well as notions of physics, thermodynamics, construction laws, human-machine interaction paradigms, etc. These 20 items should be enough to characterize designers' skill and knowledge with the required precision for the downstream steps of the DPE adoption. Nevertheless, the evaluators can add further items to customize the questionnaire due to the specific type of design activities. Designers' skill and knowledge are assessed by considering mean values. For example, if the mean value of one designer's answers to the first question is around 4, that designer is assigned the fourth level of skill. The last question (Q3) focuses on personality and comes from the Big Five Inventory (BFI) consisting of 44 items whose marks lead to the computation of a 0 to 100 score for each trait [39]. Since the scores used here develop through five levels, BFI values in the interval [0..20) correspond to level 1; those in the interval [20..40) to level 2; [40..60) to level 3; [60..80) to level 4; and [80..100) to level 5.

Once the designers have been classified based on the questionnaire outcomes, a simple algorithm is in charge of assigning all the representations available, aiming at covering as many different combinations of internal/external variables as possible. After that, the design activities to perform are described to the designers and they carry them

on. Four metrics allow quantifying the results. Among all the possibilities offered in literature, the work of Shah and Vargas-Hernandez [40] suggested the first two metrics: quantity and variety; the second two, novelty and usefulness, come from the research of Sarkar and Chakrabarti, who claim that creativity can be measured based on them [41]. These metrics are reputed as exhaustive to characterize the results because of their complementarity. Creativity, another metric quite common in these cases, does not appear explicitly because Sarkar and Chakrabarti claim that it can be easily derived from novelty and usefulness. The computation of the four metrics occurs as follows.

- (i) Quantity ( $Q$ ). It is the amount of results produced by each designer. The value can vary from 0 to  $\infty$  and is considered for each of the designer's levels of internal variables for every representation. For example, designer George is quite extrovert (extroversion level equal to 4) and very low on conscientiousness (level equal to 1) and generates seven results using the VR representation; therefore, George's values of PT1(4) and PT3(1) are both equal to 7.
- (ii) Variety ( $V$ ). It measures how much a result differs from those expressed by other designers. Each result has assigned a value ranging from 1 to 10. If all designers sharing the same levels of internal variables and exploiting a specific representation express that result, the value will be set to 1 (the lowest value of variety); if only one designer expresses it, the value will be 10 (the highest value). A simple formula allows assigning the other values in between.
- (iii) Novelty ( $N$ ). It measures how much a result does not resemble to anything known, in general. Each result has assigned a novelty value ranging from 0 to 1; the

TABLE 1: DA table related to the shape-based design activities.

Variable		Representation/metrics											
		VR				AR				PR			
		Q	V	N	U	Q	V	N	U	Q	V	N	U
S	1 (no skill)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	2	2.5	9.5	0.33	0.030	4.5	9.28	0.34	0.028	7	9.63	0.38	0.034
	3	2.4	9.65	0.34	0.021	3.9	9.35	0.35	0.033	5.2	9.53	0.35	0.051
	4	3	9.73	0.34	0.020	3.2	9.24	0.33	0.050	4	9.47	0.33	0.048
	5 (high skill)	3.8	9.77	0.39	0.026	4.6	9.63	0.39	0.032	3.8	9.23	0.32	0.056
K	1 (no knowledge)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	2	2.7	9.60	0.32	0.029	4	9.59	0.28	0.043	5.3	9.55	0.39	0.037
	3	3.1	9.78	0.33	0.026	4.2	9.56	0.32	0.035	4.5	9.51	0.37	0.041
	4	3.3	9.81	0.37	0.022	3.9	9.55	0.35	0.034	3.8	9.42	0.3	0.047
	5 (high knowledge)	4	9.36	0.4	0.027	4.3	9.58	0.39	0.038	4.7	9.48	0.34	0.042
PT1	1 (introvert)	2.3	9.62	0.33	0.043	3.2	9.49	0.32	0.059	3.9	9.57	0.35	0.072
	2	2.6	9.70	0.32	0.041	4	9.58	0.36	0.052	4.1	9.49	0.36	0.058
	3	3.1	9.81	0.35	0.039	3.7	9.51	0.37	0.051	4.6	9.65	0.31	0.065
	4	3.3	9.92	0.36	0.039	4.3	9.28	0.39	0.049	4.5	9.68	0.32	0.062
	5 (extrovert)	3.7	10	0.39	0.038	4.1	9.48	0.42	0.056	4.9	9.72	0.33	0.052
PT2	1 (disagreeable)	2.4	10	0.35	0.057	2.8	9.8	0.33	0.055	5.9	9.6	0.34	0.053
	2	2.6	9.88	0.36	0.055	3.1	9.85	0.31	0.053	4.5	9.88	0.32	0.051
	3	3.4	9.91	0.38	0.048	3.6	9.58	0.36	0.055	3.8	9.85	0.32	0.063
	4	3.5	9.75	0.37	0.041	4	9.51	0.39	0.057	3.3	9.88	0.31	0.069
	5 (agreeable)	3.3	9.71	0.4	0.035	4.4	9.69	0.37	0.058	3	10	0.33	0.065
PT3	1 (unconscientious)	2.7	10	0.34	0.039	4.2	9.6	0.35	0.047	3	8.2	0.38	0.071
	2	2.8	9.88	0.33	0.029	3.8	9.5	0.33	0.052	3.2	8.88	0.36	0.075
	3	2.9	9.93	0.35	0.031	3.9	9.69	0.34	0.057	4.5	9.43	0.36	0.079
	4	3.1	9.81	0.37	0.023	4.1	9.71	0.37	0.055	4.8	9.41	0.35	0.089
	5 (conscientious)	3.4	9.85	0.36	0.025	4	9.66	0.38	0.058	5	9.59	0.34	0.101
PT4	1 (unneurotic)	3.4	9.77	0.38	0.037	4	9.79	0.37	0.038	4.2	9.6	0.36	0.044
	2	2.9	9.65	0.37	0.042	3.8	9.68	0.34	0.034	3.8	9.48	0.34	0.043
	3	3	9.54	0.34	0.048	4.1	9.68	0.36	0.041	3.9	9.54	0.34	0.058
	4	2.7	9.48	0.35	0.053	4.3	9.65	0.4	0.055	4.5	9.59	0.33	0.055
	5 (neurotic)	3.2	9.59	0.33	0.051	4.2	9.71	0.39	0.061	4.8	9.41	0.35	0.069
PT5	1 (closed to experience)	2.9	9.86	0.34	0.04	4.8	9.68	0.34	0.045	4.4	9.51	0.37	0.051
	2	2.6	9.86	0.32	0.036	4.1	9.66	0.36	0.051	4.2	9.42	0.36	0.045
	3	2.8	9.84	0.35	0.037	3.6	9.66	0.37	0.052	4.1	9.65	0.34	0.042
	4	3.3	9.85	0.37	0.034	3.5	9.69	0.38	0.055	4.6	9.61	0.33	0.041
	5 (open to experience)	3.5	9.88	0.39	0.031	3.2	9.65	0.36	0.053	4.5	9.72	0.35	0.036
Best designer's performance		25.3	69.2	2.72	0.291	30.9	67.8	2.72	0.384	37.5	67.8	2.58	0.465
Worst designer's performance		17.8	67.5	2.3	0.206	23.9	66.4	2.25	0.29	25.4	65	2.24	0.324

computation occurs as follows. The value is equal to 0 if that result is already present in one or more existing products as it is. The value is in the range (0, 0.5) if that result is already present in some existing products as functions to perform, and the user and product behaviors during interaction are the same but the implementation (product structure) is different; the more the implementation is different, the higher the value is. The value is in the range [0.5, 1) if the result is already present in some existing product as functions to perform but the user and/or product behaviors are different. The more these behaviors are different, the higher the value is. Finally, the value is 1 if the result is not present in any existing product.

- (iv) Usefulness ( $U$ ). It represents the social value of a result; it is the product of the level of importance, the rate of popularity of usage, and the rate of duration of

benefit. The level of importance refers to the impact of the result on users' life; it can vary from 0—corresponding to unessential things, luxuries, etc.—to 1—referring to life support systems, life-saving drugs, etc. The rate of popularity of usage is the ratio between the number of designers sharing the same levels of internal variables who expressed that result and their total number. Finally, the rate of duration of benefit is the percentage of time the designer spends with the result.

After the assignment of the values to each result (to each designer, in the case of the quantity), the computation of mean values takes place, one for each cell of the DA table. For example, the mean value is computed for the variety of the results produced by all designers showing extroversion equal to 5 and having worked with VR. Finally, the best and worst designers' performance are computed. The best designer's performance

comes from summing up the highest values among the levels of each internal variable for every representation. For example, considering VR, the best designer from the quantity point of view could be the very skilled (level 5), very knowledgeable (5), much introvert (1), agreeable in average (3), conscientious (5), quite neurotic (4), and closed to experience (1) one, just because these levels show the highest quantity values regarding the internal variables. The computation of the worst designer's performance occurs in the same way, except for considering the lowest values instead of the highest ones.

**3.2.3. The Designer Record.** Once the DPE knowledge base becomes populated thanks to the presence of one or more DA tables, the characteristics of one designer allow computing his/her performance with respect to specific representations for every design activity available. Table 2 (please ignore the values in it for the moment) shows the designer record, the output data structure of the DPE that contains the results of this computation.

**3.2.4. The Procedure to Fill the Designer Records.** The filling of a designer record starts by summing up the values of the DA table corresponding to the designer's levels of skill, knowledge, and personality traits for every representation, for each type of design activity present in the DPE knowledge base. For example, consider the designer named Robert; the DPE questionnaire allows identifying him in the levels  $S=2$ ,  $K=3$ ,  $PT1=3$ ,  $PT2=1$ ,  $PT3=4$ ,  $PT4=2$ , and  $PT5=4$ . Regarding the quantity of results expressed using VR in the design activities consisting in the development of prototypes of home appliances, the corresponding values in the DA Table are 4, 3, 6, 5, 5, 3, and 8, with 34 as their sum. The values of the best and worst designers, 76 and 14, respectively, allow normalizing the performance of Robert and expressing it as a percentage using the formula  $des\_perf\_ \% = 100 * ((des\_perf - worst\_perf) / (best\_perf - worst\_perf))$ . The result is equal to 32.2%. Therefore, Robert is quite scarce in producing design results about prototype development when working with virtual reality representations.

The filled designer records represent the outcome of the DPE adoption. By summarizing, the designer record of a specific individual foresees his/her performance about specific types of design activities, with specific representations available, due to his/her skill, knowledge, and personality traits, with all of this being quantified using the four metrics described before.

**3.3. Example of DPE Adoption.** What follows describes how the DPE adoption can occur in a real context, from the filling of the DA table to that of the designer records. This real context considers shape-based design activities and the VR, AR, and PR representations. Shape-based design activities develop products by analyzing specific shapes and defining product behaviors and functions consequently [42]. These design activities are becoming more and more important due to the role of User eXperience in design. This is why they have been selected for this example. One of the main goals of

TABLE 2: Designer record of a specific individual.

Design activities	Representation/metrics							
	VR				PR			
	Q	V	N	U	Q	V	N	U
Shape-based design activities	58%	28%	55%	55%	54%	81%	18%	56%

these activities is to arouse specific emotions in the people who will interact with those products. This type of design activities is used, for example, to develop deformable interfaces for mobile devices [43] or to produce iconic objects based on the analysis of shapes generated by fashion designers, as it happens for the Italian brand Alessi, specialized in developing home appliances [44]. Here, only the first part of these design activities—the analysis of specific shapes thanks to tests where interaction exploits the sight sense—is considered. This analysis suggests to the participants specific functions to perform with products shaped that way as well as personal and product behaviors meanwhile. These suggestions are addressed in the DPE as F/B (function/behaviors) pairs. An example of F/B pair is “Contain and heat tea” (function)/“I put cold water inside the cup; what seems to be the resistor inside the cup, heats the water; when the water is hot, I put the tea bag” (behavior). Regarding the representations, this example of adoption considers only VR, AR, and PR because AV and MR require expensive tools and complex procedures not available.

**3.3.1. Filling the DA Table.** Three evaluators carry on the procedure to fill the DA table. They are experts of product development processes and shape-based design activities. The activities run from the setup of the material for the tests to their execution and to the collection and analysis of the resulting data. These steps are described in the following.

*(1) Setup of the Material for the Tests.* The material consists of the shapes used during the design activities and the documents that will help the participants meanwhile. Each participant will interact with the same two shapes, labelled as Sh1 and Sh2 in the following. There are more than one shape in order to lower the bias due to specific shape characteristics. Shape definition occurs by obeying to precise rules [42]. Table 3 summarizes the characteristics of the two shapes selected in answering to those rules.

For what concerns the VR tests, the shapes and the desk where they are placed are modeled using the CAD software package Fusion 360 by AutoDesk [45]. Thanks to the Microsoft 3D Builder software package, participants can rotate the shapes to look at them from different points of

TABLE 3: Characteristics of the shapes used to fill the DA table.

	Shape	
Rule	Sh1	Sh2
		
Combination of simpler, well-known shapes	(i) Banana (ii) Mezzaluna knife (iii) Boomerang	(i) Tape holder (ii) Hamster wheel
Presence of elements to catch the attention	(i) Asymmetry obtained thanks to the presence/positioning of the support (ii) Cavity throughout (iii) Cap placed only on the lower extremity of this cavity	(i) Cavity throughout (ii) Lower part ending with a through hole (iii) Flat surfaces (top and bottom) (iv) Solid cylinder in the center
Size concerns	(i) Color: yellow (gold and sun reminding of precious and/or spherical and/or warm objects) (i) Actual size	(i) Color: green (grass and mint reminding of natural and/or fresh/iced objects) (i) Actual size

view. For the AR tests, the shape models used in the VR tests are converted into holograms that the HoloLens device by Microsoft [46], worn by the participants, projects on a real desk. Finally, the physical models for the PR tests are built with the 3D printer Ultimaker 2 by Ultimaker [47]. Once finished, the models are placed on the same desk used for the AR tests. Figure 2 shows the three representations of Sh1 and environment as used during the VR, AR, and PR tests.

The documents for the participants describe each design activity they must perform using nontechnical language.

(2) *Execution of the Tests and Data Collection.* Once the material is available, the questionnaire reported in Appendix A is sent by e-mail to 90 possible participants, designers who have been working for years in different companies where shape-based design activities are almost known and students of university courses in mechanical engineering who have been taught about the principles of design in general and on the shape-based design activities in particular. 78 people send back the answers and the collected data allow selecting 60 participants with different levels of skill, knowledge, and personality traits and distributing them in three tests as homogeneously as possible against these characteristics. In all, 25 participants perform the VR test, 19 the AR test, and 16 the PR test. Tests take place in a university lab, one participant at a time. At the beginning, the participant receives the document describing what the evaluators expect from him/her. After that, the first shape is unveiled; the participant has ten minutes to look at it moving around (without touching it) and to write down the F/B pairs that come to his/her mind. At the end, the evaluators unveil the second shape and the participant has ten minutes again to consider it and write down the F/B pairs. Finally, the participant returns the document to the evaluators.

(3) *Data Analysis.* Once the last test comes to the end, the evaluators apply the metrics to the results, separately for each shape. Finally, the mean values considering both shapes are calculated and become the content of the DA table shown as Table 1.

Data undergo a statistical analysis using the *t*-test (test of Student). The *t*-test works by comparing two means [48]. Here, it verifies possible influences of internal and external variables on results. The computation does not appear here for space reason; nevertheless, a clear influence of both internal and external variables is detected in all the cases examined (for every shape, for every metrics). All *P* values range from 0.07 to 0.09; since the significance level here is set to 0.1 because of the low number of participants, all *P* values are lower than the significance level and the possible influence is confirmed.

3.3.2. *Filling the Designer Records.* The DA table allows filling any designer record for what concerns shape-based design activities. For example, John is a skilled designer with diverse experience in design activities. Therefore, his levels of skill and knowledge are high. He is extrovert in average, quite disagreeable, conscientious, quite neurotic, and open to experience; John’s personal characteristics, summarized in the upper part of the designer record, correspond to the tuple (4, 5, 3, 2, 5, 4, 4). The design context where John could be involved has VR and PR representations available. Under these conditions, the DPE allows filling the lower part of the designer record, containing John’s performance (Table 2). For example, the performance equal to 81% regarding *V* using PR means that John is very good, much more than the average (50%), in finding F/B pairs showing high variety when dealing with pure reality. The *V* value comes as

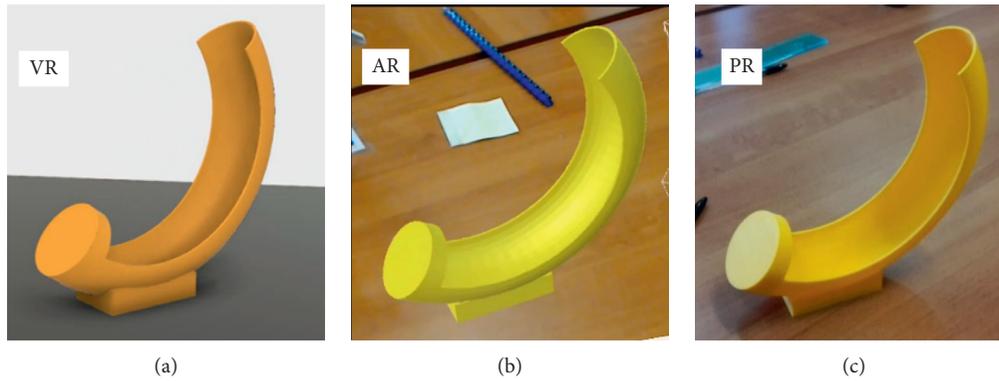


FIGURE 2: VR, AR, and PR representations of Sh1 and environment as used in the tests.

TABLE 4: Designers' characteristics in the DPE early validation.

Characteristic	Designer								
	Des1	Des2	Des3	Des4	Des5	Des6	Des7	Des8	Des9
S	4	3	4	3	3	3	3	3	5
K	5	2	2	5	4	5	4	4	4
PT1	5	1	3	4	3	2	5	4	4
PT2	4	2	5	2	5	1	5	2	3
PT3	4	4	1	3	2	3	2	1	5
PT4	5	2	3	4	1	5	2	5	5
PT5	4	2	1	5	4	2	3	5	2

follows. Considering John's tuple, the corresponding mean values in the DA table referring to  $V$  and  $PR$  are 9.47, 9.48, 9.65, 9.88, 9.59, 9.59, and 9.61, with 67.27 as their sum. Thanks to the  $V$  values in  $PR$  of the best (67.8) and worst (65) designers, it is possible to compute John's performance in percentage as  $\text{John's\_perf\_}\% = 100 * ((67.27 - 65) / (67.8 - 65)) = 81.03\%$ .

#### 4. Early DPE Validation

The early validation of the DPE adopts its current release, containing the DA table about shape-based activities, to estimate the performance of designers interacting with different shapes than those used to generate the DA table. Then, the real performances of the same designers are measured through tests. The comparison of the results starts assessing the DPE applicability and reliability.

##### 4.1. Adopting the DPE to Estimate Designers' Performance.

Four evaluators are involved. Again, all of them are expert in design processes and shape-based design activities. Nine designers are considered (Des1 to Des9); they never practiced shape-based activities; nevertheless, their variegated experiences as designers qualify them as good candidates to perform this validation. Table 4 summarizes their characteristics.

Table 5 contains the performance of the nine designers as estimated thanks to the DPE adoption (indeed, the table collects only the lower parts of the records of the nine designers to represent them compactly). This validation involves only the metrics whose computation assigns values to the F/B pairs considering individuals, i.e.,  $Q$  and  $N$ . The

TABLE 5: Designers' performance as estimated by the DPE.

Designer	Representation/metrics					
	VR		AR		PR	
	Q (%)	N (%)	Q (%)	N (%)	Q (%)	N (%)
Des1	80	64	50	89	47	29
Des2	11	26	33	17	52	65
Des3	39	31	64	32	23	62
Des4	48	60	44	62	58	38
Des5	51	57	47	57	26	29
Des6	31	26	47	57	66	65
Des7	45	60	51	60	21	32
Des8	43	45	41	53	41	38
Des9	45	50	51	79	21	21

values of the other two metrics ( $V$  and  $U$ ) would be computed considering F/B pairs found by groups made by one designer only (only nine designers are present and each of them shows different personal characteristics); thus,  $V$  and  $U$  values would be meaningless.

According to the DPE results, it seems that, using VR, Des1 could find many more F/B pairs ( $Q = 80\%$ ) than all the other designers (maximum value is that of Des5, equal to 51%). Moreover, Des4, Des5, Des6, and Des7's F/B pairs seem to be showing more or less the same novelty when using AR. All of this suggests some hypotheses to verify in the field in order to start assessing the DPE applicability and reliability. The hypotheses considered here are as follows.

- (i) *Hyp1*. Given VR and  $Q$ , Des1 should find more F/B pairs than Des3 and Des2, in this order, and the differences should be considerable.

- (ii) *Hyp2*. Given AR and  $N$ , Des4, Des5, Des6, and Des7's F/B pairs should show comparable  $N$  mean values.
- (iii) *Hyp3*. Given  $Q$ , Des8 should find similar numbers of F/B pairs independently from the representation.
- (iv) *Hyp4*. Given  $N$ , Des9 should find more novel F/B pairs with AR than with VR and PR, in this order, and the differences should be considerable.

**4.2. Performing the Tests.** Figure 3 shows the shapes used in the tests. Multiple shapes are used again to lower the bias as much as possible. They are generated following the same rules as for the shapes used to fill the DA table.

The four hypotheses lead the following associations between designers and representations. Des1, Des2, and Des3 consider all the shapes using only VR; Des4, Des5, Des6, and Des7 consider all the shapes as well but using only AR. Des8 considers Sh3 in VR, Sh4 in AR, and Sh5 in PR; Des9 does the same. Table 6 summarizes the results. The values of designers Des1 to Des7 are mean values of the results, since each of them considers the three shapes using the same representation.

**4.3. Assessing the DPE Applicability and Reliability.** No problems arose from the DPE adoption throughout this early validation. Therefore, regardless of the hypotheses verification, the DPE applicability appears as verified. For what concerns its reliability, data contained in Table 6 lead to the following considerations about the four hypotheses.

- (i) *Hyp1*: VERIFIED. Working with VR, Des1, Des2, and Des 3 expressed 7.3, 1.7, and 3.7 F/B pairs (mean values), respectively. Des1 found around the double of F/B pairs than Des3 and Des3 did the same against Des2. This matches exactly what was foreseen by the DPE regarding the metrics  $Q$ .
- (ii) *Hyp2*: VERIFIED. The F/B pairs expressed by Des4, Des5, Des6, and Des7 show mean values of  $N$  equal to 0.32, 0.33, 0.32, and 0.33, respectively. These values are very close to each other. Again, this confirms what was foreseen by the DPE.
- (iii) *Hyp3*: NOT VERIFIED. Regarding the metrics  $Q$ , Des8 found much more F/B pairs in the VR test (8) than in AR (4) and PR (3) tests; they are almost the double. This contradicts what the DPE foresaw, the independency from the representation. This misalignment could depend on the shapes used in the tests. The same designer cannot consider the same shape in the three representations because of the inevitable bias among them. Therefore, three different shapes were considered. Although the shapes have been generated by strictly following the rules, Sh3 contains five elements to catch the attention while Sh4 and Sh5 have four elements only. This difference could be the main reason for the misalignment with the DPE estimate.

- (iv) *Hyp4*: VERIFIED. Des9 found F/B pairs showing different  $N$  mean values using the three representations: 0.32 for VR, 0.55 for AR, and 0.2 for PR. As foreseen by the DPE, AR suggests more  $N$  in the F/B pairs than VR and PR (the former value is almost the double of the latter).

The verification of three hypotheses out of four starts giving positive indications about the DPE reliability.

## 5. Possible Exploitations of the DPE Results

There are several ways team performance evaluators can exploit the designer records resulting from the DPE adoption. Three possibilities are as follows:

- (i) *Situation A*. The evaluators work in a design context where few representations are available; they are called to build a small team and there are precise expectations about the design results. In this case, the designer records can help in selecting the most promising people to build the team considering the expected characteristics of the results (e.g., novel design solutions) as leading criteria.
- (ii) *Situation B*. The evaluators work in a design context short in human resources from the design point of view and time-to-market is mandatory; nevertheless, they have all representations potentially available. In this case, the designer records can help in deciding the best representation(s) to use. More in detail, the designer records rank the representations; then, the company can select the most effective ones, depending on the time-to-market constraint.
- (iii) *Situation C*. The evaluators work in a design context where they are called to suggest the most promising design team to maximize specific characteristics of the design solutions and they have almost no limits about people to involve or representations to exploit. Then, designer records of the candidates to be part of the team can help the selection of the most promising ones according to those characteristics of the design solutions; moreover, the records suggest also the best representations to use.

To go deeper in understanding these possible exploitations, the early DPE validation described in the previous section can be classified as corresponding to situation A. Consider a company having nine designers in the R&D department. The company size suggests teams of at most four people. Now, this company decides to exploit VR in design activities and aims at getting design results as novel as possible. The results of the DPE adoption, as in Table 5, help building up the required team. The situation makes the attention focus on the VR/ $N$  column. Its values allow ordering the nine designers against their performance and the result is as follows (best to worst): Des1 (64%), Des4 and Des7 (60%), Des5 (57%), Des9 (50%), Des8 (45%), Des3 (31%), and Des2 and Des6 (26%). Therefore, looking for

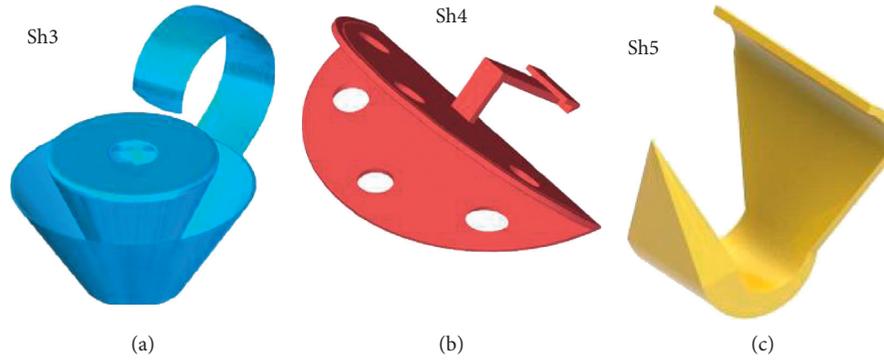


FIGURE 3: Shapes used in the tests of the DPE early validation.

TABLE 6: Results of the tests in the early DPE validation.

Designer	VR		Representation/metrics		AR		PR	
	Q	N	Q	N	Q	N	Q	N
Des1	7.3	—	—	—	—	—	—	—
Des2	1.7	—	—	—	—	—	—	—
Des3	3.7	—	—	—	—	—	—	—
Des4	—	—	—	0.32	—	—	—	—
Des5	—	—	—	0.33	—	—	—	—
Des6	—	—	—	0.32	—	—	—	—
Des7	—	—	—	0.33	—	—	—	—
Des8	8	—	4	—	3	—	—	—
Des9	—	0.32	—	0.55	—	—	—	0.2

building up a team made by four designers, the DPE suggests selecting Des1, Des4, Des7, and Des5, the first four most performing individuals.

## 6. Results and Discussion

The main result of this research is the Designer's Performance Estimator (DPE), a ready-to-use tool for everyone who needs to characterize individuals and foresee their performance in specific types of design activities, all of this in order to evaluate existing or potential teams as effectively as possible.

Among the evaluation approaches considered in this research, empirical studies, literature meta-analyses, and formal methods, as summarized in Section 2.1, the DPE shows more affinity with the last ones. The comparison with those methods highlights its peculiarities and strong points. Salgado's research [10] considered different contexts as the DPE does; nevertheless, the DPE involves more personal characteristics and external factors (representations). The work of Azadeh et al. [11] showed clear data structures and procedures like the questionnaires, their generation, and the input/output definition; the DPE does the same but it offers higher versatility (it can be applied in different contexts) and more metrics to quantify the results. Finally, the DPE presents many analogies with the approach of Lee et al. [12] like the rigorous architecture and the exploitation of existing, well-known methods and tools; nevertheless, the DPE involves again more personal characteristics and manages results individually rather than in the aggregate form only.

Although the DPE appears overcoming some lacks of the evaluation methods and tools reported before, it has drawbacks to consider as well; these drawbacks are summarized here and recalled as subjects for future perspectives in the conclusions section. Current release of the DPE allows tests as the only way to collect data to fill new DA tables or update existing ones. Only four metrics are used now and they do not consider important topics like eco-sustainability, ergonomics, user experience, etc. The knowledge base management in the current release of the DPE considers the DA table structure as fixed; adding internal/external variables is not allowed. This is quite a limit since, for example, variables referring to team working like cooperation and communication would make the DPE even more answering to the evaluators' needs. Although the DPE has proven to be applicable, its usability is quite scarce. Data collection and analysis must be performed almost manually and this makes the DPE adoption time consuming. Finally, the knowledge base is bare indeed; it contains only the DA table related to the shape-based design activities. Moreover, this table misses data referring to skill and knowledge as well as to AV and MR representations. All of this limits the DPE coverage and applicability and makes the DPE scarcely ready-to-use for practitioners now.

## 7. Conclusions

The research described in this paper aimed at helping team performance evaluators. As a result, it defined the Designer's Performance Estimator (DPE). The DPE is a ready-to-use tool

TABLE 7: The questionnaire used in the early DPE validation.

	1	2	3	4	5
(Q1) I have used/I use for my work, studies, etc. (strongly disagree = 1; strongly agree = 5):					
manual tools for measurements (e.g., calipers, multimeters, etc.)					
machines tools (manual and CNC)					
machines for additive manufacturing and rapid prototyping (3D printers, etc.)					
devices for virtual and augmented reality					
sensors for measuring speed, temperature, etc. (accelerometers, thermocouples, etc.)					
manual tools for prototype generation (potter's wheel, chisel and gimlet for wood, etc.)					
equipment for reverse engineering					
software for simple calculations and for the representation of results (e.g., Office, etc.)					
software for generating 3D models (e.g., AutoCAD, Solid Edge, etc.)					
software for programing machines and other (e.g., Matlab, etc.)					
software for simulations (e.g., Ansys, Comsol, etc.)					
software for generating and elaborating interfaces (Blender, Visual Studio, etc.)					
software for reverse engineering					
(Q2) I know from my previous studies, activities, etc. (strongly disagree = 1; strongly agree = 5):					
ways to represent technical specifications					
laws and standard for simple structural mechanics					
manufacturing techniques					
thermodynamics laws					
usability and human-machine interaction paradigms and design and evaluation procedures					
ways to manufacture specific products (from technical drawings to real final products)					
procedures for thermal-mechanical analysis					
optimization techniques of costs and resources to generate products					
ergonomics standards					
energy saving techniques					
industrial design and fashion trends					
(Q3) I see myself as someone who (strongly disagree = 1; strongly agree = 5):					
is talkative					
tends to find fault with others					
is depressed, blue					
is reserved					
is helpful and unselfish with others					
can be somewhat careless					
is relaxed and handles stress well					
is curious about many different things					
is full of energy					
starts quarrels with others					
is a reliable worker					
is ingenious, a deep thinker					
generates a lot of enthusiasm					
has a forgiving nature					
tends to be disorganized					
worries a lot					
tends to be quiet					
is generally trusting					
tends to be lazy					
is inventive					
has an assertive personality					
can be cold and aloof					
perseveres until the task is finished					
can be moody					
values artistic, aesthetic experiences					
is sometimes shy, inhibited					
is considerate and kind to almost everyone					
remains calm in tense situations					
prefers work that is routine					
makes plans and follows through with them					
gets nervous easily					
likes to reflect and play with ideas					
has few artistic interests					
is easily distracted					
is sophisticated in art, music, or literature					

for researchers/practitioners that allows describing and quantifying designers' performance considering personal characteristics and external factors together. The computation exploits a knowledge base generated thanks to the analysis of different types of design activities in different situations. Some adoptions in the field already stated the DPE applicability and started demonstrating its reliability.

For what concerns possible research perspectives, some hints, corresponding to the drawbacks described in the results and discussion section, are as follows. Other types of data sources should be allowed to fill the DA tables like scientific literature and companies' history; moreover, the way to merge pieces of information coming from heterogeneous sources needs to be investigated. Other metrics than the current four need to be made available in order to widen the DPE coverage. For example, learnability, aesthetics, and enjoyment would allow orienting the DPE towards user experience; nevertheless, it must be pointed out once again the role of the DPE as estimator of single designers' performance. Other methods, tools, competencies, and knowledge are required to perform a complete design team performance evaluation. Procedures and/or suggestions should be introduced to allow evaluators adding internal and external variables. Automatisms must be introduced, especially to collect data and fill the DA tables and the designer records, in order to lower the time required by the DPE adoption. The author is working on making the DPE more usable by implementing Google forms and developing code in Microsoft Excel workbooks. All of this should make the DPE adoption almost automatic. Finally, the knowledge base should be more populated; more DA tables should be added and the existing one would need further tests to fill the empty rows and columns. Clearly, filling empty DA tables or empty cells of existing ones is not a problem; everybody can do this by simply following the indications described in this paper. On the contrary, if fresh data should affect nonempty cells of existing DA tables, the merging policy would have to be defined time-by-time, requiring competencies and expertise.

## Abbreviations

AR:	Augmented reality representation
AV:	Augmented virtuality representation
DA table:	It contains the relationships between internal and external variables that referred to a specific type of design activities
Des1-Des9:	The nine designers involved in the early DPE validation
Designer record:	It contains the result of the DPE adoption, the estimate of the specific designer performance
DPE:	Designer's Performance Estimator, the tool developed in this research
F/B pair:	Function/behavior pair. It consists of a function suggested by the shape of a product and the related product and/or user behavior
Hyp1-Hyp4:	The four hypotheses used in the early DPE validation
K:	Designers' knowledge
MR:	Mixed reality representation

N:	Novelty, the metrics representing how much a result does not resemble to anything known
PR:	Pure reality representation
PT1-PT5:	Personality traits (extroversion, agreeableness, conscientiousness, neuroticism, and openness)
Q:	Quantity, the metrics representing the amount of results
S:	Designers' skill
Sh1-Sh5:	The shapes used in the research
U:	Usefulness, the metrics representing the social value of a result
V:	Variety, the metrics representing how much a result differs from the others
VR:	Virtual reality representation.

## Appendix

### A. Questionnaire to Assess Designers' Characteristics

Table 7 in the following contains the questionnaire used to collect data about designers in order to assess their characteristics in the early DPE validation dealing with shape-based design activities.

### Data Availability

Part of the data used to support the findings of this study are included within the article (please see the content of Tables 1 and 4). Other data, like the test results used to validate the DPE, have not been made available because a language different from English was used. Reporting these data in the original language as well as giving their translation and interpretation would be time consuming and almost useless. Nevertheless, the author thinks that the description of the DPE data structures as well as that of the procedures to fill them should be clear enough to allow the reader to replicate the research activities easily.

### Conflicts of Interest

The author declares that there are no conflicts of interest.

### Acknowledgments

The author would like to thank designers, engineers, and the students of the Mechanical Engineering courses at the University of Udine who took part in the tests.

### References

- [1] G. A. Macht and D. A. Nembhard, "Measures and models of personality and their effects on communication and team performance," *International Journal of Industrial Ergonomics*, vol. 49, pp. 78–89, 2015.
- [2] J. A. LePine, "Team adaptation and postchange performance: effects of team composition in terms of members' cognitive ability and personality," *Journal of Applied Psychology*, vol. 88, no. 1, pp. 27–39, 2003.

- [3] E. Salas, D. L. Reyes, and A. L. Woods, "The assessment of team performance: observations and needs," in *Innovative Assessment of Collaboration*, A. A. von Davier, M. Zhu, and P. C. Kyllonen, Eds., Springer International Publishing, New York, NY, USA, pp. 21–36, 2017.
- [4] S. T. Bell, "Deep-level composition variables as predictors of team performance: a meta-analysis," *Journal of Applied Psychology*, vol. 92, no. 3, pp. 595–615, 2007.
- [5] S. L. Kichuk and W. H. Wiesner, "The Big Five personality factors and team performance: implications for selecting successful product design teams," *Journal of Engineering and Technology Management*, vol. 14, no. 3-4, pp. 195–221, 1997.
- [6] S. Rothmann and E. P. Coetzer, "The big five personality dimensions and job performance," *Journal of Industrial Psychology*, vol. 29, no. 1, pp. 68–74, 2003.
- [7] M. A. G. Peeters, H. F. J. M. Van Tuijl, C. G. Rutte, and I. M. M. J. Reymen, "Personality and team performance: a meta-analysis," *European Journal of Personality*, vol. 20, no. 5, pp. 377–396, 2006.
- [8] D. R. Fitzgerald, S. Mohammed, and G. O. Kremer, "Differences in the way we decide: the effect of decision style diversity on process conflict in design teams," *Personality and Individual Differences*, vol. 104, pp. 339–344, 2017.
- [9] A. B. Bakker, M. Tims, and D. Derks, "Proactive personality and job performance: the role of job crafting and work engagement," *Human Relations*, vol. 65, no. 10, pp. 1359–1378, 2012.
- [10] J. F. Salgado, "Predicting job performance using FFM and non-FFM personality measures," *Journal of Occupational and Organizational Psychology*, vol. 76, no. 3, pp. 323–346, 2003.
- [11] A. Azadeh, M. Saberi, M. Rouzbahman, and Z. Saberi, "An intelligent algorithm for performance evaluation of job stress and HSE factors in petrochemical plants with noise and uncertainty," *Journal of Loss Prevention in the Process Industries*, vol. 26, no. 1, pp. 140–152, 2013.
- [12] A. H. I. Lee, W.-C. Chen, and C.-J. Chang, "A fuzzy AHP and BSC approach for evaluating performance of IT department in the manufacturing industry in Taiwan," *Expert Systems with Applications*, vol. 34, no. 1, pp. 96–107, 2008.
- [13] H. K. Sinangil and D. S. Ones, "Gender differences in expatriate job performance," *Applied Psychology*, vol. 52, no. 3, pp. 461–475, 2003.
- [14] H. Liao and A. Chuang, "A multilevel investigation of factors influencing employee service performance and customer outcomes," *Academy of Management Journal*, vol. 47, no. 1, pp. 41–58, 2004.
- [15] M. Ogot and G. E. Okudan, "The Five-Factor Model personality assessment for improved student design team performance," *European Journal of Engineering Education*, vol. 31, no. 5, pp. 517–529, 2006.
- [16] F. P. Morgeson, K. Delaney-Klinger, and M. A. Hemingway, "The importance of job autonomy, cognitive ability, and job-related skill for predicting role breadth and job performance," *Journal of Applied Psychology*, vol. 90, no. 2, pp. 399–406, 2005.
- [17] H. J. Keel, S. J. Motowidlo, M. P. Martin, A. L. Stotts, and C. A. Moreno, "Testing for independent effects of prosocial knowledge and technical knowledge on skill and performance," *Human Performance*, vol. 27, pp. 311–327, 2014.
- [18] S. Sonnentag, J. Volmer, and A. Spychala, "Job performance," in *The Sage Handbook of Organizational Behavior: Volume One: Micro Approaches*, J. Barling and C. L. Cooper, Eds., pp. 427–447, Sage Publications, Thousand Oaks, CA, USA, 2008.
- [19] M. Janhonen and J.-E. Johanson, "Role of knowledge conversion and social networks in team performance," *International Journal of Information Management*, vol. 31, no. 3, pp. 217–225, 2011.
- [20] S. Filippi and D. Barattin, "Influence of personality on shape-based design activities," *Advances in Human-Computer Interaction*, vol. 2019, Article ID 9651369, 9 pages, 2019.
- [21] S. Karimi and M. R. Kangavari, "A computational model of personality," *Procedia - Social and Behavioral Sciences*, vol. 32, pp. 184–196, 2012.
- [22] L. G. Barron, J. G. Randall, J. D. Trent, J. F. Johnson, and A. J. Villado, "Big Five traits: predictors of retesting propensity and score improvement," *International Journal of Selection and Assessment*, vol. 25, no. 2, pp. 138–148, 2017.
- [23] D. H. M. Pelt, D. van der Linden, C. S. Dunkel, and M. P. Born, "The general factor of personality and job performance: revisiting previous meta-analyses," *International Journal of Selection and Assessment*, vol. 25, no. 4, pp. 333–346, 2017.
- [24] R. D. Zimmerman, M. d. C. Triana, and M. R. Barrick, "Predictive criterion-related validity of observer ratings of personality and job-related competencies using multiple raters and multiple performance criteria," *Human Performance*, vol. 23, no. 4, pp. 361–378, 2010.
- [25] H. Le, I.-S. Oh, S. B. Robbins, R. Ilies, E. Holland, and P. Westrick, "Too much of a good thing: curvilinear relationships between personality traits and job performance," *Journal of Applied Psychology*, vol. 96, no. 1, pp. 113–133, 2011.
- [26] G. A. Macht, D. A. Nembhard, J. H. Kim, and L. Rothrock, "Structural models of extraversion, communication, and team performance," *International Journal of Industrial Ergonomics*, vol. 44, no. 1, pp. 82–91, 2014.
- [27] L. A. Witt, L. A. Burke, M. R. Barrick, and M. K. Mount, "The interactive effects of conscientiousness and agreeableness on job performance," *Journal of Applied Psychology*, vol. 87, no. 1, pp. 164–169, 2002.
- [28] G. M. Hurtz and J. J. Donovan, "Personality and job performance: the Big Five revisited," *Journal of Applied Psychology*, vol. 85, no. 6, pp. 869–879, 2000.
- [29] R. R. Reilly, G. S. Lynn, and Z. H. Aronson, "The role of personality in new product development team performance," *Journal of Engineering and Technology Management*, vol. 19, no. 1, pp. 39–58, 2002.
- [30] S. K. Sim and A. H. B. Duffy, "Towards an ontology of generic engineering design activities," *Research in Engineering Design*, vol. 14, no. 4, pp. 200–223, 2003.
- [31] S. Filippi and D. Barattin, "X for Design, a descriptive framework for modelling the cognitive aspects of different design activities," in *DESIGN 2016. Proceeding of the International Design Conference*, D. Marjanović, M. Štorga, N. Pavković, N. Bojčetić, and S. Škec, Eds., pp. 1265–1274, The Design Society, Glasgow, UK, 2016a.
- [32] J. S. Gero and U. Kannengiesser, "The situated function-behaviour-structure framework," *Design Studies*, vol. 25, no. 4, pp. 373–391, 2004.
- [33] S. Filippi and D. Barattin, "Classification and selection of prototyping activities for interaction design," *Intelligent Information Management*, vol. 04, no. 04, pp. 147–156, 2012.
- [34] E. M. Starkey, A. S. McKay, S. T. Hunter, and S. R. Miller, "Dissecting creativity: how dissection virtuality, analogical distance, and product complexity impact creativity and self-efficacy," in *Proceeding of the Seventh International Conference on Design Computing and Cognition DCC'16*, J. S. Gero,

- Ed., pp. 63–82, Springer International Publishing, Evanston, IL, USA, June 2016.
- [35] T. Kohler, K. Matzler, and J. Füller, “Avatar-based innovation: using virtual worlds for real-world innovation,” *Technovation*, vol. 29, no. 6-7, pp. 395–407, 2009.
  - [36] M. L. Maher, L. Lee, J. S. Gero, R. Yu, and T. Clausner, “Characterizing tangible interaction during a creative combination task,” in *Proceeding of the Seventh International Conference on Design Computing and Cognition DCC’16*, J. S. Gero, Ed., pp. 43–62, Springer International Publishing, Evanston, IL, USA, June 2016.
  - [37] R. J. Youmans, “The effects of physical prototyping and group work on the reduction of design fixation,” *Design Studies*, vol. 32, no. 2, pp. 115–138, 2011.
  - [38] T. Khan, K. Johnston, and J. Ophoff, “The impact of an augmented reality application on learning motivation of students,” *Advances in Human-Computer Interaction*, vol. 2019, Article ID 7208494, 14 pages, 2019.
  - [39] C. J. Soto and O. P. John, “Ten facet scales for the Big Five Inventory: convergence with NEO PI-R facets, self-peer agreement, and discriminant validity,” *Journal of Research in Personality*, vol. 43, no. 1, pp. 84–90, 2009.
  - [40] J. J. Shah, S. M. Smith, and N. Vargas-Hernandez, “Metrics for measuring ideation effectiveness,” *Design Studies*, vol. 24, no. 2, pp. 111–134, 2003.
  - [41] P. Sarkar and A. Chakrabarti, “Assessing design creativity,” *Design Studies*, vol. 32, no. 4, pp. 348–383, 2011.
  - [42] S. Filippi and D. Barattin, “Involving autism spectrum disorder (ASD) affected people in design,” in *Proceeding of the Joint Conference of ADM, INGEGRAF and AIP-PRIMECA*, B. Eynard, V. Nigrelli, S. M. Oliveri, G. Peris-Fajarnes, and S. Rizzuti, Eds., Springer International Publishing, New York, NY, USA, pp. 373–383, 2016b.
  - [43] J. Mikkonen and Y. T. Hsieh, “Prototyping with experience workshop,” in *DUXU 2013, Design, User Experience, and Usability. Design Philosophy, Methods, and Tools. Lecture Notes in Computer Science*, A. Marcus, Ed., vol. 8012, pp. 564–572, Springer, Berlin, Germany, 2013.
  - [44] Alessi, *The Italian Factory of Industrial Design*, [https://www.alessi.com/us\\_en/](https://www.alessi.com/us_en/), 2019.
  - [45] Fusion 360 by AutoDesk, <https://www.autodesk.com/products/fusion-360/overview>.
  - [46] HoloLens by Microsoft, <https://www.microsoft.com/en-us/hololens>.
  - [47] Ultimaker 2 by Ultimaker, <https://ultimaker.com/en/products/ultimaker-2-plus>.
  - [48] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, John Wiley & Sons, Hoboken, NY, USA, 3rd edition, 2002.