# CODING SCHEMES FOR THE ANALYSIS OF ICT SUPPORTED CO-CREATIVE DESIGN SESSIONS

N. Becattini, G. Cascini, J. A. O'Hare and C. Masclet

## Abstract

The paper presents a coding scheme for design protocol analysis of collaborative sessions supported by Augmented Reality. It maps verbal interactions during creative sessions, so as to distinguish co-designers' intentions and the related contents characterizing the characteristics of the design proposal in terms of items and related parameters. Three different co-creative sessions and a tailored metrics allowed for checking the coding scheme effectiveness, showing good mapping capabilities and versatility of application also to compare similar sessions carried out with and without AR support.

*Keywords: protocol analysis, design creativity, design cognition, collaborative design, ICT for design*

## 1. Context and relevant background

### 1.1. Role of ICT support tools in design

The adoption of Information and Communications Technology (ICT) tools for design is nowadays a consolidated reality for a lot of design studios, agencies as well as big companies, where design represents just one of the departments, as they have simplified their daily work, at least in terms of the organization, storage, conservation and reuse of contents. Different ICT tools satisfy different needs (2D/3D modelling, graphics editing…) and the market has addressed these inherent demands, mainly with software applications based on mouse and keyboard interaction, typically developed for a single user, who individually works in front of his screen. The introduction of novel ICT tools in design activities changes the interaction the designer has with the design representation and, consequently, it potentially affects its cognition. The clarification of the role and the effect of ICT tools on the cognitive processes of designers has been one of the objectives of design protocol analysis-based studies (e.g. Cuomo and Sharif, 1989; Salman et al., 2014). An improved understanding of human cognition when interacting with a computer, indeed, allows the identification of potentially different cognitive dynamics in design and, on the other hand, it informs the development of more responsive ICT systems that improve design support and its effectiveness (Tang et al., 2011). The characteristics of such ICT tools generally make them more suitable to address the later stages of the design process (e.g. embodiment and detailed stages in engineering design) as they typically allow the general layout of the solution to be modelled. ICT tools for sketching are, in turn, those more typically used in conceptual design, in that they substitute the typical pen-and-paper interaction to draw design concepts as they emerge from ideas. Nevertheless, in the last decade, Computer-Aided systems (CAx) to support idea generation have emerged, as well as studies about the interaction their users have with design variables during the formulation of the target characteristics of the design solution and their synthesis (Becattini et al., 2012). The development of new interfaces for human-computer interaction has opened avenues for the development of more effective modalities of

interaction with ICT tools, potentially removing the limitations of the traditional mouse-and-keyboard interaction that binds an individual in front of his or her terminal. Multi-touch surfaces and screens, as well as active goggles, are just a few examples of relatively new technologies that are having an impact on Human Computer Interface (HCI). Such technologies allow for a traditional use by an individual, but at the same time, they allow multiple people to actively interact together. Their integration in ICT tools to support design clearly enables designers to break away from the typical person-to-screen interaction, which still characterizes a large number of Computer Aided Engineering (CAE) collaborative design systems, whether they simply distribute the work among different designers or they more effectively allow collaborative modelling and co-modification (Fuh and Li, 2005).

## 1.2. Analysing the impact of ICT tools on cognitive processes

In order to unveil the cognitive processes of individuals interacting with a computer, the talk aloud protocol, rather than a retrospective investigation through ex-post reflections, is the typically preferred option. However, studies of human behaviour in design with the involvement of design teams (rather than individuals) are also quite typical as these design sessions require communication among the team members, making the spoken interaction one of the main sources through which to explore the cognitive processes of designers (Jiang, 2009). For instance, D'Souza and Dastmalchi (2016) explored the effect of knowledge background and disciplinary affiliation of designers, concluding that they are not necessarily critical to improve the effectiveness of these design sessions. Their findings underline that one of the most critical factors to the effectiveness of the design session (or overall process) is a proper balance between the tendency to work individually and to get involved in the multidisciplinary design team. This is also consistent with Karakaya and Demirkan (2015), as their observations show the importance of the social environment and task motivation. These factors determine a good interaction among participants. Such interaction fosters the externalization of design solutions (usually through sketching), the inherent idea generation, thus design creativity in collaborative digital environments. Along a similar thread, Tang et al. (2011) studied co-located and distributed design processes with the support of ICT tools, comparing their use in collaborative design with traditional face-to-face interaction (using pen and paper) and noting no particular differences in cognition.

New opportunities are being opened up by the emerging technologies that allow the development of potentially more effective ICT tools to support design activities. Kim and Maher (2008) explored the differences between Graphical and Tangible User Interfaces. Their conclusions show that even in small design teams (two people), a shared design representation that involves the sense of touch supported more effective spatial reasoning. Similarly, studies on cooperative design have unveiled the potential of technologies that mix reality with digital contents in collaborative design contexts. These Augmented Reality (AR) technologies span a set of various alternatives to collaborate, refine and visualize a design proposal. According to Gül et al. (2016) they appear to be very good candidates to support co-creative design, in that the engagement of participants appears to be higher. For example, participants will stand up, look around and get closer viewpoints in order to benefit from a flexible interaction with the design objects, like changing their orientation and location.

## 1.3. Analysis methods developed through the SPARK project

Within this framework, the authors are currently involved in a research and innovation project that aims at developing a responsive ICT platform to support collaborative design and check its effectiveness on creativity. The SPARK project (http://spark-project.net - SPatial Augmented Reality as a Key for co-creativity), indeed, expects that Spatial Augmented Reality (augmented reality displayed through projection) has a good potential to support and foster collaborative creative thinking in the design process. The introduction of a mixed prototype (a tangible blank 3D shape and a virtual surface rendered on the shape by projection - Figure 1) allows the participants to visualize a shared design representation of the solution and its evolution during the design process. This helps to reduce language barriers, which can arise due to the diversity of background and sketching skills of the design team member. An ICT tool to support design as above described, allows the participants to avoid wearing invasive device as active goggles, whose prolonged presence over the users' head leads to heating and eyestrain. As one of the project objectives concerns the validation of the effectiveness and efficiency of the SPARK platform

in relevant and real operational environments, the consortium planned to observe these design sessions and compare the inherent process and outcomes with those of sessions carried out with traditional means (co-creative design with pen and paper).



**Figure 1. Some key elements of the SPARK platform: A tablet with the graphical user interface (GUI) to design onto the mixed prototype (above right), as rendered by projection; Real product on the left**

To this purpose, some of the partners composing the consortium have developed various coding schemes and related metrics to map and analyse different facets characterizing the design sessions. Coding schemes for verbal and gestural interactions allow the analysis of cognition and interaction dynamics (Becattini et al., 2017), while criteria for the evaluation of generated ideas assess the effectiveness of the sessions in terms of their creative outputs (Mombeshora et al., 2017).

The first application of the coding schemes and the metrics allowed to analyse the co-creative design sessions carried out with no ICT support. The analysis of these design sessions allowed the identification of needs and demands that the different participants might feel and perceive in their design processes, including potential speculations about the scenario of use with the introduction of the Spatial Augmented Reality platform.

The project is now into the second half of the workplan: the consortium released the first version of the SPARK platform and new experiments to test it are ongoing. The aim of these experiments is to check the impact the platform has on the effectiveness (relevant quantity of ideas, which are sufficiently novel and various to allow for selection among the ones having a higher quality) and the efficiency (lower prototyping costs, reduced need of Human Resources -HR- and reworking activities) of a co-creative design process. To do this, the related experimental campaign foresees one-to-one comparisons between sessions carried out under three different conditions:

1. with the SPARK platform - an ICT tool based on Spatial Augmented Reality;
2. with a simple Augmented Reality platform, mixing digital and real contents displayed on screen (e.g. tablet) instead of projecting it onto the prototype; and
3. without any specific ICT support for designing.

This paper describes the research the authors carried out to refine the coding schemes to map spoken interactions and the creative design moves of participants when co-designing the superficial features of solutions. This, beyond the goals of the SPARK project, can also be of help for similar research that aims at understanding the role that a shared design representation plays in both design review meetings as well as in idea generation sessions (e.g. for the identification of creative stimulation precursors, how it helps to develop effective design team dynamics…). More in details, the research here described has three inter-linked objectives:

- Objective 1: Determine the capability of the coding scheme to map the verbal interactions that occur among the participants, with and without the support of ICT tools for design;
- Objective 2: Assess the capability of the refined coding scheme to map the contents of the co-creative design sessions, whether they focus on product design or package design (and in general whatever content characterizes the surface of a design representation);

- Objective 3: Evaluate the capability of the coding scheme to highlight potential differences between sessions carried out under different operative conditions (different focus of the design task, e.g. packaging vs product interface; with/without the support of ICT tools).

This preliminary verification will also allow us to start comparing the outcomes of design sessions carried out under different operating conditions, providing interesting insights about the execution of controlled experiments using ICT tools to support design and their impact.

The following section briefly presents the overall research methodology, providing additional details about the existing and the refined coding scheme, together with the definition of general hypotheses and assumptions which links the characterization of design moves for protocol analysis with creative design behaviour. Section 3 summarizes the refined and enriched coding scheme for the analysis of design protocols, which the authors estimate to be applicable to co-creative design sessions including (one or more) shared design representations rendered through traditional and ICT support. Section 4 specifies the experimental conditions of the design sessions (composition of design teams, design tasks…) to be analysed with the coding schemes to check its capabilities to fulfil the research objectives. Section 5 presents and discusses the results of the investigation. Section 6 concludes the work, summing up the main findings and the corrective actions to be implemented before reusing the coding scheme and the related metrics for the assessment cognition and the inherent design creativity in design.

## 2. Research approach: Basic assumptions and metrics for the evaluation of results

The authors carried out the research as a step-by-step process composed of five main activities, as shown in Figure 2.
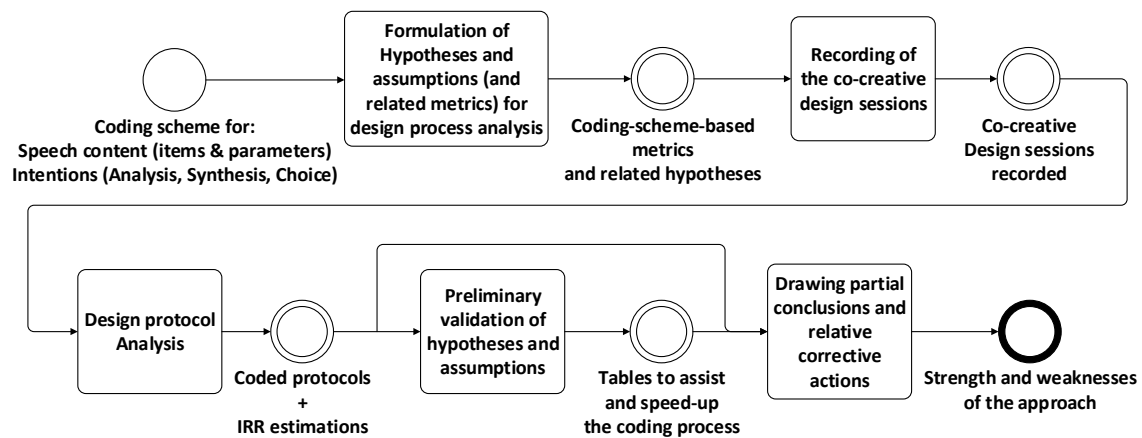


**Figure 2. Research activity diagram**

The research begins with a refined and enriched version of the coding scheme the SPARK consortium had already applied for the analysis of collaborative design sessions (Becattini et al., 2017) and more specifically with the part concerning verbal interactions. That coding scheme however, allowed just to map 'items' and 'parameters'. 'Items' describe contents (e.g. a button, a logo, a fancy tagline, a display or the icons inside…) as elements characterizing the superficial characteristics of the solution. 'Parameters' refers to the parameters the designers' change on items (e.g. position, orientation, size, colour…). This coding scheme has also been enriched so that it helps to identify where designers' utterances mention items and/or related parameters to provide their opinion and criticism on the existing design proposal (Analysis), to suggest a modification for the design proposal (Synthesis) or to clearly state a 'Choice' between/among alternative options. For a more detailed description of the refined coding schemes here mentioned, please refer to Section 3.

As a first step (Figure 2 Activity #1), the authors realized that the coding scheme for spoken interaction by itself was not sufficient to provide the required information about the effectiveness and the efficiency of a design session. In fact, the uneven distribution of codes between two different sessions serves to highlight the different cognitive styles of the participants, but they tell us nothing about the effectiveness

and the efficiency of these sessions. To this purpose, the authors also defined basic assumptions concerning some of the relevant cognitive processes in design. The following bullets summarize the main assumptions behind the collaborative design session dynamics, without distinguishing between the conditions of being supported by ICT tools for design:

- A more fluent idea generation corresponds to a larger amount of design moves characterized by synthesis, regardless if they refer to a complete idea or part of it (Torrance, 1972; Jones, 1984);
- A broader exploration of the design space corresponds to a larger set of ideas that are diverse from each other (Boden, 1994; Shah et al., 2003);
- Within a more fluent production of ideas, the chances of finding an idea of good quality is higher. On the contrary, with low fluency, the chance of finding ideas of good quality is lower (Osborn, 1953). However, it is also possible that ideas of high quality will appear, by method (Altshuller, 1984) or by chance, in design processes characterized by low fluency;
- The quality of interaction in a co-creative design session should be higher if all the participants are actively involved in the analysis of design concepts, in the synthesis of new ideas and changes to existing design, as well as in the choices between alternatives (elaborated upon the conclusions on knowledge externalization and combination as for Civi, 2000).

The above assumptions are also formulated so that it is possible to eventually run meaningful comparisons with other SPARK-related metrics, especially with reference to the final outcomes of the design process (the set of generated ideas). According to the above assumptions, the following metrics, based on the above coding scheme, appear to be reasonably suitable for the investigation of design creativity along the design process:

- *Interaction* represents the degree of involvement of participants in a co-design session. To be measured as the number of shifts in spoken interaction among the different participants.
- *Fluency* represents the quantity of new ideas generated during a co-creative session. To be measured counting the number of shifts between two subsequent design moves (source: Analysis or Choice, target: Synthesis). As a first approximation, it is assumed that subsequent 'Synthesis' segments refer to the same idea.
- *Sharing Index* provides an indirect measure of the need to share a common vision about the design proposals. To be measured as the overall amount of design moves coded as 'Analysis'.
- *Variety* represents an index of diversity of the explored solutions. To be measured as the ratio between the amount of explored item-parameter combinations (in design moves coded as 'Synthesis') and the whole set of the possible item-parameter combinations.

As mentioned above, the intentions of the designer play a central role in the measurement of creativity-related characteristics of the design process. Items and Parameters, on the other hand, are related to the extent of design space exploration. The above metrics, defined before the execution of early trials of co-creative design sessions (Figure 2 Activity #2), allow us to process the coded data emerging from design protocol analysis with no bias coming from the observation, the only bias being the coders' interpretation of segments. To address this issue, multiple coders have been involved in the application of the coding scheme before the analysis of the coded protocols (Figure 2 Activity #3). Furthermore, the outcomes of coding will be checked for an estimation of Inter-Rater Reliability (IRR). The outcomes of the IRR, together with the estimation of the capability of the coding scheme to map the design discourse with minor losses, will also provide evidence of the coding scheme strengths and weaknesses (Figure 2, Activity #5 - Objectives 1 and 2). Still with reference to the overall conclusions about the effectiveness of the coding scheme, it will be also checked against the capability to discriminate between co-creative design sessions (Objective 3) that are intrinsically and purposefully different in some characteristics (as for Section 4). This investigation is to be carried out by interaction, fluency, sharing index and variety metrics (Figure 2, Activity #4).

## 3. Coding scheme(s) refinement and enrichment

The coding scheme previously used by the SPARK project allowed us to identify the main requirement for the platform and thus steer its development, at least for what concerns the content management

system (coding scheme for items) and its main functionalities (coding scheme for parameters) to design onto the mixed prototype. However, the application of the coding scheme highlighted some flaws due to the ambiguity of some codes for both items and parameters, as noted from misalignments between coded protocols. The critical aspects concerned ambiguity of codes for items and parameters and a lack of guidance along the coding process (as different coders can assign the same meaning to a code but have a different interpretation of protocol segments).

Items concerns contents that detail the superficial characteristics of a design representation having a predefined and fixed shape, whether it is generated using ICT or traditional means. With reference to the former version of the coding scheme to map items (Becattini et al., 2017), the term 'Texture' has been rephrased into 'Background Motif'. Now it becomes more general and might better highlight needs of content repetition all over the interested portion of design representation surface (the term 'Texture' remains in the description of the code). The descriptions of all the items have been also enriched with clarifications and details of the related contents. The distinction between Image, Logo and Icon, being all of them typically sketched or computer-generated pictures or their compositions, is kept as these items typically undergo different modifications during the design of the superficial characteristics of the solution. This degree of granularity will be used for further studies, which go beyond the purpose of this paper.

**Table 1. The refined set of codes to map items in spoken interactions**

| Text | What is expressed by words. |
|---|---|
| Image | A sketched or computer-generated picture (potentially vectorial) |
| Photograph | A photograph of a real object (non-vectorial image) |
| Logo | A graphic or textual representation of the brand identity |
| Icon | A graphic or textual representation related to marks of certification or similar |
| Background Motif | A texture or a set of elements characterizing the background of the design |
| System Parts | A reference to part of the entire system |
| Whole | A reference to the design proposal as a single entity |

The codes concerning parameters previously highlighted a good reliability for most of the codes, even if one of them led to the largest set of interpretation ambiguities.

**Table 2. The refined set of codes to map parameters in spoken interactions**

| Position | The geographical location of the item on the design proposal or on one of its parts (i.e. [x,y] coordinates) |
|---|---|
| Orientation | The degree of rotation of the item (i.e. $\theta$ coordinates) |
| Size | It refers to a change in its dimensions (without changing the aspect ratio) |
| Number | The number of Items of the same kind to be added on or removed from the design proposal |
| Presence | The introduction or the removal of an item within the design proposal |
| Colour | The chromatic characteristics of an item |
| Reflectivity | The capability of an item to show glossy or matt properties |
| Material | The physical material that constitutes the item |
| Content | The subject represented by the item (being it graphical or textual) |
| Shape | The representational characteristics (dimensions, proportions…) of an item (i.e: a change in the aspect ratio) |
| Sharpness | The resolution and definition of the representation of the item. |

Therefore, the code 'Look' has been removed from the refined codes, as it generically refers to a vague perception of content and it does not directly relate to a specific variable a designer can work with. The analysis of design moves (on available relevant co-design sessions recordings) characterized by the 'Look' code and their disambiguation allowed for the identification of three additional design-related characteristics as parameters: Reflectivity, Sharpness and Shape. Each code description has been made more accurate and detailed, so as to clarify its meaning to coders and ease their interpretation.

DESIGN ORGANISATION AND MANAGEMENT

As already noted in Section 2, the above codes do not allow for mapping some basic cognitive processes in design, and therefore this gap should be filled in order to evaluate the impact of ICT tools on individuals and group creativity in collaborative design sessions. Also with reference to the large set of codes for both items and parameters, the authors decided to keep the set of codes concerning the design-related cognitive processes of co-designers small in number, in order not to overload the coders with information they have to handle at the same time. To this purpose the six main processes characterizing Gero's non-situated Function-Behavior-Structure (FBS) model (Gero and Kannengiesser, 2004) might be a too large set to be managed with the already presented ones (Tables 1 and 2). Yet, Mc Neill et al. (1998) adopted three codes to characterize the designers' intentions (they called 'micro strategies') behind design moves: Analysis of problems, Synthesis of solutions and Evaluation of solutions (plus Other for residuals). For the purpose of this research, the distinction between Analysis of problems and Evaluation of solutions is weak, as most of the analysis of problems in these co-creative sessions (e.g. accessibility, clarity in product interaction design; communication effectiveness in package design…) is driven by the shared design representation, which is in nature a (partial) description of the solution. The extremely small expected amount of design moves of pure problem analysis (abstract reflection based) would not significantly affect the results of the analysis, as emerged from previous observations. The code Synthesis keeps its original meaning. Moreover, as the SPARK platform and in general ICT tools to support design should also allow easing the selection of alternatives to quickly design more effective solutions (O'Hare et al., expected 2018), the authors decided to introduce an additional code: Choice. It characterizes the design moves where a decision is taken among two or more alternatives. Table 3 presents the codes and their description.

**Table 3. The set of codes to map co-designers' intentions in spoken interactions**

| | |
|---|---|
| Analysis | A design move characterized in that the co-designer speaking is interpreting a design representation, evaluating the extent to which it matches expectations and requirements, and providing judgement about it (positive or negative) |
| Synthesis | A design move characterized in that the co-designer speaking is proposing a solution that can address one of the issues (explicitly emerged or not) concerning the design proposal. |
| Choice | A design move characterized in that the co-designer selects between two or more alternatives about the design proposal, as it can typically be observed after design moves of synthesis or analysis. |
| Other | Other spoken interactions that do not correspond to one of the above (non-design moves) (e.g. spoken interaction describing activities that occurred before the design session, design unrelated comments….) |

## 4. Experimental setting and data processing

In order to check whether the above-described coding scheme matches the expectations set within the research objectives (1 to 3), the consortium planned the execution and the recording of tailored co-creative design sessions, which were also used to check the advancements of the SPARK platform development. The following sub-sections describe the case studies selected, the room and equipment, and the data processing methods.

### 4.1. Case studies for co-creative design sessions

The process of case study selection to run tailored co-creative design sessions followed some basic principles, which are strictly related to the capability to effectively check the fulfilment of the objectives:
- The case studies have to deal with different design domains, in order to check the versatility of the schemes to describe the superficial characteristics of a generic shared design representation;
- The case studies should allow the involvement of different stakeholders of the design process, in order to check its applicability in co-creative design sessions involving a diverse range of participant profiles;
- The case studies should be carried out with different levels of ICT support and different types of ICT support tool to check the ability to assess different types of ICT-supported design activities.

- The case studies are real cases, within the ongoing companies' projects. Expectations of both parties are high, meaning that the interactions make sense for the stakeholders.

Table 4 summarizes the main characteristics of the selected case studies.

**Table 4. Summary of the selected case studies**

| Case study ID | Design domain | Stakeholders' profile | ICT support | Duration |
|---|---|---|---|---|
| Alce Nero 2017 (AN2017) | Packaging Design | 3 Designers<br>2 Company Executives<br>1 Senior account manager | Beta version of the SPARK platform | 100 mins |
| Estel 2017 (E2017) | Product Design | 2 Designers<br>1 Final consumer | Beta version of the SPARK platform | 70 mins |
| Alce Nero 2016 (AN2016) | Packaging Design | 3 Designers<br>2 Company Executives<br>5 Final consumers | None | 40 mins |

Both of the package design case studies involved the same company of organic food producers and sellers: Alce Nero (Italy). They are one of the clients of the Italian design agency of the SPARK consortium (Artefice, brand language design) and are a relevant Spatial Augmented Reality platform end user. The case study labelled 'AN2017' concerned the conception of new packs for consumer products including baby food (see Figure 1) and a powdered preparation for baby rice. The 'AN2016' packaging design case study, concerned the redesign of a pouch-shaped pack for biscuits. The case study on product design concerns the refinement of the superficial characteristics of a sports device to measure workload-related parameters after physical activities in a gym. This case study has been led by one of the other design agencies partners of the SPARK consortium, Stimulo Design (Spain), with a potential end-user.

The designers involved in all these sessions are professionals and expert in their respective field. In ICT supported conditions, one of the designers interacts with the platform GUI and the other one shows a presentation on a large TV screen, operating from a PC and providing a more significant contribution to the facilitation of the co-design session, with respect to the one dealing with the GUI. The third designer in the Alce Nero 2017 session more intensively facilitates the session, while the one showing the presentation supports him. The final consumers that participated in the session all had a profile that matched the target audience of the solution under development. The executives were from the marketing and departments of the client company. The senior account managers from the design agency work closely with clients to understand their needs and expectations and make them clear to the creative designers.

## 4.2. Room and available equipment

All the sessions have been carried out in sufficiently large rooms, with regards to the number of participants. Apart from the Alce Nero 2016 session, the co-creative design sessions took place at Politecnico di Milano, in the room shown in Figure 3 (approx. 26,5 sqm). It also shows a typical layout for the SPARK platform, including the related hardware and software equipment: an interaction device with a SAR module installed (PC, tablet…), multiple projectors, and the prototype with markers on it. The tracking system (tracking cameras are the red devices seen top left) completes the equipment. A tailored software installed on a server, allows all these devices to interoperate and work as a system so that the projection naturally follows the position of the prototype when participants manipulate it, within the boundaries of the working area (black area on the left end of the table left). The projection renders the digital content onto the surface of the blank tangible prototype. Changes made through the GUI are shown in real time on the prototype.

The Alce Nero Session 2016 was carried out in a similar room (approx. 32 sqm). It was equipped with a long table suitable for 10 people, a large TV screen, mock-ups of some previously developed design proposals, and basic design materials (pens, paper and sticky notes).
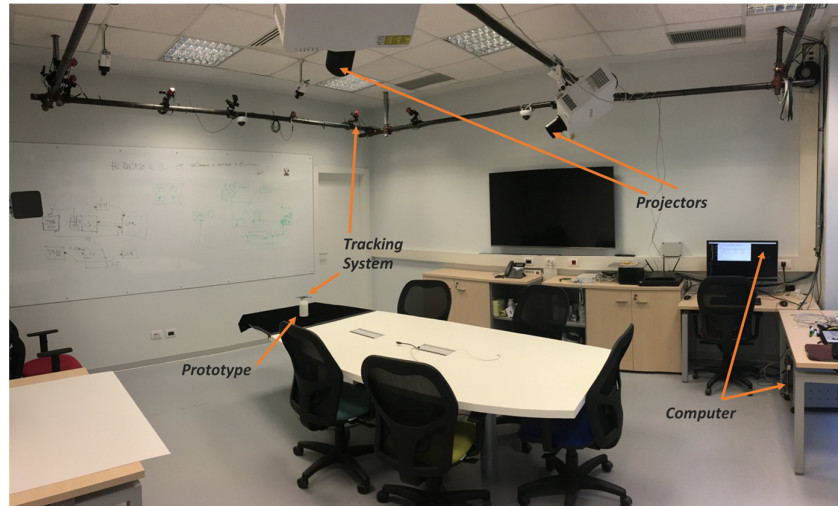
DESIGN ORGANISATION AND MANAGEMENT

**Figure 3. SPARK platform layout**

## 4.3. Data processing

All the co-creative design sessions were Audio/Video recorded with multiple cameras and individual lapel microphones to ensure a complete coverage of the scene. These protocols were transcribed and translated into English as the co-creative design sessions were in native language of participants (AN16/17 in Italian, E17 in Spanish). This allows coders from different regions to understand the contents and meaning of the verbal interactions properly and conduct their own coding of the content. A crosscheck among the authors ensured a uniform segmentation of the three protocols. Three coders, none of whom had directly observed the sessions, received instructions to work independently from each other so as not to bias the results. They were told to code each segment with just one code for the intentions, and one or more codes for both items and parameters, segments tagged as 'Other' should not be coded for items and parameters (these segments, indeed, do not refer to the design representation at all and therefore they hold no contents to be specified). The inter-rate reliability (IRR) of the three different coded protocols has been checked with Fleiss' Kappa (1971) before extracting the most recurring code(s) for each design move and analyse the protocol with the metrics described in Section 2. The next section presents and discusses these results.

## 5. Results of protocol coding, analysis and related discussion

In order to determine the capability of the coding scheme to map the verbal interactions among the co-creative design session participants, (independently from the availability of an ICT tool - Objective 1 - and from the design domain - Objective 2), it is required to:

- evaluate if coders encountered difficulties or ambiguities in the individual activity of coding (e.g. leaving transcripts segments with no code characterizing the segments);
- determine the reliability of coding, by assessing the degree of agreement among the separately coded protocols of the three sessions with Fleiss' Kappa.

**Table 5. Percentage of coded segments (C-x columns refer to the three coders)**

| Session ID | Segments [#] | Intentions (ref. Table 1) | | | Items (ref. Table 2) | | | Parameters (ref. Table 3) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | C-1 | C-2 | C-3 | C-1 | C-2 | C-3 | C-1 | C-2 | C-3 |
| AN2016 | 476 | 100% | 99,4% | 100% | 97,7% | 100% | 99,6% | 86,8% | 97,5% | 99% |
| AN2017 | 581 | 99,3% | 98,8% | 98,6% | 100% | 100% | 100% | 100% | 100% | 100% |
| E2017 | 371 | 99,5% | 97,6% | 98,8% | 93% | 100% | 99% | 83,9% | 91,7% | 95,7% |

A more careful analysis of the distribution of results show generally good performance in terms of code coverage of segments, regardless of the session (i.e. the design domain), of the specific kind of codes as

well as of the length of the session, in terms of amount of segments to be coded (Objectives 1 and 2). The best results in terms of segments coverage come from codes related to intentions (Table 1) and items (Table 2 - where relevant, i.e. not for segments coded 'Other' for intentions). Codes about parameters (Table 3), on the other hand, appear to be less clear, or their identification is not obvious, despite their overall evaluation is satisfactory, as on average less than the 5% of design moves remain without coding. However, unlike the other two cases, this happens under greater variability: coder 1 significantly affects the results as his coding performance suddenly dropped, which suggests a source of potentially ambiguous interpretations or an excessive coding workload, as this is typically the last code to be assigned for each segment and some fatigue effect can emerge due to working memory overload (segment by segment, codes of parameters were assigned last).

The results of Fleiss' Kappa estimation show the extent of agreement among coders and also highlights potentially relevant weaknesses of the coding scheme to be addressed with higher priority. Arithmetic average values of the indexes collected in Table 6 would not be appropriate, as these values come from non-linear calculations, which also depend on the number of segments to be coded and the number of involved coders.

**Table 6. Values of Fleiss' Kappa to estimate IRR for the coded protocols**

| Session | Intentions | Items | Parameters |
|---------|------------|-------|------------|
| AN2016 | 0,91 | 0,54 | 0,39 |
| AN2017 | 0,40 | 0,65 | 0,51 |
| E2017 | 0,20 | 0,60 | 0,53 |

Nevertheless, Table 6 show very contradictory results, which are however consistent with the degree of development of the coding scheme and its maturity of application. The recently introduced set of codes to map intentions (Table 1) provide reliable results with low repeatability: this very low IRR in product design (E2017) sessions is compensated by an extremely high IRR in the field of packaging design (AN2016). This preliminary identification of challenges in product design cannot be conclusive, as also the second package design session got a quite low IRR score for this class of codes. For what concerns the other two families of codes (items and parameters, respectively as for Table 2 and 3), their large variability, together with the possibility to characterize each design moves with multiple codes, intrinsically complicates the achievement of high scores of Fleiss' Kappa. This also partially explains the existing gap between codes for items and parameters, since the latter ones have more different variants (approx. +35/40%). To this perspective, the authors believe that the set of codes presented in Section 3 is correct and sufficiently comprehensive, as significantly lower values of IRR should be expected after one of the first applications. Suggestions for how to improve the IRR performance with this coding scheme are presented in the discussion section.

In order to evaluate the effective capability of the coding scheme (as for Tables 1, 2 and 3 – Figure 2, Activity #4) to highlight potential differences between design sessions carried out under different conditions (Objective 3), Table 7 provides figures for the metrics described in Section 2 (both observed and standard scores, which have been normalised according to the overall number of segments in the protocol).

**Table 7. Summary of the values for the selected metrics for co-creative assessment organized by case study (left: observed figures, right: standardized figures)**

| | AN2016 | | AN2017 | | E2017 | |
|---|--------|-----|--------|-----|-------|-----|
| | Observed | Std | Observed | Std | Observed | Std |
| Interaction | 293 | 0,57 | 418 | 0,72 | 294 | 0,79 |
| Fluency | 2 | 0,004 | 61 | 0,10 | 26 | 0,07 |
| Sharing Index | 450 | 0,95 | 315 | 0,54 | 230 | 0,61 |
| Variety | NA | | 0,23 | | 0,08 | |
| SEGMENTS | 476 | | 581 | | 371 | |

DESIGN ORGANISATION AND MANAGEMENT

At first glance, the lack of empty cells in the table suggests that the coding scheme allows for metrics application without particular drawbacks, with the only exception of the Alce Nero 2016 session. Here, most of the design moves are about Analysis with extremely poor or null activity of Synthesis (as confirmed by the Fluency index), which results in no meaningful result for Variety for this session.

As, during analysis, the exploration of the design space concerns the identification of requirements and constraints to select more promising opportunities for idea development, the authors expect to split the evaluation of variety in the next applications of the coding scheme. The distinction between variety during analysis and synthesis will also help to differentiate the exploration of the design space between the problem and the solution space (Dorst and Cross, 2001). Values about interaction show a very strong similarity between the two sessions carried out by means of ICT tools to support design (a prototype version of the SPARK platform) and a marked difference with reference to the co-creative design session of 2016, carried out with traditional means. With the introduction of a real-like shared design representation, which probably facilitates the understanding and knowledge harmonization among co-designers, the various stakeholder involved reduce their inhibition and more actively participate the session (around +30% of participation).

## 6. Conclusions

This research paper describes the refinement and testing process of a coding scheme for the analysis of co-creative design protocols, which potentially include the use of ICT technologies, e.g. Augmented Reality, to visualize design concepts as a mixed prototype (partially tangible, partially virtual) that plays the role of shared design representation for co-designers.

A repeatable research method for coding scheme refinement is proposed and applied within the context of the SPARK project with three case studies, each describing a full design session. With reference to the research objectives, the paper clearly states the overarching assumptions behind the research and the evaluation of creative behaviour.

The results of the analysis, as emerged from the three case studies under investigation, show that the most recent updates introduced in the coding scheme probably require further time to be internalised and proficiently used by coders for design protocol analysis. On the other hand, this also suggests that coders should receive a more solid training before starting the coding activity to achieve more convergent results if compared to each other. Additional coding instruments, such as more detailed coding guidelines and a coding book showing examples of real content and the agreed approach to coding them, are currently under development and they will be explored in next publications (as for Figure 2, Activity #5).

For what concerns the capability to define the similarities and differences between session sharing some characteristics, the paper suggests some original indexes to assess the interaction and information sharing by session participants. On the other hand, the results show the need of improving the way variety is computed, as the current results omit the exploration of the problem space, focusing just on the solution space. Very similar results for verbal exchanges among co-designers emerge in sessions where ICT tools to support design are introduced suggesting that this metrics has a higher chance of providing accurate results.

The team of the SPARK consortium involved in the development and refinement of coding schemes for verbal interaction plans to run further trials on existing recordings, involving different coders to evaluate the effectiveness of a more comprehensive training-before-coding on creating a shared understanding of what each code means and why it has to be associated with a design move (as for Figure 2, Activity #6). Once the coding scheme and training procedure has been refined through further trials, we hope that it will provide a robust method for researchers to analyse the verbal content and interactions amongst participants of co-creative sessions. This will enable greater insight into the role that ICT support tools play in enabling creative design moves.

### Acknowledgements

# References

Becattini, N., Borgianni, Y., Cascini, G. and Rotini, F. (2012), "Experience protocol analysis of computer-guided design tasks", *Proceedings of DESIGN 2012 / 12th International Design Conference, Dubrovnik, Croatia*, The Design Society, Glasgow, pp. 1821-1830.

Becattini, N., Masclet, C., Ben-Guefrache, F., Prudhomme, G., Cascini, G. and Dekoninck, E. (2017), "Characterisation of a co-creative design session through the analysis of multi-modal interactions", *International Conference on Engineering Design 2017*, The Design Society, Glasgow, pp. 479-488.

Civi, E. (2000), "Knowledge management as a competitive asset: a review", *Marketing Intelligence & Planning*, Vol. 18 No. 4, pp. 166-174. https://doi.org/10.1108/02634500010333280

Cuomo, D.L. and Sharif, J. (1989), "A study of human performance in computer-aided architectural design", *International Journal of Human-Computer Interaction*, Vol. 1 No. 1, pp. 69-107. https://doi.org/10.1080/10447318909525958

Dorst, K. and Cross, N. (2001), "Creativity in the design process: co-evolution of problem–solution", *Design Studies*, Vol. 22 No. 5, pp. 425-437. https://doi.org/10.1016/S0142-694X(01)00009-6

D'souza, N. and Dastmalchi, M.R. (2016), "Creativity on the move: Exploring little-c (p) and big-C (p) creative events within a multidisciplinary design team process", *Design Studies*, Vol. 46, pp. 6-37. https://doi.org/10.1016/j.destud.2016.07.003

Fleiss, J.L. (1971), "Measuring nominal scale agreement among many raters", *Psychological Bulletin*, Vol. 76 No. 5, pp. 378-382. https://doi.org/10.1037/h0031619

Fuh, J.Y. and Li, W.D. (2005), "Advances in collaborative CAD: the-state-of-the art", *Computer-Aided Design*, Vol. 37 No. 5, pp. 571-581. https://doi.org/10.1016/j.cad.2004.08.005

Gero, J.S. and Kannengiesser, U. (2004), "The situated function–behaviour–structure framework", *Design Studies*, Vol. 25 No. 4, pp. 373-391. https://doi.org/10.1016/j.destud.2003.10.010

Gül, L.F., Halıcı, M., Uzun, C. and Esengün, M. (2016), "Understanding the Impact of Mobile Augmented Reality on Co-design Cognition and Co-modelling", *International Conference on Cooperative Design, Visualization and Engineering*, Springer, pp. 362-370. https://doi.org/10.1007/978-3-319-46771-9_47

Jones J.C. (1984), "A method of systematic design", In: Cross N. (ed.), *Development in Design Methodology*, John Wiley & Sons Ltd, Chichester, pp. 9-31.

Karakaya, A.F. and Demirkan, H. (2015), "Collaborative digital environments to enhance the creativity of designers", *Computers in Human Behavior*, Vol. 42, pp. 176-186. https://doi.org/10.1016/j.chb.2014.03.029

Kim, M.J. and Maher, M.L. (2008), "The impact of tangible user interfaces on spatial cognition during collaborative design", *Design Studies*, Vol. 29 No. 3, pp. 222-253. https://doi.org/10.1016/j.destud.2007.12.006

Mc Neill, T., Gero, J.S. and Warren, J. (1998), "Understanding conceptual electronic design using protocol analysis", *Research in Engineering Design*, Vol. 10 No. 3, pp. 129-140. https://doi.org/10.1007/BF01607155

Mombeshora, M., Dekoninck, E., O'Hare, J., Boujut, J.F. and Cascini, G. (2017), "Applying multiple metrics in the performance measurement of design sessions in industry: a co-design case study", *Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 2: Design Processes, Design Organisation and Management, Vancouver, Canada, 21 - 25.08. 2017*, The Design Society, Glasgow.

Osborn, A.F. (1953), *Applied imagination*, Scribner'S, Oxford.

Salman, H.S., Laing, R. and Conniff, A. (2014), "The impact of computer aided architectural design programs on conceptual design in an educational context", *Design Studies*, Vol. 35 No. 4, pp. 412-439. https://doi.org/10.1016/j.destud.2014.02.002

Shah, J.J., Smith, S. M. and Vargas-Hernandez, N. (2003), "Metrics for measuring ideation effectiveness", *Design Studies*, Vol. 24 No. 2, pp. 111-134. https://doi.org/10.1016/S0142-694X(02)00034-0

Tang, H.H., Lee, Y.Y. and Gero, J.S. (2011), "Comparing collaborative co-located and distributed design processes in digital and traditional sketching environments: A protocol study using the function–behaviour–structure coding scheme", *Design Studies*, Vol. 32 No. 1, pp. 1-29. https://doi.org/10.1016/j.destud.2010.06.004

Torrance, E. (1972), "Predictive validity of the Torrance tests of creative thinking", *The Journal of Creative Behavior*, Vol. 6 No. 4, pp. 236-262. https://doi.org/10.1002/j.2162-6057.1972.tb00936.x

Dr. Niccolò Becattini, Post Doc
Politecnico di Milano
Via Giuseppe La Masa 1, 20156 Milano, Italy
Email: niccolo.becattini@polimi.it