21ST INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED17
21-25 AUGUST 2017, THE UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, CANADA

APPLYING MULTIPLE METRICS IN THE PERFORMANCE MEASUREMENT OF DESIGN SESSIONS IN INDUSTRY: A CO-DESIGN CASE STUDY

Mombeshora, Mendy (1); Dekoninck, Elies (1); O'Hare, Jamie (1); Boujut, Jean-François (2); Cascini, Gaetano (3)
1: University of Bath, United Kingdom; 2: Université Grenoble Alpes, France; 3: Politecnico di Milano, Italy

Abstract

The recently launched SPARK project (www.spark-project.net), aims to understand how Spatial Augmented Reality (SAR) technology can be used to support the co-design process. The project aims to develop, build and test a SAR technology for co-design. This paper presents the testing of a suite of metrics to assess the performance of co-design sessions in industry. These metrics will be needed in the SPARK project to help iteratively develop that technology, as well as to provide insights into the overall effects of using SAR tools in co-design. The aim of the study reported in this paper was to design, develop and test a process for applying a suite of co-design session metrics. During the study, the metrics application process was applied to two industry cases and based on the outcomes, refinements were made. This resulted in a proposed process for the evaluation of the effectiveness of co-design sessions. The work presented in this paper will be of particular interest to other researchers seeking to evaluate the impact of a product development tool on a specific design session.

Keywords: Collaborative design, Spatially Augmented Reality, Case study, Research methodologies and methods, Performance measurement

Contact:
Dr. Elies Dekoninck
University of Bath
Mechanical Engineering
United Kingdom
e.a.dekoninck@bath.ac.uk

Please cite this paper as:
Surnames, Initials: *Title of paper*. In: Proceedings of the 21st International Conference on Engineering Design (ICED17), Vol. 2: Design Processes | Design Organisation and Management, Vancouver, Canada, 21.-25.08.2017.

1 INTRODUCTION

Consisting of projecting information coming from the digital world directly in physical space, Spatial Augmented Reality (SAR) has the ability to produce immersive contents by overlapping virtual and real-world environments; going beyond conventional display methods based on screens and planar projections. Through a combination of dynamic projection mapping, multiple perspective views and device-less interaction, it supports face-to-face interaction with 3D virtual objects; opening up new perspectives for future applications in a diverse range of areas. One such area is the product development (PD) process. Co-design, a recent and growing trend within product development, is defined as "the activity of designers and people not trained in design working together in the design development process". In co-design, non-designers such as users, customers and clients are playing an increasingly important role in the PD process; shifting their role from being the 'subject' of study (through approaches such as being usability testing) to being a 'partner' in the process (Sanders and Stappers, 2008). This transition has primarily been driven by the belief that if designers can engage with users in a more profound, expansive and co-operative manner during the design process then the chance of delivering a commercially successful final product that effectively fulfils user requirements increases. Responding to the co-design trend, the research community has begun to propose tools and approaches aimed at improving the effectiveness and efficiency of co-design activities (Sanders and Stappers, 2014). SAR has been noted as a technology that seems particularly relevant as a basis for co-design support tools. It is this potential that served as the starting point for the on-going SPARK project (www.spark-project.net); whose aim is to understand how SAR can be used to support the co-design process and develop, build and test a SAR-based support tool for co-design.

The SPARK project focuses particularly on the co-design session, which is defined as 'a pre-arranged session that involves designers and people not trained in design working together in the design development process'. This definition allows for the inclusion of a broad range of activities such as idea generation, reviewing and selecting concepts, creating plans for testing, etc. However, some features of the co-design session, such as lack of a shared design vocabulary and designers acting as sole creative experts (Taffe, 2005) present new challenges for creative practitioners trying to facilitate this type of session. Lee (2008) identified several of these challenges, these include the mixed terminology used in the field and some design skills – such as aesthetics – not being transferrable to non-designers. Co-design sessions seem to be important, but potentially challenging, activities within the PD process. The development of support tools for co-design is important as they can determine the likelihood of success of a design activity and the users' willingness to participate in future co-design activities (Füller et al., 2009). As the SPARK project aims to address some of these challenges, a key starting point for the research was understanding how the performance of co-design sessions can be measured. Without adequately measuring its performance, it is impossible to evaluate if any proposed SAR-based tool enhances the performance of co-design sessions or not. Resultantly, a performance measurement system for co-design sessions was required.

1.1 Developing a Performance Measurement System

When regarding organisations and business processes, a myriad of interpretations of the term 'performance' exist; making the concept particularly ambiguous in a product development context. Upadhaya et al. (2014) define performance measurement as the process of collecting, analysing and/or reporting information regarding the performance of an individual, group, organisation, system or component. When the subject is the product development process (or 'the design process'), complexities arise due to its creative and multidisciplinary nature, and variation in contexts. To satisfy customer needs, the development of a product can be initiated, executed and concluded in numerous ways, making the evaluation of the product development process' performance difficult. Cedergren et al. (2010), for example, suggest characterising the product development process through the categorisation of its composite activities and using those as a basis for performance measurement. Activities are the basic components of the product development process and through them, 'performance' can be observed when inputs are transformed into outputs (O'Donnell and Duffy, 2002). There is a widespread use of the terms 'effectiveness' and 'efficiency' to describe performance, however, these have not been consistently defined within a product development context and their application has yielded a variety of interpretations. While effectiveness is generally related to the development of the right product and efficiency the development of the product in the right way, O'Donell and Duffy (2002) emphasise the importance of clearly distinguishing them. In this work: effectiveness measures relate to the extent to which a particular co-design session - within a product development process - meets the objectives and/or expectations of the stakeholders involved. Efficiency is viewed in terms of how well a design organisation uses resources, such as funding and staff, to achieve its objectives.

Taking this broad view of effectiveness and efficiency meant a heterogeneous family of measures would be needed to capture before- and-after changes in both. Other authors have also taken similar approaches; for example, Chidamber and Kemerer (1994) suggest that it is possible to use a suite of metrics to measure the impact of the technology on business processes and Behn (2003) proposes that a collection of performance measures with the various characteristics necessary to help you (directly and indirectly) achieve your (research) purposes is needed. The research purposes in the case of the SPARK project are: to help iteratively develop a SAR technology-based co-design support tool; as well as to provide insights into the overall effects of using SAR technology-based tools during co-design. Bhuiyan (2011) highlights that NPD process metrics typically focus on a single activity or the entire process; however, as no one function is the sole contributor to the product development process, she emphasises the importance of measuring the performance of stages of the NPD process in an interdependent fashion. Succeeding a review of academic literature on product development metrics and an industry metrics survey, an approach for the measurement system for the SPARK project was established. The approach aims to provide a suite of metrics for measuring the performance of specific co-design sessions as well as the whole product development process. These metrics will be used to help iteratively develop the SAR technology based tool, as well as provide insights into the overall effects of using SAR technology-based tools in co-design. Using the right type of measures is paramount as a potential conflict between the efficiency of the process and the quality of design outputs exists. The developed metrics were split into three different categories: design process efficiency; quality of outputs and effectiveness of co-design sessions; effectiveness of interaction between participants and with prototypes. This paper reports on the metrics within the measurement system related to the second category (not those related to interaction or the overall process).

1.1.1 The Quality of Outputs and Effectiveness of Co-Design Sessions Metrics Suite

Initially proposed in in O'Hare et al. (2016), the metrics related to quality of outputs and effectiveness of co-design sessions are intended to be used together to assess the effectiveness of a range of different types of co-design sessions: from idea-generation focused sessions; to review-focused sessions (this range of different co-design sessions was previously identified from interviews with 13 design companies). Some of these metrics were based on established metrics but adapted for application within the SPARK project, whereas some are entirely new. One of the novel approaches proposed in O'Hare et al. (2016) was to use a 'before and after' morphological chart to collect data for some of the proposed metrics; namely the 'quantity', 'variety' and 'Filtering Effectiveness' metrics. The novel suite of quantitative metrics that can be used to evaluate the performance of co-design sessions, along with a description of each of the metrics and how to calculate them, is presented in Figure 1.

Effectiveness of Idea Generation	Quantity	New ideas generated during session *No. of Sub-Solutions in After Morphological Chart - No. of Sub-Solutions in Before Morphological Chart*
	Quality	New ideas generated that are taken forward at the end of the session for further development *No. of New Ideas – No. of Rejected New Ideas*
	Variety	How varied the ideas generated during the session are *(No. of Features Rows with New Sub-Solutions ÷ No. of Features Rows in 'Before' Chart) + (No. Of New Features Rows ÷ No. of Features Rows)*
	Novelty	How novel the ideas generated during the session are *Mean Average Novelty Score from Three Experts for Each Idea Generated*
Idea Evaluation and Task Progress	Filtering Effectiveness	The success of the idea filtering activities during the session *No. of Ideas Rejected ÷ (No. of Ideas Considered – Desired No. of Ideas to Keep)*
	Task Progress	How the participants in the session progress through tasks *Weighted Scoring for Tasks Created or Completed*

Figure 1. Metrics suite – including metric descriptions and how they are calculated

The aim of the study reported in this paper was to design, develop and test a process for applying the proposed metrics suite in a real industrial case study. This paper reports on the development of an application process for the metrics and an initial trial to apply the suite to 'normal' co-design sessions where design consultants were working with non-designers but without SAR technology-based support tools. Resultantly, the results of the work presented in this paper will be of interest to other researchers wanting to evaluate the impact of a product development tool on a specific design session.

2 THE RESEARCH DESIGN AND METHODOLOGY

With the aim of designing, developing and testing a process for applying a suite of metrics to a co-design session during the product development process, this study had three research objectives. The first objective was to establish how information required for the metrics application can be generated from collected evidence; the next objective was to design and test a process for applying the metrics suite to co-creative sessions; and the last objective was to make recommendations to support and improve how metrics suites can be applied to design sessions. A pragmatic approach - similar to the first action research cycle (McNiff and Whitehead, 2006) – was broken down into three stages of the research. The first stage– process design – involved the proposal of a metrics application process. The second stage – process testing – applied the proposed process to real-life cases. As it was likely that many shortcomings of the process would be identified during the testing phase, the third stage – process refinement – involved enhancing the initially proposed process.

2.1 The Application Cases

Two product development projects were selected as test cases for the study; both within the same organisation – Stimulo – a Spanish product design consultancy with international clients. These projects were carefully selected for inclusion as they were deemed reasonable and information-rich for the purposes of this study in particular. The co-creative sessions within the projects that were selected as the focus of the of the metrics application offered a complexity and variation in focus that would test the proposed process under differing conditions. Case 1 was a co-creative session whose main aim was to collaboratively review previously elaborated ideas regarding product 'look and feel' with the client. The session was part of a long project to design and develop a communication device for personal rescue. The session occurred between the 'ID definition' and 'technical approach' phases of their product design and development process. Case 2 occurred during the 'technical' phase of the design and development process of an outdoor cooking device. The main aim of that session was to review, with the client, previously discussed ideas regarding target users, cost, and product assembly. It is to these two cases, that the proposed metrics application process was applied during testing phase of the study.

3 METRICS APPLICATION PROCESS DESIGN

The aim of the process design phase was to establish a practical metrics application process which could subsequently be tested and refined through live application. Before the process could be designed, it was essential to establish the type of evidence that would need to be collected in order to generate the information required for the application of the metrics and how it could be collected. Following the establishment of the evidence to be collected, the three-phase metrics application process, presented in Figure 2, was developed. It is this process that was to be applied to the two application cases.

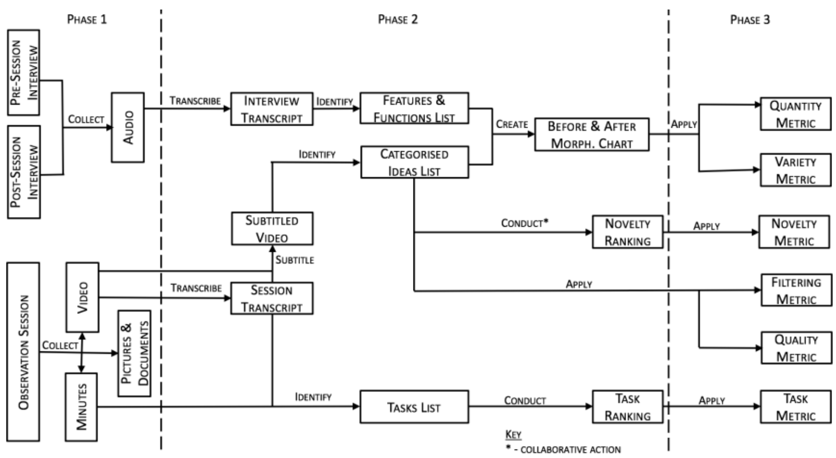

Figure 2. Initially proposed metrics application process

It is this process that was to be applied to the two application cases within the study. The first phase is the evidence collection phase, during this phase evidence required for the application is collected through recording the co-creative observation session and conducting pre- and post-observation session interviews with creative practitioners involved in the observation session. The pre- and post-observation session interviews take the form of semi-structured interviews, directed by interview protocols, whose aim is to gain a holistic understanding of what goes into the observation session and what comes out of it, respectively. This is followed by the processing and analysis phase where the information required for the application of the metrics is generated from the collected evidence. This phase mainly involves transcribing the recorded sessions, identifying ideas and tasks within the transcripts which constitute the information required for the metrics application, creating morphological charts and completing rankings. The last phase is where the information generated is input into the metrics calculations giving the performance measurement of the observed session.

4 EVIDENCE COLLECTION, PROCESSING AND DATA ANALYSIS

Following the initial proposal in the process design phase, the process testing phase – where the proposed application process was applied to the two industry cases – was undertaken. This section reports on that part of the process from evidence collection to information generation in the initially proposed metrics application process. As directed by the proposed application process, observation sessions were recorded and pre- and post-session interviews with the creative practitioners involved in their co-creative sessions were undertaken and recorded. Figure 3 provides details of session and interview participants and duration for each of the application cases. The observation session for Case 1 was conducted in Catalan, with the other application case and all the interviews conducted in English.

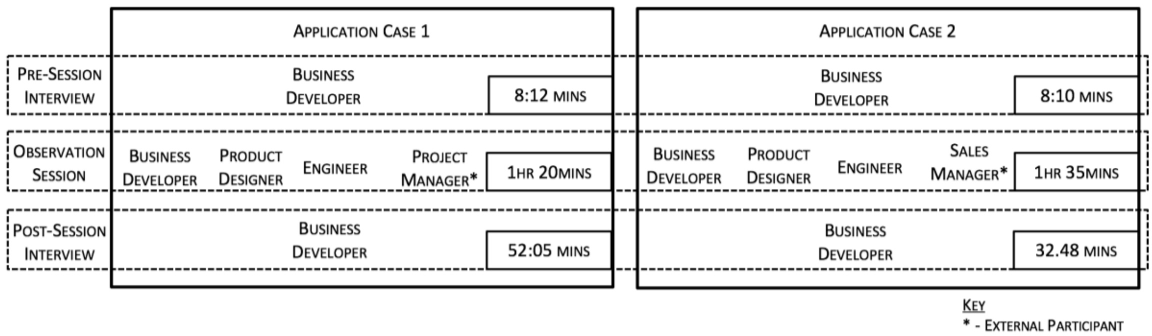

Figure 3. Evidence collected from application cases

During evidence collection, audio recordings of the interviews and video recordings of the observation sessions were made. The observation sessions were video recorded, in particular, on the basis that analysing the data as a subtitled video would make the process of identifying ideas easier than analysing the transcript due to the added visual aspect. While the transcribing process was straight-forward, with the Catalan session transcript being translated into English, the video subtitling proved challenging. This was particularly true for Case 1 in Catalan as the subtitling process was not only time-consuming but synchronising text to audio was made difficult due to language differences. Additionally, due to the camera angles, the video recordings did not offer a sufficient visual advantage to make the effort of subtitling worthwhile. Resultantly, instead of analysing the video recordings, the transcripts of the sessions formed the main data set to be analysed. In addition, to the audio and video recordings, pictures of the design representations that were used during the sessions were taken and other supporting documents, such as session presentations and technical drawings, were collected.

4.1 Identifying Ideas and Tasks

Through coding, the session and interview transcripts and session minutes were systematically searched to identify and categorise specific content of interest. The coding process was supported by the computer aided data analysis software Nvivo and resulted in identified and categorised ideas, features, functions and tasks needed for the metrics. For the purpose of identifying ideas, the definition of an idea as "a creative output and generative proposition linked to the function, behaviour or structure of a concept" by Howard (2008) was adopted. As a co-creative session is viewed as a meeting within which tasks are carried out with the aim of achieving a goal, therefore the definition of a task as "an activity the task doer performs in order to accomplish a goal" put forth by Vakkari (2003) was adopted. In this case, a task is similar to an 'action item' typically identified in business meetings, except that 'action items' normally have to be carried out after the meeting, whilst tasks can be 'cleared' during co-creative design meetings. During the analysis, it became evident that while the interviews gave an informative overview of the sessions, they were not as rich as anticipated and did not contain detailed enough information compared to the observation sessions. This was mainly due to the interviewees not going into enough detail and the facilitator not having sufficient knowledge of the project and the session which would allow them to probe for more detailed information. Resultantly, the observation session transcripts formed the focal point of the data analysis.

4.2 Creating the Ideas Chart

Following the identification and categorisation of ideas and the features and functions they relate to through coding, the 'ideas chart' was created as a means of presenting them in a digestible and easy to use form. The ideas chart is a table that is used to generate the required information that is related to the metrics by categorising the identified ideas accordingly. It is based on a combination of the 'categorised ideas list' and the 'features and functions list' shown the Figure 3. The headings of the chart are as follows: Feature/Function, Idea Summary, Idea Number, Location, Relevant to Session (Yes/No), Type of Idea (New/Old), Idea Status (Rejected/Not Rejected). Through the use of the chart, each identified idea is summarised, the feature/function it relates to is recorded, along with its relevance to the session, if it is a newly generated idea, and its status at the end of the session. As the chart is designed to be used in conjunction with an annotated copy of the transcript, the location of the idea in the transcript is also recorded; this creates a quick way of referring back to source of the idea. Different methods of completing the ideas chart were used for each of the application cases. The first method, used for Case 1, had the researcher completing the ideas chart before checking through the categorisations with the creative practitioners involved in the session. This method proved to be very complicated and time consuming as an in-depth knowledge of the product was required to adequately understand and summarise the ideas that were being discussed - knowledge that the researcher lacked. This resulted in the development of the second method, the method adopted for Case 2. The second method was more collaborative in nature, with the facilitator filling in an empty ideas chart together with a creative practitioner who was involved in the session; with the annotated transcript guiding the process. While this method makes the process of completing the ideas chart easier, it requires more input from the creative practitioners. The use of both methods resulted in completed ideas charts, while the method used for Case 2 was quicker and less demanding on the researcher, the one used for Case 1 forced the researcher to gain an in-depth understanding of the content of the sessions putting them in a more informed position. The preferred approach would be to learn as much as possible about the project, product and session before analysing the data and then adopting the method used in Case 2, completing the ideas chart with a practitioner involved in the session.

4.3 Creating the Morphological Charts

The metrics suite was created in such a way that a morphological chart is used as the base framework for analysis. Specifically, during the pre-session interviews the creative practitioner would explain the concepts that were to be presented in the observation session and this information, along with pictures of the design representations used within the session would be used to create a 'before' morphological chart representing the start of the session. Similarly, an 'after' morphological chart would be constructed based on the ideas still under consideration at the end of the session. The difference between the two charts would show the evolution that occurred over the course of the session. However, due to insufficient information being collected during the pre- and post-session interview and adopting the method where the facilitator and practitioner worked collaboratively to create the ideas chart, the use of the morphological charts was changed. Instead, they too were completed collaboratively by the facilitator and the practitioner. As there was insufficient information in the interviews, instead of figuring out what went into the session and creating a 'before' morphological chart, the opposite occurred where details of what happened during the session were used as the starting point of the morphological chart creation. Following the completion of the ideas chart, the newly generated ideas were filled into an empty morphological chart based on the features and functions they relate to. For each of the ideas in the morphological chart, the practitioner would provide information on old ideas that were being considered that relate to the same feature/function, where no prior solution existed a new features row was created. The resultant session morphological chart is a combination of the ‘before’ and ‘after’ morphological charts; ideas that correspond to before and after were highlighted to differentiate them. Figure 4 presents an example of what the session's morphological chart would look like and the information that you can get from it. After the discussion, the morphological chart includes all the ideas (old and new) that are related to the product and that are of interest to the session, then it can be used to generate information that can be used to apply the filtering effectiveness metric. This is the filtered morphological chart; an example of this chart is presented in Figure 4. It is important to note that in this case the old ideas include those that have been identified in the session transcript and also those that have been supplied by the collaborating practitioner. Similar before-and-after comparison approaches are used to get the 'variety' metric.

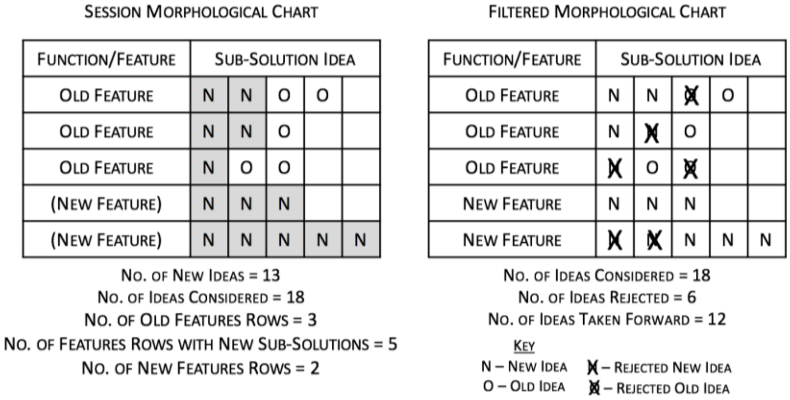

Figure 4. Session and filtered morphological chart examples

4.4 Novelty and Task Rating

Two of the metrics, novelty and task progress, require ratings to be carried out. As a panel of experts was to rate the novelty of each of the new ideas, it was essential to present the ideas in a digestible form. This resulted in the development of the novelty rating chart using the ideas summaries in the ideas chart. For the exercise, novelty was defined as “how unusual or unexpected the idea is” and each idea’s novelty was to be rated out of 10; where 1 = ordinary, 5 = neither ordinary nor unusual/unexpected and 10 = unusual/unexpected. In this case, the experts who carried out the rankings were creative practitioners from Stimulo who were not involved in the creative session that was being assessed. Due to the complexity of the ideas and language barrier, instead of solely relying on the idea summaries, a creative practitioner who was part of the session explained the ideas and provided supporting contextual information. To control the novelty rating as much as possible, it was essential to ensure that the explainer remained as neutral as possible and that there was no conferring between the experts as they rated the ideas. The amount of explanation required depends on the nature of the ideas. Following the identification of tasks from the session minutes and transcript, a task ranking chart containing the following headings: Task Summary, Status (Open, Created, Closed) and Rating (1, 2, 3) was created. The chart was populated with summaries of the tasks identified from the session minutes and transcript. Initially, all the sections of the chart were to be completed by the researcher and checked by the practitioner. For better accuracy, this was changed and the ranking was performed by the practitioner instead. The ranking is weighted based on the strategic importance of the task to the project success (1 = low, 2 = medium, 3 = high).

5 METRICS CALCULATIONS AND RESULTS

Continuing with the process testing corresponding to Phase 3 (shown in Figure 3) – the calculation phase – this section reports on how the metrics' values were calculated for each of the cases from the information generated to give a measure of the performance of the observed sessions.

As a measure of new ideas generated during the session, it was initially proposed that the quantity metric be tracked by comparing the number of sub-solution ideas in the ‘before’ and ‘after’ morphological charts. With the use of the session morphological chart, this would correspond to the number of new ideas in that chart. While this method of counting ideas in the morphological chart is adequate, a simpler method was developed as a consequence of having the ideas chart; the new method involves tallying the number of new ideas in the ideas chart. The quality metric is simply applied by identifying the number of new ideas generated that are taken forward at the end of the session. Initially, it was proposed that the metric be applied by calculating the difference between the number of new ideas rejected and the number of new ideas generated; however, the creation of the ideas chart means that the number of new ideas taken forward has already been populated and this information is available in the chart. The initially proposed calculation for the variety metric involved creating 'before' and 'after' morphological charts and gleaning the required information from there; however, this process was simplified through the use of the session morphological chart (described in Section 4.3). In this case, the value for Variety is calculated using the following equation: *(No. of Old Features Rows ÷ No. of Features Rows with New Sub-Solutions) + (No. Of New Features Rows ÷ No. of Old Features Rows)*. The novelty metric is calculated as the mean average score from a panel of experts for each of the ideas generated during the session. This score was simply calculated from the novelty ranking charts completed by the experts.

The filtering effectiveness metric is a measure of the success of the idea filtering activities within the co-design session. To calculate its value, the no. of ideas rejected and no. of ideas considered are attained from the filtered morphological chart. The desired number of ideas is attained by asking the practitioner; where there is no desired number of ideas specified - such as was the case with both sessions in this study - it can be assumed that for each function one idea is desired. The information required to apply the task progress metric, a weighted score for each task created or resolved, was taken from the task ranking chart that was completed by the creative practitioner. The calculation was straightforward. For interest, the results from the metrics application are presented below in Table 1; it shows the performance of the each of the observed sessions.

Table 1. Performance of the sessions based on metric suite application

	Quantity	Quality	Variety	Novelty	Filtering Effectiveness	Task Progress
Case 1	30	29	2.125	6.07	63	0.29
Case 2	13	7	1	5.73	24	0.65

6 EVALUATION AND DISCUSSION OF APPLICATION PROCESS

Corresponding to the last research phase – the process refinement phase – this section details how the process initially proposed in the first phase – the process design phase – was enhanced following the second phase – the process testing phase. Upon completion of the two case studies in industry, reflection was undertaken to understand of the application process and lead to new learning that can aid in the refinement of the proposed metric application process.

During the evidence collection phase, time and effort was put into recording the observation sessions. The initial intention was to transcribe, subtitle and then directly analyse the videos as it was posited that the visual aspect afforded by the video recording would facilitate a better understanding of the sessions, compared to working with transcriptions alone. However, during the processing and analysis phase it became evident that this was not productive. Not only was the process of adequately subtitling the videos time-intensive - particularly for the non-English session - the visual aspect did not offer any discernible advantages compared to using just the transcriptions as, in these cases, it was hard to clearly see what the participants were doing despite various camera angles. In general, if the video recordings can be from a higher vantage point, then they could be a valuable supplement to the session transcripts during the analysis phase. As the observation sessions that made up the application cases were both in the middle of on-going projects, they proved to be exceptionally difficult to understand from the perspective of a researcher not intimately familiar with the products and projects. This made accurately identifying and categorising ideas nearly impossible. The pre- and post-session interviews with creative practitioners were undertaken as a means of, not only, providing information that would be used directly used in the application of certain metrics, but also to inform the researcher's understanding of the product and project. Upon analysing the contents of the interviews, it became apparent that they did not provide sufficient details corresponding to the contents of the observation sessions. During the interviews, the participants gave very general information regarding the products and projects, however, without an intimate knowledge of the projects and products, it was not possible for the researcher to probe for more detailed responses. Collaborating with the practitioners on identifying and categorising the ideas proved to be the better approach. On occasion the practitioners also found it difficult to understand what was going on during the session and to identify and categorise ideas. This was due to the collaborative analysis occurring a couple of months after the initially session observations. Consequently, it is paramount that this activity occurs as soon as possible after observation session. While working collaboratively with the practitioners during evidence analysis has a myriad of benefits, it requires a lot of contact time with the practitioners and introduces a risk of lowering the validity of the data due to bias from the practitioner. In addition to working collaboratively with the practitioners, examining prototypes, images and other relevant documents proved to be valuable in establishing an understanding of the products and projects. Testing the metrics application process on live industrial projects proved to be particularly challenging. Ascertaining the source of difficulty proved to be problematic. It was not clear if the difficulties encountered were due to the metrics suite, the specific observed sessions, lack of knowledge of the project and company context, or the application process. An alternative and simpler approach would have been to initially test the proposed process within a controlled experiment and refine it before applying it to live industrial projects.

6.1 Main finding: the refined Metrics Application Process

The reflections presented above, informed the refinements made to the initially proposed application process resulting in the process presented in Figure 5. It is this process and metrics suite that is put forward as a means of measuring the performance of co-design sessions.

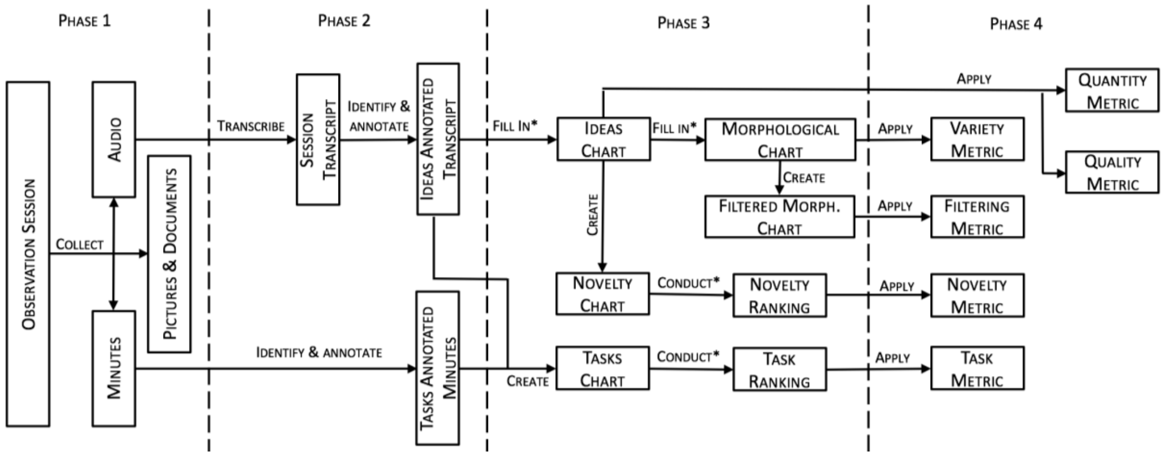

Figure 5. Refined metrics application process

The refined metrics application process is an evolution of the initially proposed application process presented in Figure 2. Some of the main changes and their influences are as follows:

- To accommodate the analysis work that has to be carried out collaboratively with the creative practitioners, an extra phase was added to the process. It is composed of the evidence collection phase, the processing phase, the collaborative analysis phase and the calculation phase.
- To aid in the categorisation of ideas, the ideas chart was developed. This chart forms the basis of the application of five of the metrics in the suite and making the overall application process simpler.
- The addition of the ideas chart allowed for the development of the session and filtered morphological charts; these are simpler to create than the 'before' and 'after' charts.
- Capturing video evidence was removed from the process as analysing it proved to be challenging and audio evidence was found to be adequate.
- The pre- and post-session interviews were removed as working collaboratively with the practitioners negates their need.

7 CONCLUSIONS AND FURTHER WORK

Conducted as part of the SPARK project –which aims to develop, build and test a SAR-based support tool for co-design – the aim of the study presented in this paper was to design, develop and test a process for applying a suite of metrics to a co-design session in a real-life industry-based environment. This paper has reported on an initial trial where the suite of metrics was applied to 'normal' co-design session that involved design consultants working collaboratively with non-designers without the support of a SAR-based tool. The process of assessing the performance of a design session through the use of a suite of metrics in a real-life industrial context can be a particularly challenging one. It may be more beneficial to start with the adoption of participatory action research, with the researcher more embedded in the case company and involved in the product development process before the sessions of interest are observed as also suggested by Ottosson and Björk (2004). This would minimise the amount of collaboration required between the researchers and the creative participants during the analysis of the evidence collected during the observed sessions. Within the context of this study, the complexity of the products within the application cases posed particular challenges to the application process. It may be the case that when the product is simple then the required collaboration can be scaled back, making it more in line with the initially proposed process. However, these cases have suggested that all the analysis performed by the researcher need to be checked by the practitioners to ensure that it is an accurate representation of the observed co-creative session.

In the immediate short term, further work for the SPARK project team involves applying the metrics suite using the process proposed in this paper to a different set of sessions from industry for which data has already been collected where the products are simpler and co-creative sessions appear to be more focused. Acting as further testing of the applicability of the proposed process, the outcome will result in the confirmation or further refinement of the proposed process. Ultimately, any specified metrics application process aims to enable repeatable and robust of the metrics suite by a variety of researchers. Once the process has been used in several cases it becomes possible to document it in a form that presents sufficient detail to allow others to use it. In the case of the SPARK project, this will take the form of researchers in the team who were not involved in the development of the application process applying the documented process to co-design sessions.

REFERENCES

Behn, R.D., (2003), Why measure performance? Different purposes require different measures. Public Administration Review, 63(5), pp.586-606.

Bhuiyan, N., (2011), A framework for successful new product development. Journal of Industrial Engineering and Management, 4(4), pp.746–770.

Cedergren, S., Wall, A. & Norstrom, C., (2010), Evaluation of performance in a product development context. Business Horizons, 53(4), pp.359–369.

Chidamber, S.R. & Kemerer, C.F., (1994), A metrics suite for object oriented design. IEEE Transactions on Software Engineering, 20(6), pp.476–493.

Füller, J., Mühlbacher, H., Matzler, K., & Jawecki, G. (2010), Consumer empowerment through internet-based co-creation. Journal of Management Information Systems, 26(3), 71–102.

Howard, T.J., (2008), Information management for creative stimuli in engineering design (Doctoral dissertation, University of Bath Library).

Lee, Y., (2008), Design participation tactics: the challenges and new roles for designers in the co-design process. Co-Design, 4(1), pp.31-50.

O'Donnell, F.J. & Duffy, A.H., (2002), Modelling design development performance. International Journal Operations Production Management, 22, p.1198.

O'Hare, J., Mombeshora, M., Dekoninck, E. A. and Cascini, G. (2016), Developing metrics to assess technology-enabled creative co-design sessions, ICDC2016: The 4th International Conference on Design Creativity. Atlanta, GA, USA.

Ottosson, S. & Bjork, E., (2004), Research on dynamic systems: Some considerations. Technovation, 24, 863-9

McNiff, J. & Whitehead, J.(2006), All you need to know about Action Research, London, Sage Publications Ltd.

Sanders, E. B.-N., & Stappers, P. J. (2008), Co-creation and the new landscapes of design. CoDesign, 4(1), 5–18.

Sanders, E. B.-N., & Stappers, P. J. (2014), Probes, toolkits and prototypes: three approaches to making in codesigning. CoDesign, 10(1), 5–14.

Taffe, S. (2015), The hybrid designer/end-user: Revealing paradoxes in co-design. Design Studies, 40, 39–59.

Upadhaya, B., Munir, R., & Blount, Y. (2014), Association between performance measurement systems and organisational effectiveness. International Journal of Operations & Production Management, 34(7), 2-2.

Vakkari, P. (2003),Task-based information searching. Annual Review of Information Science and Technology, 37, 413-464.

ACKNOWLEDGEMENTS

The work reported in this paper is part of the SPARK project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No.688417. This paper reflects only the authors' views and the European Commission is not responsible for any use that may be made of the information it contains. The authors would like to thank all the partners of the SPARK Consortium for the extensive support and valuable contribution provided in the development of this activity.