




# INVESTIGATING THE IMPACT OF SCALE IN DESIGN SESSIONS SUPPORTED BY A SPATIAL AUGMENTED REALITY (SAR) TOOL

L. Giunta , E. Dekoninck and J. Gopsill

University of Bath, United Kingdom

 [lg413@bath.ac.uk](mailto:lg413@bath.ac.uk)

## Abstract

Spatial Augmented Reality (SAR) differs from other forms of AR by allowing the projection of digital images onto a model. This allows the AR to be more tangible and for interaction to be more realistic. The scale of the model plays a role in the realism but may be constrained by technical factors. This study attempts to understand the influence scale has on a design session by analysing the concept generation process, the ease of designing and the design behaviour. Understanding how these factors are influenced by the model scale better the understanding of how SAR can influence design.

*Keywords: spatial augmented reality, scale, collaborative design, co-design, augmented reality (AR)*

## 1. Introduction

Spatial Augmented Reality (SAR) differentiates itself as an augmenting technology by applying environmental projection onto physical scenes that does not require attachment to the user. This has shown to be advantageous when compared to Virtual Reality, which requires headsets or screens, as it has been proven to be more comfortable for long design sessions, provides a greater field of view and supports collaboration to a greater extent (Caruso et al., 2015). SAR is becoming increasingly prevalent in designers' toolboxes as it allows them to edit and iterate a digital skin which has been overlaid onto a physical object. The ease with which this digital layer can be edited has attracted interest from the fields of packaging and product design, as multiple iterations of the surface design can be explored using the same physical prototype. This is especially important near the end of the design process where modifications are less centred around physical shape (Ong et al., 2011). The technology is of particular interest for the packaging and advertising industry as well as any design whose colour, material, and finish is being evaluated (Caruso et al., 2016).

SAR inherently supports transdisciplinary design through the projection onto a single physical object forming an intermediary object for collaborative design. As a result, SAR-led design sessions can involve a wide variety of stakeholders (end-users, clients, etc) with different disciplinary expertise leading to a transdisciplinary design activity. This affords designers the opportunity for improved feedback during design sessions as well as enhancing co-creative opportunities. The act of projecting onto a physical object, rather than using a screen, reduces the burden on participants to mentally translate the object into a 3D representation and ensures a common understanding of the design is maintained between participants.

The nature of SAR utilizing a physical object for projection means that, although SAR allows participants to better place an object into context, some objects will need to be scaled due to the projection envelope and resolution of the digital projection technology used in the SAR system. For objects that are too big to

fit within the projection envelope of the SAR system or too small to show the digital surfaces well enough, the physical model will need to be scaled. This may limit the applicability of SAR systems should the need to scale influence the ability of designers to perform their tasks and raises questions as to the effects that scale may have on the ability and behaviour of participants to design.

To investigate this, this paper presents the results of a pilot study that evaluates the impact of scale on the generation of concepts within a collaborative design review session. This research contributes to the development of SAR systems for design by showing how designers and their collaborators respond to a scenario where the physical model was scaled up to overcome the limitations of the resolution of the digital projection. Some of the behaviours explored in this research include: how the design assets (graphics) are used at smaller scales vs the larger scale; and, whether there is a difference in the number of concepts created at the two different scales. Highlighting these and any other differences in behaviour between conditions will provide the field with a better understanding of how scale influences designers. The paper continues by defining Spatial Augmented Reality, the parameters that limit the size of model and studies that have demonstrated SARs potential to support transdisciplinary design activities. This is followed by a description of the study that has been conducted to investigate the effect of scale on a collaborative design review activity. The results of the study are presented and followed by a discussion to the underlying theory that could be developed; limitations in the results; and, future work. The paper then concludes by presenting the key findings from the study.

## 2. Spatial Augmented Reality

SAR is one of three categories in [Bimber and Raskar's \(2005\)](#) categorisation of Augmented Reality (Figure 1). It is the category that requires no input from the on-looker; be it in terms of holding a device that creates an overlay or wearing a piece of equipment that provides an overlay. Figure 1 shows the two main methods in which SAR is achieved and that is through either projection on the physical object or the provision of a transparent display in between the on-looker and object where information is overlaid.

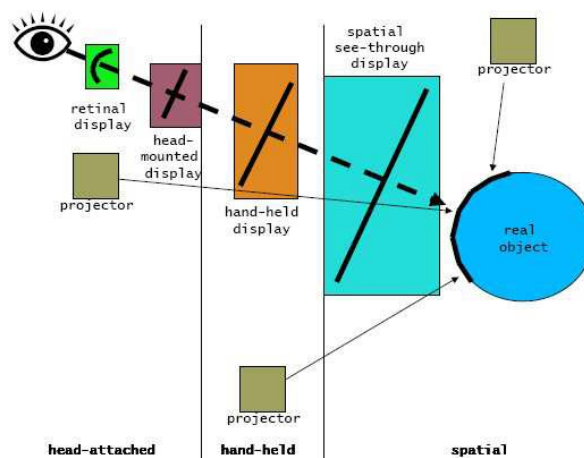
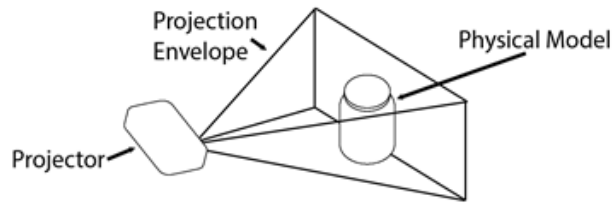


Figure 1. Categorisation of AR technologies ([Bimber and Raskar, 2005](#))

This paper focuses on projection-based SAR that enables a digital overlay to be mapped onto a physical object's geometry. Projection-based SAR can be achieved through single/multi projector systems, as shown in Figures 2 and 3. This technology has been used in multiple studies with a wide range of applications in design, such as the prototyping of both interfaces and products ([Porter et al., 2010](#); [Park et al., 2015](#); [Dey et al., 2018](#); [Morosi et al., 2018](#)).

SAR's use of projector(s) enforces a projection envelope that limits the maximum size of the augmented model (Figure 2). Parameters that limit this envelope size include the projector's technical specifications (such as, but not limited to: throw, aspect ratio, focal length, resolution, and luminosity) and distance between the model and projector (limited by the space available for the SAR setup). Varying this distance will increase the projection envelope's size yet decrease the luminosity and resolution of the projection on the model. Therefore, it is non-trivial to determine the maximum and minimum projection envelope to produce an effective design environment.



**Figure 2. Projection SAR set-up**

Figure 3 shows an augmented prototype side by side with a regular object with the projector creating a ‘spotlight’ effect on the augmented prototype. The black cloth denotes the projection envelope of this system. While an object this size easily fits within the projection envelope at a 1:1 scale, a larger object, such as a suitcase or a car, would struggle. Additionally, a smaller object, such as a watch, would suffer from decreased resolution.



**Figure 3. A real life object (left) next to a physical model with a digital overlay (right) the tablet shows the interface used to control the SAR projection (Becattini et al., 2018)**

Remaining within the projection envelope is thus imperative to SAR’s function. When this area is exited the model is no longer augmented. Additionally, when a model is too small, the resolution of the projector may be insufficient to correctly display images and/or be too finicky for a designer to arrange and move graphics. As a result, scaling the model will be necessary if one wishes to use SAR, thus leading to question the potential impact that scale may have on a transdisciplinary team’s design behaviour in generating solutions for a given design problem.

### 3. Study and research methodology

To investigate the role that scale plays in collaborative design sessions, a controlled experiment was created to evaluate the generation of concepts for the same product at two different scales. This particular design scenario was chosen as this is representative of industry practices (O’Hare et al., 2018). The scenario consisted of two designers working to design a number of concepts for a watch. In the scenario, the designers were provided with a:

- SAR platform to create, develop and evaluate designs on a physical model of a watch;
- inspiration board to present the market and provide greater context to the activity; and,
- set of digital graphics that could be applied to the watch face.

The participants were all students from the Faculty of Engineering and Design. All participants had experience in CAD and were either in their final year of MEng or were pursuing a doctorate degree. Ages ranged from 20 to 38 and the gender ratio was 16 male to 8 female.

The SAR platform, shown in Figure 4, consisted of a projector mounted to an arm which was subsequently attached to a base that held the physical model. The platform could track the model as it rotated; however, for the purpose of the experiment the participants were not informed of this feature and thus did not attempt to do so. The distance from the projector to the model was kept constant across the two conditions. Only the model was replaced when switching between the conditions. The digital overlay was set to project and fully cover the model in both conditions with no “spillage” thus projecting only onto the model.

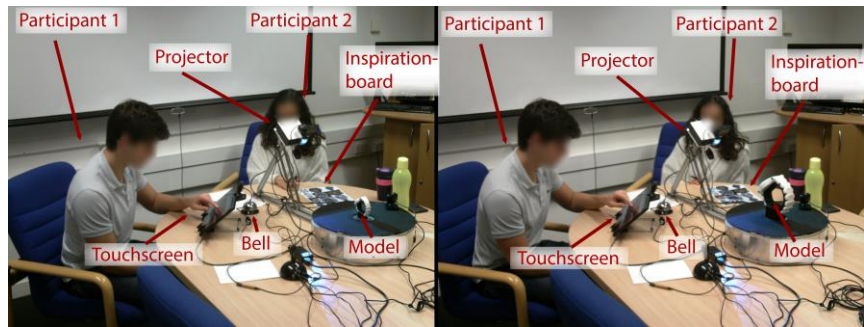


Figure 4. Experimental setup. 1:1 scale on the left, 2:1 scale on right<sup>1</sup>

The inspiration board consisted of 17 watch images showing the watch face and strap. Only one image showed the watch being worn by a male hand. All watches were analog with varying numbers of indicators and hands, however no watch had less than one additional feature (calendar, date, chronometer, etc.). The prevalent colours for the watch face and casing were: blue, black and gold. The watch straps showed both leather and metal straps of various colours, predominately brown and black. All images were sourced through Google Image Search.

The graphics provided to the participants are shown in Figure 5. Twenty graphics were provided in total, some (WeekCount, SecHand, MinLines, Logo, DayTime, DayCount, and Circle) are unique assets, where the others vary in line weight and text size. The SAR platform's tablet interface, shown in Figure 3, allows users to place, scale, rotate, and translate the assets on a digital representation of the watch. These changes are then displayed in real time on the physical model. It is worth emphasising that while all assets can be scaled to the desired dimensions by the participants, only some assets are available with thicker or thinner line weights and font sizes.

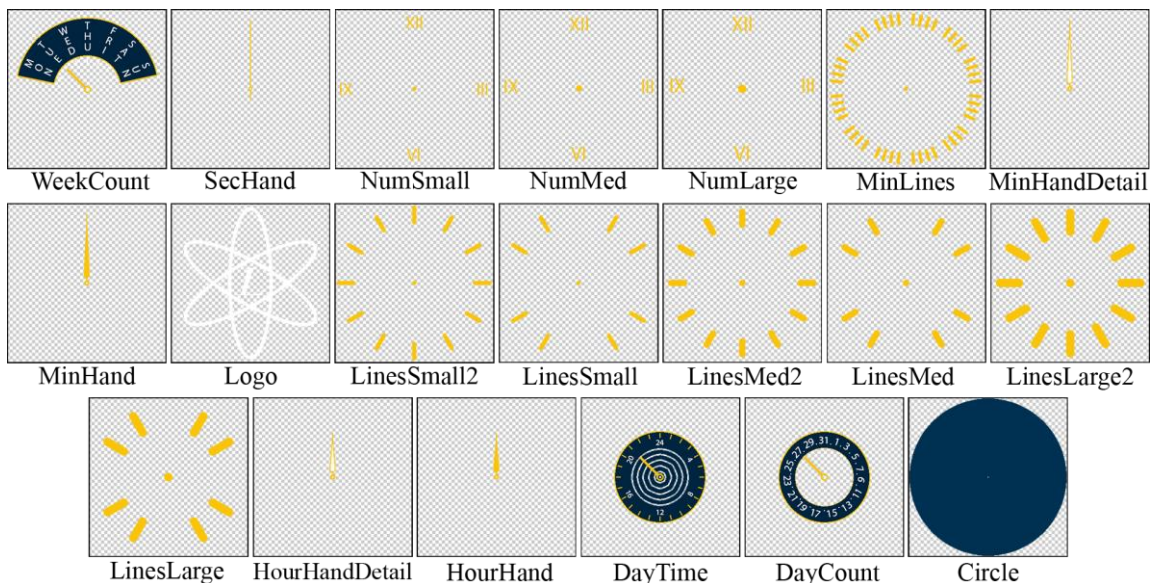


Figure 5. Graphics provided to participants for watch design

The experimental conditions were constant with the exception of the scale of the SAR model used for projection; Condition One used a 1:1 scale model (external  $\varnothing$  45mm), while Condition Two was 2:1 (external  $\varnothing$  90mm) as shown in Figure 4. The hypothesis was that the scaled-up model of the watch would allow more comprehensive use of the most appropriate design assets but might, at the same time, affect the design process due to the unnatural scale of the object.

Each session was structured as follows: the participants (two in every session) were briefed on the task they needed to complete by means of a pre-prepared statement. They were told to imagine themselves

<sup>1</sup> Participants were only allowed to participate in one of the two conditions. These images were taken after completion of all experiments

as designers tasked to design a number of watches for a client in line with the inspiration-board. It was made clear that the participants could create as many designs as they felt necessary but each one would need to be reviewed with their “client”. Each time they felt they had developed a concept they would later like to present to their client they would have to hit a bell to record the design. It was also stressed that the “client” would be reviewing the proposed watches by means of the SAR system. Thus, in the event of any discrepancies between the view on the interface and the SAR system, the SAR system would take primacy. With the scene set, one of the participants was given the interface that controlled the SAR system and allowed to familiarise themselves with it. During the familiarization process the other participant was provided with a text description of all the assets that had been preloaded onto the system allowing them to know what assets they could request without having to look at the other designer’s interface. This was done to ensure the physical model of the watch was used as much as possible as the ‘shared design representation’. The session started once the participants felt confident in using the interface and finished whenever the participants felt they had developed as many concepts as they felt necessary or when the one-hour mark was reached. The duration of each session was timed from start to whenever the participants claimed to have been satisfied with the number of concepts generated. The participants were randomly assigned to either Condition 1:1 or 2:1.

Once the session was complete the participants were asked to review the concepts they had saved, describing whether they still wished to take them to their “client” or not. This resulted in the “concepts taken forward” score. The method used was the novelty metric described in [Dekoninck et al. \(2018\)](#). Each participant was also asked to score the SAR platform using the Creativity Support Index (CSI) ([Cherry and Latulipe, 2014](#)) as recommended by [O’Hare et al. \(2018\)](#) in order to establish the designers’ scores for the usability of the SAR system itself. The CSI survey was developed to analyse ICT tools and is based on the well-established NASA Task Load Index (TLX) developed by [Hart and Staveland \(1988\)](#). Unlike the NASA TLX, the CSI survey better supports the evaluation of creative activities with a score between 0-100.

In addition to the data collected at the end of the session, a log of all the interactions with the SAR interface was captured allowing for an analysis of the graphics used during the session. These assets can be tracked and catalogued to reveal which condition favoured which assets as well as highlight which percentage of the available assets were used and placed onto the model. Lastly the log file provides a chronological overview of all the additions and removals of assets that resulted in a concept.

## 4. Results

The analysis of scale and its effect on the concept generation behaviour of the participants was investigated from three perspectives:

1. Concept Generation process: Total time taken to complete the exercise, Number of concepts generated
2. Ease of Designing: CSI survey results
3. Design Behaviour: Percentage of assets used, Individual asset usage across conditions, Asset use and Interaction

These provide a comprehensive understanding of the events in the design sessions, combining both qualitative and quantitative data to examine how the model’s scale has influenced the design sessions. Twelve sessions were completed, six for the 1:1 condition and six for the 2:1 condition. This culminated in over nine session hours with video, audio and log recording.

### 4.1. Concept generation

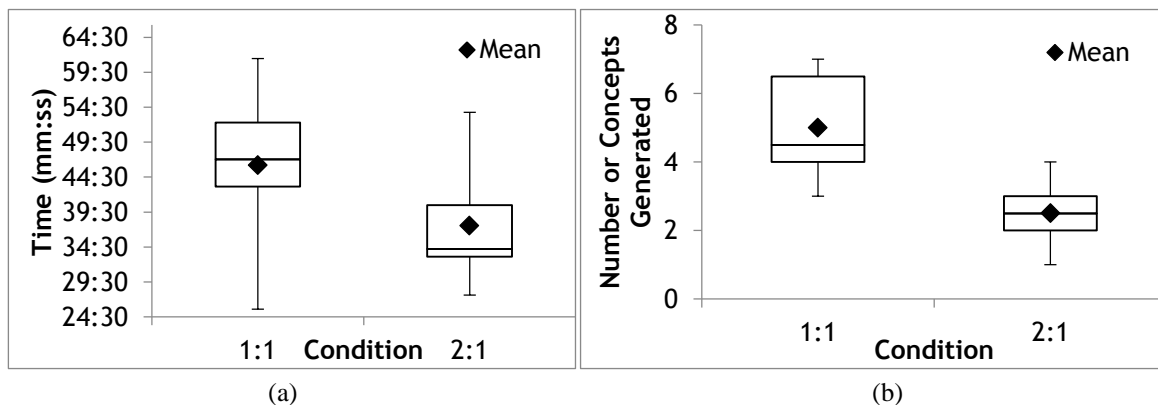
Tables 1 and 2 show the time taken for each session to complete. This is further summarized in a boxplot in Figure 6a where it can be seen that the average differed by more than 10 minutes and the distributions featured little overlap. This provides evidence to suggest that increasing the scale reduces the time spent generating concepts. In addition, participants in condition 2:1 seemed to be more consistent in the amount of time they required to complete the activity.

**Table 1. Average time taken per concept and total session time 1:1 scale**

Session Nr.	Total time taken (mm:ss)	Number of concepts generated	AVG Time per concept (mm:ss)	Concepts Taken Forward	Percentage Taken Forward
1	48:35	4	12:09	4	100
2	42:23	5	8:29	3	60
3	53:31	7	7:39	6	85.71
4	25:36	3	8:32	3	100
5	61:26	4	15:21	3	75
6	45:29	7	6:30	6	85.71
Average	46:10	5	9:47	4.166667	84.40

**Table 2. Average time taken per concept and total session time 2:1 scale**

Session Nr.	Total time taken (mm:ss)	Number of concepts generated	AVG Time per concept (mm:ss)	Concepts Taken Forward	Percentage Taken Forward
1	34:40	3	11:33	3	100
2	53:45	4	13:26	4	100
3	32:56	2	16:28	2	100
4	42:26	2	21:13	2	100
5	27:38	3	9:13	3	100
6	33:43	1	33:43	1	100
Average	37:31	2.5	17:36	2.5	100



**Figure 6. Comparison of time taken to complete task and number of ideas generated across experimental conditions**

Furthermore, the average time taken per session was higher for Condition 1:1 yet, the time taken per concept was lower. In addition, it is important to note that while the Condition 1:1 did generate more concepts (Figure 6b) in less time, the participants were less satisfied with their concepts by the end of the session with 84.4% of concepts generated being taken forward. This is in contrast to Condition 2:1 where, despite the lower number of concepts generated, 100% of concepts were taken forward.

#### 4.2. Ease of designing

The results of the CSI survey (Figure 7) demonstrate that the participants did not feel any less comfortable using either model. While this may initially seem somewhat counterintuitive, as one would expect the larger scale model to be easier to view and use, it may reflect that the 1:1 scale model is better at conveying the object due to its realistic proportions and thus renders the design process less abstract and mentally intensive. One other explanation is that the participants reflect more

on the SAR system as a whole when reviewing their experience and thus, the user interface plays a larger role in their experience when compared to the models. Additionally, the increase in scale appears to reduce the variance across participants.

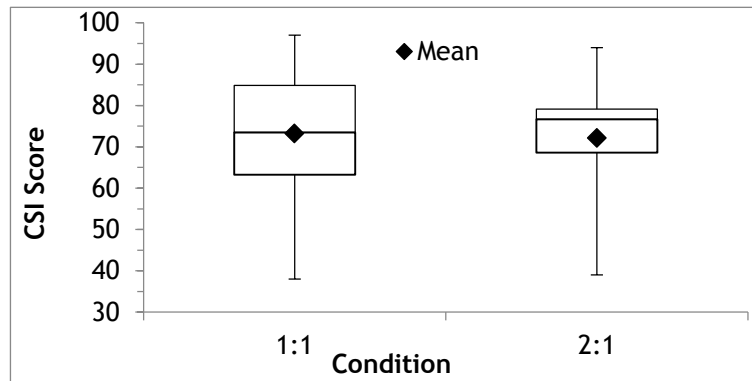


Figure 7. CSI results

### 4.3. Design behaviour

Figure 8 reveals the use of assets across the two conditions and within sessions. Figure 8a shows the usage of assets within each scenario and reveals that Condition 1:1 tended to cover more of the asset library than Condition 2:1 although there is some variance and overlap present. This may imply that Condition 1:1 featured more experimentation and iteration. This would aid with explaining the higher concept generation and lower concepts taken forward observed in Condition 1:1. The visual difficulty of using the 1:1 model may invite more experimentation and exploration but at the expense of more detailed design development and thus satisfaction with the finished results.

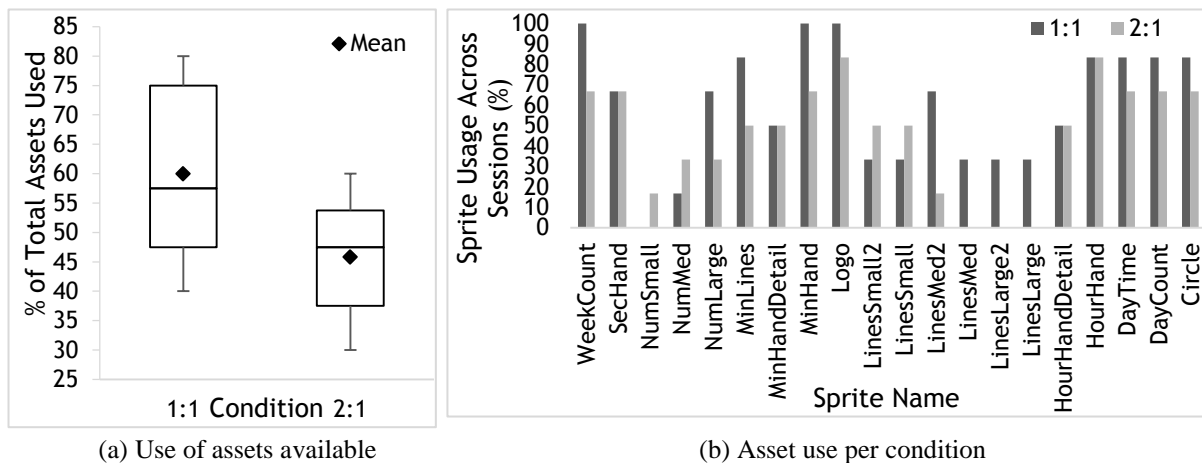
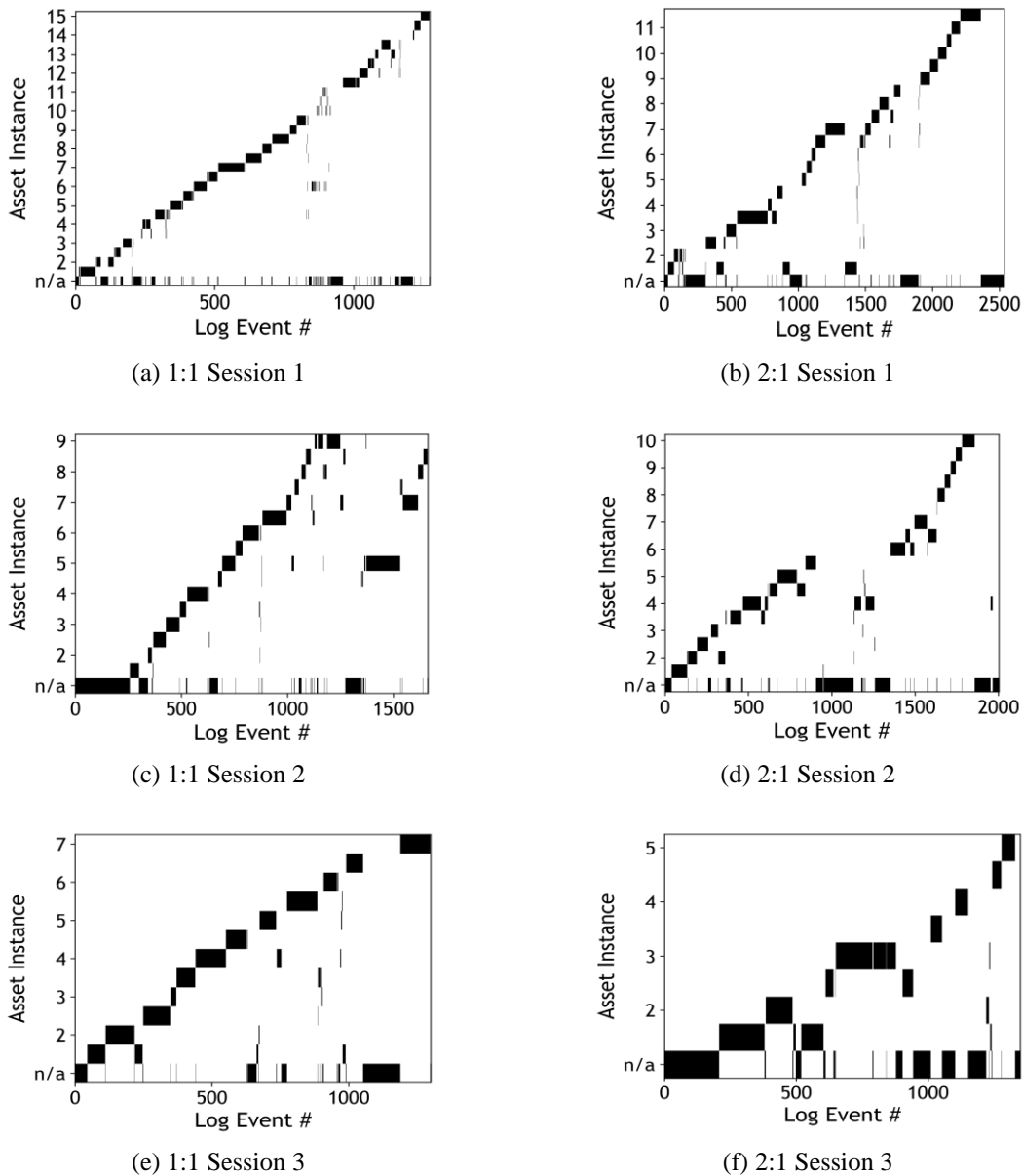


Figure 8. Asset use across conditions

Figure 8b shows the assets evaluated by the designers in each condition; with the x-axis indicating the percentage of sessions that used that asset for a particular condition. It can be seen that all assets were used by at least some of the participants in Condition 1:1; conversely there were some assets that were never trialled by any of the participants in Condition 2:1, such as LinesMed, LinesLarge2, LinesLarge. Figure 9 shows a sample of the interaction behaviour of three teams from each condition. The sequence features the events captured, mapped to the y-axis of the graphs and increasing by one with each event. These are plotted against the instances of assets from the asset library on the x-axis (n.b. designers could have multiple instances of the same asset within the scene. For example, two hour hands). Asset “n/a” relates to events that do not map to a specific asset and are events that affect the entire scene (e.g. reset scene, UI button presses).



**Figure 9. Asset use and interaction during sessions for a random sample of both conditions**

The striking result from visualising the designers' behaviour on the assets is the waterfall effect where designers will focus on a particular asset at a time but, once finished, they rarely revisit or iterate the position of the asset again. What one is seeing here is the creation of concepts on an asset-by-asset basis with the assets' in the scene forming additional constraints for next asset when it is placed. The designers will place a new asset and adjust until they are happy it works with the current set of assets or remove it from the scene. It is interesting to see that the designers rarely challenge the previous assets' placement. This phenomenon is exhibited independent of the scale used by SAR. The variance between the axes of the different (random) samples does also show how much variance there was between teams in terms of the number of assets manipulated and for how long.

## 5. Discussion

The concept generation analysis (Figure 6) shows that the time taken to complete the task decreased when the model scale was increased. What is noteworthy however is the number of concepts generated and the number of concepts taken forward. Despite taking comparable amounts of time to complete the task, participants in Condition 2:1 generated half as many concepts as participants in Condition 1:1.



Additionally, the Condition 2:1 participants chose to take all their generated concepts forward. In contrast, the participants in Condition 1:1 took only 84.40% of concepts forward. This, coupled with the time taken and number of concepts generated, implies that participants found the 2:1 scale model supported a more methodical approach. Further supporting this argument are the findings displayed in Figure 8, which show Condition 1:1 participants explored a wider variety of assets. This is what may have caused participants in Condition 2:1 to be slower in their concept generation; by conducting a more detailed design development they were subsequently more satisfied with the results of their work when compared to participants using the 1:1 scale model. Future analysis should thus attempt to capture more information into participant behaviour by means of either post-session questionnaires or more detailed analysis of the session video and audio recordings.

The results of the CSI surveys provide some interesting insights into how the participants enjoyed and felt supported in their task by the SAR platform. The results of the surveys show no significant difference between how the participants experienced the SAR platform across the two conditions. This indicates that the 2:1 model performs to the same standard, from the users' perspective of the SAR system as a whole, as the 1:1 model and thus offers no detraction. One alternative theory is that while the users are influenced by the scale of the projection model this is balanced out in some other way. For example, the 2:1 model may be easier to view but may be causing the participants to put more effort into bridging the difference between the scaled model and the real-life product they are creating. Future work should focus on the analysis of more scales and the use of a questionnaire which clearly separates feedback on the shared design representation (the model) and the user interface on the tablet.

The log analysis revealed that participants in Condition 1:1 explored the digital asset library to a greater extent. While the results show a link between model size and asset usage, increasing the number of sessions may be beneficial to highlight these differences further. It is also the case that an increase in the size of the digital asset library would diminish the designers' ability to explore it fully within the session time, therefore preventing it from influencing the emerging distribution. Accepting these limitations, Condition 1:1 showed a higher frequency of asset usage across sessions. In the majority of cases, Condition 1:1 participants used the same asset across more sessions compared Condition 2:1. Future work could build on this to investigate the types of assets present in the asset library and whether scale is more appropriate for certain types of design. It may be that modifying the experiment to have the assets exhibit greater variance, such as in resolution, contrast, line thickness, size, detail, etc. may aid in better understanding which types of assets are preferred at which scale and thus better gauge the influence of these characteristics on the participants. Additionally, this may yield a more detailed understanding of what drives the asset selection process and how the scale of the physical model may influence this.

Looking into the analysis of the interaction behaviour revealed a "waterfall" pattern across all sessions. This shows the design team would place and iterate an asset's position, which they then determined to keep or reject. A concept is then built on an asset-by-asset basis where the design team is evaluating the application of a new asset relative to the current layout of assets on the model. Once an initial concept is created, very few changes to pre-existing assets are made to generate a new concept. The addition of assets to a concept acts as a method to constrain the design to a point where no new assets will aid in the design team's objective of making a valid concept and at this point, the team determines a concept has been reached. Being able to see when a new concept is generated in the log itself and mapping this within the graph would improve the understanding of how the shared design space is populated. This would provide a better understanding of the steps preceding the generation of a concept and thus perhaps illustrate which factors influence concept generation.

While very few changes to existing assets were observed in the logs, this does not necessarily mean that the design teams did not discuss or challenge the position of existing assets. Further processing of the audio transcripts could provide an insight into this, such the apparent barriers that designers encounter when changing existing assets during a concept generation process.

## 6. Conclusion

The paper has presented a pilot study to identify whether scale has an influence on the design process when using a SAR system. The findings reported in this paper appear to show that the scale of the model does not play a major role in how the designers experience the SAR platform and their experience whilst

designing. However, the scale of the model appears to have played a role in the number of concepts generated, the designers' satisfaction with these concepts and their usage of the digital assets provided. Additionally, the paper identified a consistent "waterfall" asset-by-asset layering design behaviour when generating concepts with SAR.

In conclusion, this paper indicates that the physical model scale used in a SAR system plays a role in how design session participants accomplish their task; both in how they utilize the assets provided and in their ability to generate satisfactory concepts thus meriting further exploration. Additional studies exploring the influence of larger scales will need to be performed to be able to draw stronger conclusions as to the impact of scale. Furthermore, due to the pilot nature of the study, this research did not consider the impact that scaling down could have on an SAR session. This too will be explored before more definitive conclusions can be drawn. Lastly it may be of interest to explore how the participants feel about their concepts once these are returned to their proper scale as an additional metric to evaluate the impact of scale.

## Acknowledgements

The work reported in this paper has been undertaken as part of the EPSRC Trans-disciplinary Engineering (EP/R013179/1) grant.

## References

- Becattini, N. et al. (2018), "Coding schemes for the analysis of ict supported co-creative design sessions", *Proceedings of International Design Conference, Design*, Vol. 2, pp. 533-544. <https://doi.org/10.21278/idc.2018.0532>
- Bimber, O. and Raskar, R. (2005), *Spatial Augmented Reality Merging Real and Virtual Worlds*, A K Peters, Wellesley.
- Caruso, G. et al. (2016), *Technologies and Techniques – State of the Art Updates*, Milano.
- Caruso, G., Carulli, M. and Bordegoni, M. (2015), "Augmented Reality System for the Visualization and Interaction with 3D Digital Models in a Wide Environment", *Computer-Aided Design and Applications*, Vol. 12 No. 1, pp. 86-95. <http://doi.org/10.1080/16864360.2014.949579>
- Cherry, E. and Latulipe, C. (2014), "Quantifying the creativity support of digital tools through the creativity support index", *ACM Transactions on Computer-Human Interaction*, Vol. 21 No. 4. <http://doi.org/10.1145/2617588>
- Dekoninck, E.A. et al. (2018), "Exploring ways to speed up the application of metrics to assess co-creative design sessions", in *Proceedings of the Fifth International Conference on Design Creativity*, pp. 1-8.
- Dey, A. et al. (2018), "A Systematic Review of 10 Years of Augmented Reality Usability Studies : 2005 to 2014", *Frontiers in Robotics and AI*, Vol. 5 No. April. <http://doi.org/10.3389/frobt.2018.00037>
- Hart, S.G. and Staveland, L.E. (1988), "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research", *Advances in Psychology*, Vol. 52 No. C, pp. 139-183. [http://doi.org/10.1016/S0166-4115\(08\)62386-9](http://doi.org/10.1016/S0166-4115(08)62386-9)
- Khemlani, L. et al. (1997), "Semantically Rich Building Representation", *Design and Representation ACADIA '97 Conference Proceedings*, pp. 207-227.
- Morosi, F. et al. (2018), "Analysis of Co-Design Scenarios and Activities for the Development of a Spatial-Augmented Reality Design Platform", in *International Design Conference - Design 2018*, pp. 381-392. <http://doi.org/10.21278/idc.2018.0504>
- O'Hare, J. et al. (2018), "Defining requirements for an Augmented Reality system to overcome the challenges of creating and using design representations in co-design sessions", *CoDesign. Taylor & Francis*, Vol. 00 No. 00, pp. 1-24. <http://doi.org/10.1080/15710882.2018.1546319>
- O'Hare, J.A. et al. (2018), "Exploring the Performance of Augmented Reality Technologies in Co- Creative Sessions : Initial Results from Controlled Experiments", *International Design Conference - Design 2018*, pp. 405-416. <http://doi.org/10.21278/idc.2018.0391>
- Ong, S.K. et al. (2011), "Augmented Reality in Product Development and Manufacturing", in *Handbook of Augmented Reality*. Springer New York, New York, NY, pp. 651-669. [http://doi.org/10.1007/978-1-4614-0064-6\\_30](http://doi.org/10.1007/978-1-4614-0064-6_30)
- Park, M.K. et al. (2015), "Spatial augmented reality for product appearance design evaluation", *Journal of Computational Design and Engineering. Elsevier*, Vol. 2 No. 1, pp. 38-46. <http://doi.org/10.1016/j.jcde.2014.11.004>
- Porter, S.R. et al. (2010), "Validating spatial augmented reality for interactive rapid prototyping", in *2010 IEEE International Symposium on Mixed and Augmented Reality*, IEEE, pp. 265-266. <http://doi.org/10.1109/ISMAR.2010.5643599>