Show me your Living Room: Investigating the Role of Representing User Environments in AR Remote Consultations

Nicolas Kahrl[†]
Human-Centered Information Systems
TU Clausthal
Clausthal-Zellerfeld, Germany
nicolas.kahrl@tu-clausthal.de

Michael Prilla
Human-Centered Information Systems
TU Clausthal
Clausthal-Zellerfeld, Germany
michael.prilla@tu-clausthal.de

Oliver Blunk
Human-Centered Information Systems
TU Clausthal
Clausthal-Zellerfeld, Germany
oliver.blunk@tu-clausthal.de

ABSTRACT

The study reported here investigates AR based support for remote consultations, in which an on-site user is supported by a remote helper. In such situations, it is important for the remote helper (or, in our case, the consultant) to see the environment of the person asking for support in order to relate to it. Based on literature, we created and tested different mechanisms using a 2D video stream with a captured 2D/3D texturized virtual model of the room. In addition, we compared the often-used way of fixing the remote helper's view to the view of the on-site user with the possibility to move around freely in the 2D/3D model. The aim of the study was evaluating how to support an on-site user wearing an AR HMD. The study tested four conditions composed from these differences and with nine real furniture consultants. In the study, we compared four mechanisms in which the consultants were able to place furniture in the living room of a customer and advise the customer on their purchase. We found that there were hardly any differences in task load, social presence or perceived support between the four different conditions. However, participants had clear preferences for certain conditions and aspects of them. From our analysis, we provide recommendations for the design of mixed reality support for remote consultations.

CCS CONCEPTS

•Human-centered computing~Human computer interaction (HCI)~Interaction paradigms~Mixed / augmented reality
•Human-centered computing~Human computer interaction (HCI)~Interaction paradigms ~ Virtual reality
•Human-centered computing~Human computer interaction (HCI)~Empirical studies in HCI

KEYWORDS

Augmented Reality, Virtual Reality, Remote Support, Collaboration, Consultancy, Furniture

ACM Reference format:

Nicolas Kahrl, Michael Prilla and Oliver Blunk. 2020. Show me your Living Room: Investigating the Role of Representing User Environments in AR Remote Consultations. In *Mensch und Computer 2020 (MuC'20), September 6–9, 2020, Magdeburg, Germany.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3404983.3405520

1 Introduction

Remote support using Augmented Reality (AR) is a field of interest for researchers and practitioners. In such cases, an on-site user of an AR device looks through the device and receives virtual information and support such as pointers from a remote helper (e.g., [10, 14, 21]). It provides both economic and ecological potentials: While avoiding travel saves time and money, it also makes the provision of support more flexible and avoids pollution. For these and other reasons, ways of supporting users remotely have been explored in maintenance [10], repair [7], crime scene investigation [4], construction [14] and many other areas. The majority of such solutions use a video stream created by an AR device of the on-site user and share it with the remote helper. The helpers can then point to areas of the stream, which is transmitted to the on-site user, or verbally refer to what they see. While this setup has shown to work in many areas, it also has shortcomings: With video, spatial understanding of the situation and immersion into it are limited [17, 28], and by fixing the remote helper view to the camera stream of the on-site user, freedom of perspectives and awareness of actions of the on-site user are limited [8]. To cope with these challenges, several approaches have been suggested to allow the remote helper to move freely in the on-site environment and to represent this environment. These approaches have shown to improve the remote support situation.

The work presented here investigates a specific remote support situation. As part of a research project, we are creating means to enable customers to receive remote consultations while buying furniture. For this we are employing augmented reality. Customers are wearing a Microsoft HoloLens as a head mounted display (HMD) to view recommended couches as virtual objects in their living room. This enables customers to try out various configurations which might not be on display in the furniture

[†] Corresponding Author

MuC'20, September 6–9, 2020, Magdeburg, Germany

ACM ISBN 978-1-4503-7540-5/20/09...$15.00
https://doi.org/10.1145/3404983.3405520

store [2]. However, it is still an open question how to best represent the customer environment for the remote consultant. As we have found in a prior study [3], using the video stream offered by the HMD as described above has various shortcomings for the consultant. In this study, we want to explore how to improve this situation by providing alternative representations of the customer environment to the consultant who acts as the remote helper.

Furniture stores and online providers have already implemented AR and VR solutions to represent their products inside the home of their customers. However, to the knowledge of the authors there is no application that allows consultants to support the user to select furniture for their home based on AR representations in their homes. This affords a collaborative augmented reality application in which the consultation process can be conducted digitally. As the solutions for representing the environment and moving freely in it have not been investigated for consultancy settings, in the study presented here we compared different solutions in order to find tool setups consultants can benefit from. To do so, we compared a video consultation as mentioned by most research, a 2D desktop view with either fixed or free views around the room of the customer and a Virtual Reality (VR) environment in which the consultant can be placed in the room of the customer. Our results show that despite the novelty and difficulty in handling virtual and VR environments, they may provide support for consultancy.

2 Related work

Recent research on remote cooperation supported by AR often focuses on a remote helper who supports a worker through a live video streamed from the HMD the worker is wearing. Examples for these scenarios have been studied for construction tasks [7, 14], the impact of non-verbal communication in remote settings [19] or even for expert support in crime scene investigations [4, 22]. A particular challenge in these situations is the difference in tools used by the worker and the helper. While workers often used (AR) HMDs and have a (real world and augmented) 3D view on the setting, helpers often use (2D) computer screens or mobile devices to watch the video stream and provide support. Research has found that this creates challenges for pointing to certain areas of the stream (e.g., [1, 6]) as well as perspectives when looking at the environment of the worker (e.g., [20, 25, 27]). In a comparison of techniques to overcome these problems, Brown and Prilla [3] conducted a study on remote consultations. They compared manual depth pointers that had simplified shadows on walls as depth cues and assisted depth pointers, for which the pointer snaps to the environments. Regarding perspectives, they compared exocentric views as a top down view and egocentric views with the camera being fixed to the local user's view. Their results showed that the remote helpers preferred the assisted depth pointer and that for placing objects the exocentric view was preferred by remote helpers, while for specifying positions the on-site clients preferred the egocentric view. As a result, they recommended a tool with both mechanisms as the best choice.

As mentioned above, most systems used to support on-site users via AR, use the camera stream created by the AR device and make it available or editable for the remote helper (e.g., [1, 3, 4, 6, 7, 14]. While this is a reasonable approach which has shown to provide good support, it also has shortcomings. First, using the video stream of the AR device binds the view perspective of the remote helper to the view perspective of the on-site user. For HMDs, this means that the helper can only look at the area of a room the on-site user looks at, which may result in drawbacks such as low awareness and presence [8, 17]. Second, viewing a video stream on a 2D monitor device does not represent the spatiality of the setting as visible to an on-site viewer, and it does not immerse the remote helper much into the on-site situation [17, 18, 30]. This was quantified by Tachakra [28], who conducted a study with 235 patients in telemedicine and found that the depth perception with a camera was less than 90% of binocular vision and needs to be improved by camera rotations, light adjustments, comparisons with the other side and other strategies.

To cope with the latter challenge of representing the on-site environment to the user in a spatial and immersive way, several approaches have been tested, which include cave projections of the area [16, 18], mixed realities by allowing the remote helper to view the scene in VR [21, 31] and virtual representations of the on-site environment on 2Dmonitor devices [29]. Among this work, Tecchia et al. [30] created a remote collaboration tool that captures and renders the remote workspace, as well as the helpers hands. The worker's space is captured with a 3D camera and an additional screen shows the captured data and augmentations added to the scene. The helper uses a VR HMD to see the worker's environment, while head and hand movement is tracked and augmented as a means to help the worker. Tecchia et al. conducted a pilot study with four participants that had knowledge of their previous 2D prototype. The participants had the task to help the worker assemble a Lego toy. Early findings suggest that the 3D system is an improvement to their earlier 2D system. More recently, Gao et al. [9] built a tool to capture entire scenes from the worker's environment for a remote collaboration. They captured the scene with a depth sensor to create a 3D point cloud that was reconstructed on the helper's VR system. An initial pilot study was conducted in which the participants played the role of the helper. The worker needed to find Lego models in a specific order that were predefined for the helper. The helper had to guide the worker to find each model. The prototype was tested for an oriented view interface, where only the current view frustum was rendered into a 3D point cloud for the helper and a shared view interface where the scene was reconstructed at the beginning of the session and where the helper could see the view frustum of the worker. Participants were able to find objects faster in the shared view. Both modes had advantages and disadvantages and it was concluded that the system helps remote workers to understand the spatial relationship between items in a fast way.

Teo et al. [31] conducted a user study to test remote collaboration with a combination of 360-degree video and 3D reconstruction. Both users were in different rooms that were connected to each other with a door. This study was split into two parts, where the first part included the exclusive use of the 360-degree view or the 3D view and the second part had the ability to switch between both. For the local user they used an actor to minimize the difference of proactive and passive local users. The users had to find objects of interest and manipulate them collaboratively. They found that participants performed better and felt higher social presence on the 360-degree video. However, the results indicated that the 3D mode provided similar user experience and benefits from independent navigation. They recommend a system that implements the ability to switch between both modes to benefit from each system's unique advantages. In addition, Kolkmeier et al. [17] found that a VR setup compared to a 2D-desktop setup increases spatial presence and perceived message understanding.

Regarding the former challenge of (not) fixing the view perspective of the remote helper to the view perspective of the on-site user, several approaches such as scanning the room digitally and providing exocentric perspectives [3, 26], using 3D cameras to capture and stream the on-site environment and allowing the remote user to move in this stream [10, 18, 30, 31] and telepresence robots [24] have been suggested. Kolkmeier et al. [17] created a remote mixed reality collaboration system with different setups. They compared what they called dependent and independent viewports for the two users. For the on-site operator they used a HoloLens with several attached RGBD cameras to capture the on-site environment with the highest amount of detail. The remote operator used a VR-HMD in one setup and a desktop client in the other. The on-site operator was able to see the hands and the head of the remote operator, even in the desktop setup. For the experiment the on-site operator had to escape a room and therefore successfully complete three tasks with the help of the remote operator. Their experiment showed that view independency increases the quality of the collaboration. These findings are in line with the findings of Tait and Billinghurst [29], who found that an independent view increased communication from remote participant to local participant, higher focus on same objects or areas and increased feelings of joy for the remote user. While these approaches often work with 3D cameras, Dong and Höllerer [5] created a system for the HoloLens that uses the HoloLens' built-in camera to project the cameras texture on the HoloLens internal triangulated spatial map in real time. They used a multi pass shader to save the color data in a 4096x4096 pixel texture, with an adaptive mapping scheme to adapt to geometrical data changes. While they do not provide a user evaluation, they show that their approach creates a decent quality of the representation.

Figure 1: Furniture with multiple configuration options

3 Background: Remote Consultations for Furniture Sales

Traditionally, furniture sale is a brick & mortar business. Customers come to the store to receive a consultation and to figure out which furniture fits best in their room. To give the customer the best experience and a preview of how each furniture will look in their own environment in the desired place, digitalization is needed [2]. In this paper, the term *customer* is used synonymously with the term *on-site users*, as they are located in their home environment and equipped with a Microsoft HoloLens. They can then use the HMD to configure the virtual couch as depicted in Figure 1. In this study, we use our living room lab as a representation of the local environment of an on-site user (see Figure 2). Furniture sales is a complex area, with an astonishing number of customization options for each couch, for example: footrest, extensible arms or backrests, reading lights, USB ports, etc. as well as options like fabrics, cushioning, colors, etc. As such, a consultant is highly recommended when buying furniture. In this paper, the term *consultant* will be used as a synonym for *remote helper*, as they consult from their workspace on a remote online basis.

Figure 2: Our living room lab, which was built for multiple user studies in the context of remote furniture consultations.

In prior work [2], we conducted an ethnographic study to analyze the domain further and to identify characteristics of furniture consultancy. Findings include: There is a knowledge gap between consultants (who know more about furniture) and customers (who know more about their local environment) [15], and a lack of preparation from both groups prior to a consultation session. Additionally, consultants stated that customers often lack imagination when discussing different configuration options. Based on our ethnographic work we conclude that digitalization can aid this consultation process. In this study, we evaluate different environments as means for consultants to help customers in their local environment, which will be a foundation for further studies in this area. This study addresses multiple findings from the ethnographic study: A well-designed consultation environment for the consultant might aid in preparing for customers' needs, as the consultant can see the local environment e.g. in preparation for subsequent meetings. Customizable digital furniture pieces can be used to aid the customers' imagination. This study works towards the goal of creating a well-designed remote consultation environment.

4 Study Design

4.1 Research Questions and Open Issues

As our literature analysis shows, several ways of remote support for on-site users have been investigated and found to provide support in previous work. This includes using a video stream provided to the remote user as the most popular solution, but also includes providing the remote users with a 2D or 3D virtual environment simulating the environment of the user. In addition, to provide freedom in viewing the on-site environment, approaches for providing the remote user with independent view perspectives have been shown. While this work was done in different areas such as remote consultations, maintenance and others, it only covers parts of the problem we are looking at here: There is only little work available on remote consultations such as laid out above, and for this work only video streaming has been tested. Whether this option is a good choice for remote consultations and whether 3D environments allow for better remote consultations remain questions to be answered in order to create support for such remote consultations.

As the on-site user is in physical proximity and touch of their environment, the main challenge lies in supporting the remote helper (or, as in our case, the consultant working from remote) to achieve spatiality, presence and awareness of the situation. Therefore, we focus on support for the consultant in the study presented here. In particular, we focus on the way the environment of the on-site user is presented to the remote helper and its effects on the support provided. As little is known about this support for remote consultations, our research question driving this work is: *"How does the way the environment of the on-site user is represented affect the consultation provided by the remote consultant"*.

4.2 Setting: Living Room Laboratory

To create a realistic online consultation scenario, we built a living room lab in one of the facilities of our university to simulate a real living atmosphere, as seen in Figure 2. The lab is furnished with an armchair, a coffee table, bookshelf, low board and is missing only a couch (bookshelf and low board are not visible in Figure 2). It offers enough space to place a variety of different sized furniture in multiple locations, to enable a versatile consultation.

4.3 Implementation

To answer the abovementioned question, we implemented four different mechanisms for this representation, which were taken from literature (see Table I. for an overview). This includes the popular means of using the video stream of the on-site user for remote support, providing the remote helper with a digital representation of the on-site environment with their view fixed to the on-site users view or a freely adaptable view, as well as providing the remote helper (consultant) with a 3D representation of the on-site environment in which they can move freely in virtual reality. We use the video stream of a remote user as a baseline as it is often found in literature. As can be seen from Table I., except for mechanism 1, we kept the differences between the other three scenarios to one variable (from mechanism 2 to 3 and from mechanism 3 to 4).

Table I. Mechanisms implemented for the comparison of representations of customers environment in the study.

Mechanism	View	Customer Environment
1	Fixed to customer	Video
2	Fixed to customer	2D virtual
3	Free	2D virtual
4	Free	3D virtual

We implemented the mechanisms using the 3D game engine Unity3D and parts of the Mixed Reality Toolkit for the HoloLens. For the 3D environment we used SteamVR. For the synchronization of room positions, furniture positions and pointing positions for on-site and remote users we used the sharing service of the Mixed Reality Toolkit.

To mimic a furniture consultation, in each of the mechanisms, the consultant could choose between a fixed set of couches to be offered to the customer and place the furniture freely in the living room environment of the customer. For an easy recognition, the consultant had an additional sheet with the visual representation of the furniture. While mechanism 1 uses a video stream, mechanisms 2 to 4 used a previously captured model of our test lab living room. Analyzing the technological approaches taken for capturing, we decided to forgo 3D cameras, as our ultimate goal is to provide means for remote consultations for customers and we considered the setup procedure as not feasible for customers. The

Figure 3: The four different mechanisms implemented and compared in the study. Shared video stream (top left), 2D virtual environment fixed view (top right), 2D virtual environment free view (bottom left), 3D virtual reality (bottom right)

Microsoft HoloLens not only helps displaying the augmented reality content during consultations but is also scanning the room anyway. Therefore, for the capturing implementation we used a method similar to Dong and Höllerer [5]: First, we scanned the room by using features of the HoloLens and saved the resulting spatial map into a mesh that was then used for texture mapping. For texturing we created a shader that is drawn by a second camera, remaps each triangle on a 4096x4096 pixel sized texture and places the corresponding image from the HoloLens video camera onto each triangle. The Unity camera then saves the rendered image into a “RenderTexture”. The mesh, image and triangle count is sent to the consultant client. We decided to use this method over a continuous stream of data to minimize the amount of data needed, as we could not assume furniture sellers to have stable high-speed internet connections.

Mechanism 1: Shared Video Stream

For the first mechanism we used the front camera of the HoloLens as video stream. In this mechanism the viewport of the consultant was the same as the customer’s (see Figure 3 top left; the field of view of the consultant was fixed and resulted in a 45 degree field of view of the client). For this setup we used a prototype from earlier studies [3]. Unlike the other mechanisms, the rotation of the couch was done with the left and right arrow keys and placing the furniture further inside the room or closer to the client was done with the mouse wheel. To fix the couch on the current position, the user had to press the left mouse button, unlocking this state was also done with the left mouse button. A list of different furniture was shown on a separate view on the top right of the screen, as seen in Figure 3. As the provision of a video stream to the remote helper is most popular among remote collaboration scenarios with AR, we used mechanism 1 as the baseline for our tests.

Mechanism 2: 2D Virtual Environment, Fixed View

For the second and third mechanisms we used the previously recorded room data (see above). For mechanism 2, we then synchronized the viewport of the consultant with the customer’s field of view, resulting in the field of view seen by the customer displayed as a virtual environment on the screen of the consultant (see Figure 3 top right). Compared to the video stream, this resulted in a wider (110 degree) field of view.

Most controls for operating the furniture were set to the keyboard: To get a preview of the selected couch, the consultant had to press the “o” key. To fix it to the chosen position and to make the furniture visible for the client the “p” button had to be pressed. The position was calculated with a ray cast (a straight ray that returns every object it hits) from the camera towards the mouse pointer, returning the position of the first surface hit by the ray. This way users could place the couch on specific positions on the floor. As seen in Figure 3 (bottom left), the menu for furniture was shown in a 3x3 grid with the ability to choose the furniture by pressing “k” or “l” respectively. After pressing a menu button, the menu stayed up for three seconds. The selected furniture was highlighted by displaying it in larger size. To enable proper placement of the couch, it could be rotated with the “,” or “.” keys.

Mechanism 3: 2D Virtual Environment, Free View

The third mechanism used the same virtual environment as the second mechanism and allowed the consultant to freely move anywhere inside the environment (instead of being fixed), thus decoupling the views of consultant and customer. The keyboard layout stayed the same as in mechanism 2, with the addition of movement with the arrow keys and moving up or down with the page buttons. Additionally, the consultant was enabled to look around, which was realized with mouse movement while the right mouse button was pressed. The free perspective allowed the user to choose the desired view port or a top down view. To be able to see what the client was looking at, a frustum indicating the field of view of the customer was shown (see Figure 3, bottom left).

Mechanism 4: 3D Virtual Reality

For mechanism 4, we used the Oculus Rift S virtual reality headset to display the virtual environment we had captured before (see Figure 3, bottom right). To move freely inside the room of the customer, the consultant had the ability to teleport anywhere on the displayed room's floor by pressing the button of the controller stick connected to the headset. For pointing, a ray cast was realized by pushing the trigger button at the back of the stick and casting forward in a 45-degree angle. In Virtual Reality (VR) the menu was placed above the right hand and was used with pressing the trigger button up or down with at least |0.9| strength on the y-axis. The user can look at the menu from different angles, while moving his hands, allowing them to see the furniture from every side. Rotating the furniture was possible in every state and was realized with an input strength of at least |0.9| on the x-axis to move the object left or right. Showing the furniture in the targeted place is realized with pressing the "A" button and fixating it at the desired position with the "B" button.

5 Study

5.1 Course of the study

Each run of the study consisted of four scenarios for every consultation session, in each of which one of the mechanisms shown in Table I. was used. The scenarios were inspired by an ethnographic study on on-site furniture consultations [2]. The order of scenarios and the use of mechanisms in the respective scenarios where pseudo randomized for equal distribution in order to avoid sequence effects.

For the consultant, we set up a room in the back office of the local furniture store where all four cases could be tested without interference e.g. by real customers or colleagues. The customer was played by a researcher to minimize the variance each different person brings into a consultation and to avoid differences occurring because of different customer behavior or preferences. The customer actor followed guidelines on what kind of furniture they wanted to buy and what kind of aspects they did not want. The customer acted in the furniture laboratory described above in section 4.2.

For each scenario the consultant had 18 minutes time to find a couch the customer liked and afterwards fill in the corresponding questionnaire. Combined with an introduction where we explained the context of the study and the interview at the end of each session, each consultant needed one and a half hours. All scenarios allowed the consultant to place the same set of furniture anywhere in the room. The selectable 3D objects were taken from the real furniture catalogue available in the furniture store so that additional information such as different fabric, color or seat quality were known by the consultants and they could refer to it. As an additional means of communication, the consultant could point to positions in the room, which was represented with an arrow or sphere visible for both sides. Voice communication between the consultant (our participant) and the customer (acted by a researcher) was done via mobile network.

5.2 Methods

To compare the mechanisms described above, we measured the social presence created by each mechanism as a measure of social connection between consultant and customer, the task load created by the four mechanisms as a criterion on how much effort the consultants needed to use the mechanisms, and perceived quality of support provided with each of the mechanisms.

For social presence, we used the co-presence, attentional allocation and perceived message understanding categories of the networked minds social presence inventory [11], using a seven item Likert scale. This helps to measure to what extend the consultants perceive the affective understanding and the understanding of their messages [11].

Task load was measured using the well-known raw TLX questionnaire [12, 13], which uses a 21 item scale to measure six dimensions of task load. As the overall goal is creating a solution for remote consultation, we measure task load as well, as a lower cognitive load for remote users should help them to concentrate better on their main task: advising the customer.

For perceived support quality, we used a set of custom questions similar to [14], in which we asked the consultant about perceived consultancy quality ("I could consult the customer well"), ability to recognize the environment of the client properly ("With regard to size, I could recognized well which piece of furniture the customer can put in their living room"), support by the representation of the environment ("The representation of the room helped me to consult the customer") and perceived customer satisfaction ("The customer was satisfied with my consultancy"). All of the latter questions were to be answered on Likert scales from 1 (no agreement) to 7 (full agreement).

In addition, we added short demographic questions asking for age and existing experience with AR/VR technology. Overall, our questionnaire consisted of 29 questions.

During each test we took notes of questions regarding positioning, room scale and other questions that exceed the normal consultation or could contain valuable information for our experiment. This excludes questions about preferred furniture color, size of people in household etc., as we interpreted these as normal parts of consultancy rather than effects of the mechanisms. After the participants had completed all tasks, we conducted a brief interview asking them for the impressions and a ranking of the mechanisms.

5.3 Participants

We worked with a local furniture reseller to conduct the study. We recruited nine participants for the study, who all worked as furniture consultants for the couches we used in the study. The reseller allowed each participant one and a half hour off from their normal work to participate in our experiment. Four participants were female while five were male. Age groups were mixed, with three participants being younger than 30, two younger than 50 and four older than 50. Five of the nine participants worked in this field for over 10 years, with only two participants working for less than three years. Only one participant used a VR-HMD before with a usage time of less than one hour.

6 Results

In this section, we present our experiment findings. As the results from the Likert scales in the questionnaires are non-parametric data, we used a Friedman test to check for differences between the values for each condition. In case we found significant differences, we used a Wilcoxon Signed Rank test as a post hoc test to identify which conditions different from each other. To account for multiple pairwise tests, we applied Holm-Bonferroni correction for the α value (using α = 0.05).

From the interviews, we derived advantages and shortcomings of the mechanisms and perceived differences between them. We analyzed the rankings our participants stated for the mechanisms.

6.1 Statistical results

For task load, we found that only the values for the mental demand item in the raw TLX questionnaire showed significant differences (x(3)=7.838, p=0.043, see Table II). Unsurprisingly the mental requirement for VR was the highest (M = 11.67, SD = 4.822) compared to video (M = 8.11, SD = 3.887), fixed view (M = 8.78, SD = 3.801) and free view (M = 9.89, SD = 5.667). However, none of the pairwise Wilcoxon Signed Rank tests showed a significant difference between values with Holm-Bonferroni correction (α = 0.0083).

Table II. Results for the raw TLX questionnaire on task load for all conditions. Effect size and p values shown were computed with a Friedman test.

Mental Demand (x(3)=7.838, p=0.043)			
	mean	StDev	Mean Rank
Video	8.11	3.887	1.94
Fixed view	8.78	3.801	2.22
Free view	9.89	5.667	2.39
VR	11.67	4.822	3.44
Physical Demand (x(3)=2.475, p=0.502)			
	mean	StDev	Mean Rank
Video	6.78	5.239	2.22
Fixed view	6.22	3.456	2.56
Free view	7.78	5.142	2.22
VR	8.67	4.359	3.00
Temporal Demand (x(3)=5.524, p=0.137)			
	mean	StDev	Mean Rank
Video	6.11	3.180	1.67
Fixed view	7.56	3.358	2.72
Free view	8.33	4.500	2.78
VR	7.56	4.586	2.83
Performance (x(3)=3.727, p=0.137)			
	mean	StDev	Mean Rank
Video	16.56	2.833	3.00
Fixed view	13.78	4.790	2.00
Free view	13.78	4.790	2.44
VR	15.00	4.359	2.56
Effort (x(3)=4.172, p=0.243)			
	mean	StDev	Mean Rank
Video	6.00	3.464	2.06
Fixed view	6.33	3.428	2.78
Free view	6.33	3.742	2.22
VR	8.44	5.028	2.94
Frustration (x(3)=4.720, p=0.198)			
	mean	StDev	Mean Rank
Video	9.00	6.481	2.50
Fixed view	6.67	3.674	2.22
Free view	7.56	4.216	2.06
VR	10.33	5.220	3.22

The Networked Minds questionnaire did not show any significant differences (see Table III). This is noteworthy, as one would expect 3D immersive virtual reality to create more presence. Likewise, for the perceived support quality, we did not find any significant differences between the values (see Table IV).

Table III. Results for the co-presence, attentional allocation and perceived message understanding categories of the Networked Minds questionnaire. Effect size and p values were computed with a Friedman test.

Co-Presence (x(3)=2.797, p=0.45)			
	mean	StDev	Mean Rank
Video	6.30	0.978	2.67
Fixed view	6.46	0.644	2.72
Free view	6.43	0.830	2.61
VR	6.07	1.267	2.00
Attentional Allocation (x(3)=3.0, p=0.437)			
	mean	StDev	Mean Rank
Video	3.76	.961	2.89
Fixed view	3.52	.549	2.67
Free view	3.48	.437	2.33
VR	3.46	.519	2.11
Perceived Message Understanding (x(3)=3.333, p=0.358)			
	mean	StDev	Mean Rank
Video	4.72	.677	2.94
Fixed view	4.81	.437	2.72
Free view	4.65	.444	2.28
VR	4.67	.312	2.06

Notably, most mean values for the Networked Minds questionnaire as well as the questions we aksed the consultants showed hardly any differences at all. For example, mean values for the question “My partner found it easy to understand me” showed very similar values for the video (M =5.78, SD = 0.972), fixed view (M = 5.89, SD = 1.453), free view (M = 5.89, SD = 0.928) and VR (M = 5.78, SD = 1.202) conditions. Similarly values for attentional allocation were very similar for the video (M = 3.76, SD = 0.962), fixed view (M = 3.52, SD = 0.549), free view (M = 3.48, SD = 0.437) and VR (M = 3.46, SD = 0.519) conditions.

While it is hard to interpret these results, it could be the case that the consultants regarded these questions as measures for their ability to consult clients in the scenarios rather than as measures for the support provided by the different media. Our interview data supports this assumption.

Table IV. Results for the questionnaire on perceived support quality. Effect size and p values shown were computed with a Friedman test.

I could consult the customer well. (x(3)=1.545, p=0.727)			
	mean	StDev	Mean Rank
Video	6.00	1.323	2.44
Fixed view	6.11	1.054	2.72
Free view	6.00	1.323	2.56
VR	5.78	1.481	2.28
With regard to size, I could recognize well which piece of furniture the customer can put in their living room. (x(3)=3.286, p=0.364)			
	mean	StDev	Mean Rank
Video	5.22	1.394	2.22
Fixed view	5.56	1.333	2.33
Free view	6.0	1.0	3.06
VR	5.78	1.481	2.39
The representation of the room helped me to consult the customer. (x(3)=1.881, p=0.62)			
	mean	StDev	Mean Rank
Video	5.56	1.236	2.39
Fixed view	5.44	1.509	2.17
Free view	5.89	1.453	2.67
VR	6.11	1.364	2.78
The customer was satisfied with my consultancy. (x(3)=2.132, p=0.617)			
	mean	StDev	Mean Rank
Video	6.44	0.726	2.83
Fixed view	6.11	1.054	2.33
Free view	6.11	1.269	2.33
VR	6.22	1.093	2.50

It is surprising that there is no difference for the statement “With regard to size, I could recognize well which piece of furniture the customer can put in their living room.”. As we thought that VR could have an edge in that regard due to the addition of depth perception. While the mean value for VR is slightly higher than the other conditions, there was no significant difference. This supports our assumption that the consultants' perception regarding their general skills had a major influence on the values while the medium used had a minor influence.

6.2 Interview results

The interviews revealed huge difference in perception and preferences among the consultants. This can be seen in their preferences for mechanisms: We asked each participant to rank the tools from one to four, with one being the mechanism they prefer. Table V shows the ranking for each mechanism. In this ranking, the video stream was ranked worse than the other mechanisms, with most people rating it the least favorable. The fixed view mechanism received mixed rankings with a slight tendency towards a negative perception. Virtual Reality and the free view mechanisms got mixed rating as well but were the two mechanisms ranked first and second most often. Overall, the rankings of the cases (excluding video stream) were close together so that no significant difference could be identified. However, this ranking showed a tendency against the video condition and for the free view and VR conditions. Below, we analyze the statements of the consultants regarding reasons for their rankings and their preferences.

Table V: Amount each mechanism was placed on respective rank. In case participant placed multiple mechanisms on the same rank they are counted as the lower rank.

Mechanism/Rank	**1**	**2**	**3**	**4**
Video	0	2	0	7
Fixed view	2	1	5	1
Free view	4	2	0	3
Virtual Reality	3	3	2	1

Shared Video Stream

Five of the participants who rated this mechanism low primarily stated that the clients view was unwieldy and the field of view too small to work with efficiently. They mentioned that placing furniture was a hassle while the client was moving, and that even when the client tried to hold still, slight head movement was perceived disturbing by the consultant.

2D Virtual Environment, Fixed View

The two participants who rated the second mechanism highest mentioned that it is most useful to have the clients view. They argued that clients will focus on the parts of the room where they want the furniture to be placed. On the other hand, five of the participants who rated both the first and second mechanism low argued that the client view was mostly distracting and limiting. They valued the possibility to freely look around in the room to get a better understanding of the room structure and furniture composition. Much like the first mechanism, they stated that movements of the client made it difficult to place the furniture on the intended position.

2D Virtual Environment, Free View

The participants who rated the free view mechanism as highest mostly valued the ability to look around freely and the usability of features to move and look around, while two of the participants who rated this mechanism lower criticized the keyboard layout. They mentioned that other tools they work with primarily use the mouse for interactions, so that these kinds of interactions feel more intuitive for them. One participant who rated this mechanism high also mentioned that they liked the visibility of the client's position and field of view. On a side note, two participants mentioned that an option, to see the room in a bird view would have been useful.

3D Virtual Reality

Two of the three participants that rated VR as the best mechanism listed the 3D vision as the main reason for rating it the highest. Participants that rated VR lower mentioned that it was distracting to work with but argued that this could be due to it being new and that a longer period of familiarization could mitigate this feeling. It was also mentioned that the usability of the VR was easy to understand, however as mentioned in the upcoming observations paragraph this stands in contradiction to our observations.

Representation

Seven of the participants agreed that the captured view of the room in mechanisms 2, 3 and 4 was helpful. Additionally, two participants mentioned that the spatial understanding helps in the consultation. While the view was received positively, six of the participants mentioned that the pixelation and the abrupt change of color was bothering them and that it was hard to identify small objects like fruit baskets. Nevertheless, six of the nine participants thought that this kind of view would be a helpful extension to a furniture consultation, while the other three were undecided about the helpfulness with one leaning towards it being not helpful for his work.

Observations

Our observations suggest that the controls of mechanism one was the easiest to handle. This is explainable by the fact that the tools that are already used in the store have a similar type of input handling and therefor needed the least adaption and familiarization period. In general, our observations revealed that the participants had problems with the controls of each tool and would have needed a longer testing phase, to remember each available control element. This was also shown in the number of tools that were used in the conditions. Only a few used the option to point at areas of interest or used the ability to fly or teleport to different positions in the room. In nearly all cases the participants chose the starting position or slightly adapted the starting position at the beginning and left it as it was over the tests period.

7 Discussion

The results of our experiment show no significant differences in the metrics we applied. However, from the interviews with participants, we both learned about possible reasons for the lack of differences as well as about preferences of the consultants for further usage of the tools. Both enable us to draw conclusions for further work and design.

7.1 The Influence of Representations on Remote Consultation

On first sight, contrary to previous studies our experiment seems to suggest that there is no additional benefit when using virtual representations of the environment or virtual reality for furniture consultation. This is indicated by our data, as each mechanism showed similar results. However, from the interviews we can also see that most consultants liked the virtual representation of the environment and that they would potentially have a more positive view towards VR if they had more time to learn how to operate the tool in the respective conditions (see below for a discussion of this). In addition, although our results were not significant towards this point, from our interviews we gathered that there was a tendency towards a free view, as the fixed view was mostly perceived as disruptive. In summary, while virtual environment and free view demanded additional effort from the participants, at the same time they created similar results for perceived support quality and were ranked slightly better than other conditions. Therefore, we may assume that with more exposition to these mechanisms, free view and virtual environments (even VR) may provide positive effects on online consultancy.

7.2 The Influence of Familiarity with the Medium on Remote Consultation

Our evaluation showed hardly any differences in the TLX, Networked Minds social presence and perceived quality of support measures that we applied. The only significant difference is in the TLX mental demand item, which was rated significantly higher for VR. As described above, statements of participants in the interview indicate that this could be due to the unfamiliarity to such systems. This effect was probably strongest for VR, as it included free view, stereoscopic view and controls: None of the participants had ever used the sticks typically used to navigate in VR. For VR, some of the participants even stated that they would need more time to learn how to deal with these controls, and that this could diminish the distraction that they perceived.

This may also be an explanation of the overall very close results from the questionnaire. Results for all items were also somewhat high overall. This could be explained by the fact that the familiarization phase with the new tools was very limited due to the overall time a consultant was available for the study and the low experience most consultant had with such tools. This also aligns with our observation and the comments on troubles with some of the controls. As mentioned above, another reason may be found in the fact that for consultants, their self-assessment towards client satisfaction tends to be high, as it would otherwise mean that the consultants did not understand the client and could not fulfill their jobs properly.

7.3 Potentials and Needs of Virtual Representations for Remote Consultation

Despite our findings on increased mental effort to interact with VR devices and comments on lacking familiarity with the operation of virtual representations, our work also did not show significant disadvantages of virtual representations for remote consultancy. In addition, consultants also stated that they liked the spatial understanding that the 3D view in VR gave them, and they rated all conditions using virtual representations higher than the video stream condition. Given that using technologies for the first time (e.g., VR goggles, virtual representations on screens) creates a considerable bias when compared to more familiar technologies (see also [23]), we conclude that virtual representations are as good as video streams for remote consultations, and that they may even show advantages once people get used to them.

In this context, it is also important to take a closer look at study participants. In our case, furniture consultants are domain experts, but they showed very little experience in dealing with virtual representations and interacting with them. This resulted in the difficulties described above. Other studies (e.g., [21] and many others) use students and research staff, who potentially have more experience with operating the devices. This may be one of the reasons why these studies revealed immediate benefits of virtual representations. We conclude from this that it is very important to integrate potential later users into studies like ours, as this integration reveals what needs to be done to implement the corresponding technology in practice.

7.4 Design Ideas and Recommendations

Given the mixed statements of consultants and our data, for the support of furniture consultation we recommend combining different mechanisms and let the user choose. Starting with the 2D free view and fixed view application, clients may switch between both modes freely at any given time, to enable a smooth control for overview and placement in free view and the ability to focus on specific areas the customer is looking at in fixed view. It could also be useful to create the ability to switch into a bird eye view with a simple click like in [3], as it was mentioned as an improvement by two consultants .

Since Virtual Reality did not have an instant impact on the consultation but could have the potential with more practice in the future, this option should also be offered. However, as our results suggest, this would at least need some training and, to make sure that there are positive effects, some more research. Either way, we would discourage a mixed application between 2D and 3D, as putting on the VR-HMD might interrupt the flow and lower the quality of the consultation.

We also recommend an improved room representation. While it was possible for the consultant the get the gist of the room, it would improve the consultation even further when consultants are able to clearly identify even small items like fruit baskets etc.

8 Limitations

The toolset of the study was limited. With the ability to select one of nine different couches the variety of the consultation suffered, as it only represented a transect of the furniture store's catalogue. An implementation of all available options would change the interface considerably and would expand the possibility of the consultation. Additional features such as a virtual tape measure would allow for an even better consultation. Running an experiment under these conditions might have amplified some of the problems mentioned by the consultants, including the difficulties to place couches in fixed view conditions. In addition, it may have amplified the positive effect of spatial recognition.

With our choice of choosing an actor as the on-site user, we could not assess the perception of the customer. This was done to keep the conditions consistent, but also let us miss some insights: It would be interesting to see different customer behavior and the impact this would have on the toolset of the consultation, and vice versa. Further studies with real customers are required.

It is important to note that our study was conducted with only nine participants and thus the generalizability is limited. While the knowledge of the participants as field experts was extremely high, this study can only indicate preferences of this domain.

9 Conclusion

In this paper, we implemented and evaluated four different mechanisms for the consultant's application. We compared a 2D video stream with a captured texturized 2D or 3D virtual model, as well as a view fixed to the client and a free view. We tested our mechanisms with nine furniture consultants and found hardly any differences between the results. From our interviews we deduct that the handling of the 2D and 3D environment was similarly difficult for the consultants, as it needed navigation and command techniques they were not used to. This suggests that with more familiarity with these tools, they may provide a benefit for furniture consultancy. We also found from the interviews that different mechanisms showed benefits for different consultants, and we recommend a mix of fixed and free views in a virtual environment for the design of tools.

In the future, we want to adjust the consultants' interface so that it more intuitive to use. We will include the ability to switch between both customers view and free view with the additional ability to switch into a bird view with just one click. We want to expand our tool further with more options and a real expanded client tool. With further tests we want to investigate how to support the consultant and the customer in a remote consultation. This includes testing immediately after the exposure with these tools and after consultants have familiarized with them.

REFERENCES

[1] Bier, E.A. 1987. Skitters and Jacks: Interactive 3D Positioning Tools. *Proceedings of the 1986 Workshop on Interactive 3D Graphics* (New York, NY, USA, 1987), 183–196.

[2] Blunk, O., Brown, G., Osmers, N. and Prilla, M. 2020. Potentials of AR technology for the digitalization of consultancy intensive sales processes on the example of furniture sales. *Entwicklungen, Chancen und Herausforderungen der Digitalisierung* (Postdam, Mar. 2020), 57–72.

[3] Brown, G. and Prilla, M. 2019. Evaluating Pointing Modes and Frames of Reference for Remotely Supporting an Augmented Reality User in a Collaborative (Virtual) Environment. *Mensch & Computer 2019* (2019).

[4] Datcu, D., Lukosch, S.G. and Lukosch, H.K. 2016. Handheld Augmented Reality for Distributed Collaborative Crime Scene Investigation. *Proceedings of the 19th International Conference on Supporting Group Work* (Sanibel Island, Florida, USA, 2016), 267–276.

[5] Dong, S. and Hollerer, T. 2018. Real-Time Re-Textured Geometry Modeling Using Microsoft HoloLens. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Reutlingen, Mar. 2018), 231–237.

[6] Duval, E. 2011. Attention Please!: Learning Analytics for Visualization and Recommendation. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (New York, NY, USA, 2011), 9–17.

[7] Fakourfar, O., Ta, K., Tang, R., Bateman, S. and Tang, A. 2016. Stabilized Annotations for Mobile Remote Assistance. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2016), 1548–1560.

[8] Fussell, S.R., Setlock, L.D. and Kraut, R.E. 2003. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. *Proceedings of the SIGCHI conference on Human factors in computing systems* (2003), 513–520.

[9] Gao, L., Bai, H., Lindeman, R. and Billinghurst, M. 2017. Static local environment capturing and sharing for MR remote collaboration. *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications on - SA '17* (Bangkok, Thailand, 2017), 1–6.

[10] Gauglitz, S., Nuernberger, B., Turk, M. and Höllerer, T. 2014. World-stabilized Annotations and Virtual Scene Navigation for Remote Collaboration. *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2014), 449–459.

[11] Harms, C. and Biocca, F. 2004. Internal consistency and reliability of the networked minds measure of social presence. (2004).

[12] Hart, S.G. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 50, 9 (Oct. 2006), 904–908. DOI:https://doi.org/10.1177/154193120605000909.

[13] Hart, S.G. and Staveland, L.E. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*. P.A. Hancock and N. Meshkati, eds. North-Holland. 139–183.

[14] Johnson, S., Gibson, M. and Mutlu, B. 2015. Handheld or Handsfree?: Remote Collaboration via Lightweight Head-Mounted Displays and Handheld Devices. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (New York, NY, USA, 2015), 1825–1836.

[15] Jungermann, H. and Fischer, K. 2005. Using Expertise and Experience for Giving and Taking Advice. 157–173.

[16] Kasahara, S. and Rekimoto, J. 2014. JackIn: Integrating First-person View with Out-of-body Vision Generation for Human-human Augmentation. *Proceedings of the 5th Augmented Human International Conference* (New York, NY, USA, 2014), 46:1–46:8.

[17] Kolkmeier, J., Harmsen, E., Giesselink, S., Reidsma, D., Theune, M. and Heylen, D. 2018. With a little help from a holographic friend: the OpenIMPRESS mixed reality telepresence toolkit for remote collaboration systems. *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology - VRST '18* (Tokyo, Japan, 2018), 1–11.

[18] Komiyama, R., Miyaki, T. and Rekimoto, J. 2017. JackIn space: designing a seamless transition between first and third person view for effective telepresence collaborations. *Proceedings of the 8th Augmented Human International Conference* (2017), 1–9.

[19] Müller, R., Helmert, J.R., Pannasch, S. and Velichkovsky, B.M. 2013. Gaze transfer in remote cooperation: is it always helpful to see what your partner is attending to? *Quarterly Journal of Experimental Psychology (2006)*. 66, 7 (2013), 1302–1316. DOI:https://doi.org/10.1080/17470218.2012.737813.

[20] Park, K.S., Kapoor, A. and Leigh, J. 2000. Lessons learned from employing multiple perspectives in a collaborative virtual environment for visualizing scientific data. *Proceedings of the third international conference on Collaborative virtual environments - CVE '00* (San Francisco, California, United States, 2000), 73–82.

[21] Piumsomboon, T., Lee, G.A., Irlitti, A., Ens, B., Thomas, B.H. and Billinghurst, M. 2019. On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), 228:1–228:17.

[22] Poelman, R., Akman, O., Lukosch, S. and Jonker, P. 2012. As if Being There: Mediated Reality for Crime Scene Investigation. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2012), 1267–1276.

[23] Prilla, M., Jan\s sen, M. and Kunzendorff, T. 2019. How to Interact with Augmented Reality Head Mounted Devices in Care Work? A Study Comparing Handheld Touch (Hands-on) and Gesture (Hands-free) Interaction. *AIS Transactions on Human-Computer Interaction*. 11, 3 (2019), 157–178.

[24] Rae, I., Mutlu, B. and Takayama, L. 2014. Bodies in motion: mobility, presence, and task awareness in telepresence. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014), 2153–2162.

[25] Salamin, P., Thalmann, D. and Vexo, F. 2006. The Benefits of Third-person Perspective in Virtual and Augmented Reality? *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, 2006), 27–30.

[26] Schafer, W.A., Bowman, D.A. and Schafer, W. *Evaluating the Effects of Frame of Reference on Spatial Collaboration Using Desktop Collaborative Virtual Environments*.

[27] Shepard, R.N. and Metzler, J. 1971. Mental Rotation of Three-Dimensional Objects. *Science*. 171, 3972 (Feb. 1971), 701–703. DOI:https://doi.org/10.1126/science.171.3972.701.

[28] Tachakra, S. 2001. Depth Perception in Telemedical Consultations. *Telemedicine Journal and e-Health*. 7, 2 (Jun. 2001), 77–85. DOI:https://doi.org/10.1089/153056201750279575.

[29] Tait, M. and Billinghurst, M. 2015. The effect of view independence in a collaborative AR system. *Computer Supported Cooperative Work (CSCW)*. (2015).

[30] Tecchia, F., Alem, L. and Huang, W. 2012. 3D helping hands: a gesture based MR system for remote collaboration. *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry - VRCAI '12* (Singapore, Singapore, 2012), 323.

[31] Teo, T., Lawrence, L., Lee, G.A., Billinghurst, M. and Adcock, M. 2019. Mixed Reality Remote Collaboration Combining 360 Video and 3D Reconstruction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk, May 2019), 1–14.