Hybrid User Interface for Audience Feedback Guided Asymmetric Immersive Presentation of Financial Data

Matt Gottsacker* Mengyu Chen† David Saffo‡ Feiyu Lu§ Blair MacIntyre¶

Global Technology Applied Research, JPMorgan Chase & Co.

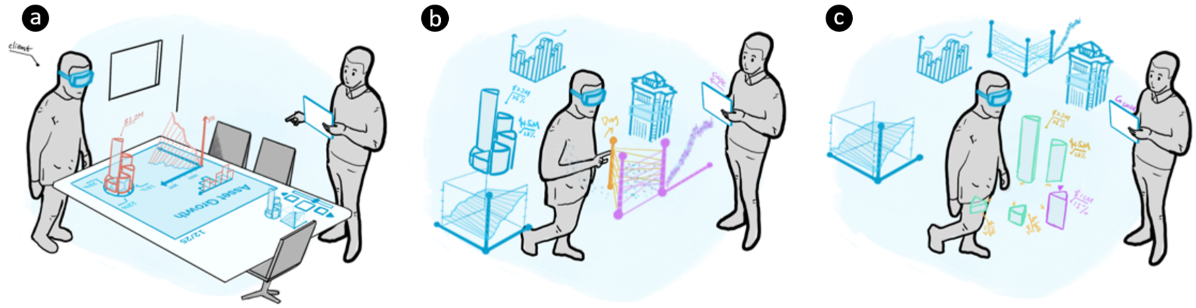

Figure 1: An design illustration of the system being developed: (a) presentation content is grounded in a physical environment with the additive third dimension; (b) the audience can explore and interact with data visualizations through embodied interactions; (c) the presenter can draw the viewer's attention and control when, where, and how the content is displayed throughout the presentation.

ABSTRACT

This paper presents the design of a system for giving engaging, immersive presentations based on financial data in an immersive environment. The system consists of a presenter controlling the presentation flow through a tablet interface and an audience experiencing the presentation content through an augmented reality (AR) head-worn display (HWD). The system leverages various sensors on the HWD to detect the audience's cognitive state and provide actionable and context-aware suggestions to the presenter about how to adjust their presentation delivery. We discuss the user scenario motivating our system design, its associated design considerations, the features of our system, and how our system can be extended to other presentation contexts.

Index Terms: H.5.1 [INFORMATION INTERFACES AND PRESENTATION (e.g., HCI)]: Multimedia Information Systems—Artificial, augmented, and virtual realities; K.4.3 [COMPUTERS AND SOCIETY]: Organizational Impacts—Computer-supported collaborative work

1 INTRODUCTION

Facilitating the dynamic of non-experts seeking services from experts, or professional-client relationships, is vital for services requiring expert-level domain knowledge, such as health care, engineering, and financial services. To this end, presentations are a common mechanism professionals utilize to communicate and educate clients on complex domain-specific subject matter. Take, for example, financial service professionals such as bankers, financial advisors, and certified public accountants, who often present financially related material to educate clients on financial goods and services. A presentation in finance often includes content of various forms, such as regional market data, product price charts, sector-specific statistics, and economic outlook, in addition to details specific to the client's finances. Based on the knowledge level and interest of the audience, professionals often need to prepare and customize content to help the audience understand certain economic concepts and gain meaningful insights from the financial data.

To improve the communication of complex financial information, various methods have been explored, such as adding extra dimensions to data visualizations (3D line charts, 3D data points, etc.) [6, 29], illustrating geospatial relationships of regional market information in 3D [2, 20], and displaying content in a virtual environment [8, 25, 40]. While existing methods provide novel ways to make it easier for the audience to get insights from the data, these systems are often self-exploratory and lack guidance from an expert, making them less useful or applicable in a real-world scenario.

We are interested in making the presentation of financial information during professional-client meetings more engaging and interactive so that the audience may get less overwhelmed by the complex information presented to them and get better financial literacy education. Mixed reality offers huge opportunities as it provides higher interaction affordances for the client to engage with the content as well as the capacity to display the content in a more expressive way. In this paper, we explore the design considerations of an immersive presentation system in the context of financial services and professional-client dynamics. We then discuss our early system implementation of an asymmetric presentation system that utilizes a hybrid user interface across mixed reality and mobile devices. The system uses implicit and explicit audience feedback to provide a more customizable presenter and audience experience.

2 PRIOR WORK

2.1 Multi-modal Augmented Presentation

Blending visual elements such as texts, graphs, and illustrations with a live presenter's gestures and verbal content is a popular technique widely used in commercial product presentations, educational videos, and TV productions. The augmented visual information on the screen enriches the audience's perceptual experience and enables an easier understanding of the presented content [22]. Studies have

*e-mail: matt.gottsacker@jpmchase.com
†e-mail:mengyu.chen@jpmchase.com
‡e-mail:david.saffo@jpmchase.com
§e-mail:feiyu.lu@jpmchase.com
¶e-mail:blair.macintyre@jpmchase.com

shown that presentations and tutorials with more infographic and visual cues can make the audience maintain higher emotional and cognitive engagement over a longer period of time [19, 49]. Besides the most common way of presenting visual slides with mouse and keyboard (e.g., Powerpoint, Keynote), recent works have explored using different interactive cues, such as speech [24], body gestures [38], and sketch-drawing [32] to trigger and place real-time augmented visual contents in a presentation. Such improvisation techniques provide additional customizability during the presentation and allow the presenter to introduce highly personalized materials to the audience on the fly.

2.2 MR Collaboration Systems

There has been a wide range of applications and research works of mixed reality collaboration in domains such as education [7, 30, 33], manufacturing [12, 41, 48], medical service [26], architecture design [17, 18], remote assistance [15, 16], etc. Schäfer [39] in a recent survey, categorized these use cases into mainly three categories: meeting, design, and remote expert. The interactive features and benefits brought by these mixed reality collaboration systems are mainly about shared 3D object manipulation, multi-media content sharing, 2D/3D drawing, natural gesture control, annotation, and AR view sharing. Recently, there is also an increasing interest in asymmetric collaboration where individuals of a collaboration group "have different means to visualize and interact with virtual content." [14] A recent study showed that assigning different roles and tasks to users in asymmetric collaboration groups will have a direct influence on their sense of presence and overall experience [23]. In our work, as the roles of expert (presenter) and non-expert (audience) are clearly defined, we leverage the asymmetry to better accommodate the different needs of different end-users in our immersive presentation system design.

3 Design Considerations

There are three categories of end-user needs and requirements that we are considering: 1) social interaction, 2) interactive content, and 3) shared awareness.

3.1 Social Interaction

Although the COVID pandemic has had an impact on how the service industry operates and made it shift to a hybrid mode of in-person and remote meetings, the professionals in financial services may still prefer in-person meetings with their clients to maintain a high level of social engagement through direct eye contact and body languages. Having face-to-face meetings may help them gain more trust [46, 50] from the clients and may lead to better outcomes [5, 42] as prior studies have shown. Therefore, it is desirable for an XR presentation system to support interaction feedback between the users and approximate natural conversations via multiple modalities (e.g., speech, gesture, gaze). At the same time, in day-to-day business operations, the presenter, as an expert service provider, may spend an extensive amount of time communicating and interacting with many clients. For the foreseeable future, it will be difficult for the presenter to wear an XR headset for an entire day [3], and such devices complicate interpersonal communication, e.g., by occluding some of the wearer's facial expressions and preventing effective eye contact [27, 47]. Providing solutions to alleviate the burden of wearing an XR headset or finding alternatives to address the eye contact needs of the presenter in both co-located and remote virtual environment is a challenge that needs to be addressed.

Remote methods of communication have their distinct benefits for professional-client relationships as well. Perhaps the biggest advantage of these benefits is it allows professionals and clients to connect largely without the limitation of physical location. This allows experts to expand the number of potential clients they can interact with and reduces the time and resources spent traveling for both parties. However, these benefits are costly, as they may lower overall social interaction and interpersonal engagement in professional-client relationships. Immersive approaches have the potential to alleviate these challenges allowing professional-client relationships to leverage remote communication while painting high levels of social interaction and interpersonal engagement by enabling shared virtual spaces.

In both of these scenarios, distributed or co-located, there remains an open question of how professional-client relationships can best share virtual and physical space. Furthermore, it remains an open question of how occupying different ends of the reality-virtuality continuum (RVC) [28, 43] affects overall social interaction and interpersonal engagement [31, 44]. When co-located, fully immersing both professionals and clients in VR potentially lowers interpersonal engagement by isolating users from reality. AR presentation might help mitigate this by encouraging physical face-to-face interactions in the shared physical space; however, without further interventions, each user's view of the other will still be obstructed by the physical headset and, potentially, the virtual content. In the distributed context, these constraints are exacerbated as users no longer share the same physical constraints, and as such, the ability to see each other becomes completely reliant on technology interventions. With current approaches, these interventions take the form of either video-streaming or 3D avatars [13, 36], the particular design of which can greatly affect the dynamics of professional-client relationships. Interpersonal engagement will benefit from commercial technology improvements in eye-gaze, facial, and full-body sensing.

3.2 Interactive Content

Financial services require a highly personalized approach for the presenter to prepare materials for the audience. Every client may have very specific needs based on their personal financial situation and plans. The presenter may also need to introduce financial literacy content during the conversation with the client and help them better understand the presented materials. Therefore, the presenter needs to be able to easily select and manipulate the presentation content based on the audience's feedback during the presentation. In some cases, the presenter may also need tools to improvise and create original content (e.g., charts and graphs) that cannot be prepared ahead of the presentation due to the uncertain nature of the client's financial situation. Providing support to create data visualization and enabling custom and accurate input for visual content (e.g., sketching, annotating) can be useful for the presenter. On the other hand, the audience may wish to have a more engaging experience during their financial literacy education. Allowing them to interact with the material and increase their domain knowledge while keeping a high level of engagement will be beneficial.

3.3 Shared Awareness

Unlike 2D content on a physical screen visible to both the presenter and the audience, immersive 3D content in an HWD can be viewed differently from different perspectives and based on the capabilities of different devices. Given that immersive technology is relatively new to the general public, the audience of a financial presentation may not have sufficient experience with consuming detailed financial content in a virtual or augmented space. Therefore, the presenter and the audience should have a shared awareness of what each other sees so that the expert can spatially guide the non-expert audience during the experience and help them better navigate in the space [21, 35].

In addition to spatial navigation, the perspective difference may also challenge the presenter to spatially present virtual content at places where the audience can clearly see it. Especially in a room-scale immersive space where the audience may walk around, the presenter needs to ensure they place presentation content not too far or close to the audience. Awareness of where the audience is located or facing can be useful for a more effective content presentation.

On the other hand, during a presentation on financial literacy education, the presenter may present advanced financial materials that are complex and difficult for non-expert audiences. The audience may find it difficult to reference such materials via just verbal or gestural cues. The presenter may need additional means (e.g., learn from audience gaze, shared annotation) to locate the content being referenced by the audience and be on the same page with them to reduce communication costs.

4 System Design

Our goal is to support financial experts in giving their audience an engaging and easy-to-understand presentation of financial materials in an immersive space. Building on the design considerations above, the system supports a variety of content, including 2D text, images, videos, and 3D objects, as well as interactive 2D and 3D data visualizations, allowing the presenter to explain complex financial data while keeping the audience engaged. Technically, the system will let us explore different spatial content presentation techniques, such as multi-view object placement and automatic content arrangement. The system is being designed to facilitate on-the-fly content customization by the presenter so that the experience can be adjusted based on audience interests and needs.

The immersive presentation system is specifically designed as an asymmetric experience to accommodate the different tasks and needs of the presenter and the audience. The presenter, as a domain expert and service provider, will use a handheld tablet to control the presentation content while not being encumbered by an HWD. The audience will be placed in a mixed-reality HWD to fully engage with the presented virtual content and consume it in their own physical space. Fig. 1 shows a concept storyboard of our initial target scenario.

4.1 Immersive Content Presentation

The system has both authoring and presentation components. A financial expert can prepare content elements similar to those available to common 2D presentation programs, such as text, images, and videos, in addition to 3D content elements and interactive 2D and 3D data visualizations. The presenter prepares content and arranges it in a series of *scenes* (analogous to slides in a 2D presentation program) before the presentation begins. These scenes store a set of virtual objects that have been annotated with *spatial layout and context labels* (e.g., main content, title, supporting data) to help the system adapt to different presentation environments, and adjust content dynamically as the presentation proceeds. The presenter is able to add additional prepared content elements to a scene as needed throughout the presentation. During the presentation, the presenter will need to accurately perceive the layout of the current in the spatial environment through intuitive views on the handheld tablet and interact with it by highlighting, arranging, adding, or removing content elements. Most of the spatial layout of the content elements will be handled automatically by the system using spatial labels and reacting to input from the presenter, as discussed below.

4.1.1 Spatial Object Views

Placing virtual content in the presentation scene presents challenges because the presenter must be able to specify positions in the 3D presentation scene from the 2D screen of the tablet. To support the efficient adjustment or placement of virtual objects at presentation time, the tablet interface includes multiple views for the presenter.

Top-down View - The top-down view displays a virtual camera's rendering of the virtual scene. The presenter uses the touch screen to browse and select different virtual content elements and drag and drop them into the positions shown in the top-down view. The objects will then be instantiated and displayed as augmented content in the audience's AR headset. The system may adjust other content (e.g., its scale and position) in response to this new element being added.

Shared Audience View - The system supports the presenter viewing the presentation environment from the audience's point of view (by rendering the scene via a virtual camera placed in the audience's head position). The virtual camera tracks the audience's movements so that the presenter can be aware of where the audience is looking and help them navigate or guide their interaction with the virtual objects in the space. The presenter can use this view to verify how the presentation content looks to the audience. In both this view, and the top-down view, the presenter can point to elements in the scene to support using deictic phrases to refer to virtual content in the presentation space.

Feedforward View Similar to the slide previews in common 2D slide presentation software, the presenter is able to view subsequent presentation scenes before they are displayed to the audience. This preview is shown in both the top-down view and the audience view. In the audience feedforward view, a virtual camera is placed in the audience's head position, but it is rotated to look toward the center of the next scene.

4.1.2 Presentation Improvisation

During the presentation, the presenter may improvise by adding virtual content to the presentation that was not part of the originally authored scenes. For example, the presenter may create backup data visualizations to support a particular point they are presenting, or they may create data visualizations on-the-fly to illustrate basic concepts. Positioning content elements on the fly in 3D can be challenging because the presenter controls the 3D presentation through a 2D application, but the spatial object views described above coupled with dynamic content arrangement reduce the impact of this problem. For instance, the presenter can drag virtual objects into surfaces shown in the top-down view. Then, the virtual content will be visible to the presenter in the feedforward view. If the audience is not looking in the direction where the virtual content will be placed, the audience feedforward view's virtual camera will move to show the content. The presenter can perform additional fine-grained manipulations (i.e., adjust the content's position, rotation, and scale) through controls near the feedforward view. When the presenter confirms the placement, the content will be made visible to the audience. Any further manipulations (e.g., the presenter rotating the content so the audience can get a different view) can be done through controls that are visible directly in the audience feedforward view.

4.1.3 Automatic Content Arrangement

A presentation could be given in a variety of structurally different physical and immersive spaces. To reduce the friction of the presenter authoring and setting up the presentation, we are designing the system to intelligently lay out content in the physical space. The goals of the automatic arrangement are two-fold: (1) to maximize the visibility and spatial referencing of the presentation content for both parties; (2) to minimize the cost of interacting with the virtual content for both. Recent research has demonstrated the potential of adaptive content placement for productivity work [10, 11]. However, little has been explored in multi-user scenarios in asymmetrical settings. In our design, we plan to author the presentation for an abstract space with semantically meaningful content zones, and then adapt that space to the real space when the presentation begins. The presenter can make quick adjustments to the generated layout to enhance their control and preferred content style.

4.1.4 Data Visualization and Manipulation

Data visualization techniques are a core component of financial data presentations. For example, time series visualizations, such as line charts, are commonly used to illustrate the change in the value of assets over time due to changing market conditions, compounding interest, or other economic factors. However, these visualizations are often presented as static images accompanied by text on presentation slides or captions and body text in reports—limiting the overall engagement of audiences.

Utilizing immersive analytics presentation and data storytelling techniques provides the opportunity to implement several interventions to increase the level of engagement of the presentation. Our system aims to leverage spatial immersion and situated visualization to embed interactive 3D visualizations into mixed-reality financial presentations. To this end, the presenter interface can be used to place pre-authored data visualizations into the virtual space. Presenters can then manipulate the data visualizations in several ways, such as changing its facet and position to change the audience's perspective, annotating or highlighting points of interest, and animating transitions of the visualization through changes in its data, such as time range, or through changes in its encoding channels or visualization type. These visualization techniques will enable presenters to change and manipulate data visualizations alongside the theme and narrative of the presentation.

4.2 Audience Feedback Driven Content

Our system is designed to both restore and enhance the interpersonal communication cues that may be negatively affected or outright blocked by the asymmetric nature of our system and the AR headset. For example, the weight and form factor of the AR headset might cause the audience to emote or express themselves differently [27], and the headset also prevents the presenter from observing the audience's full face and therefore making judgments about their cognitive and emotional states [47]. Our system leverages the multimodal sensors on the audience's AR headset to track the audience's behaviors (e.g., gaze and head pose) and provide useful insights to the presenter to help them better navigate this social interaction. In this way, the presenter can adjust their presentation pace or change materials accordingly and provide a more personalized presentation experience for the audience.

4.2.1 Audience Feedback Tracking

In addition to the presenter being able to view the audience's view of the presentation content, the system also tracks the behaviors and speech of the audience to provide guidance to the presenter. The system can leverage sensors commonly in consumer AR headsets, such as head motion trackers, eye trackers, and microphones. Head motion can provide insight into the audience's emotional state [37]. The audience's gaze can be used to detect the virtual (and real) objects to which the audience is attending, as well as associated measures such as emotional state [9] and cognitive load [4]. Last, recording the audience's speech with the headset's microphone and analyzing it with natural language processing tools can be useful for detecting the audience's emotional state [1].

4.2.2 Intelligent Content Recommendation

Crucial to maintaining the audience's privacy, our system does not provide details about the audience's emotional state (or any other personal measures) to the presenter. Rather, the system uses the multimodal sensor outputs to compute the audience's cognitive state associated with the target of their attention, and then it makes context-aware suggestions to the presenter. For example, at some point during a presentation, the system may detect that the audience feels bored and is not looking at the object that the presenter is speaking about. In this case, the system can recommend the presenter move on from their current topic and discuss the virtual content that draws the audience's gaze. In an alternate scenario, the system could detect when the audience feels engaged with some particular presentation content. The presenter's interface will suggest data visualizations that are related to the one causing excitement to give the audience another example in which they might be interested.

5 Discussion

While our system primarily focuses on supporting the physically co-located context, we are considering extending our system's capabilities to support distributed participants in the future. However, there are a number of challenges to extending our approach in this way that will require different and novel design interventions.

First, our motivation for a cross-platform asymmetric system was driven by the need to support face-to-face interpersonal interaction. We handle this naturally in the co-located case by leaving the presenter's appearance unobstructed to allow the audience to see them clearly while in mixed reality. Achieving this same level of face-to-face interaction will be more difficult to achieve while users are distributed. We believe this challenge will have to be addressed by designing virtual avatars that enable a high level of telepresence and simulated eye contact. However, this means that we will need to provide an effective way of tracking the presenter's gaze and head pose in order to provide a high fidelity eye representation to help the audience maintain a high level of sense of authenticity and social co-presence [45]. Possible solutions would be to set up a controlled environment for the presenter with eye trackers or cameras, or let the presenter wear a head-worn display with eye-tracking capability. The former solution may place constraints on the presenter's mobility and flexibility in choosing space for a presentation. In a real-world scenario, there may also be various restrictions (privacy, cost, etc.) regarding a custom setup. The latter will require a symmetric system design to have both users situated in an immersive presentation space, which may lead to the question of whether the presenter will still use a tablet interface to control the presentation. In our target scenario, a hybrid interface with a tablet is highly desirable as it provides mobility along with easy input for tasks such as drawing and accurate annotation, as well as tactile feedback upon interacting with the presentation control interface. However, it is also not impossible to convert a table-top surface into a portable virtual control interface that can be anchored anywhere with a similar level of tactile feedback. How well-received such a virtual control interface versus a hybrid interface with a tablet is by the presenter during a presentation is worth investigating especially.

Second, our system leverages situated analytics by placing visualizations in context with the physical environment. This is primarily enabled by the presenter interface that lets them control where in the environment to place what content. However, if the presenter no longer shares the same physical environment as the audience, placing content situated in the audience's environment becomes much more difficult due to the diminished awareness of the audience's environment. This will become less of a challenge if we utilize virtual reality presentations. However, it is possible that even in a distributed scenario, audiences might prefer mixed-reality presentations. In order to keep our current approach of mixed-reality presentation, new interventions are needed to increase the shared awareness of audiences' physical environment. One possible approach would be to disembody presenters' physical motions from their avatars to allow them to perform the same action while being presented in different ways [34].

6 Future Work and Conclusion

Presenting financial materials in an immersive environment provides unique research opportunities in social interaction and role-based collaboration. Our research explores the design space and interaction affordances by integrating asymmetric spatial object control and audience behavior tracking into the presentation system. By providing audience feedback information to the presenter in real-time, we have the potential to enable more engaging and dynamic ways of presentation and provide more effective communication between experts and non-expert in the professional workplace. In the future, we plan to conduct user studies to validate our system design and gain more insights about real-world end users.

DISCLAIMER

This paper was prepared for information purposes with contributions from the Global Technology Applied Research center of J.P. Morgan Chase & Co. This paper is not a product of the Research Department of J.P. Morgan Chase & Co. or its affiliates. Neither J.P. Morgan Chase & Co. nor any of its affiliates make any explicit or implied representation or warranty and none of them accept any liability in connection with this paper, including, but not limited to, the completeness, accuracy, reliability of information contained herein and the potential legal, compliance, tax or accounting effects thereof. This document is not intended as investment research or investment advice, or a recommendation, offer or solicitation for the purchase or sale of any security, financial instrument, financial product or service, or to be used in any way for evaluating the merits of participating in any transaction.

REFERENCES

[1] M. B. Akçay and K. Oğuz. Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers. *Speech Communication*, 116:56–76, 2020.

[2] A. Arleo, C. Tsigkanos, C. Jia, R. A. Leite, I. Murturi, M. Klaffenböck, S. Dustdar, M. Wimmer, S. Miksch, and J. Sorger. Sabrina: Modeling and visualization of financial data over time with incremental domain knowledge. In *2019 IEEE Visualization Conference (VIS)*, pp. 51–55, 2019. doi: 10.1109/VISUAL.2019.8933598

[3] V. Biener, S. Kalamkar, N. Nouri, E. Ofek, M. Pahud, J. J. Dudley, J. Hu, P. O. Kristensson, M. Weerasinghe, K. Č. Pucihar, et al. Quantifying the effects of working in vr for one week. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3810–3820, 2022.

[4] E. Bozkir, D. Geisler, and E. Kasneci. Person independent, privacy preserving, and real time assessment of cognitive load using eye tracking in a virtual reality setup. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1834–1837. IEEE, 2019.

[5] E. Bradner and G. Mark. Why distance matters: Effects on cooperation, persuasion and deception. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, CSCW '02, p. 226–235. Association for Computing Machinery, New York, NY, USA, 2002. doi: 10.1145/587078.587110

[6] R. Brath and P. Macmurchy. Sphere-based information visualization: Challenges and benefits. In *2012 16th International Conference on Information Visualisation*, pp. 1–6, 2012. doi: 10.1109/IV.2012.11

[7] D. Calandra, F. G. Pratticò, M. Migliorini, V. Verda, and F. Lamberti. A multi-role, multi-user, multi-technology virtual reality-based road tunnel fire simulator for training purposes. pp. 96–105, 01 2021. doi: 10.5220/0010319400960105

[8] R. Castellano, A. Rinaldi, and G. Tieri. Financial literacy and education: How technology can help involving the children. In *INTED2019 Proceedings*, 13th International Technology, Education and Development Conference, pp. 3262–3266. IATED, 11-13 March, 2019 2019. doi: 10.21125/inted.2019.0852

[9] W. Celniak and P. Augustyniak. Eye-tracking as a component of multimodal emotion recognition systems. In *International Conference on Information Technologies in Biomedicine*, pp. 66–75. Springer, 2022.

[10] Y. Cheng, Y. Yan, X. Yi, Y. Shi, and D. Lindlbauer. Semanticadapt: Optimization-based adaptation of mixed reality layouts leveraging virtual-physical semantic connections. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, p. 282–297. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474750

[11] J. a. M. Evangelista Belo, M. N. Lystbæk, A. M. Feit, K. Pfeuffer, P. Kán, A. Oulasvirta, and K. Grønbæk. Auit – the adaptive user interfaces toolkit for designing xr applications. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, UIST '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3526113.3545651

[12] P. Galambos, Á. B. Csapó, P. Zentay, I. M. Fulop, T. Haidegger, P. Z. Baranyi, and I. J. Rudas. Design, programming and orchestration of heterogeneous manufacturing systems through vr-powered remote collaboration. *Robotics and Computer-integrated Manufacturing*, 33:68–77, 2015.

[13] M. Garau, M. Slater, V. Vinayagamoorthy, A. Brogni, A. Steed, and M. A. Sasse. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, p. 529–536. Association for Computing Machinery, New York, NY, USA, 2003. doi: 10.1145/642611.642703

[14] J. G. Grandi, H. G. Debarba, and A. Maciel. Characterizing asymmetric collaborative interactions in virtual and augmented realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 127–135, 2019. doi: 10.1109/VR.2019.8798080

[15] P. Gurevich, J. Lanir, and B. Cohen. Design and implementation of teleadvisor: a projection-based augmented reality system for remote collaboration. *Computer Supported Cooperative Work (CSCW)*, 24:527 – 562, 2015.

[16] P. Gurevich, J. Lanir, B. Cohen, and R. Stone. Teleadvisor: A versatile augmented reality tool for remote assistance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, p. 619–622. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2207676.2207763

[17] S. Hong, A. El Antably, and Y. Kalay. Architectural design creativity in multi-user virtual environment: A comparative analysis between remote collaboration media. *Environment and Planning B: Urban Analytics and City Science*, 46:239980831773326, 10 2017. doi: 10.1177/2399808317733267

[18] T.-W. Hsu, M.-H. Tsai, S. V. Babu, P.-H. Hsu, H.-M. Chang, W.-C. Lin, and J.-H. Chuang. Design and initial evaluation of a vr based immersive and interactive architectural design discussion system. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 363–371, 2020. doi: 10.1109/VR46266.2020.00056

[19] J. King and J. Markant. Individual differences in selective attention and scanning dynamics influence children's learning from relevant non-targets in a visual search task. *Journal of Experimental Child Psychology*, 193:104797, 2020. doi: 10.1016/j.jecp.2019.104797

[20] S. Ko, R. Maciejewski, Y. Jang, and D. S. Ebert. Marketanalyzer: An interactive visual analytics system for analyzing competitive advantage using point of sale data. *Computer Graphics Forum*, 31, 2012.

[21] T. Kurata, N. Sakata, M. Kourogi, H. Kuzuoka, and M. Billinghurst. Remote collaboration using a shoulder-worn active camera/laser. In *Eighth International Symposium on Wearable Computers*, vol. 1, pp. 62–69, 2004. doi: 10.1109/ISWC.2004.37

[22] S. Lackmann, P.-M. Léger, P. Charland, C. Aubé, and J. Talbot. The influence of video format on engagement and performance in online learning. *Brain Sciences*, 11, 2021.

[23] J. Lee, M. Kim, and J. Kim. Rolevr: Multi-experience in immersive virtual reality between co-located hmd and non-hmd users. *Multimedia Tools Appl.*, 79(1–2):979–1005, jan 2020. doi: 10.1007/s11042-019-08220-w

[24] J. Liao, A. Karim, S. S. Jadon, R. H. Kazi, and R. Suzuki. Realitytalk: Real-time speech-driven augmented presentation for ar live storytelling. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, UIST '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3526113.3545702

[25] A. Lugmayr, Y. J. Lim, J. Hollick, J. Khuu, and F. Chan. Financial data visualization in 3d on immersive virtual reality displays. In N. Mehandjiev and B. Saadouni, eds., *Enterprise Applications, Markets and Services in the Finance Industry*, pp. 118–130. Springer International Publishing, Cham, 2019.

[26] A. Luxenburger, A. Prange, M. M. Moniri, and D. Sonntag. Medicalvr: towards medical remote collaboration using virtual reality. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, 2016.

[27] G. McAtamney and C. Parker. An examination of the effects of a wearable display on informal face-to-face communication. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 45–54, 2006.

[28] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: a class of displays on the reality-virtuality continuum. In

H. Das, ed., *Telemanipulator and Telepresence Technologies*, vol. 2351, pp. 282 – 292. International Society for Optics and Photonics, SPIE, 1995. doi: 10.1117/12.197321

[29] B. Mirel, P. Goldsmith, R. Brath, and B. Cort. Visual analytics for model-based policy analysis: Exploring rapid changes in commodities markets. In *Proceedings of the 8th Annual International Conference on Digital Government Research: Bridging Disciplines & Domains*, dg.o '07, p. 312–313. Digital Government Society of North America, 2007.

[30] T. Monahan, G. McArdle, and M. Bertolotto. Virtual reality for collaborative e-learning. *Computers & Education*, 50(4):1339–1353, 2008. doi: 10.1016/j.compedu.2006.12.008

[31] N. Numan and A. Steed. Exploring user behaviour in asymmetric collaborative mixed reality. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, VRST '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3562939.3565630

[32] K. Perlin, Z. He, and K. Rosenberg. Chalktalk : A visualization and communication language – as a tool in the domain of computer science education, 2018.

[33] D. N. E. Phon, M. B. Ali, and N. D. A. Halim. Collaborative augmented reality in education: A review. In *2014 International Conference on Teaching and Learning in Computing and Engineering*, pp. 78–83, 2014. doi: 10.1109/LaTiCE.2014.23

[34] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173620

[35] T. Piumsomboon, G. A. Lee, A. Irlitti, B. Ens, B. H. Thomas, and M. Billinghurst. On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–17. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300458

[36] D. Roth, J.-L. Lugrin, D. Galakhov, A. Hofmann, G. Bente, M. E. Latoschik, and A. Fuhrmann. Avatar realism and social interaction quality in virtual reality. In *2016 IEEE Virtual Reality (VR)*, pp. 277–278, 2016. doi: 10.1109/VR.2016.7504761

[37] A. Samanta and T. Guha. On the role of head motion in affective expression. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2886–2890. IEEE, 2017.

[38] N. Saquib, R. H. Kazi, L.-Y. Wei, and W. Li. Interactive body-driven graphics for augmented video performance. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–12. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300852

[39] A. Schäfer, G. Reis, and D. Stricker. A survey on synchronous augmented, virtual, andmixed reality remote collaboration systems. *ACM Comput. Surv.*, 55(6), dec 2022. doi: 10.1145/3533376

[40] K. Schroeder, B. Ajdadilish, A. P. Henkel, and A. Calero Valdez. Evaluation of a financial portfolio visualization using computer displays and mixed reality devices with domain experts. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–9. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376556

[41] Y. Shen, S. Ong, and A. Nee. Augmented reality for collaborative product design and development. *Design Studies*, 31(2):118–145, 2010. doi: 10.1016/j.destud.2009.11.001

[42] V. K. Singh, A. Mani, and A. Pentland. Social persuasion in online and physical networks. *Proceedings of the IEEE*, 102(12):1903–1910, 2014. doi: 10.1109/JPROC.2014.2363986

[43] R. Skarbez, M. Smith, and M. C. Whitton. Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2, 2021. doi: 10.3389/frvir.2021.647997

[44] M. Slater, R. Lotto, M. Arnold, and M. Sanchez-Vives. How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicologia*, 40, 09 2009.

[45] W. Steptoe and A. Steed. High-fidelity avatar eye-representation. In *2008 IEEE Virtual Reality Conference*, pp. 111–114. IEEE, 2008.

[46] T. Tøth. Trust in client-vendor relations: An empirical study of collaboration across national and organizational boundaries. In *Proceedings of the 5th ACM International Conference on Collaboration across Boundaries: Culture, Distance & Technology*, CABS '14, p. 5–14. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2631488.2631491

[47] M. Viola. Seeing through the shades of situated affectivity. sunglasses as a socio-affective artifact. *Philosophical Psychology*, pp. 1–25, 2022.

[48] P. Wang, S. Zhang, X. Bai, M. Billinghurst, W. He, M. Sun, Y. Chen, H. Lv, and H. Ji. 2.5dhands: a gesture-based mr remote collaborative platform. *The International Journal of Advanced Manufacturing Technology*, 102, 06 2019. doi: 10.1007/s00170-018-03237-1

[49] X. Wang, L. Lin, M. Han, and J. M. Spector. Impacts of cues on learning: Using eye-tracking technologies to examine the functions and designs of added cues in short instructional videos. *Computers in Human Behavior*, 107:106279, 2020. doi: 10.1016/j.chb.2020.106279

[50] J. M. Wilson, S. G. Straus, and B. McEvily. All in due time: The development of trust in computer-mediated and face-to-face teams. *Organizational Behavior and Human Decision Processes*, 99(1):16–33, 2006. doi: 10.1016/j.obhdp.2005.08.001