

## Comparing the Effect of Audio and Visual Notifications on Workspace Awareness using Head-Mounted Displays for Remote Collaboration in Augmented Reality

Cidota, Marina; Lukosch, Stephan; Datcu, Dragos; Lukosch, Heide

**DOI**

[10.1007/s41133-016-0003-x](https://doi.org/10.1007/s41133-016-0003-x)

**Publication date**

2016

**Document Version**

Final published version

**Published in**

Augmented Human Research

**Citation (APA)**

Cidota, M., Lukosch, S., Datcu, D., & Lukosch, H. (2016). Comparing the Effect of Audio and Visual Notifications on Workspace Awareness using Head-Mounted Displays for Remote Collaboration in Augmented Reality. *Augmented Human Research*, 1(1). <https://doi.org/10.1007/s41133-016-0003-x>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

# Comparing the Effect of Audio and Visual Notifications on Workspace Awareness Using Head-Mounted Displays for Remote Collaboration in Augmented Reality

Marina Cidota<sup>1</sup> · Stephan Lukosch<sup>1</sup> · Dragos Datcu<sup>1</sup> · Heide Lukosch<sup>1</sup>

Received: 2 July 2016 / Accepted: 14 September 2016 / Published online: 10 October 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** In many fields of activity, working in teams is necessary for completing tasks in a proper manner and often requires visual context-related information to be exchanged between team members. In such a collaborative environment, awareness of other people's activity is an important feature of shared-workspace collaboration. We have developed an augmented reality framework for virtual colocation that supports visual communication between two people who are in different physical locations. We address these people as the remote user, who uses a laptop and the local user, who wears a head-mounted display with an RGB camera. The remote user can assist the local user in solving a spatial problem, by providing instructions in form of virtual objects in the view of the local user. For annotating the shared workspace, we use the state-of-the-art algorithm for localization and mapping without markers that provides “anchors” in the 3D space for placing virtual content. In this paper, we report on a user study that explores on how automatic audio and visual notifications about the remote user's activities affect the local user's workspace awareness. We used an existing game to research virtual colocation, addressing a spatial challenge on increasing levels of task complexity. The results of the

user study show that participants clearly preferred visual notifications over audio or no notifications, no matter the level of the difficulty of the task.

**Keywords** Workspace awareness · Augmented reality · Virtual colocation · Optical see-through HMD

## Introduction

Awareness refers to actors' taking heed of the context of their joint effort [40]. Even though it seems to be more a question of observing and showing certain modalities of action, information sharing is crucial to develop awareness, as it allows teams to manage the process of collaborative working, and to coordinate group or team activities [15]. Awareness information therefore plays a mediating role for collaboration and creating shared understanding [24]. However, several different types of awareness can be distinguished [40]: general awareness [21], collaboration awareness [32], peripheral awareness [8, 22], background awareness [11], passive awareness [15], reciprocal awareness [17], mutual awareness [7], workspace awareness [25].

Workspace awareness is defined “as the up-to-the-moment understanding of another person's interaction with the shared workspace” [25]. For workspace awareness, people need to gather information from the environment, understand it and predict what this means for the future. Shared visual spaces provide situational awareness [16] and facilitate conversational grounding [18, 19]. In collaborative environments, visual information about team members and objects of shared interest can support successful collaboration and enables greater situational awareness [23]. Situational awareness is thus crucial for

---

✉ Marina Cidota  
m.a.cidota@tudelft.nl

Stephan Lukosch  
s.g.lukosch@tudelft.nl

Dragos Datcu  
d.datcu@tudelft.nl

Heide Lukosch  
h.k.lukosch@tudelft.nl

<sup>1</sup> Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

fluid, natural and successful collaboration to adjust, align and integrate personal activities to the activities of other distributed actors [25]. Designers of a collaborative system need to take many different aspects into account in order to support awareness, although this is often not a primary goal in developing a system of this type [25]. This means that generally the major goal is not just to provide and maintain awareness, but to complete certain tasks in the environment.

In many domains, a quick and adequate exchange of visual context-related information to establish a common ground is necessary in order to make proper decisions and to avoid costly mistakes that cannot be easily undone. Augmented reality (AR) systems allow users to see the real world, with virtual objects superimposed upon, or composited with the real world [5, 6] where virtual objects are computer graphic objects that exist in essence or effect, but not formally or actually [36]. AR systems are not limited to the use of HMDs and mainly have to combine real and virtual objects as previously described, be interactive in real-time and register objects within 3D [5]. AR systems can be used to establish a common ground during cross-organizational collaboration in dynamic tasks [37]. AR systems have also been used to increase social presence in video-based communication [14] or to help in complex assembly tasks [30]. They can further be used to establish the experience of being virtually colocated. Virtual collocation entails that people are virtually present at any place of the world and interact with others that are physically present in another location by using AR techniques. Example of such virtual collocation can, e.g., be found in the field of crime scene investigation [39], inflight maintenance [13] or information exchange in the security domain work [34]. Such new approaches create new collaborative experiences and allow distributed users to collaborate on spatial tasks, create a shared understanding and establish a common ground.

In previous work, we have developed AR systems for virtual collocation [34, 39]. A local investigator wears an head-mounted display (HMD) with an integrated camera. By streaming the video captured from the camera, a remote colleague can see what the local investigator is seeing and both can interact in AR. Usability studies of our AR system with employees from different security organizations and reenacted scenarios show that such AR systems are suitable for information exchange, as well as distributed situational awareness and collaboration of teams in the security domain. The usability studies, however, also revealed some issues. Local users had a limited workspace awareness and experienced discomfort when virtual content added by the remote expert unexpectedly popped up in their view without any prior notification. This problem appeared only on the local user's side. Remote users reported no problems

with understanding the local users' activities within the shared virtual space. In order to address the limited workspace awareness of the local users, we implemented an automatic audio/visual notification mechanism to inform local user of remote user's activities.

In this article,<sup>1</sup> we evaluate and compare three conditions (i.e., no, audio and visual notifications) to improve workspace awareness for the local user while performing collaborative tasks in AR. We also investigate how much this additional information influences the local user's personal focus in performing the tasks. The paper is organized as follows: "Related Work" section presents the related work on AR systems that support collaboration between a remote expert and a local worker. It also covers literature on workspace awareness in different types of collaborative systems. "User Study" section presents the user study, with details about the task design, questionnaires and the system architecture of our AR environment. "Results" section discusses the results and relevant findings. The paper ends with a critical section in "Critical Observations" section and conclusions and future work in "Conclusion and Future Work" section.

## Related Work

### Remote Collaboration

There are many recent examples from different domains in which AR systems successfully enable collaboration between a local user and a remote user.

Kurata et al. [31] built the Wearable Active Camera/Laser (WACL) system that allows a remote expert to observe and point at targets in the real workplace around the field-worker, by controlling the WACL with a mouse click on the video image received from the local operator through a wireless network. The system enables the remote collaborators not only to independently set their viewpoints into the local user's workplace but also to point to real objects directly with a laser spot.

In a project on improving future crime scene investigation, technological foundations for users interacting with and within a virtual collocation environment were developed [39]. Here, remote expert investigators can interact with one local investigator who wears a video see-through HMD. By using simple spatial tools, the remote user can place restricted area ribbons, highlight objects of interest, or analyze a bullet trajectory. The virtual objects placed cannot be edited or moved around.

Gauglitz et al. [20] implemented a markerless tracking system that uses no prior knowledge about the scene

<sup>1</sup> This article is an extended version of previous work [12].

modeled, to enable a remote user to provide spatial annotations that are displayed to a local user with a handheld device.

Adcock et al. [2] built a prototype that allows remote collaboration between two people in which only the remote user is able to add virtual content. The environment of the local worker is captured using depth cameras and transmitted to the remote expert. The remote user makes annotations via a touch screen, creating lines, markers and symbols to assist the local user. The annotations are then displayed remotely with a data projector.

Huang et al. [30] developed a system for remote guidance that allows the expert to be immersed in the virtual 3D space of the worker's workspace. The system fuses the hands of the expert in real time, captured in 3D, with the shared view of the worker, creating a shared virtual interaction space.

Following the same line of thought but different hardware, Sodhi et al. [41] developed BeThere, a proof-of-concept system designed to explore 3D gestures and spatial input which allow remote users to perform a variety of virtual interactions in a local user's physical environment, using a self-contained mobile smart-phone with attached depth sensors.

Lukosch et al. [34] developed an AR system to support visual communication between a local and a remote user in the security field. Users can interact in the shared work space, adding virtual content using both a 2D graphical user interface (for the remote expert) and a 3D user interface with hand gestural input (for the local user who wears an optical see-through HMD).

These examples illustrate the use of AR to support collaboration between a local worker and a remote expert in various domains. Though user studies have shown that virtual colocation is possible and effective, the same studies have shown that local users feel remotely controlled by the experts, diminishing their abilities. Additionally, the experts feel that they miss something when they are not physically present at the scene [9, 39]. It remains an open research question about how users can become and stay aware of each other's activities in an AR-based collaboration environment for virtual colocation. In the following section, we present different solutions proposed for improving workspace awareness in systems that allow collaboration between members of a team.

## Workspace Awareness

The concept of workspace awareness has been researched in the field of computer-supported cooperative work to address a variety of coordination problems [14, 44]. Gutwin and Greenberg [25] distinguish three different information categories that contribute to workspace awareness:

1. *Who*: This category provides information on the presence of others, identity and authorship.
2. *What*: This category provides information on users' activities, their intentions and the affected artifacts.
3. *Where*: This category provides information on the location of activities, the gaze direction of users, the view of users and the reach of users.

As workspace awareness is maintained by different types of notifications of other people's activities during synchronous or asynchronous collaboration delivered to the users, the interruptions caused by these notifications can become a problem for the personal focus of the users. The effects of interruptions on people's activities have been thoroughly studied in literature: it has been repeatedly noted that an interruption has a disruptive effect both on user's task performance and emotional state [35].

Ardissono et al. [4] described a context dependent organization of the awareness information and presented an analysis of interruption and notification management in the collaboration environments of heterogeneous Web applications. Notifications were delivered in a graphical form as pop-up windows in the low-right corner of the screen.

Dourish and Bellotti [15] analyzed four collaborative writing systems to explore three different approaches to the critical problem of group activity awareness: the informational approach, the role-restrictive approach and the shared feedback approach. The results show the usefulness of information on progress and joint activities in increasing the collaborators' awareness. Furthermore, the results outline the tension between group awareness and personal focus.

In most collaborative systems, awareness is maintained by visual clues, which may lead to the overloading of the visual channel for conveying information. In addition, audio notifications have been studied in human-computer interaction (HCI) as an alternative to provide awareness. This has been studied using symbolic sounds (named "earcons") played to indicate particular events [10]. These are useful for systems that require few earcons. Users can quickly associate an earcon with its representation, but must still remember a mapping. However, for a large number of earcons the message can become very sophisticated.

Gutwin et al. [27] developed a granular synthesis engine that produces realistic chalk sounds for off-screen activity in a groupware. The experiments they conducted proved the effectiveness of synthesized dynamic audio that accurately reflect the type and quality of a digital action for providing information about the other's activity.

Hancock et al. [28] conducted two experiments with an interactive multitouch, multiuser tabletop display on the use of different non-speech audio notifications. First, they

used affirmative sounds to confirm user's action and negative sounds to indicate errors. Secondly, they tested two conditions: localized sound, where each user has their own speaker and coded sound, where users share one speaker, but the sounds are personalized for each user. The results show an improvement of group awareness but also reiterate the presence of tension between group awareness and individual focus as discovered by Dourish and Bellotti [15].

Compared to the related work presented above, in our current study we use automatic audio/visual notifications that are generated whenever the remote user interacts with the system. Thus, we provide information on activities performed by the remote user (the "what" information category, as identified by [25]). The user study focuses on comparing three conditions (No notifications, Audio notifications and Visual notifications) under two different workload levels of the task. We hypothesize that adding the automatic notifications affects the user's ability to focus on the task, but we consider this as part of the trade-off between improving the workspace awareness and the personal focus in a collaborative system.

## User Study

In order to explore the effect of audio and visual notifications on the collaboration process, a user study was conducted. Having in mind the tension that exists between group awareness and personal focus of the individual, our goal was to find out which communication channel is more suitable to receive notifications in our system and what is the proper information to be sent to the local user [26].

Two variables were used in our study: the type of notifications the system generates and the workload level of the task (two levels of difficulty). We included both variables in the design of the experiment as well as in the evaluation instrument of our AR framework.

## Task Design

We aimed at developing a clean research environment to study the use of AR to support workspace awareness in a spatial task. Games in general are useful as research tools, as one scenario can be repeated under the same conditions several times and thus offers the opportunity of comparable test situations. Games as research instruments do not focus on players' knowledge creation or adaption, but instead allow researchers to investigate elements, such as actors and processes, in a controlled environment [43].

Learning transfer in these games does not take place between game and the player, but between game and an outside observer [38]. The observer can intervene in the

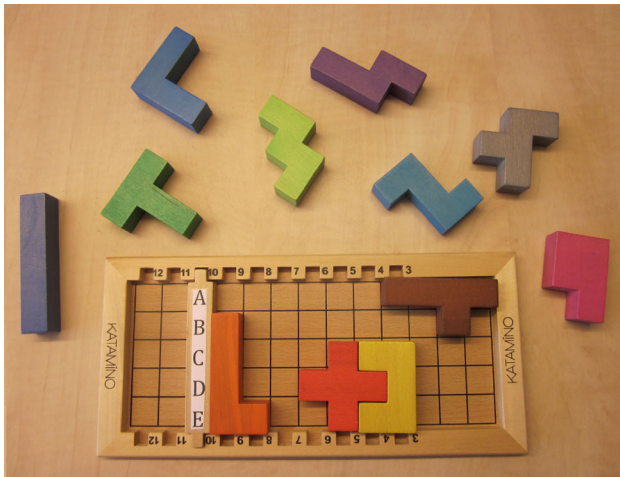
gaming process according to his or her research aims. In our case, using a game allowed us to change the level of task complexity during the game play. In general, when games are used for research, the validity, or degree of correspondence between the reference system and the simulated model thereof, is crucial [38].

Following the conceptualization of [38], a researcher starts with one or several questions about the reference system, with the handicap of being unable to collect data about it. As valid simplification of the reference system, a (simulation) game is developed or chosen, and played, and data about it are gathered. After data collection, analysis and interpretation, the researcher has to translate his or her findings back to the reference system, in order to make a difference to it. Therefore, the game should show a high level of realism, or fidelity, to make sure that the reference system is represented with all necessary roles, rules, actions, and decisions included. In our case, we found a game that provided a realistic representation of a spatial task with different levels of task complexity, allowing us to study virtual colocation.

We designed a scenario that fits into the general situation where a remote expert provides information to a local worker to accomplish a certain task. We created a user case that allowed for controlled experiments, in which the local user receives assistance in solving a 2D assembly puzzle named Katamino<sup>2</sup> (see Fig. 1). The Katamino puzzle was chosen for following reasons:

- The concept of a puzzle supports the design of scenarios in which a remote "expert" provides instructions to a local "worker".
- Katamino offers different levels of difficulty. Depending on the position of the slider on the boardgame, there are 9 levels increasing in difficulty (between 3 and 11). For example, setting level 10 (as in Fig. 1) implies that the local player has to use 10 pieces to fully cover the rectangle area on the board game defined in the right part of the slider. With such different levels of difficulty it is possible to evaluate whether the notifications affect the focus of the local "worker" under different task loads.
- For each level of difficulty, there are several different solutions, depending on the pieces chosen at the beginning of the game. For each level of difficulty, the local users were asked to solve the puzzle 3 times, each time using different sets of pieces. We consider that solving the puzzle for one set of pieces will not have an immediate effect on the ability of the player to solve it with a different set of pieces. Thus, we tried to minimize the bias related to the learning effect of repeating a task.

<sup>2</sup> <http://boardgamegeek.com/boardgame/6931/katamino>.



**Fig. 1** Katamino assembly puzzle

By superimposing different virtual objects over the pieces of the puzzle, the remote user provides instructions that lead to certain solutions chosen beforehand. Using a 2D GUI, the remote user can communicate to the local user different actions: to remove a piece that is not part of the current solution, to place a piece or to rotate it on the board game. In order for the remote player to identify certain squares on the board game using text messages, we added the letter A–E on the slider (e.g., A3 means the intersection of the row A with the column 3).

In our experimental design, we have not included audio communication in between the users, as this might have influenced the workspace awareness for the local user. Instead we chose to focus only on providing automatically generated awareness information, preventing interference with other information sources that could have affected our observations on workspace awareness.

## Participants

Twelve participants played the game as local users in the experiment, each rewarded with a 10 EUR gift card. The gift card was provided as motivation to attend the experiment. In order to ensure the same conditions for providing instructions to each participant playing the role of local user, the remote user was kept the same person for all of the 12 local users. There were 4 females and 8 males, with an age between 18 and 44 years ( $M = 26.7$ ;  $SD = 6.8$ ), all with academic connections (6 bachelor students, 1 master student, 4 Ph.D. candidates and 1 lecturer).

## Measures

Although awareness to users is an essential feature in a collaborative system, evaluation is not straightforward and

little research has been done for the assessment of the quality of the awareness provided by a system [3].

In [25], the authors identified specific elements that characterize the workspace awareness in a system. Some of them are related to the present (e.g., presence, identity, action, intention, location and gaze) and other elements are related to the past (e.g., action history, event history, presence history and location history). Antunes et al. [3] considered that there are three important issues associated with workspace awareness: tasks characterized by who, what, when and how they are accomplished; interaction that defines how the group interacts with the workspace and what information is necessary to sustain it and finally the level of task interdependence perceived by the group.

Starting from the work of [25] and [3], we created a list of questions that are relevant for our AR system and applicable in the current hardware configuration. For instance, we do not have a sensor for eye-tracking and thus cannot provide information on the gaze of the remote user to the local user. Considering the different levels of difficulty in the Katamino puzzle and the potentially disruptive effect of the provided notifications, we added the NASA-TLX questionnaire [29] to assess the task load of the local users. Since awareness is not the only goal to be achieved in designing a collaborative system, we were also interested to find out if the local users were able to focus on the task they were supposed to do and also if they were overloaded with too much information. The resulting set of questions is shown in Tables 1 and 2. All responses in questionnaire 1 were on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree).

## Equipment

We have developed a framework named DECLARE (DistributEd CoLLaborative Augmented Reality Environment) based on a centralized architecture for data communication to support virtual collocation of users.

DECLARE is a multimodal, multiuser, highly scalable parallel framework that integrates a shared memory across the running components. The data communication is ensured via a modular framework integrating a shared memory mechanism (being part of DECLARE) which decouples the data transfer in time and space. The decoupling in time mechanism of DECLARE implies that if either the local user or the remote user disconnects temporarily, the video and AR data related to the current work session is automatically transferred as an initial update at the next work session of the user. Practically the remote expert and the local users resume the activity according to their roles within the scenario without losing track of the activities occurring while they were offline.

**Table 1** Questionnaire 1—questions after each condition

---

Workspace awareness
Q1.1 I was able to focus on the activity
Q1.2 I was aware of what the remote user was doing
Q1.3 I was aware of where the remote user was working or viewing
Q1.4 I was aware of when the remote user was doing an action
Q1.5 I could follow the remote user's instructions
Q1.6 I was distracted by the automatic notifications about the remote user's activity
Q1.7 I was overloaded with too much notifications about the remote user's activity
Q1.8 I was overloaded with too much augmentation from the remote user
Q1.9 I felt that I received help from the remote user
Q1.10 I felt that I was dependent on the remote user in completing the task
Task load (TLX)
Q1.11 I felt that the task was mentally demanding
Q1.12 I felt that the task was physically demanding
Q1.13 I felt that the pace of the task was rapid
Q1.14 I felt that I had to work hard to accomplish the task
Q1.15 I felt that I was stressed by the task
Q1.16 I felt that I was successful in accomplishing the task

---

**Table 2** Questionnaire 2—questions after all conditions for each level of difficulty

---

Q2.1 Classify your preference for the 3 conditions (No notifications, Audio, Visual) (“most preferred /average preferred /least preferred”)
Q2.2 I would have preferred less automatic notifications of the remote user's actions (“yes/no”)

---

**Fig. 2** The optical see-through META SpaceGlasses

DECLARE can be adapted to different types of display devices, depending on the specific limitations of each device. For instance, a video see-through HMD has the advantage of a higher field of view, the possibility to display solid virtual objects and has a good alignment of the virtual content superimposed over the real world. However, due to the processing requirements capturing through a camera, then streaming the video images displayed in the headset is a process often leading to a higher latency. This creates a problem of getting dizzy. On the other hand, with an optical see-through HMD, we encounter a difficult problem in aligning the virtual content to the real world. For our present study, we used the optical see-through META SpaceGlasses<sup>3</sup> (see Fig. 2) which have an integrated depth

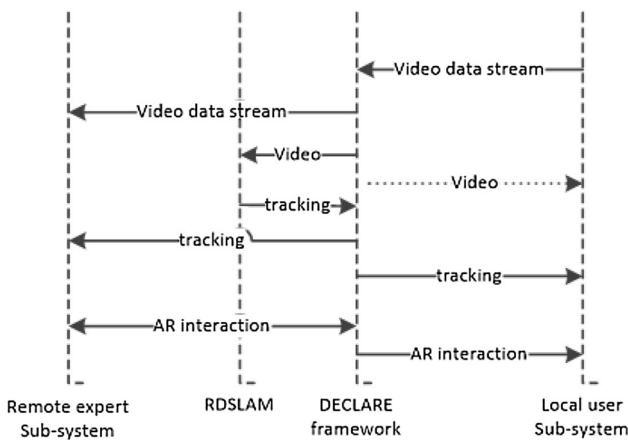
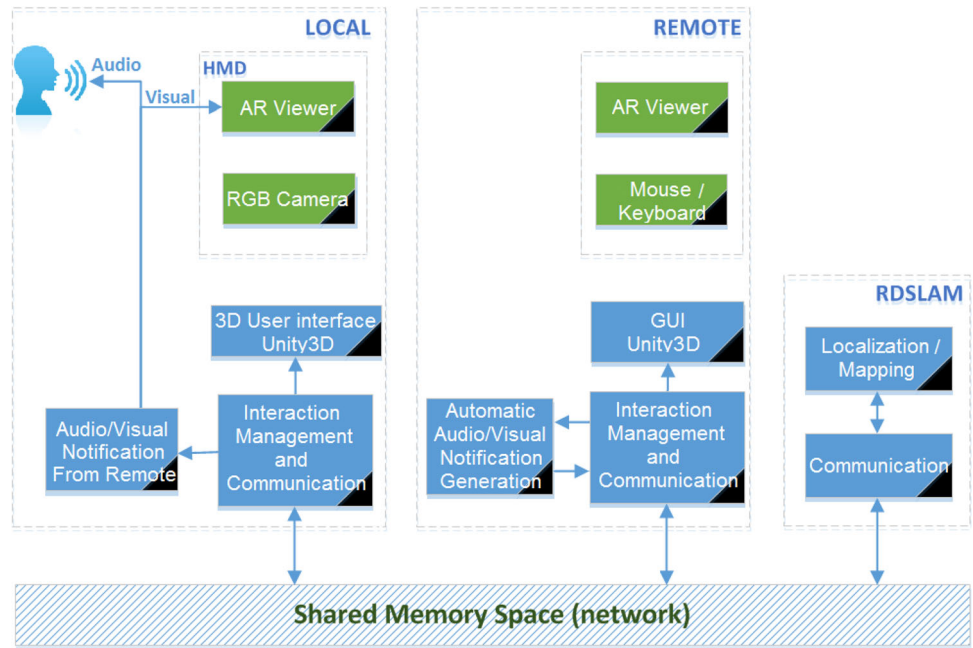
<sup>3</sup> <https://www.spaceglasses.com/>.

sensor and RGB camera. Technical specifications are: display resolution  $960 \times 540$  pixels (qHD), aspect ratio 16:9, FOV  $23^\circ$ ; RGB Camera  $1280 \times 720$  (MJPEG), 30 fps.

All components of DECLARE communicate through a shared memory space via wireless or wired connections (see Fig. 3), using a data and event notification approach (see Fig. 4). The video frames captured by the RGB camera of the headset worn by the local user are streamed in real time to the remote expert. For placing the virtual objects, we use the robust dynamic simultaneously localization and mapping (RDSLAM), a state-of-the-art markerless tracking model, with its implementation as provided by authors [42]. This module receives the video frames from the local user's HMD camera. Based on the input video frames, RDSLAM computes for each frame the parameters of the camera position and orientation together with a sparse cloud of 3D tracked points. These tracked points are used to attach the virtual content in AR. In order to make the RDSLAM ready for tracking, a prior calibration has to be done by the local user, by moving the HMD horizontally (see Fig. 5 right). This is done once, at the beginning of the play session and later during the play session, if the tracking becomes unstable.

In addition to tracking the HMD position and orientation, RDSLAM performs a mapping of the physical environment (of the local) and further generates an internal representation of the world. The result of markerless tracking relates to

**Fig. 3** Diagram of the main components of DECLARE



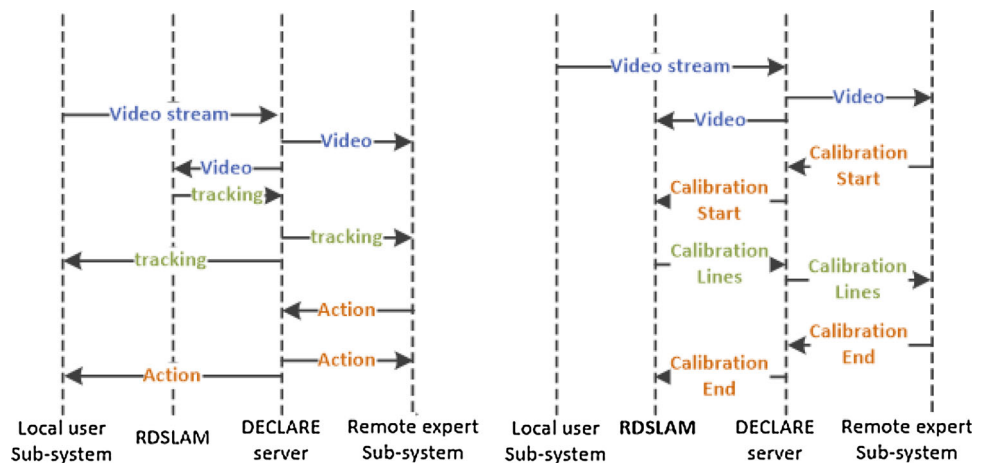
**Fig. 4** Diagram of data and event notifications for user actions across DECLARE-based modules and subsystems

information on the HMD camera location and orientation while the mapping result relates to a representation of the physical world in form of a sparse cloud of 3D points.

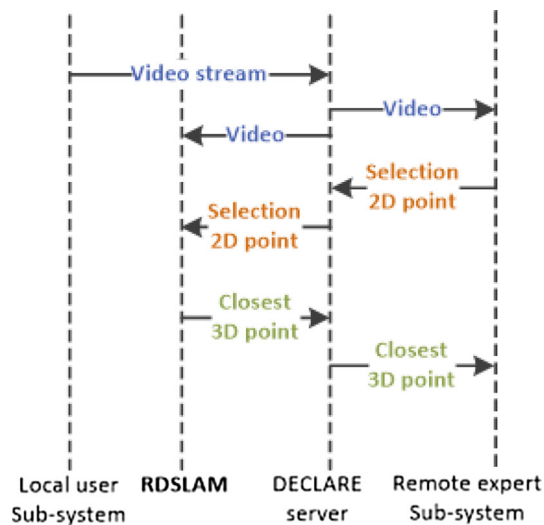
The sparse cloud of 3D points represents visual key-points which connect the augmented world to the physical world and further act as virtual anchors supporting annotation by AR markers. Once attached to a key point (see Fig. 6), a 3D-aligned object is correctly rendered in the next video frames on both conditions during run time, being consistent with the HMD camera motion (see Fig. 5 left).

In DECLARE, the network communication is implemented using both TCP/IP and UDP standard. TCP/IP is used for data transfers between the server and software modules running on hardware devices linked via network cables. UDP is used for data transfers between software modules connected via wireless links. In case of the

**Fig. 5** Diagram of events for the local user tracking at runtime (left) and for tracking calibration (right)







**Fig. 6** Diagram of data notifications for the selection event

UDP-based network communication, each frame from a video sequence is encoded as a compressed image into a UDP packet. The VGA resolution ( $640 \times 480$  pixels/frame) and JPEG compression quality of approximately 85 % proved to fit below the size limit of the UDP data-grams.

Updates are automatically sent to each software module or client subsystem (subsystems such as the software application of the local user or the software application of the remote expert), with regard to the new data available, by using a notification and push system of events and data (see Figs. 4–6).

The automatic processing of events and data notifications is only done for the software modules and subsystems that register for the specific type of data. In this way, the subsystem of the local user (see Figs. 4–6) does not receive video data from its own camera via the network. This ensures an optimal use of the network bandwidth, especially in wireless data communication.

Consistency of actions is a critical aspect implemented in DECLARE. Local updates in the graphical user interface with regard to direct user actions in the system of the local and on the remote experts system, are taken into consideration only when data and events related to the action are received as feedback from the server. This ensures availability of the updates on all the subsystems and modules. The flow of data and event notifications is illustrated in the diagram in Fig. 4. For the video stream, a synchronization mechanism is implemented in the shared memory to ensure the same video frame is played for the local, remote and the RDSLAM module at the same time.

Considering the technical specifications of DECLARE framework, we further implemented specialized functionalities for both local user and remote user, and created different user interfaces using Unity3D game engine.



**Fig. 7** Icons used for providing instructions to the local user

### *Local User Support*

The video captured by the RGB camera mounted on top of the HMD worn by the local user is sent to the other components in the system. In order to align the virtual content (see Fig. 7) with the real objects in the view of the local user, we had to clip an area of the image captured by the RGB camera of the headset. This area corresponds to what the local user sees through the display of the HMD.

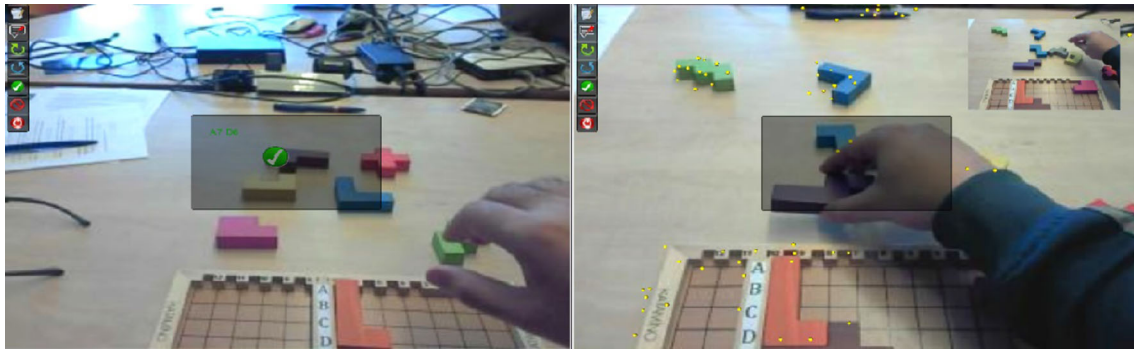
### *The Remote User*

The remote user can view on a laptop the video captured by the local user’s HMD camera and can interact with the system by using the keyboard and a standard mouse device through the means of a classic 2D GUI. In the left part of the screen (see Fig. 8), a menu with buttons that allow the remote user trigger different actions is shown. Besides text messages, the remote user can attach virtual objects as shown in Fig. 7 (from left to right: the “approval” sign indicates a piece that has to be used by the local player, the “reject” sign indicates a piece that has to be removed from a configuration and signs that indicate that a piece should be rotated by  $90^\circ$  clockwise or counterclockwise).

Regarding the positioning of the virtual content, there are two types of virtual objects: the fixed objects that stay in the same position relative to the camera (on the screen) and the virtual objects that are linked to the points tracked by RDSLAM algorithm. The remote user has the option of “freezing” the video stream by pressing the “F” key. While doing this, he/she is still connected to the local user, seeing the live streaming of the camera in the right upper part of his view (see Fig. 8 right). The full image for live streaming reappears when the remote user presses the “U” key. The transparent rectangle in the middle of the image (see Fig. 8) represents the area of the HMD used to display additional information. When objects are within this area, the remote user knows they are visible to the local user.

### *The Automatic Notifications*

In order to support workspace awareness in our AR framework, we implemented different cues for the local user, every time the remote user is interacting with the system. Each notification is automatically sent to the local user to inform them about the action taken by the remote



**Fig. 8** The view of the remote user



**Fig. 9** Icons used for visual notifications

user. These cues are presented to the local user as either audio or visual notifications:

- Audio (speech) notifications: every action taken by the remote is automatically indicated with an audio message. The audio notifications are short spoken messages generated by an online text-to-speech synthesizer,<sup>4</sup> to distinguish between different actions of the remote user. The message described the action taken (e.g., “Remote added an object,” “Remote selected an object,” “Remote deleted an object,” “Remote freezes the image,” “Remote plays video stream again”) and is played only once to the local user.
- Visual notifications: Every action taken by the remote user is indicated in the right lower corner of the local user’s view as a small icon that blinks twice and then disappears. We have chosen suggestive icons for each action of the remote user (see Fig. 9 from left to right: adding, selecting and deleting virtual objects, pausing the video stream and playing it again).

## The Experiment

At the beginning of the experiment, each local user spent about 10 min to be informed about the rules of the game; the conditions in which the game would be played; the type of augmentation that can be provided by the remote expert and also about the notifications generated by the system.

Then the local user solved the puzzle 6 times, i.e., under 3 conditions, each for 2 levels of difficulty (7 and 10). The

exact same 6 configurations of the puzzle chosen beforehand were solved by all of the 16 participants, who were asked to find each solution in approximately 5–7 min.

Each participant needed between 75 and 90 min to do the experiment. Solving the puzzle in all conditions took about 40 min. The rest of the time was spent on filling in the questionnaires and on a debriefing session at the end to collect additional observations and suggestions that were not included in the responses of the questionnaires.

The local user wore an optical see-through HMD and was sitting in front of a table on which the pieces of the puzzle were spread (see Fig. 10). The remote user was sitting in a separate room, in front of a laptop, having written on a paper the solutions for all the 6 configurations of the puzzle, found by using a Katamino solver application.<sup>5</sup> The instructions were provided using only visual communication between users.

Each round started by removing the pieces that are not needed to build a certain solution (which is known only by the remote user). In this stage, the remote player marked with the virtual label “reject” from Fig. 7 each piece that the local user has to put aside (5 pieces to be removed for level 7 and 2 pieces for level 10). After that, the local user received information on the correct position of 2 pieces for level 7 (and 3 pieces for level 10). A piece was indicated using the virtual label “approval” in Fig. 7. The position on the board game was given as a text message with squares that should be covered by that piece (e.g., “A1 B3”). After that, the local user was supposed to solve the puzzle alone. Considering that the solution had to be found in 5–7 min, instructions continued to be provided as long as the remote user considered this to be necessary.

After each condition, the local users were asked to answer the questions in Questionnaire 1. After performing under all conditions, they filled in twice Questionnaire 2, once for each level of difficulty.

<sup>4</sup> <http://www.fromtexttospeech.com/>.

<sup>5</sup> [http://www.edwin.nl/katamino/katamino\\_solver.html](http://www.edwin.nl/katamino/katamino_solver.html).



**Fig. 10** Pictures taken during the experiments (*left*) the remote user and (*right*) the local user

**Results**

**Results from the Questionnaires**

We analyzed the Likert scale responses (Q1.1–Q1.16) based on two factors namely the game level, i.e., *L1*: level 7 and *L2*: level 10, and the conditions: No, Audio, Visual.

In order to see whether the participants perceived differently the two games levels *L1* and *L2* in terms of task load, we run the parametric paired sample *t* test for the pairs (*L1*, *L2*) under each condition, for the NASA-TLX questionnaire (Q1.11–Q1.16). The values in Table 3 indicate that for most of the comparisons, the null hypothesis was not rejected at a significance level  $\alpha = .05$  (bold values are the *p*-values lower than  $\alpha$ ). That means that in these situations there are no statistical differences between the two levels of the game *L1* and *L2*. But for Q1.11, the paired *t* test rejected the null hypothesis for no and audio conditions, while for the video condition the p-value is above but very close to the threshold value 0.05. Comparing the mean values in Table 4, we conclude that the users perceived the task in game level *L2* as being more mentally demanding than the task in game level *L1*.

The Likert scale ratings between the conditions were checked using the Friedman test ( $\alpha = .05$ ). For the situations for which the null hypothesis was rejected (i.e., there were statistical differences between the three conditions), a

**Table 3** The *p* values for the paired sample *t* test for the pairs (*L1*, *L2*) under each condition

	Q1.11	Q1.12	Q1.13	Q1.14	Q1.15	Q1.16
No	<b>0.027</b>	1	0.137	0.180	<b>0.034</b>	0.866
Audio	<b>0.002</b>	0.275	0.674	0.236	0.368	0.305
Video	0.067	1	0.136	<b>0.034</b>	0.166	0.674

**Table 4** Mean/SD values for Q1.11 under each level (*L1*, *L2*) and under each condition

	<i>L1</i>	<i>L2</i>
No	2.917/0.793	3.5/0.904
Audio	3.083/0.9	3.917/0.669
Video	3/0.953	3.583/0.9

pair-wise Wilcoxon signed-rank test was applied. All these results are obtained using MATLAB toolbox for statistics.

The values in Table 5 resulted by applying Friedman test. For game level *L1*, the questions with Likert ratings significantly different between conditions are (Q1.4, Q1.5,

**Table 5** Friedman test ( $\alpha = 0.05$ )

	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$
Q1.1	0.422	1.727	0.387	1.9
Q1.2	0.589	1.059	0.163	3.63
Q1.3	0.594	1.04	0.562	1.152
Q1.4	<b>0.044</b>	<b>6.258</b>	<b>0.002</b>	<b>12.235</b>
Q1.5	<b>0.039</b>	<b>6.5</b>	0.417	1.75
Q1.6	<b>0.026</b>	<b>7.294</b>	<b>0.007</b>	<b>9.929</b>
Q1.7	<b>0.003</b>	<b>11.4</b>	<b>0.001</b>	<b>13.613</b>
Q1.8	0.325	2.25	1	0
Q1.9	0.296	2.437	0.33	2.214
Q1.10	<b>0.005</b>	<b>10.64</b>	0.468	1.52
Q1.11	0.651	0.857	0.552	1.188
Q1.12	0.606	1	0.472	1.5
Q1.13	0.368	2	0.834	0.364
Q1.14	0.852	0.32	0.882	0.25
Q1.15	0.956	0.091	0.191	3.308
Q1.16	<b>0.017</b>	<b>8.176</b>	<b>0.024</b>	<b>7.448</b>

The left (*p*,  $\chi^2$ ) values relates to game level *L1*, the right (*p*,  $\chi^2$ ) values relates to game level *L2*

Q1.6, Q1.7, Q1.10 and Q1.16) and for game level L2, (Q1.4, Q1.6, Q1.7 and Q1.16).

Further on we performed pair-wise comparisons between conditions. The results are displayed in Figs. 11 and 12 and can be summarized as following:

*Audio Versus No*

Audio helped the local users to be more aware of when the remote users were doing an action (Q1.4) in L1 ( $Z = 2.41, p = 0.016, Md = 4.5$  vs. 4) and L2 ( $Z = 2.69, p = 0.007, Md = 4$  vs. 3); Audio resulted to distract more the local users (Q1.6) in both L1 ( $Z = 2.44, p = 0.014, Md = 3.5$  vs. 1) and L2 ( $Z = 2.46, p = 0.014, Md = 2$  vs. 1); Audio determined the local users to be more overloaded with notifications about the remote user’s activity (Q1.7) in both L1 ( $Z = 2.74, p = 0.006, Md = 4$  vs. 1) and in L2 ( $Z = 2.69, p = 0.007, Md = 2.5$  vs. 1) and Audio made local users feel less successful in accomplishing the task in L2 (Q1.16:  $Z = -2.16, p = 0.031, Md = 2.5$  vs. 4).

*Visual Versus Audio*

Local users could follow the remote user’s instructions (Q1.5) better in Visual (L1:  $Z = 2.33, p = 0.019, Md = 4.5$

vs. 4); the Visual also caused less distraction with information about the remote user’s activity (Q1.6) (L1:  $Z = -2.28, p = 0.023, Md = 2$  vs. 3.5) and Visual caused less information overload to the local user about the remote user’s activity (Q1.7), in both L1 and L2 (L1:  $Z = -1.98, p = 0.048, Md = 2$  vs. 4; L2:  $Z = -2.14, p = 0.033, Md = 2$  vs. 2.5).

*Visual Versus No*

Visual helped local users to be more aware of when the remote user was doing an action (Q1.4) (L2:  $Z = 2.36, p = 0.018, Md = 4$  vs. 3); visual notifications caused more distractions (Q1.6) (L2:  $Z = 2.27, p = 0.023, Md = 2$  vs. 1) and were more overloading on the local user (Q1.7) in both L1 ( $Z = 1.99, p = 0.046, Md = 2$  vs. 1) and L2 ( $Z = 2.33, p = 0.019, Md = 2$  vs. 1); with visual notifications, local users felt less dependent on the remote user in completing the tasks (Q1.10) (L1:  $Z = -2.71, p = 0.007, Md = 3$  vs. 4); visual notifications made local users feel more successful in accomplishing the tasks (Q1.16) in both L1 ( $Z = 2.35, p = 0.019, Md = 5$  vs. 3) and L2 ( $Z = 2.23, p = 0.026, Md = 4$  vs. 2.5).

As a conclusion, the analysis of the Likert responses shows strong statistical evidence that visual notifications cause less overload and less distraction than Audio. Visual and Audio notifications are more successful for making the local user aware of the remote users’ activities and for accomplishing tasks compared to No notifications. Alongside the statistics performed above, Fig. 13 comes to strengthen the position of the visual notification as the most preferred condition for providing workspace awareness.

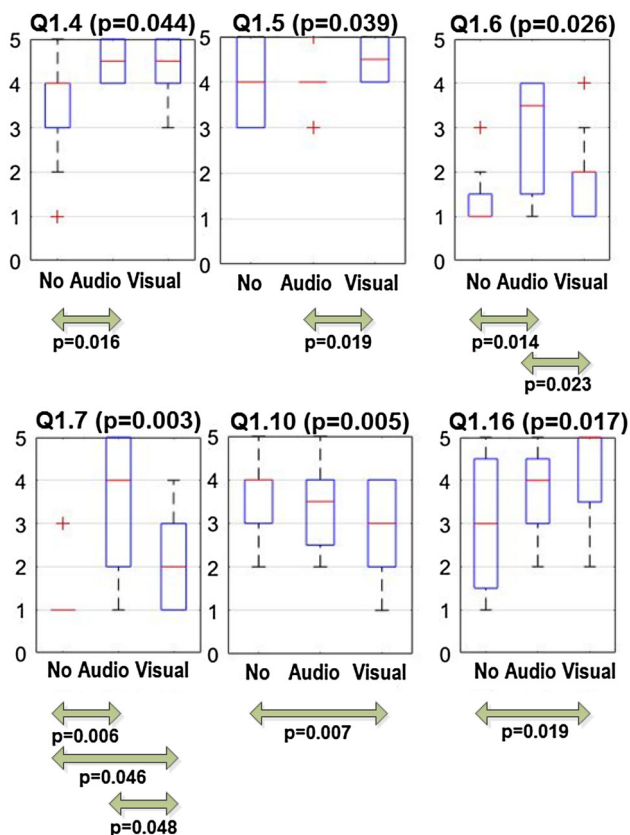
**Results from Discussions**

From the debriefings, we received positive feedback on the overall performance of our AR environment, but we also got some interesting suggestions for future developments.

Most of the participants considered that the visual notifications, as currently implemented, are the best way to be informed about remote users actions, compared to audio or no notifications.

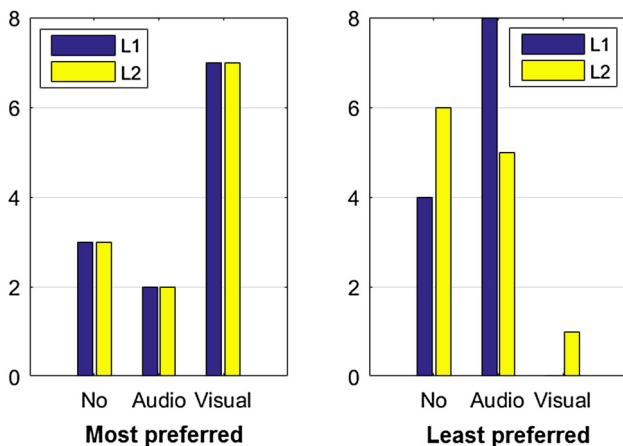
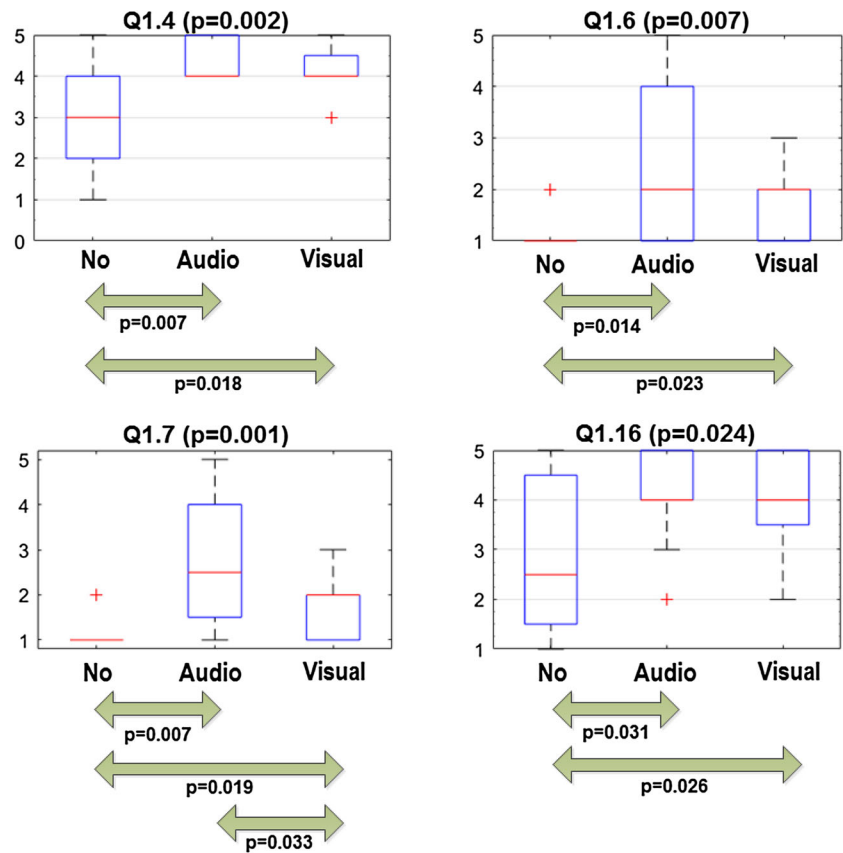
Some participants mentioned that certain notifications (like those for selecting and deleting an object or freezing/playing live streaming) were not useful for them in completing the task. An interesting idea we received was to implement a customizable notification system, in which the local user could choose the actions to be notified on.

In some cases, the participants complained about the auditory clutter caused by overlapping of the audio notifications. This happened when the remote user performed many actions in a short amount of time and an audio notification was generated before the previous one was



**Fig. 11** Results of Wilcoxon signed-rank test for game level L1

**Fig. 12** Results of Wilcoxon signed-rank test for game level L2



**Fig. 13** Most preferred (left) and least preferred (right) condition for game levels (L1, L2), according to the subjective responses given to Q2.1

completed. But in the same time, they admitted that speech notifications are necessary if they had to make distinction between many actions of the remote user. For maximum two or three actions, a non-speech audio would be preferred instead of spoken words.

Another suggestion was that the visual notification for freezing the image should be displayed all the time until

the remote user plays live streaming again. Three of the participants said that in a small workspace no notifications are needed, but in the same time they admitted the benefit of notifications in a bigger environment. One participant said that it would be very helpful to have visual notifications that indicate the position where an object was added (e.g., using arrows for 4–8 directions).

Another idea was to use audio only in few cases (considered more important), for the rest of the notifications visual being a better option, or to have visual notifications for all the remote users action but for important action (that can be defined in a priority list) these visual icons should be accompanied by a non-speech audio signal.

### Critical Observations

#### Limitations of the User Study

Our user study was conducted in a “controlled” environment, with no serious consequences for participants who did not manage to successfully complete the tasks. Although we consider that our findings are valid in many real life situations, there is an open question how difficult circumstances may affect a person’s experience when using our system.

## Limitations of the Current Setup

In our experiment, the local user was sitting in front of a table and had to solve a 2D assemble puzzle, a task that required a short range distance alignment between the virtual and the real world. As the real objects in focus are close to the local user's eyes and to the RGB camera, the movements of the head induce a great variation of the position of the area in the RGB image that matches the display area in the HMD view of the local user. Through many different empirical trials in which we altered the position and the dimensions of this area, we managed to choose a fixed area that worked well for the majority of the participants in our experiment (see the transparent rectangle in the middle of the image in Fig. 8). However, the chosen fixed area did not support high accuracy for the alignment of virtual and real objects. For instance, we were not able to precisely place virtual annotations for real objects that were small and close to one another (e.g., keys on the keyboard of a laptop or squares on the board game in our experiment). This limitation led us to text messages instead of visual icons to indicate squares on the board game.

## Conclusion and Future Work

A quick and adequate exchange of visual context-related information to establish a common ground is necessary in order to make proper decisions and to avoid costly mistakes that cannot be easily undone. AR systems have successfully be used to establish such a common ground via virtual colocation. User studies showed that the workspace awareness of a local user needs to be improved during virtual colocation. For that purpose, we explored in this article on how to increase the workspace awareness of a local user who is connected to a remote user. The remote user provides instructions to solve puzzle tasks by using an AR system for virtual colocation. We implemented automatic audio as well visual notifications that are generated whenever the remote user interacts with the system. Each notification is sent to the local user to inform them on the action that has just been taken by the remote user.

We reported on a user study to explore the impact of audio and visual notifications about the remote user's actions on the workspace awareness of the local user. We used a game as research instrument for the experiment in order to set up a valid, repeatable and observable experiment. Although a well-grounded method in game research is still lacking [1], and requirements for games as research instruments are not very well defined yet [33], the game we used allowed us to study virtual colocation in a spatial task. In future research, we would also use different types of

games for being able to better understand and define the role of distinct game elements in the research process.

The results of our study show that the local users prefer visual notifications over audio and no notifications. It is interesting to see that the visual notifications cause less overload for the test persons. This could be explained by the fact that the task in the game we used requires visual attention already, and an audio signal would mean that participants would have to divide attention between two cues (audio and visual), instead of staying focused on only one (visual). For AR systems this would mean that limiting cues to one mode would mean a benefit for the user. Future research should contribute to a consolidation of this aspect by including further experiments with different modes of notifications. Also, we consider extending the area of awareness for the local user by adding the "where" category as described by [25]. This means providing information on the location of remote user's activities. For that purpose, we will make use of an inertial measurement unit (IMU) mounted the HMD of the local user to be able to determine the current position of the local user relative to the previous position when the remote user "froze" the image in order to perform an activity.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Aarseth E (2003) Playing research: methodological approaches to game analysis. In: Proceedings of the digital arts and culture conference, pp 28–29
2. Adcock M, Anderson S, Thomas BH (2013) Remotefusion: real time depth camera fusion for remote collaboration on physical tasks. In: 12th ACM international conference on virtual reality continuum and its applications in industry, VRCAI 2013, Hong Kong, 17–19 November 2013, pp 235–242. doi:10.1145/2534329.2534331
3. Antunes P, Herskovic V, Ochoa SF, Pino JA (2014) Reviewing the quality of awareness support in collaborative applications. *J Syst Softw* 89:146–169. doi:10.1016/j.jss.2013.11.1078
4. Ardissono L, Bosio G, Goy A, Petrone G, Segnan M (2009) Managing context-dependent workspace awareness in an e-collaboration environment. In: Proceedings of the 2009 IEEE/WIC/ACM international conference on web intelligence and international conference on intelligent agent technology—workshops, Milan, Italy, 15–18 September 2009, pp 42–45. doi:10.1109/WI-IAT.2009.226
5. Azuma RT (1997) A survey of augmented reality. *Presence Teleoper Virtual Environ* 6:355–385
6. Azuma RT, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B (2001) Recent advances in augmented reality. *Comput Graph Appl IEEE* 21(6):34–47. doi:10.1109/38.963459

7. Benford S, Bowers J, Fahlén LE, Greenhalgh C (1994) Managing mutual awareness in collaborative virtual environments. In: Proceedings of the conference on virtual reality software and technology. World Scientific Publishing Co., Inc., River Edge, VRST '94, pp 223–236
8. Benford S, Greenhalgh C, Rodden T, Pycocock J (2001) Collaborative virtual environments. *Commun ACM* 44(7):79–85. doi:[10.1145/379300.379322](https://doi.org/10.1145/379300.379322)
9. Billinghurst M, Thomas B (2011) Mobile collaborative augmented reality. In: Alem L, Huang W (eds) Recent trends of mobile collaborative augmented reality systems. Springer, Berlin, pp 1–19
10. Blattner M, Sumikawa DA, Greenberg R (1989) Earcons and icons: their structure and common design principles. *Hum Comput Interact* 4(1):11–44. doi:[10.1207/s15327051hci0401\\_1](https://doi.org/10.1207/s15327051hci0401_1)
11. Bly SA, Harrison SR, Irwin S (1993) Media spaces: bringing people together in a video, audio, and computing environment. *Commun ACM* 36(1):28–46. doi:[10.1145/151233.151235](https://doi.org/10.1145/151233.151235)
12. Cidota M, Lukosch S, Datcu D, Lukosch H (2016) Workspace awareness in collaborative AR using HMDS: a user study comparing audio and visual notifications. In: Proceedings of the 7th augmented human international conference. ACM Press, New York
13. Datcu D, Cidota M, Lukosch S, Oliveira DM, Wolff M (2014) Virtual co-location to support remote assistance for inflight maintenance in ground training for space missions. In: Proceedings of the 15th international conference on computer systems and technologies, CompSysTech '14. ACM Press, New York, pp 134–141. doi:[10.1145/2659532.2659647](https://doi.org/10.1145/2659532.2659647)
14. de Souza Almeida I, Oikawa MA, Carres JP, Miyazaki J, Kato H, Billinghurst M (2012) Ar-based video-mediated communication: a social presence enhancing experience. In: 14th symposium on virtual and augmented reality, SVR 2012. IEEE Computer Society, pp 125–130
15. Dourish P, Bellotti V (1992) Awareness and coordination in shared workspaces. In: Proceedings of the ACM 1992 conference on computer supported cooperative work, Canada, Toronto, pp 107–114
16. Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. *Hum Factors J Hum Factors Ergon Soc* 37:32–64
17. Fish RS, Kraut RE, Chalfonte BL (1990) The videowindow system in informal communication. In: Proceedings of the 1990 ACM conference on computer-supported cooperative work, CSCW '90. ACM Press, New York, pp 1–11. doi:[10.1145/99332.99335](https://doi.org/10.1145/99332.99335)
18. Fussell SR, Kraut RE, Siegel J (2000) Coordination of communication: effects of shared visual context on collaborative work. In: Proceedings of the 2000 ACM conference on computer supported cooperative work, CSCW '00. ACM Press, New York, pp 21–30. doi:[10.1145/358916.358947](https://doi.org/10.1145/358916.358947)
19. Fussell SR, Setlock LD, Kraut RE (2003) Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '03. ACM Press, New York, pp 513–520. doi:[10.1145/642611.642701](https://doi.org/10.1145/642611.642701)
20. Gauglitz S, Lee C, Turk M, Höllerer T (2012) Integrating the physical environment into mobile remote collaboration. In: Mobile HCI '12, Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services, San Francisco, CA, USA, 21–24 September 2012, pp 241–250. doi:[10.1145/2371574.2371610](https://doi.org/10.1145/2371574.2371610)
21. Gaver WW (1991) Sound support for collaboration. In: Proceedings of the second conference on European conference on computer-supported cooperative work, ECSCW'91. Kluwer Academic Publishers, Norwell, pp 293–308
22. Gaver WW (1992) The affordances of media spaces for collaboration. In: Proceedings of the 1992 ACM conference on computer-supported cooperative work, CSCW '92. ACM Press, New York, pp 17–24. doi:[10.1145/143457.371596](https://doi.org/10.1145/143457.371596)
23. Gergle D, Kraut RE, Fussell SR (2013) Using visual information for grounding and awareness in collaborative tasks. *Hum Comput Interact* 28(1):1–39. doi:[10.1080/07370024.2012.678246](https://doi.org/10.1080/07370024.2012.678246)
24. Gerosa M, Fuks H, Raposo A, Lucena C (2004) Awareness support in the AulaNet learning environment. In: Proceedings of the IASTED international conference on web-based education—WBE 2004. ACTA Press, Innsbruck, pp 490–495
25. Gutwin C, Greenberg S (2002) A descriptive framework of workspace awareness for real-time groupware. *Comput Support Coop Work* 11(3):411–446. doi:[10.1023/A:1021271517844](https://doi.org/10.1023/A:1021271517844)
26. Gutwin C, Greenberg S, Roseman M (1996) Workspace awareness in real-time distributed groupware: framework, widgets, and evaluation. In: Sasse M, Cunningham R, Winder R (eds) People and computers XI (Proceedings of the HCI'96). Springer-Verlag, Imperial College, London, pp 281–298
27. Gutwin C, Schneider OS, Xiao R, Brewster SA (2011) Chalk sounds: the effects of dynamic synthesized audio on workspace awareness in distributed groupware. In: Proceedings of the 2011 ACM conference on computer supported cooperative work, CSCW 2011, Hangzhou, China, 19–23 March 2011, pp 85–94. doi:[10.1145/1958824.1958838](https://doi.org/10.1145/1958824.1958838)
28. Hancock MS, Shen C, Forlines C, Ryall K (2005) Exploring non-speech auditory feedback at an interactive multi-user tabletop. In: Proceedings of the graphics interface 2005 conference, 9–11 May 2005, Victoria, British Columbia, pp 41–50. doi:[10.1145/1089508.1089516](https://doi.org/10.1145/1089508.1089516)
29. Hart SG, Staveland LE (1988) Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) Human mental workload. North Holland Press, Amsterdam
30. Huang W, Alem L, Tecchia F (2013) HandsIn3D: supporting remote guidance with immersive virtual environments. In: Kotzé P, Marsden G, Lindgaard G, Wesson J, Winckler M (eds) Human-computer interaction—INTERACT 2013. Lecture notes in computer science, vol 8117. Springer, pp 70–77
31. Kurata T, Sakata N, Kouroggi M, Kuzuoka H, Billinghurst M (2004) Remote collaboration using a shoulder-worn active camera/laser. In: 8th International symposium on wearable computers (ISWC 2004), 31 October–3 November 2004, Arlington, pp 62–69. doi:[10.1109/ISWC.2004.37](https://doi.org/10.1109/ISWC.2004.37)
32. Lauwers JC, Lantz KA (1990) Collaboration awareness in support of collaboration transparency: requirements for the next generation of shared window systems. In: CHI '90 conference on human factors in computing systems. Special issue of the SIGCHI bulletin, Seattle, Washington, pp 303–311
33. Lukosch HK, Groen D, Kurapati S, Verbraeck A (2015) Using simulation games as research instruments: lessons learned from the transportation domain. In: Toyoda Y, Kanegae H (eds) Proceedings of the 46th international simulation and gaming association annual conference, ISOCARP, Kyoto, Japan, 17–21 July 2015, pp 177–190
34. Lukosch S, Lukosch H, Datcu D, Cidota M (2015) Providing information on the spot: using augmented reality for situational awareness in the security domain. *Comput Support Coop Work (CSCW)* 24(6):613–664. doi:[10.1007/s10606-015-9235-4](https://doi.org/10.1007/s10606-015-9235-4)
35. McFarlane DC, Latorella KA (2002) The scope and importance of human interruption in human-computer interaction design. *Hum Comput Interact* 17(1):1–61. doi:[10.1207/S15327051HCI1701\\_1](https://doi.org/10.1207/S15327051HCI1701_1)
36. Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays. *IEICE Trans Inf Syst E77-D(12):1321–1329*

37. Nilsson S, Johansson B, Jonsson A (2009) Using ar to support cross-organisational collaboration in dynamic tasks. In: 8th IEEE international symposium on mixed and augmented reality, 2009. ISMAR 2009. IEEE Computer Society, pp 3–12. doi:[10.1109/ISMAR.2009.5336522](https://doi.org/10.1109/ISMAR.2009.5336522)
38. Peters V, Vissers G, Heijne G (1998) The validity of games. *Simul Gaming* 29(1):20–30. doi:[10.1177/1046878198291003](https://doi.org/10.1177/1046878198291003)
39. Poelman R, Akman O, Lukosch S, Jonker P (2012) As if being there: mediated reality for crime scene investigation. In: CSCW '12: proceedings of the 2012 ACM conference on computer supported cooperative work. ACM Press, New York, pp 1267–1276. doi:[10.1145/2145204.2145394](https://doi.org/10.1145/2145204.2145394)
40. Schmidt K (2002) The problem with 'awareness': introductory remarks on 'awareness in csw'. *Comput Support Coop Work (CSCW)* 11(3–4):285–298. doi:[10.1023/A:1021272909573](https://doi.org/10.1023/A:1021272909573)
41. Sodhi RS, Jones BR, Forsyth D, Bailey BP, Maciocci G (2013) Bethere: 3d mobile collaboration with spatial input. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '13. ACM Press, New York, pp 179–188. doi:[10.1145/2470654.2470679](https://doi.org/10.1145/2470654.2470679)
42. Tan W, Liu H, Dong Z, Zhang G, Bao H (2013) Robust monocular slam in dynamic environments. In: IEEE international symposium on mixed and augmented reality (ISMAR), pp 209–218. doi:[10.1109/ISMAR.2013.6671781](https://doi.org/10.1109/ISMAR.2013.6671781)
43. van den Hoogen J, Lo J, Meijer S (2016) Debriefing research games: context, substance and method. *Simul Gaming*. doi:[10.1177/1046878116651023](https://doi.org/10.1177/1046878116651023)
44. van Knippenberg D, De Dreu CKW, Homan AC (2004) Work group diversity and group performance: an integrative model and research agenda. *J Appl Psychol* 89(6):1008–1022. doi:[10.1037/0021-9010.89.6.1008](https://doi.org/10.1037/0021-9010.89.6.1008)