ABSTRACT
Teams and groups that meet in 3D virtual worlds to work or learn together choose the medium for its distinct features and advantages over other online media. While these advantages are often mentioned by the academic and educational communities, empirical evidence is rare. Furthermore, it is still unclear how the distinct features of the medium can be best put to use. In this paper we present an extensive empirical study of collaboration in virtual worlds, investigating the design of environment and activities. The results of the research yield new insights on the design of effective and memorable collaboration experiences. Based on these findings, we present a set of guidelines for the design of collaboration in virtual worlds and discuss possible implications for the design of collaboration in general.

KEY WORDS
Experience design, spatial configuration, embodiment, virtual worlds, avatars, virtual teams, 3D CVE, CSCW

1. Introduction
Teams and groups that meet in 3D virtual worlds to work together choose the medium for its distinct features and its potential advantages over those of other online media like text chat, video conferencing, or computer-supported cooperative work (CSCW) software and Web services. The academic and educational community often points to features like immersion, embodiment, and spatiality, which are claimed to lead to positive effects that are assumed to facilitate collaborative work, including enhanced (co-)presence, improved team awareness, and higher engagement and greater motivation [2,9,10]. Scientists have investigated collaboration from a process perspective (e.g., [15]) but empirical studies to evaluate these advantages are rare [1,8].

In a previous study we empirically compared the medium virtual world against simple text chat, using three generic collaboration tasks that implemented information sharing, grounding, and decision-making. The results indicated that using virtual worlds for collaboration tasks improves retention: members of virtual world teams recalled more of the information shared and created in the meetings than those who collaborated using traditional 2D text chat [12].

The research presented here is a continuation of our previous study. It was designed to systematically evaluate the benefits and possibly identify and understand how to avert perils of virtual worlds as a collaboration tool. In particular, we investigate how effective collaboration experiences can be designed for.

2. Research Question and Research Design
The overarching research question for this study is: How should virtual world environments and collaborative activities within them be designed in order to best support and foster team collaboration?

This question addresses both the design of virtual world environments – including space, landscape, structure, tools, instruments, ambience – and the design of activities in these environments – types of activities, use of embodiment, group configurations, roles, relations, rules, steps, timing. While the investigation at hand addresses the design of both environment and activity (i.e., the design of whole experiences, see [7,13]), its scope does not include all mentioned aspects, but focuses on embodiment [5] and spatiality [11,6], two of the main features that distinguish virtual worlds from other online collaboration media [3]. By embodiment we mean the possibility of using customized avatars as means of communication, by spatiality the possibility of using an infinite abundance of 3D space for the design of collaboration tasks and activities.

3. Experimental Design
We designed an experiment in order to evaluate different collaboration tasks and modes during a fictitious project meeting in a simulated authentic collaboration context. We used three types of tasks that are typical for an early-stage project meeting. These three tasks were evaluated across three different collaboration modes in a virtual world. We implemented a 3x3 mixed-factorial design with task type as the within-subject factor: (1) information sharing, (2) brainstorming, and (3) decision-making, and...
collaboration mode as the between-subject factor: (A) collaboration in a localized environment using ‘traditional’ 2D tools and objects, (B) collaboration in a localized environment using 3D objects and tools, and (C) collaboration in a separated environment using 3D objects and tools. Table 1 illustrates this experimental design; the tasks and collaboration modes are described in detail in their respective subsections below. The dependent variables are objective task performance (here: retention of information from the meeting), subjective performance, and task evaluation. The following sub-sections describe our hypotheses for the experiment and elaborate on the collaboration tasks and modes.

Table 1
The three experiment conditions manipulating the between-subject factor “collaboration mode” with varying levels of spatiality and embodiment.

<table>
<thead>
<tr>
<th>Collaboration Environment</th>
<th>Objects and Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control condition (A)</td>
<td>2D (text-based, interactive walls)</td>
</tr>
<tr>
<td>Experimental condition (B)</td>
<td>3D (3D objects and tools)</td>
</tr>
<tr>
<td>Experimental condition (C)</td>
<td>3D (3D objects and tools)</td>
</tr>
</tbody>
</table>

3.1 Hypotheses

Our main hypotheses for this research were:

H1 Structuring the virtual world environment (connecting spatially separated tasks through a directed path, inspired by the mnemonic technique of loci [4]) has a positive impact on the structuredness of collaboration and thus ultimately the objective task performance (retention), as compared to having all tasks at one and the same location. (C > B, A)

H2 Utilizing 3D objects for collaboration tasks makes participants feel they accomplish work and has thus a positive impact on subjective task performance compared to using presentation slides and 2D walls. (C, B > A)

H3 Making active use of the virtual embodiment (the representation as avatars) in collaboration activities has a positive impact on engagement and hence the subjective rating of the task in the particular collaboration mode, which we measure as task evaluation. (C > B, A)

3.2 Collaboration Tasks

We conducted the experiment with randomly assigned teams of five participants. Each of the participants was then randomly assigned one out of five fictitious personas, which they were asked to personify during the collaboration experiment in the virtual world. Each of the personas was represented by a specific avatar (3 male, 2 female). The scenario of the collaboration tasks was built around a virtual project team that meets for the first time, online, in a virtual world that was specially-designed for the kick-off meeting of a fictitious project they work on together. The goal of the project was to create a website that explains the financial crisis to non-experts. We implemented a hidden profile situation: each participant initially only knows the profile information of the persona that was given to them – knowledge about others needs to be learned (through information sharing) and created during the meeting [14]. The agenda of the kick-off meeting consisted of these three collaboration tasks:

1) Information sharing

The participants had to present themselves (or rather, their persona) to their team members, using the information on the profile sheet they received. Information to share includes age, profession, working history, mentionable expertise, and hobbies.

2) Brainstorming

In this group creativity task the team members had to discuss the project and collaboratively develop basic ideas and concepts concerning the project approach and goals. This task is approached using a brainstorming technique called ‘brainwriting’: ideas about the project are written on a set of cards that are open for all team members to amend. This activity aims to result in a number of refined ideas, structured by topics on separate cards.

3) Decision-making

The team members discuss and collaboratively decide on a role assignment: Each of the team members has to be assigned to one or more of three available project roles: Content, development, and marketing. In addition to this, a project manager must be appointed. While the project roles can be shared by more than one team member, the project manager role can be assigned to only one team member.

Brief instructions for the tasks and information about the fictitious project were given to the participants along with the profile of the assigned persona on a one-page information sheet, just before logging into the virtual world. As the five personas were designed to have skills that give clear indications on how to build an optimal team structure for the project, sharing the right information (Task 1) helps assign the project roles in an optimal way (Task 3). Task 2 creates the project context.
3.3 Collaboration Modes

The experiment was designed to compare three different setups for collaboration in a virtual world; we call these ‘collaboration modes’. These variations in the use of spatiality and embodiment are summarized in Table 1 above, and explained in detail here. The purpose of the design of the three collaboration modes is to identify design aspects and approaches that support and foster the process and outcome of collaboration. To this end, we looked to investigate both the spatial setup of the work environment and the character of the tools in it.

**Condition A – Localized environment, 2D objects**

This condition – the control condition – confined the team to a relatively small area, enclosed in an invisible ‘glass cube’ to prevent the participants (or more precise, their avatars) from going astray. The tools for the team to use for the three tasks are placed in juxtaposition with each other, as shown in figure 1. The three tools for task completion each come with a sign post beside them, which explain the particular task and how to use the tool in brief phrases.

For Task 1, a screen displays profile presentation slides for each persona. Upon single left click the slide changed to the next team member. The slides showed a head shot of the avatar and summed up the profile information in keywords, exactly like on the profile sheets.

Task 2 features a board composed of brainwriting cards, on which participants could sketch their ideas in the form of keywords and short phrases. Upon left click on any line of a card the subsequent written statement of the participant appeared on it instead of in regular text chat.

In Task 3, the role assignment was accomplished by (re-)positioning name tags in a triangle on a board, the corners of which symbolize the three roles (the corners are marked textually). The project manager (PM) to be assigned could mark their name red, with a single click. Figure 2 shows a team during role assignment.

**Figure 1**
A persona presenting herself using her profile slide.

**Condition B – Localized environment, 3D objects**

In this experimental condition, the team is confined to a small area as well, the arrangement of the task spaces is the same as in the control condition, merely the tools differ. Brief descriptions of how to use the tools are written on sign posts also here.

For Task 1, a table puts 3D objects on display – personal items that describe the personas’ backgrounds (figure 3). The table pops up objects associated with the persona’s profile that clicks on it, as a visual, three-dimensional form of a slide presentation.

For Task 2 in this collaboration mode, the cards were modified in that they move towards the avatar who writes on it (as seen in the background in figure 3), then jump back into formation. In comparison to the brainwriting cards in the control condition, this design aimed to make use of the space and make the cards more readable (and writable) to the particular writers by ‘zooming’ in to them. An additional advantage was believed to be the assignment of a card to one distinct avatar at a given time.

**Figure 2**
Team assigning roles using the 2D interactive board.

**Figure 3**
Condition B: discussing and brainwriting.
For Task 3, instead of the 2D interactive board, the participants were provided with a table that allows to position flags with their owners' head shots on them within a triangular area similar to the one in Condition A. Instead of text labels, the triangle was spanned by symbolic 3D objects that depict the available roles (a megaphone for marketing, bricks for development, and a canvas for content creation). To assign the PM role a top hat – the ‘Project Manager Hat’ – had to be attached to one of the flags. This setup is shown in figure 4. Also this Condition featured sign posts that explain each task and associated tool in brief phrases.

*Figure 4*
Team assigning roles using 3D objects in Condition B.

Following the approach of embodied interaction, in this condition, Task 3 was done by ‘voting by feet’, that is the team members had to position their avatars in the decision-making triangle, which was now scaled up to walk-in size (figure 7). The avatars’ proximity to symbols communicated their preferences or decisions toward roles.

*Figure 7*
Team assigning roles ‘by feet’ in the role assignment triangle in Condition C.

**Condition C – Structured environment, 3D objects**

This more advanced experimental condition was implemented with an entirely different spatial setup: the three tasks have separate locations, cannot be seen from one another, and are connected only through a path that winds around a mountain. This design aimed at separating the tasks not only mentally but also spatially. The participants thus had to move around a whole island in the course of the meeting. Distances were short however, walking from one task to the next took about ten seconds.

In Task 1, each team member had a personal table with personal items, to present themselves (figure 5). Curtains hiding the items faded away whenever their respective owners wrote in chat. This way, emphasis was put on the one table in focus at the time. In addition, clicking on drawers in the tables executed personalized gestures, an element that aimed for offering more visual support in order to influence the memorability of profiles.

Task 2 was now an area in which the nine cards were dispersed between different trees, bushes, and coves (figure 6). Each card was thus located in a different setting with a different background, which was intended to influence the memorability of the ideas created during the task and written on the boards.

*Figure 5*
Condition C: a team member presenting himself using personal items (3D objects) and gestures.

*Figure 6*
Team discussing ideas between the brainwriting cards in the spatial setup of the second experimental condition.
4. Method

4.1 Participants

We conducted the experiment with 16 teams in total. Due to some incomplete questionnaires, our fully usable data set comprises N=67 participants. Participants were recruited in the United States, Switzerland, and Israel. The research protocol and procedure was approved by the Internal Review Board (IRB) at PARC, California.

4.2 Measures (dependent variables)

**Objective task performance**
We measured the retention (i.e., the participants’ recall of information shared and created in the meeting) by counting the information items they remembered. For Task 1 we counted the items of the different persona profiles (i.e., age, profession, work experience, hobbies). For Task 2, we counted the number of generated ideas the participants recalled from the meeting. For Task 3 we measured how well the participants recalled the role assignment decided upon in the meeting (i.e., which roles were assigned to which team members). Here we counted the assignments that were recalled in concordance by different participants.

**Subjective task performance**
In order to measure the subjective performance of the collaborating teams, we asked the participants to rate the collaboration for the three tasks separately. For Task 1 we asked if the team shared their profile information well and listened carefully, for Task 2 we asked if the brainwriting was done in an organized way and if ideas were jointly created, while for Task 3 we asked if the role assignment process was clear and effective, and whether its state was visible at all times.

**Task evaluation**
With the aim of not only getting an evaluation of the collaboration performance, but also of the collaboration tasks in the various modes, we asked participants to evaluate each of the tasks on the basis of nine attributes rated on a 5-point Likert scale: easy, interesting, satisfying, successful, engaging, immersive, efficient, enjoyable, and inspiring.

**Further measures**
We further measured the time to completion of the three tasks, logged all text chat, and logged all navigation information and object interaction of the five avatars.

4.3 Procedure

A questionnaire of demographic data and personality profiling had to be filled in prior to the virtual world activity (here we also measured the expertise with virtual worlds, the covariate we used in the analyses). A second questionnaire, including subjective reports and ratings, was administered right after logging out of the virtual world. After completing the second questionnaire, the participants were given a third questionnaire to test the recall (i.e., how much of the information the participants could remember). In addition, we developed a tracking module for the OpenSim platform that logs all (text) chat, avatar navigation data, and information about object interaction. The combination of such diverse objective and subjective data allows for in-depth analyses of team interaction processes and outcomes.

5. Results

We conducted analyses of covariance (ANCOVA). We used virtual world expertise as a covariate in order for the results to be independent from possible imbalances in virtual world proficiency among the participants. In general, we found that also first-time users quickly understood how to navigate and communicate in the OpenSim virtual world environment, especially since the experiment was setup for simple interaction (e.g., only left mouse clicks to interact with the collaboration tools); nonetheless, we deemed the inclusion of the covariate to be correct. In the following we describe the results of the different analyses.

**Objective task performance (retention)**
We conducted ANCOVA analyses for each task separately, using the three conditions as the independent variable, retention scores as the dependent variable, and virtual world expertise as the covariate. For Task 1, the ANCOVA did not reveal any statistically significant differences between the conditions, $F(2, 61) = 1.45$, $p > .05$. For Task 2, the ANCOVA was statistically significant, $F(2, 60) = 9.73$, $p < .001$. A Tukey Post-hoc test showed that the difference was due to a significantly better retention score in Condition C compared to Conditions A and B (C > A, B); that is, more ideas were remembered correctly in Condition C than in the other conditions. The same pattern of result was found for Task 3 (decision-making). The ANCOVA was statistically significant, $F(2, 61) = 7.03$, $p = .002$. A Tukey Post-hoc test showed that retention was significantly higher in Condition C, as compared to the other conditions (C > A, B). Figure 8 illustrates the means and standard deviations of the retention results for Tasks 2 and 3.

**Subjective task performance**
We conducted ANCOVA analyses for each task separately, using the three conditions as the independent variable, subjective task performance ratings as the dependent variable, and virtual world expertise as the covariate. The conditions yielded high measures but did not significantly differ regarding subjective task performance, as measured for Task 1, $F(2, 63) = .61$, $p > .05$, and Task 3, $F(2, 63) = .29$, $p > .05$. For Task 2 the ANCOVA was
statistically significant, \( F(2, 63) = 6.97, p = .002 \). A Tukey Post-hoc test showed that participants perceived the team’s performance significantly higher in Conditions B and C than in Condition A (C, B > A). Figure 9 illustrates the results for the subjective task performance for Task 2.

**Task evaluation**

We conducted ANCOVA analyses for each task separately, using the three conditions as the independent variable, task evaluation as the dependent variable, and virtual world expertise as the covariate. No statistically significant differences were found between the conditions regarding evaluation for Task 1, \( F(2, 71) = .58, p > .05 \), nor Task 3, \( F(2, 71) = .60, p > .05 \), which is why we omit those graphs again. The ANCOVA for Task 2 was statistically significant, \( F(2, 71) = 4.74, p = .01 \). A Tukey Post-hoc test revealed that participants rated their evaluation in Task 2 significantly more positive in Condition C compared to Condition A (C > A). Condition B did not significantly differ from A or C. Figure 10 shows the Task 2 results.

**Time to task completion**

As a second objective measurement we looked at the time to task completion (figure 11). Work on the tasks took between 7 and 16 minutes. While there is a tendency of collaboration in Condition C taking slightly longer than in A it seems to weak to have caused the superior retention.

Figure 9
Subjective task performance: means and standard deviations for Task 2. C, B > A is significant.

Figure 10
Task evaluation: means and standard deviations for Task 2. C > A is significant.

Figure 8
Objective task performance: means and standard deviations for Task 2 (top) and Task 3 (bottom). C > A, B is significant for both tasks.

Figure 11
Time for task completion: means and standard deviations for Task 1 (left), Task 2 (center), and Task 3 (right). The figure shows the measures for the three collaboration modes (A, B, C) clustered according to the three tasks.
6. Discussion

The results provide partial support for our hypotheses; they were accepted for some tasks and rejected for others. This outcome highlights the task-dependency of design evaluations. Although we attempted to manipulate the design factors (as specified for each condition) in equivalent ways, they resulted in different (objective, subjective) performance and task evaluation results. For Task 2 we see a significant supremacy of Condition C in both objective and subjective measures, while for Task 3 the objectively measured supremacy of Condition C over Conditions A and B is reflected only as a tendency in the subjective measures.

We can confirm that structuring the collaboration tasks into separated locations (hypothesis H1) improves the memorability of the information shared and created, for the brainwriting and decision-making tasks (Tasks 2 and 3). For the information sharing task H1 did not hold true, which could be due to the fact that the pre-defined information to be shared had no relation to the collaboration environment; in Tasks 2 and 3 in contrast, all information was created in-place and had an innate spatial relationship to the collaboration environment. Thus, as a first finding we can bring to book that separating collaboration tasks spatially helps the participants memorize the information when it is part of the spatial context of the environment in which the collaboration takes place. We therefore conclude it to be beneficial to separate those tasks spatially which create knowledge and information. For the memorization of information that exists prior to the collaboration, a spatial separation of collaboration tasks does not seem to lead to an added value.

H2 was accepted for Task 2. That is, for a brainstorming task, the use of 3D objects made collaboration participants feel more performant. When sharing information or making decisions on the other hand, all conditions yielded the same (high) measures for perceived performance. This result indicates that forcing avatars to write on a static 2D wall does not feel right, when equipped with an avatar in an interactive 3D virtual world. The same objects implemented in a more responsive way or in a more open spatial arrangement made the team members perceive their team work as significantly more performant. We therefore conclude that for collaboration in a virtual world, the medium’s offerings concerning 3D space and responsive and interactive objects need to be appropriately harnessed in order to make collaborating participants feel successful in their tasks and meetings.

Hypothesis H3 was partially accepted for Task 2: the active use of avatars as a virtual embodiment had a significantly more positive impact on the task evaluation in the spacious setup of Condition C – which required constant moving around between idea cards – than in the static brainwriting board setup of Condition A. The ‘intermediate’ setup of Condition B led to an intermediate result, indicating a possible positive relationship between the level of the active use of embodiment and positive task evaluation (a possible proportional trend is also visible in figure 10). We therefore conclude that making active use of embodiment increases people’s satisfaction with task work.

7. Emergent Design Guidelines

From the results of this first investigation of different virtual world collaboration modes for different task types, the following guidelines for designing for effective memorable collaboration experiences emerge:

1) **Separate knowledge creation tasks spatially to structure thoughts.** Dedicating separate places for separate activities where information or knowledge is created can support memorizing their outcomes. Task types that are likely to benefit from spatial separation include decision-making tasks, creativity tasks, planning tasks, and learning activities.

2) **Make active use of embodiment to engage people.** Using positioning and other movement as a communication channel can improve important aspects like satisfaction, inspiration and enjoyment.

3) **Use simple interaction design.** Designing the users interaction in a simple and intuitive way (e.g., only left clicks on objects instead of complex menus) allows everyone to participate in the collaboration.

4) **Don’t over-design.** Scaffolding collaboration processes is good, but predicting and constraining people’s actions too much can lead to frustration.

5) **Don’t under-design.** Restraining collaboration by forcing avatars to sit or stand still and write on 2D boards despite of the abundant space and interaction possibilities the medium offers can lead to apathy. If tasks need to be text-based they can be implemented using responsive objects or spatial arrangements.

6) **Don’t set artificial boundaries.** Set natural ones. Designing an island is a more natural way of restricting people’s movement than creating walls (even if the walls are invisible).

7) **Don’t stuff.** Putting too many things in too little space can jam participation and bring frustration.

8. Conclusion and Outlook

We have presented an experimental investigation of the design of collaboration in virtual worlds. The results suggest that the approach of structuring tasks and making use of features like embodiment and spatiality have a positive impact on collaboration performance and satisfaction, if they are applied (a) for the right situations, and (b) in the right dosages. Different task types benefit
from the virtual world features of space and embodiment in different ways. Elaborating on this, we have discussed the experiment results and presented emergent findings. We have also presented a set of design guidelines that have emerged from the findings of the research.

For the design of collaboration experiences in virtual worlds, a field that is only little investigated so far, the implications of this research offer first indications as to which design approaches are superior to others in order to reach certain goals (e.g., recall, satisfaction, engagement).

The fact that our hypotheses were accepted for some tasks and rejected for others has confirmed the value of the general approach of making more explicit use of distinct features of the medium virtual world, but has also highlighted that the application and the situation is crucial in the process of selecting a collaboration design approach, and that the dosage of applying design approaches and implementing aspects is a pivotal factor. The decisions of which design approach to follow and how 'stringent' to design the collaboration activities and tools have to be made on a case-to-case basis, separately for every collaboration task.

The research at hand is therefore mere but a first step in understanding how to best design virtual worlds and activities and tools in them, when planning for fruitful and memorable online collaboration. Various possible directions for future research are opened up by this first step. An investigation of more diverse tasks and generic task types can lead to more insight regarding the task-design-approach-fit, more thorough studies of interaction tools and collaboration arrangements can add to the syntactical knowledge of collaboration design, and finally thinking about the concept of patterns one could imagine the development of a hierarchical pattern language, resulting in a complex and powerful blueprint for designing virtual world collaboration experiences.

How far the design guidelines for virtual world collaboration can be transferred and applied to designing collaboration activities in the physical world is still very unclear (and would be another future research direction). We believe, however, that future virtual world research will ultimately be able to improve 'real' collaboration activities and thus entire meetings also in the physical world. This belief is motivated especially by the advancements in ubiquitous computing and the trend of physical and virtual worlds merging more and more.

ACKNOWLEDGMENTS
The research was partly funded by the Swiss National Science Foundation (grant No. PBTIP1-130959) and the European Union (grant No. FP7-PEOPLE-2009-IEF-254277). The authors would like to thank the participants of the research for their time and engagement and for their valuable comments.

REFERENCES