

OPEN ACCESS

EDITED BY

Nick Bryan-Kinns,
Queen Mary University of London,
United Kingdom

REVIEWED BY

Gavin Sim,
University of Central Lancashire,
United Kingdom
Yomna Abdelrahman,
Munich University of the Federal
Armed Forces, Germany

*CORRESPONDENCE

Riccardo Bovo
rb1619@ic.ac.uk
Daniele Giunchi
d.giunchi@ucl.ac.uk

SPECIALTY SECTION

This article was submitted to
Human-Media Interaction,
a section of the journal
Frontiers in Computer Science

RECEIVED 25 April 2022

ACCEPTED 17 August 2022

PUBLISHED 04 October 2022

CITATION

Bovo R, Giunchi D, Costanza E,
Steed A and Heinis T (2022) Mitigation
strategies for participant
non-attendance in VR remote
collaborative experiments.
Front. Comput. Sci. 4:928269.
doi: 10.3389/fcomp.2022.928269

COPYRIGHT

© 2022 Bovo, Giunchi, Costanza,
Steed and Heinis. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Mitigation strategies for participant non-attendance in VR remote collaborative experiments

Riccardo Bovo^{1*}, Daniele Giunchi^{2*}, Enrico Costanza²,
Anthony Steed² and Thomas Heinis¹

¹Computer Science, Imperial College London, London, United Kingdom, ²Computer Science, University College London, London, United Kingdom

COVID-19 led to the temporary closure of many HCI research facilities disrupting many ongoing user studies. While some studies could easily move online, this has proven problematic for virtual reality (VR) studies. The main challenge of remote VR study is the recruitment of participants who have access to specialized hardware such as head-mounted displays. This challenge is exacerbated in collaborative VR studies, where multiple participants need to be available and remotely connect to the study simultaneously. We identify the latter as the worst-case scenario regarding resource wastage and frustration. Across two collaborative user studies, we identified the personal connection between the experimenter and the participant as a critical factor in reducing non-attendance. We compare three recruitment strategies that we have iteratively developed based on our recent experiences. We introduce a metric to quantify the cost for each recruitment strategy, and we show that our final strategy achieves the best metric score. Our work is valuable for HCI researchers recruiting participants for collaborative VR remote studies, but it can be easily extended to every remote experiment scenario.

KEYWORDS

collaborative experiment, non-attendance, virtual reality (VR), cost function, recruitment procedures

1. Introduction

Traditionally the virtual reality (VR) research community relies on in-lab experiments because of ease in controlling the environment and the specialized hardware required. However, the COVID-19 pandemic has forced many VR researchers to adapt their research experiments from in-lab to remote. This change comes with opportunities and challenges. For example, in-lab studies usually promote local recruitment of students inside universities (Fröhler et al., 2022) while remote user studies offer the possibility to recruit a wider and more diversified population.

On the other hand, such a remote context imposes additional complexities. For example, researchers need to leverage consumer equipment and online delivery of experimental materials (Steed et al., 2021). In any case, it is necessary to re-evaluate experiment procedures, with multiple approaches in recruiting participants, running

sessions, and collecting results (Steed et al., 2020). Researchers developed several toolkits that have features that could be used to develop remote VR experiments: experiment design frameworks (Watson et al., 2019; Bebko and Troje, 2020; Brookes et al., 2020), immersive questionnaires (Feick et al., 2020; Bovo et al., 2022). In addition, researchers have explored methodologies to run unsupervised VR studies remotely (Mottelson et al., 2021; Radiah et al., 2021a). While previous efforts focused on remote VR user studies with individual participants, fewer frameworks or protocols for collaborative VR studies have been formalized to date.

Two exceptions are Radiah et al. (2021b) and Saffo et al. (2021) which explore the use of social VR platforms to recruit and run experiments by accessing large pools of available users without requesting them to install additional software developed by researchers. However, VR social platforms (ie., as VRChat Inc., 2014; AltspaceVR, 2015; Rec Room Inc., 2016), have considerable limitations such as the impossibility of serializing data and heavy constraints on the experimental design. While Radiah et al. (2021b) and Saffo et al. (2021) report a workaround by recording information in videos and post-processing after the experiment with OCR, these solutions are limited to the framerate of the video and also limit the size of the data that can be saved. As long as social VR platforms will not support experiments, accessing their user community for easy recruitment of participants may not often be useful. Our study explores recruitment procedures for collaborative experiments that cannot leverage VR social platforms.

Non-attendance, also known as no-shows, is a real problem in-lab studies, and, in remote studies, researchers should expect even higher no-show rates (Scott and Johnson, 2020). We propose a recruitment procedure that mitigates common issues in remote collaborative experiments, such as non-attendance. While unsupervised individual studies can be carried out asynchronously, studies that involve collaborative scenarios demand several participants and, very frequently, a moderator. In particular, synchronous collaborative studies need participants, experimenters, and confederates to be available simultaneously and connected via the network. In-person laboratory collaborative studies rely on the fact that participants' presence is a mandatory requirement, the experiment setup is managed by the experimenter, and additional conditions are explained to the participant before the experiment starts. In general, running VR remote collaborative studies has a higher risk of failure due to the low degree of control that the experimenter has over the VR devices, the physical space, and the differences in connectivity. An in-person experiment enables the researcher to set up both hardware/software and laboratory with almost the same environmental conditions that are not altered (same hardware involved, same room, same network, etc.).

A remote experiment requires the researcher to carry out this setup for each participant so to make sure that the variance

of the conditions is within an acceptable range. For example, network conditions can be different, the places where the experiment is taken may vary, and it is harder to keep control of interruptions. These differences can be managed in advance but require greater effort from the experimenter, and it needs to be considered in the experiment protocol. We propose to keep this phase part of the recruitment protocol to evaluate in advance if the participant's conditions are within the acceptable variance.

In addition, the experimenter needs to spend a considerable amount of time negotiating via email time slots and the availability of participant groups. During such phase, we noticed that the worst-case scenario happens when the participants give their availability and confirm their attendance for the collaborative session but then do not show up. Such behavior leads to a waste of experimenter time and remuneration due to any attending participant(s). Our experience with studies demonstrates that when recruiting participants external to the community of researchers, the risk of non-attendance is higher. Thus, we create a protocol that helps us identify participants who will not attend before investing time in mediating availability for the collaborative session. This way, we avoid wasting the experimenter's and the other participant's time. Usually, in single-user experiments, participant screening is realized in one phase, where both requirements and availability are considered. We introduced a better approach for remote collaborative experiments aimed at identifying non-attendance participants as early as possible during the recruitment process.

First, we propose to abstract the experiment's stages as a sequence of activities for the experimenter and the participants, such as experiment screening, recruitment, and experiment running. This process of abstraction allows us to propose to the research community a model for calculating the cost of an experiment. Second, we designed a generic cost function to calculate the cost for each stage, achieving the total with a simple sum. Finally, with such analysis, we are able to study a better approach for that sequence of activities, and we defined different protocols.

We propose to divide the screening into two phases: a self-assessed one (i.e., an online questionnaire) and a mediated one (i.e., teleconference with the experimenter) as a phase preceding the collaborative experiment. Such a procedure allows us to significantly reduce the number of non-attendance participants by moving them from the collaborative phase to the screening phase, reducing their impact on the cost of the experiment. Moreover, this mediated phase has other benefits besides reducing the non-attendance rate in the collaborative phase. First, preliminary troubleshooting can be carried out face-to-face rather than email-mediated. Second, during this teleconference session, the experimenter can verify details from the initial screening and evaluate the environment where the user will perform the experiment, the internet connection speed, the language proficiency level, etc.

We show results that support the benefits of this mediated phase between the experiment mediator and each participant. We run two VR remote collaborative studies recruiting participants internally to the research community and externally to the broader community of VR users. We engaged with 94 possible participants across three different iterations of our recruitment process. We classified each individual engaged with our recruitment procedures as follows: a dropout, a non-attendance, and concluded. We developed a cost function that allows us to understand the cost of each participant that concluded the experiment while considering the resources that were also spent for those that dropped out or did not attend the experiment. We demonstrate how our proposed recruitment procedure reduces costs per external participant when recruiting for a collaborative VR experiment.

2. Methods

2.1. Participants

We advertised each of the two remote VR collaborative user studies both to the research community (i.e., within our universities and through mailing lists related to the academic research community) and externally by advertising the experiment on multiple online platforms (see Section 2.3). We interacted with 94 possible participants overall, and at the end of the recruitment process, only 42 participants (21 pairs) took part in the experiment, corresponding to 45%. The rest either dropped out or booked the experiment but did not attend. We noticed a considerable difference across the communities: of 35 potential participants from the research community, 28 took part in the experiment (80%). In contrast, out of 59 possible participants outside the research community, only 14 took part in the experiment (24%). Eleven pairs of participants took part in the first experiment and ten pairs in the second one; none of the participants performed both experiments. For the first experiment $MAge = 33$, $SDAge = 8.3$ and the 88% of the participants were male, and 12% female. For the second experiment $MAge = 30.5$, $SDAge = 5.9$ and 70% of the participants were male and the 30% female.

2.2. Cost estimation

We determine a metric that assigns a score to each protocol to evaluate the average cost per participant in the experiment. This cost function primarily depends on the time spent by participants and experimenters in each phase and the remuneration for participants. In addition, the metric depends on the number of participants, the number of phases in the protocol, the probability of dropout before the actual study, and the probability of non-attendance at the actual study. We

calculate the probabilities as frequentist statistics do, considering the fraction where the numerator is the number of highlighted events and the denominator is the total number of events. The final score represents the average cost per participant for that specific protocol. This effort considers all the resources spent in allocating time slots never used, payment carried out without completing the experiment, and the time spent by the researchers to communicate with potential participants. These situations can happen for two main reasons: the absence of one or more participants or even the experimenter or some technical difficulties at the time of the experiment. Each protocol consists of several phases (Figure 1) that form a temporal sequence, where each phase to start needs to wait until the end of the previous one. These phases have a cost that can be temporal or monetary. We convert all temporal costs to monetary costs using the London Living Wage (Klara Skrivankova, 2022). Our cost function (Equation 1) considers the number of different phases, for each phase quantifies the costs per participant and also highlights the cost differences between the two identified categories of participants: external group and internal group. The variables are described in Table 1. The cost function scores the cost per participant and is given by the following formula:

$$Cost = \sum_{i=1}^n (P_{D,i} * C_{D,i} + P_{NA,i} * C_{NA,i} + P_{C,i} * C_{C,i}) / N_C \quad (1)$$

$$Conversion_{time,money} = M/60 \quad (2)$$

$$C_D = Conversion_{time,money} * T_{run} \quad (3)$$

$$C_{NA} = Conversion_{time,money} * T_{run} + V_D \quad (4)$$

$$C_C = Conversion_{time,money} * T_{run} + V_F \quad (5)$$

2.3. Advertising experiment

All recruiting protocols started with an advertising phase. To recruit internally to the university, we circulated an informative text within the local academic context introducing the experiment's broad aim, the time investment, and the remuneration. We invited participants to contact the experimenter if interested. We used internal mailing lists across departments and universities, highlighting the professors and researchers involved to meet the ethics requirements. We also advertised the studies externally because internal recruitment often is limited to researchers/students who have access to a headset for research reasons or those that belong to the same group. Our target population constraints were not related to academic or student status. We opened and advertised

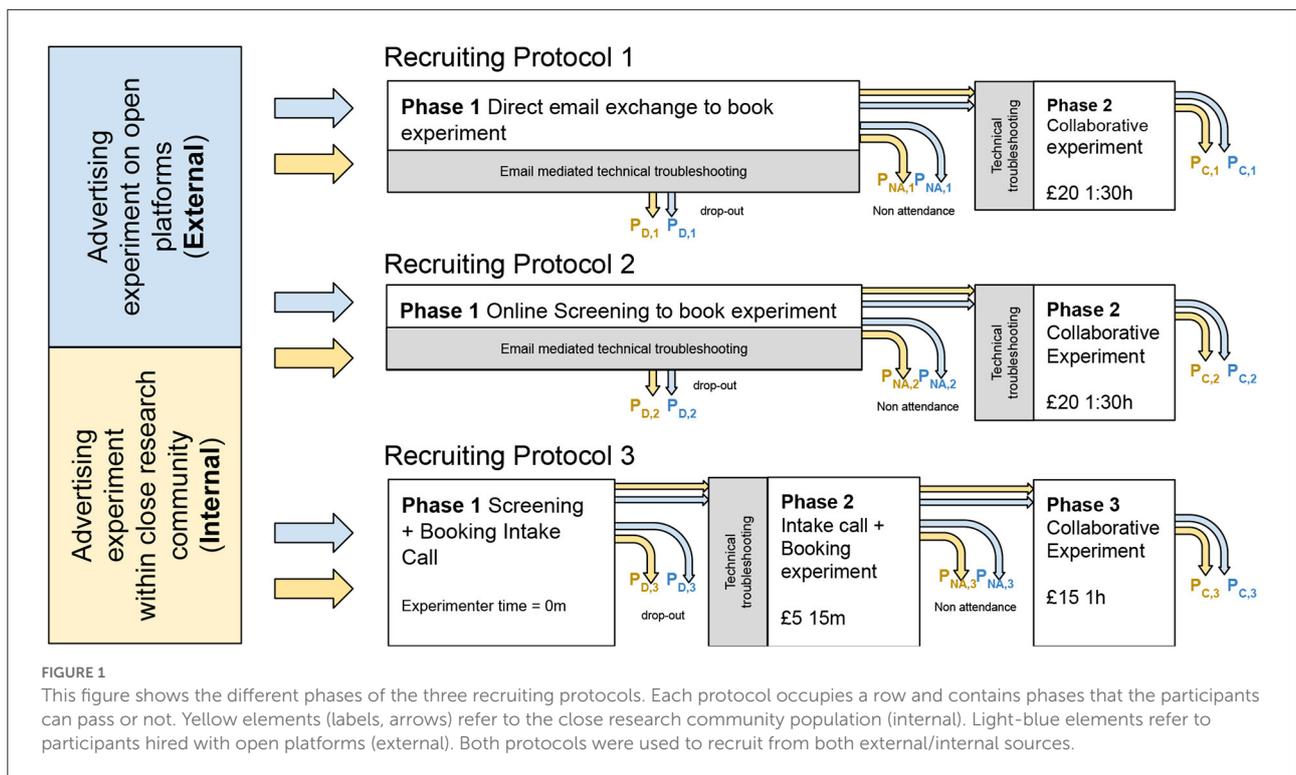


FIGURE 1
 This figure shows the different phases of the three recruiting protocols. Each protocol occupies a row and contains phases that the participants can pass or not. Yellow elements (labels, arrows) refer to the close research community population (internal). Light-blue elements refer to participants hired with open platforms (external). Both protocols were used to recruit from both external/internal sources.

our studies to any English adult speaker with access to an Oculus VR HMD.

The advertisement text was posted on multiple platforms dedicated to experiments with online recruitment, such as [Immersive Experiences Working Group \(2022\)](#), [Kruusimägi et al. \(2022\)](#), and [Prolific \(2022\)](#). As these platforms are not necessarily popular among VR users, we also advertised our experiment on channels dedicated to open exchanges between VR users, such as the Facebook oculus channel ([Meta, 2004](#)) and Reddit ([Advance Publications, 2005](#)).

2.3.1. Advertising experiment task

The first study consisted of two participants carrying out a collaborative visual search task across two datasets consisting of a 3D terrain map and a puzzle. Participants were asked to collaboratively identify specific features of the dataset (i.e., tallest mountain peak, the largest settlement, matching puzzle block).

The second study also consisted in two participants carrying out a collaborative visual analysis task across a series of visualizations (i.e., scatter plots and bar charts) of a movie dataset and a car dataset. Participants were asked to collaboratively extract insight from the dataset visualizations. The experiment description during the recruitment phases was of a similar length and did not describe the experiment in depth but rather described the task as a collaborative visual analysis task.

2.3.2. Advertising experiment length

For both experiments, the advertised length was 90 min in total, and both effective experiment lengths were very similar; the first experiment MDuration = 83 m, SDDuration = 7 m, while the second experiment MDuration = 85 m, SDDuration = 10 m. The duration was measured with the collected data timestamps and the post-interview recordings. In the cost function, we estimate the duration to be 1 h and 30 m for both experiments, as highlighted in [Figure 1](#).

2.4. Recruiting protocol 1

The first phase of protocol 1 consisted of a direct email exchange. Participants get in contact after reading the advertisement. Then the experimenter emails them asking to specify three dates where they are available out of a set of candidate dates. Once participants replied, they were matched to other participants with equivalent availability and a calendar event was sent to both to confirm the collaborative experiment. To speed up emailing procedure, we use a template text that requires manual updates of experiment dates. Often this exchange was repeated multiple times as available dates expired without matching candidates.

The calendar event contained all information to prepare for the collaborative experiment, such as the download and installation instructions for the Android Package (APK) and the informed consent form to send back before the experiment. If

TABLE 1 Equation variables.

Notation	Definition
$Cost$	cost per participant
n	is the number of phases of the protocol P
$P_{D,i}, P_D$	is the probability of dropping out at a phase before the collaborative session
$P_{NA,i}, P_{NA}$	is the probability of non-attendance for the collaborative session phase
$P_{C,i}, P_C$	is the probability of completing the experiment successfully
$C_{D,i}, C_D$	is the cost per dropout phase (Equation 3) that takes into account the time spent by the experimenter when a participant is reported missing [time to money conversion provided by Equation (2)]
$C_{NA,i}, C_{NA}$	is the cost per non-attendance phase (Equation 4) that takes into account the time spent by the experimenter and the participant that attends when the other participant does not show up [upon monetary conversion with Equation (2)], and the Amazon voucher (5€ only for protocol 3) from the previous phase
$C_{C,i}, C_C$	is the cost of completion (Equation 5) that takes into account the final Amazon voucher (20€ for protocol 1 and 2, 15€ for protocol 3) and the time spent by the experimenter [upon monetary conversion with Equation (2)]
N_C	is the total number of participants that took part in the collaborative experiment
M	is the cost per hour for a Ph.D. student in London, all the times are calculated in min, and costs are calculated in British currency (GBP)
T_{run}	is the time required to run such a phase
V_D	is the voucher of the previous dropout phase if present
V_F	the final voucher
$Conversion_{time,money}$	is calculated from Equation (2)

both participants accepted the calendar event at least 2 days before, then the experiment would go ahead. If either of the two participants did not confirm, the event was canceled, and a notification was sent with a rescheduling option. Until this point, the experimenter's time invested in such a procedure was, on average, 15 min per participant. Such *time* multiplied by the M gives us the value of $C_{D,1}$, which is the cost related to a participant dropping out at this stage of this protocol 1 and is equivalent to £2.7.

The second phase consisted of the collaborative experiment. On the day of the experiment, if both participants attended, the experiment occurred; if only one participant showed up, he was offered the agreed payment and an option to reschedule. In case both the participants were missing, an email was sent to both of them for rescheduling. Moreover, researchers would explain the difficulty in organizing collaborative experiments and thank him

in case of an agreed reschedule. If both participants attended, they were given detailed instructions and asked to perform a quick test of the VR environment to ensure both could connect, and APK was installed correctly. Upon completing the experiment, a £20 voucher was given to both participants.

$C_{NA,1}$ represent the cost related to a participant not attending the collaborative phase. We calculated this as the cost of the email negotiating mediation plus the time the experimenter allocated for the experiment, which added up to 90 min and was equivalent to 19.3€. $C_{C,1}$ represent the cost related to a participant attending and completing the experiment and is equal to $C_{NA,1}$ plus the voucher value (£20) used to compensate a participant completing the experiment and is equal to £39.3.

2.5. Recruiting protocol 2

The second protocol was created to reduce the overhead related to the email exchange. The second protocol differs from the first only in relation to phase 1. To reduce the overhead related to an email exchange, we developed an online intake form using Microsoft Forms Office 360. The intake form consisted of questions related to experiment requirements such as internet speed connectivity, access to a VR headset, and English language proficiency.

Moreover, also availability slots were chosen in the online intake form rather than via email, therefore, reducing the number of emails exchanged from phase 1. Once the form was submitted, an automated email was received by researchers. If the participant's availability matched any other participant's availability, a calendar invitation with instructions was sent to both. The available days were manually updated in the online form to get new participants to match already chosen slots of previously submitted intake forms. The second phase was the same as protocol 1.

The first phase reduced the amount of *time* that the experimenter spent from 15 to 10 min, therefore, reducing the cost of this phase to $C_{D,2} = 1.8€$ as $C_{NA,2}$ and $C_{C,2}$ and are based on the cost of the first phase too their cost variation was minimal compared to protocol 1 ($C_{NA,2} = £18.4$, $C_{C,2} = £38.42$).

2.6. Recruiting protocol 3

Despite the improvements in terms of email exchange overheads, the second protocol still had the major issue related to non-attendance. We designed protocol 3 to mitigate such problems starting from an insight obtained from previous experiences: we observed that all participants who attended without their collaborator opted for rescheduling rather than collecting the payment, and they always attended the second collaborative experiment. Therefore, we model an intermediate phase (i.e., phase 2 or "intake call"), with the hypothesis that if

TABLE 2 The number of participants across the protocols and experiments.

Exp.	Protocol	Total	Cost	Internal				External			
				Dropout	Non-attend	Attend	Cost	Dropout	Non-attend	Attend	Cost
1	p1	31	£851	4	1	9	£385	7	15	4	£467
1/2	p2	28	£711	1	1	10	£405	4	10	3	£367
2	p3	25	£358	1	1	6	£203	11	2	3	£156

they attend this phase, they will attend the collaborative session. Such intake calls had a much lower risk/cost associated with it. This phase was taken between the experimenter and one participant at a time and lasted only 15 min. This design prevents wasting time required for the experiment and losing money to pay the participant who attended.

For this procedure, phase 1 consisted in booking the intake call. We further optimize the booking and calendar event process by using an online tool called Calendly (Calendly, 2022), which allows participants to self-book for predefined slots and automatically generate calendar events sent to both participants and the experimenter. The participant received instructions to install an APK and test the internet connection as part of the calendar event. With Calendly, participants were also required to fill a form consisting of an initial set of screening checks in which responses were embedded in the calendar event. If the screening checks were not matching the requirements, the experimenter would ask for confirmation and eventually delete the event. Screening checks consisted of self-evaluation of the experiment requirements by answering multiple-choice questions.

The second phase consisted of the intake call, which lasted no more than 15 min (often just 5 min). Participants were given the option of preloading the APK on their headset or installing it during the intake call with the experimenter's help. Furthermore, the intake call allowed the experimenter to test the participant's internet connection speed and ensure that all networking APK capabilities were fully supported. This phase also aimed to give an overview of the experiment and the VR environment without requiring participants to read the instructions. We provided them with a direct experience and examples of what sort of task they would be asked to perform during the collaborative experiment. After the intake call, each participant received the initial voucher payment ($V_{intake} = £5$) independently of their choice of booking the next experiment phase (collaborative phase). Therefore, the dropout cost in terms of *time* that the experimenter spent was 15 min, and the cost of participants not attending this intake call was modeled to be equivalent to the time the experimenter allocated for the intake call multiplied by M and equivalent to $C_{D,3} = £2.8$. However, all participants chose to book the next phase. The call participant received the calendar invitation to the collaborative phase at the end of the intake.

The last phase was the collaborative experiment. Thanks to the intake call, which covered instructions and technical tests, this phase was shorter and, therefore, paid according to the same rate. No technical issues occurred during the experiment. $C_{NA,3}$ represents the cost related to a participant not attending the collaborative phase of protocol 3. We calculated this as the $C_{D,3}$ plus the V_{intake} plus the time the experimenter allocated for the experiment, which added up to 60 min and was equivalent to £18.8. $C_{C,3}$ represent the cost related to a participant attending and completing the experiment and is equal to $C_{NA,3}$ plus the $V_3 = £15$ used to compensate a participant completing the experiment, making $C_{C,3}$ equal to £31.05.

3. Results

3.1. Number of participants per protocol and experiment

During experiment 1 we used both p1 and p2. During experiment 2 we used both p2 and p3. Both external and internal sources were used across all three protocols. A breakdown is shown in Table 2.

3.2. Costs per participants

Using our participants' data, we calculated the dropout probability, the non-attendance probability, and the completion probability for each pair protocol (p1, p2, p3) and group (external, internal). We use these probabilities with the costs outlined in each protocol section to identify the average cost for each participant who completed the experiment across our defined conditions. Results can be seen in Figure 2C.

3.3. Non-attendance

A two-way ANOVA was performed to analyze the effect of internal/external sources of participants (Section 2.1) and recruitment procedure on non-attendance. A two-way ANOVA revealed that there was a statistically significant interaction between the effects of internal/external sources of participants

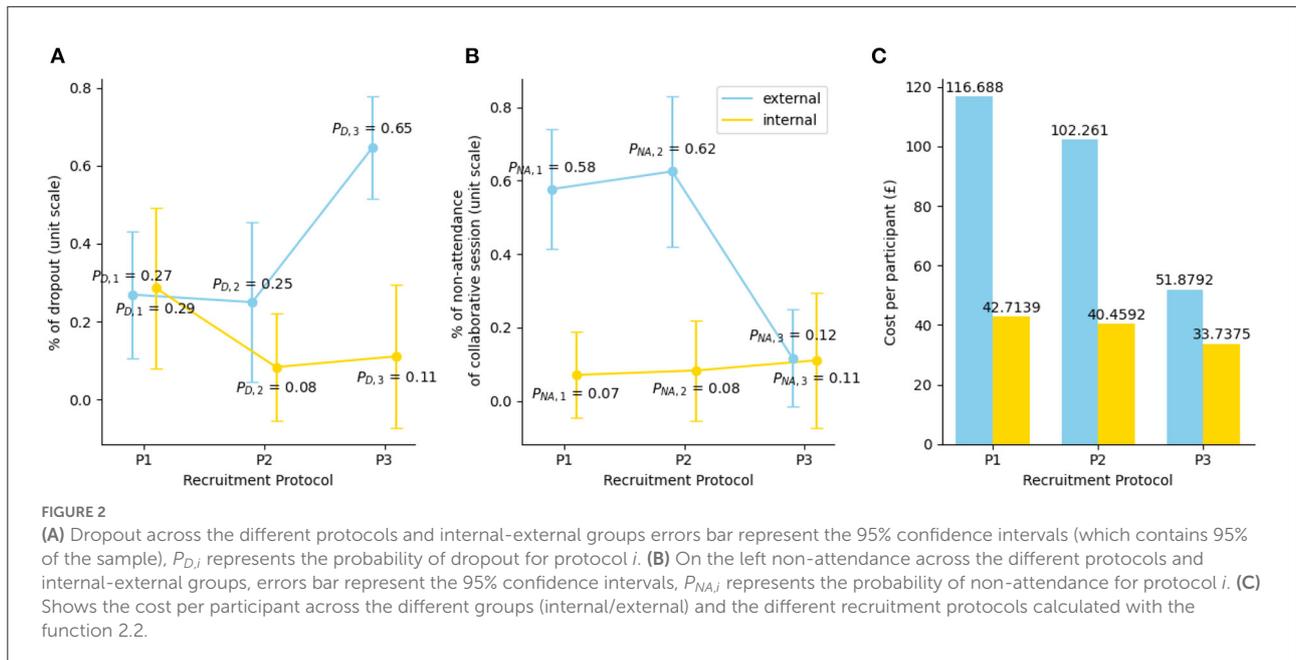


TABLE 3 Post-hoc comparisons - protocol * internal/external.

		Non-attendance				Dropout			
		Mean dif.	SE	t	P_{tukey}	Mean dif.	SE	t	P_{tukey}
p1, ext.	p2, ext.	-0.048	0.129	-0.372	0.999	0.019	0.138	0.139	1.000
	p3, ext.	0.459	0.127	3.616	0.006**	-0.378	0.136	-2.784	0.070
	p1, int.	0.505	0.135	3.744	0.004**	-0.016	0.144	-0.114	1.000
	p2, int.	0.494	0.142	3.473	0.010*	0.186	0.152	1.224	0.824
p2, ext.	p3, int.	0.466	0.158	2.957	0.045*	0.158	0.168	0.940	0.935
	p3, ext.	0.507	0.142	3.577	0.007**	-0.397	0.152	-2.620	0.103
	p1, int.	0.554	0.149	3.714	0.005**	-0.036	0.159	-0.224	1.000
	p2, int.	0.542	0.156	3.483	0.010**	0.167	0.166	1.003	0.916
p3, ext.	p3, int.	0.514	0.170	3.028	0.037*	0.139	0.181	0.766	0.972
	p1, int.	0.046	0.147	0.314	1.000	0.361	0.157	2.301	0.205
	p2, int.	0.034	0.154	0.223	1.000	0.564	0.164	3.436	0.011*
	p3, int.	0.007	0.168	0.039	1.000	0.536	0.179	2.988	0.041*
p1, int.	p2, int.	-0.012	0.160	-0.074	1.000	0.202	0.171	1.182	0.844
	p3, int.	-0.040	0.174	-0.228	1.000	0.175	0.186	0.939	0.935
p2, int.	p3, int.	-0.028	0.180	-0.155	1.000	-0.028	0.192	-0.145	1.000

* $p < 0.05$, ** $p < 0.01$.

and the recruitment procedure [$F_{(2,1)} = 3.446, p = 0.036$], as well as the main effect of internal/external sources on non-attendance ($p = <0.001$). No main effect was found for the recruitment procedure on non-attendance ($p = 0.079$). A Tukey post-hoc test (Table 3) revealed significant pairwise differences between p1/p2, external and the rest of the conditions. The analysis shows how for the external group, the third recruitment protocol has a statistically significant effect in reducing the number of non-attendance participants (Figure 2B, Table 3).

To perform the ANOVA analysis, we used the JASP statistical software package (JASP Team, 2022).

3.4. Dropout

A two-way ANOVA was performed to analyze the effect of internal/external sources of participants and recruitment procedure on dropout. A two-way ANOVA revealed that there

was not statistically significant interaction between the effects of internal/external sources of participants and the recruitment procedure [$F_{(2,1)} = 2.894, p = 0.061$]. However a Tukey *post-hoc* test (Table 3) revealed significant pairwise differences between p3 external and p1/p2 internal. The analysis shows how for the external group the third recruitment protocol has a statistically significant effect in reducing the number of dropout participants (Figure 2A, Table 3). Simple main effects analysis showed that internal/external sources of participants did have a statistically significant effect on dropout ($p = 0.018$). A simple main effects analysis showed that the recruitment procedure did not have a statistically significant effect on dropout ($p = 0.225$).

4. Discussion

From the comparison between internal and external participants, we noticed that a potential participant from the internal group is less likely to not attend or dropout during the process. Nevertheless, while internal groups of potential participants are more reliable internal participants' pools are often not big enough to collect the necessary data for the experiment. Within the academic community of the researchers involved, there were not enough VR HMDs users; therefore, external participants were required.

While reaching out to external participants via advertisement is not complicated, thanks to several platforms for remote experiment recruitment and social network channels tailored to VR users, many of these possible external participants were likely to not attend. Non-attendance increased the experiment's cost of time and the experimenter's frustration. To reduce the costs of our recruitment process, we changed our protocol twice to try and reduce the time invested in the recruitment process.

Protocol 1 required a considerable number of email messages between participants and experimenters to identify a matching date and time for the collaborative experiment. Protocol 2 therefore aimed at reducing the overhead of these email exchanges by removing the first two emails, the email sent by a participant that communicates his/her interest in the experiment and the first email sent by the experimenter offering a list of available dates for the experiment. Instead, the experimenter would receive a notification of the intake form completion, which allowed to save some steps. Nevertheless, the experimenter still experienced high levels of non-attendance.

The changes from protocol 2 to 3 aimed at reducing the negative impact of non-attendance. We noticed during both these protocols (2 and 3) that participants who are coupled with non-attending participants always chose to reschedule and always attended. Therefore, we modified the protocol by introducing an initial inexpensive phase that did not require negotiating multiple participants' time availability and did not require experimenter time to be set

up. This phase would allow potential participants to self-book a short intake meeting. Results show that with the third protocol, we shifted non-attendance participants from the collaborative phase to the intake call. In the intake call, such non-attendance behavior was less damaging because no time was invested in negotiating/matching experiment time and participant availability. Moreover, offsetting part of the reward/payment at this stage (without increasing the total compensation) could have created a bond of trust between participant and experimenter. By looking at the costs for each recruitment protocol and in relation to each group, we can see how the protocol changes were most beneficial for the external recruitment group. Therefore, we recommend the use of this protocol only when researchers are recruiting external participants for the remote collaborative experiment. The cost changes for the internal group were negligible as the impact of potential participants' non-attendance and dropout were minimal for this group.

Ultimately, the changes implemented in protocol 3 had additional benefits: carrying out technical troubleshooting in a face-to-face modality rather than via email or during the collaborative session, potentially wasting the other participant's time. Moreover, in the case of troubleshooting mediated by email, this can potentially be converted into wasted time if, afterwards, the participant would not attend. Phase 2 of protocol 3 allows not only to identify non-attending participants early in the process but also to avoid investing time for technical troubleshooting for participants that may not attend. A final benefit is that the experimenter can verify details from the initial screening and evaluate the environment where the user will perform the experiment, the internet connection speed, the language proficiency level, etc. During the online screening, these details are self-reported. At the same time, participants report them in good faith, and they might under/overestimate their skills correctly (i.e., language proficiency) or make measurement errors (i.e., internet connection). Double-checking this information at the beginning of a collaborative experiment may cancel the collaborative experiment and waste resources instead. If these checks happen on an initial inexpensive phase, then the damaging effects of this miss reported information could be mitigated as well.

4.1. Generalizability of of cost function

In our study, we show how it is possible to apply a procedure of abstraction of the experiment as a sequence of not-overlapping stages required from the experimenter and the participants. Stopping one of such activities results in a loss of time and money. Such abstraction is a critical aspect of our analysis: each experiment could be split into a sequence of activities and visualized as a pipeline. Our method is generic and

can be applied to different experiments that can be described as a sequence of stages with some costs.

With this analysis, we compared three different pipelines, defined in our study as protocols, designing a cost function that is applied to the stages pipeline. Such cost function takes into consideration generic parameters important for defining the cost of each stage (time to complete the activity, cost per participant, salary of the experimenter). Thus, it represents a common template for VR collaborative remote study. In addition, this cost function definition is modular as the final score is the sum of scores from stages. Thus, it can be modeled easily according to specific experimenter requirements. For example, the experimenter can improve the accuracy of the cost function by considering costs per activity where specific materials to rent, rooms or personnel are required. Or can add or remove stages from its protocol. With such a metric, we showed that by adding one simple control activity at the beginning of the protocols, we are able to manage issues such as non-attendance in the collaborative phase and set up technical problems before they lead to a failure during the experiment stage.

5. Limitations and future works

To our knowledge, this is the first study that model the cost of remote VR experiments by introducing a mathematical formalism such as a cost function. This methodology helped propose a preliminary stage of the experiment workflow to reduce costs (time and resources). However, our cost function considered a limited number of variables that can impact the final cost of an experiment. Other variables such as study duration, period of the year, and participants' language may influence the study, and additional parameters could be present in the cost function. However, it is possible to inject in such function additional terms or multiplying factors that consider the impacts on the cost. Such terms can represent new independent elements of the process. New factors show a correlation between terms in the cost function. Another possible limitation is the accuracy of the cost function for those stages whose time can be estimated roughly (e.g., email exchanges) as it depends on uncontrollable factors such as the daily availability of the experimenter and participant or possible differences in the time zones. A better formulation could be achieved with a statistical study that focuses on those specific stages. An alternative is to create a more complex mathematical model to improve the accuracy for the element that requires more accuracy.

6. Conclusion

In this paper, we reported three recruiting protocols that show the evolution from first naive recruitment to a final

protocol with mitigation phases to reduce non-attendance events in the collaborative stage and limit technical problems in remote collaborative scenarios. We highlight how different protocols for recruiting people for a remote collaborative user study lead to different costs per participant in terms of time and money. The lesson learned from the recruitment protocols is that a short pre-experiment one-to-one teleconference allows experimenters to minimize the cost of last-moment cancellation. We demonstrate that such a phase offsets the no-show participant from the collaborative session to the intake call. A collaborative experimental session requires greater organization and time allocation than a short one-to-one intake call, so the overall experiment cost is reduced.

We propose a generalizable/modular cost function that describes the average cost per participant for a completed experiment, and we apply it to our user studies across all the defined protocols. We believe that the research community could benefit from such a modular approach to model and estimate the cost of any experimental process. The COVID-19 pandemic impacted many human activities, including research that utilized human-centered studies. Running a remote experiment seems to overcome the impossibility of gathering participants in running an in-person study within academic premises. On the other hand, such a solution can raise multiple issues, especially in the case of collaborative experiments. We believe that this work will help researchers avoid wasting time and money due to the failure of an experiment session as well as quantitatively compare different recruitment approaches.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: the data from the two previous studies used in this article is waiting for publication. Requests to access these datasets should be directed to RB, rb1619@ic.ac.uk.

Ethics statement

This study has been approved by the UCL Interaction Centre (UCLIC) Research Department's Ethics Chair. The patients/participants provided their written informed consent to participate in this study.

Author contributions

RB ran the experiments and collected the data. DG created the cost function. RB and DG wrote the paper. EC, AS, and TH

reviewed and edited the paper. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by Grant No. 739578 RISE and ICASE with award reference: EP/T517690/1.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Advance Publications (2005). *Reddit*. New York, NY.
- AltspaceVR (2015). *AltspaceVR*.
- Bebko, A. O., and Troje, N. F. (2020). bmltux: design and control of experiments in virtual reality and beyond. *i-Perception* 11:2041669520938400. doi: 10.1177/2041669520938400
- Bovo, R., Giunchi, D., Steed, A., and Heinis, T. (2022). "MR-RIEW: an MR toolkit for designing remote immersive experiment workflows," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (Christchurch: IEEE), 766–767.
- Brookes, J., Warburton, M., Alghadier, M., Mon-Williams, M., and Mushtaq, F. (2020). Studying human behavior with virtual reality: the unity experiment framework. *Behav. Res. Methods* 52, 455–463. doi: 10.3758/s13428-019-01242-0
- Calendly (2022). *Calendly*.
- Feick, M., Kleer, N., Tang, A., and Krüger, A. (2020). "The virtual reality questionnaire toolkit," in *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology, UIST '20 Adjunct* (New York, NY: Association for Computing Machinery), 68–69. doi: 10.1145/3379350.3416188
- Fröhler, B., Anthes, C., Pointecker, F., Friedl, J., Schwajda, D., Riegler, A., et al. (2022). A survey on cross-virtuality analytics. *Comput. Graph. Forum* 41, 465–494. doi: 10.1111/cgf.14447
- Immersive Experiences Working Group (2022). XRDRN. Available online at: <https://www.xrdrn.org>
- JASP Team (2022). *JASP (Version 0.16.1)* [Computer software].
- Klara Skrivanekova (2022). *London Living Wage*.
- Kruusimägi, M., and Terrell, M., and Ratzinger, D. (2022). Available online at: <https://callforparticipants.com>
- Meta (2004). *Facebook*.
- Mottelson, A., Petersen, G. B., Lilija, K., and Makransky, G. (2021). Conducting unsupervised virtual reality user studies online. *Front. Virt. Real.* 2:681482. doi: 10.3389/frvir.2021.681482
- Prolific (2022). *Prolific*.
- Radijah, R., Mäkelä, V., Prange, S., Rodriguez, S. D., Piening, R., Zhou, Y., et al. (2021a). Remote VR studies: a framework for running virtual reality studies remotely via participant-owned HMDs. *ACM Trans. Comput. Hum. Interact.* 28, 1–36. doi: 10.1145/3472617
- Radijah, R., Zhou, Y., Welsch, R., Mäkelä, V., and Alt, F. (2021b). "When friends become strangers: understanding the influence of avatar gender on interpersonal distance in virtual reality," in *Human-Computer Interaction – INTERACT 2021. INTERACT 2021. Lecture Notes in Computer Science, Vol. 12936* (Cham: Springer), 234–250. doi: 10.1007/978-3-030-85607-67
- Rec Room Inc. (2016). *Rec Room*.
- Saffo, D., Bartolomeo, S. D., Yildirim, C., and Dunne, C. (2021). "Remote and collaborative virtual reality experiments via social VR platforms," in *CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. doi: 10.1145/3411764.3445426
- Scott, S., and Johnson, K. (2020). *Pros and Cons of Remote Moderated Testing: Considerations for Ongoing Research During Covid-19*. BOLD Insight.
- Steed, A., Archer, D., Congdon, B., Friston, S., Swapp, D., and Thiel, F. J. (2021). Some lessons learned running virtual reality experiments out of the laboratory. arXiv:2104.05359. 5–7. doi: 10.48550/arXiv.2104.05359
- Steed, A., Ortega, F. R., Williams, A. S., Kruijff, E., Stuerzlinger, W., Batmaz, A. U., et al. (2020). *Evaluating Immersive Experiences During Covid-19 and Beyond*. Technical Report 4.
- VRChat Inc. (2014). *Vrchat*.
- Watson, M. R., Voloh, B., Thomas, C., Hasan, A., and Womelsdorf, T. (2019). Use: an integrative suite for temporally-precise psychophysical experiments in virtual environments for human, nonhuman, and artificially intelligent agents. *J. Neurosci. Methods* 326:108374. doi: 10.1016/j.jneumeth.2019.108374