Case

Running Behavioral Studies With Human Participants Online

Frank E. Ritter[1] ✉ Sarah Ricupero[1]

[1]College of IST and Huck Institute, Pennsylvania State University

Correspondence: Frank E. Ritter, frank.ritter@psu.edu

https://doi.org/srmdro-2022-3901

Abstract

Many studies have been transitioned to be run online so that researchers can continue to collect data during the pandemic. This transition has been supported by the usage of video conferencing and file-sharing tools. We describe four case studies of experiments that we would have run in person and are now running or have run online and highlight the advantages, disadvantages, and how-to's of these studies. In summary, what made our online studies possible was software that could be downloaded and run locally, a strong protocol for the experimenters, and piloting the online version. This approach may be useful in the future because it decreases participant travel time and supports a wider geographic and temporal range of participants.

Learning Outcomes

By the end of this case study, readers should be able to:

- Know some of the advantages, disadvantages, and risks arising in running studies online.
- Evaluate online research studies that are run via videoconferencing.
- Convert an appropriate in-person study to be an online study.
- Develop a new online research study run remotely via videoconferencing.

Introduction

The ability to collect data from human participants through an online medium has become an increasingly popular and necessary method of research during the local COVID-19 pandemic. Due to the COVID-19 pandemic, we had to modify several studies to run entirely through virtual participation. We share our reflections to show that this technique has significant advantages that could make running online an important, and perhaps dominant, approach for running studies after the pandemic is over.

So, we present what we have learned from moving studies online and running them using video conferencing software (e.g., Zoom) as a way to communicate with participants. We first note the studies and then summarize the changes within the sections of a methods section. This approach is an extension to our work on how to run behavioral studies (*RBS*) with human participants (Ritter et al., 2012). The approach to developing studies taught in *RBS* includes piloting and being mindful of ways to reduce the risks of study failure.

This work is also related to how to run studies with the Mechanical Turk platform, a service that allows individuals to participate in online surveys and studies through a web browser (e.g., Paolacci et al., 2010; Ghafurian et al., 2020). The difference is that here, the participants are known to the experimenter, the sessions are more complex and may include multiple pieces of software or multiple sessions, and sessions are run in real time.

Example Cases

We start by providing summaries of four examples of running a study online. We have run three studies using this approach during the pandemic, and a master's student ran a study entirely online that also provides lessons. We present them as examples and to generalize from.

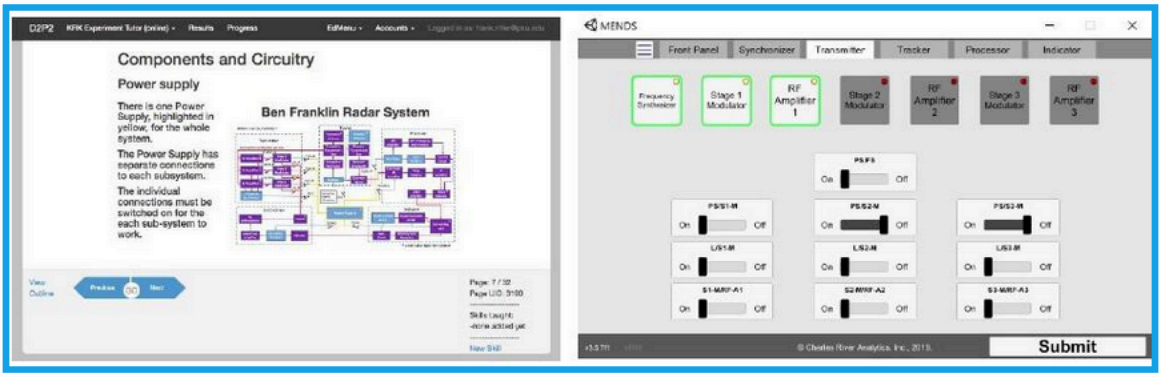
Case 1: KRK Study on Learning Trouble-Shooting

We have been developing a theory of learning and retention, named for the authors as the KRK theory (Kim et al., 2013). To test this theory with a single complex task, we have successfully run a large study (> 400 sessions and > 100 participants) on the effects of learning on retention. The topic, measurements, and basic method of this study are similar to studies in psychology, cognitive psychology, human-computer interaction (HCI), and human factors. It was designed to be run in person and in parallel (e.g., 1–7 participants in a room of computers running the simulation) because of the large number of participants needed and the need for multiple sessions to study learning and retention. Prior to the pandemic, it was piloted (Ritter et al., 2019).

This study includes an online tutor, a simulator in a computer application, drawing on paper, and web-based assessments. Figure 1 shows the tutor and the simulator that participants used. During piloting, participants came into a lab space to use allocated computers.

Figure 1 Citation Count: 2 CITE NOW REPLACE VERSIONS

Figure 1 A screenshot of the tutor (left) and the simulation (right) that participants used.



Alternate Text 95 (maximum 255 characters)

A screenshot depicts two window screens with components and circuity and the corresponding meds

Long Description

The details from the screen on the left reads as follows.

Components and circuitry

Power supply: There is one Power supply, highlighted in yellow, for the whole system.

The power supply has separate connections to each subsystem.

The individual connections must be switched on for the each sub-system to work.

An integrated network depicts the Ben Franklin radar system.

The screen along the right reads as follows.

The title bar reads Mends.

The transmitter menu indicates multiple panels on the screen that reads as follows.

Frequency synthesizer, stage 1 modulator, R F amplifier 1, stage 2 modulator, R F amplifier 2, stage 3 modulator, R F amplifier 3. Stage 2 modulator, R F amplifier 2, stage 3 modulator, R F amplifier 3 are indicated in shaded square boxes.

From left to right the panels depicted in three rows are as follows.

From top to bottom in row 1 the panels read as follows.

P S / S 1 – M: On – Off

L / S 1 – M: On – Off

S 1 – M R F – A 1: On – Off

From top to bottom in row 2 the panels read as follows.

P S / F S: On – Off

P S / S 2 – M: On – Off

L / S 2- M: On – Off

S 2 – M R F – A 2: On – Off

From top to bottom in row 3 the panels read as follows.

P S / S 3 – M: On – Off

L / S 3 – M: On – Off

S 3 – M R F – A 3: On – Off

The transition to remote running was not very complicated for this study. The tutor was adapted to record demographic information instead of a paper form. In person, tasks (such as drawing a schematic of the device to be troubleshot) were converted to an online format (the participant holds the paper up to their web cam for a screenshot). Timers were added in the tutor for tasks that were previously timed by the experimenter. Participants downloaded a simulation and uploaded it including their data file.

This study had a complex, mixed design, requiring participants to return for 1–4 sessions after their initial session, over the course of up to 14 days. Participants were to be compensated with cash or a gift card. After we moved online, participants were compensated with payments into their student account or online gift cards. While piloting, we found that in person participation was limited to about 10 weeks during a semester because of longest condition required an 18 day span of sessions. When running online, we were able to run through the semester and even into breaks. We continue to lose some time at the beginning of the semester because scheduling for everyone is chaotic.

This study ran a total of 134 participants, with 111 providing complete data. Six participants had data loss and 17 dropped out during the process of running the multiple sessions. This data loss rate and dropout rate seem similar to in-person studies with repeated visits. A preliminary report on this project is available (Ritter et al., 2022).

Case 1 Summary

- This complex study on learning and retention used computer-based instruction and assessment, which allowed it to be transitioned to be run online without much difficulty.
- Running live allowed monitoring and in person tasks such as drawing.
- Studies with an extended timeframe (i.e., number of sessions) for participation can be more easily run online to avoid the time restrictions of a standard semester.

Case 2: STRUDEL Study on Simulation Fidelity

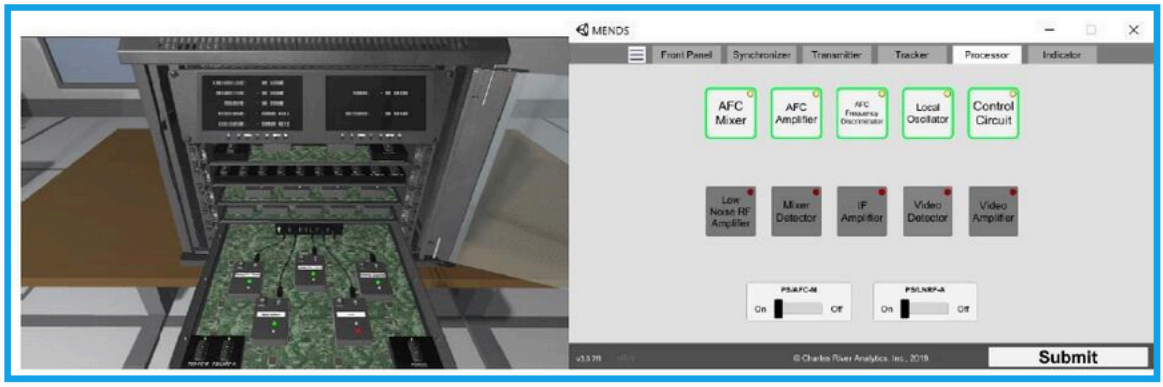
We are working on a project to explore the effects of simulation fidelity on learning outcome, called STRUDEL (Simulating Training Results to Understanding Differing Effects of fidelity on Learning). We have generated a preliminary analysis of the time to use a high- and a low-fidelity interface, suggesting that an appropriate low fidelity simulation in training can lead to greater learning overall (Ritter & McDermott, 2020; Ritter et al., in press).

To test and illustrate these results, we have also run a human study. In this study, participants are run through the Zoom platform. Participants first learn about how to troubleshoot in general and about the simulations they will use through an online tutor (training in high- and low-fidelity interfaces as appropriate). They access the tutor while they and their screen are being observed in real time by the experimenter. This tutor was similar to the one in the KRK study shown in Figure 1.

They then download a simulator to run on their PC. They learn to troubleshoot through repeated practice using either a relatively slow to use though sophisticated simulation of a radar (high fidelity) or learn to troubleshoot using a faster to use but less sophisticated version of the same radar (low fidelity). These simulations are shown in Figure 2. After training, both groups are tested on the sophisticated version. This study is thus a simple between-subjects design.

Figure 2 Citation Count: 1 CITE NOW REPLACE VERSIONS

Figure 2 The high-fidelity simulation (left) and the low-fidelity simulation (right) that participants used.



Alternate Text 117 (maximum 255 characters)

A screenshot of two window screens depicts the hardware components of a simulation model and the corresponding mends.

Long Description

The screen along the right reads as follows.

The title bar reads Mends.

The processor menu indicates multiple panels on the screen that reads as follows.

A F C Mixer, A F C amplifier, A F C frequency discriminator, local oscillator, control circuit.

Low noise R F amplifier, mixer detector, I F amplifier, video detector, video amplifier.

P S / A F C - M: On – Off

P S / S L N R F – A: On – Off

This study was initially based on a physical task with hardware that is reflected in the software for the study (the hardware remains unbuilt because of the pandemic). To be run online, this study simply used two different versions of the simulation (high- and low-fidelity) rather than a software simulation and a hardware apparatus. We have run 50 participants. This study is similar to the KRK study in that the software is basically the same, and the measurements are very similar. The online format does not provide as much contrast between the two different training systems, but it is easier to run because the hardware does not have to be maintained, and recruiting is easier because participants do not have to travel to the hardware.

Case 2 Summary

- Instruction and assessment that are computer-based can be used for multiple similar studies (i.e., the KRK study and the STRUDEL study).
- Physical tasks can have virtual counterparts that can allow for online running.
- Screensharing provides a usual way to monitor and support participants.

Case 3: VITAMMINS Study With Trauma Nursing Tutor and Simulator

We previously conducted a small study examining tutors for teaching nurses about trauma care. This project builds upon previous work testing tutors in person for teaching trauma care (Garrison et al., 2020). We have recently moved an extension to this study online, called Vitammins (Virtual Intelligent Tutor for the Andragogy of Military Medicine Integrated Skills).

The current study has had about 18 student nurses either use an online tutor on the initial steps of receiving a patient into an emergency department (shown in Figure 3) or read the equivalent material in a book on trauma nursing (Emergency Nurses Association, 2014), both at their own pace. Time on task was recorded either by the tutor or the participants in a log when they read the book. They then took a quiz administered online using a learning management system and used a simulation. This was followed by a post-study focus group and debriefing meeting conducted on Zoom to understand the learners' reactions to the material. This was a between-subjects design. There were three dropouts at the end of the main condition and nine dropouts at the one-month post-debriefing.

Figure 3 The online tutor for the VITAMMINS study (left) and the simulation (right).

Alternate Text 120 (maximum 255 characters)

A screenshot of two window screens depicts an assessment level of consciousness (L O C) quiz and the Stat for Vitamins.

Long Description

The screen along the left reads as follows.

Assess level of consciousness (L O C) quiz

Question 4 / 7

What is this patient's level of consciousness?

*Alert

*Vertical

*Pain

*Unconsciousness

Submit button

Educator Navigation (Edit question set)

The screen along the right reads as follows.

The title bar reads STAT: Vitamins v 2.47

Assessments

(Talk to patient

Palpate for...

Auscultate for...

Visually inspect for...

Percuss

Measure vitals

Interventions

*Place C-spine collar

*Remove C-spine collar

*Manual spine stabilization

*Transfer to the operating room

*Use oral suction device

*Use forceps to clear airway

*Perform mouth

*Perform

Reports

*Report AVPU score

*Report G C S score.

The center panels reads as follows.

A patient lying in dorsal position

The text reads the patient's moaning means that they are verbal on the A V P U scale. Click on this box to continue

Select a specific region for Talk to patient followed by ellipses

Talk to the patient menu lists the drop down options as follows.

· Ask the patient name (tick mark)

· Ask patient how they feed

· Ask patient what happened

· Ask patient to squeeze your hand

· Ask patient what day it is

The screen on the right reads as follows.

Vitals

Log

01:25: Performed 'Talk to patient…", Ask patient name.'

00:45: Performed: use extremity tourniquet' on region 'right lower leg'.

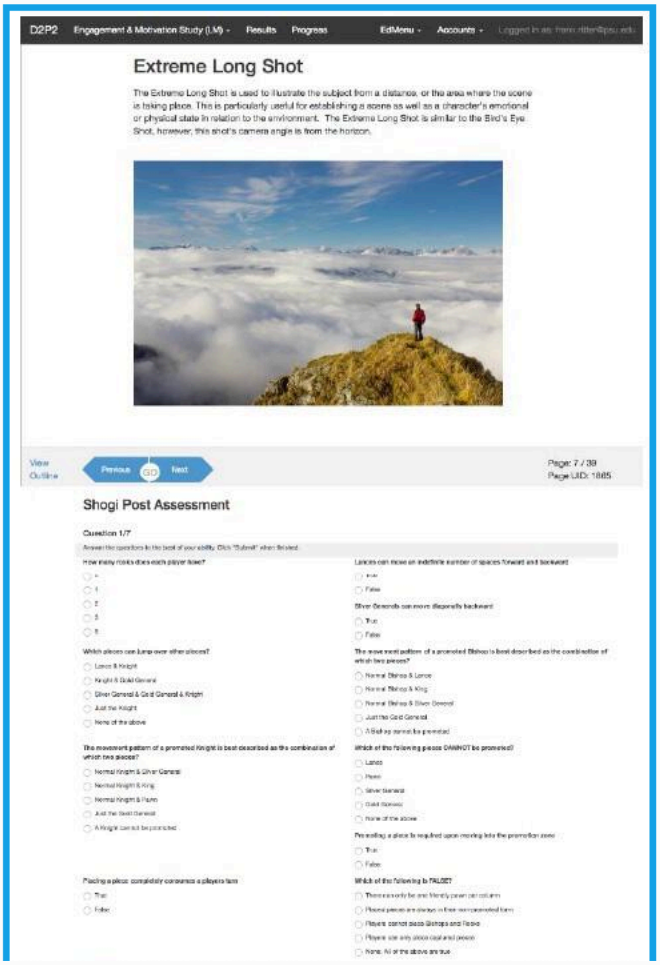
Case 3 Summary

- Virtual instruction that was previously used for an in-person study can be transitioned for online studies.
- Virtual studies can be in many subject areas, including trauma nursing.
- Online studies can provide for participants to work at their own pace when observation is not required.

Case 4: Testing Flow in Using Tutors

Metaxas (2018) as part of his master's thesis examined flow (Csikszentmihalyi & LeFevre, 1989; Csíkszentmihályi, 1990) in online tutors. In this project, Metaxas created two online tutors and measured performance on the tutors (shown in Figure 4). He also had in the tutors a demographic survey before learning and a survey after the learning session about flow.

Figure 4 Two pages from the tutors used to study flow.



Art Title

Alternate Text 169 (maximum 255 characters)

A screenshot of a window screen depicts a panoramic view of an individual standing on top of the cliff amid clouds in an extreme long shot and the Shogi post assessment.

Long Description

The questions of the Shogi post assessment reads as follows.

Questions 1 / 7

Answer the questions to the best of your ability. Click 'submit' when finished.

How many rooks does each player have?

- 0
- 1
- 2
- 3
- 9

Which places can jump over other pieces?

- Lance and Knight
- Knight and Gold general
- Silver general and Gold general and Knight
- Just the knight
- None of the above

The movement pattern of a promoted knight is best described as the combination of which two pieces?

- Normal Knight and Silver General
- Normal Knight and King
- Normal Knight and Pawn
- Just the Gold General
- A knight cannot be promoted

Placing a piece completely consumes a players turn

- True
- False

Lenses can move an indefinite number of spaces forward and backward

- True
- False

Silver Generals can move diagonally backward

- True
- False

The movement pattern of a promoted Bishop is best described as the combination of which two pieces?

- Nominal Bishop and Lance
- Norma; Bishop and King
- Normal Bishop and Silver General
- Just the Gold General
- A Bishop cannot be promoted

Which of the following pieces CANNOT be promoted?

- Lance
- Pawn
- Silver General
- Gold General
- None of the above

Promoting a piece is required upon moving into the promotion zone

- True
- False

Which of the following is FALSE?

- There can only be one friendly pawn per columns
- Players cannot place Bishops and Rooks
- Players can only place captured pieces

None. All of the above are true.

Metaxas recruited 75 participants through social networks including email and social media posts. He ran them over one weekend. In this within-participants design, five sets of data were lost due to incompleteness. Participants ran themselves by accessing the link to the tutors.

Case 4 Summary

- An online study can make data collection for a degree easier, such as for a master's thesis.
- Online studies can allow participants to complete a task in their own time frame.
- Running online can make recruiting easier because participants do not have to travel to the study.

How to Run a Study Online Organized by Method Subsections

While we have developed methods for running studies remotely as a result of the pandemic, these methods can still be useful in future environments. We have discovered the benefit of reaching populations that are further from the university in time or space, which can allow a more diverse subject pool. This benefit has also been observed in remote nursing studies (Cantrell & Lupinacci, 2007). Additionally, we have been able to continue running through typical school breaks (i.e., summer and spring break), which is beneficial when participants must return for multiple sessions.

This section introduces the changes we made to our typical method to accommodate running online. The changes are grouped by the method section where they would apply.

Participant Recruitment and Contact

For case studies one and two, we made similar adjustments to participant recruitment. All contact with participants occurred via email. While the use of telephone calls was also discussed, email was the easiest contact information to collect, and it allowed the recruitment process to occur over the course of several days and weeks. This form of contact also allowed experimenters to remind participants of their session time and reschedule if necessary and provided a record. After initial contact, we emailed potential participants twice. Running this study did require extending our Institutional Review Board (IRB) protocol to note that we were recording participant emails and, in some cases, ID numbers (when they were paid through their registrar account).

While many studies recruit by posting fliers throughout campus, the pandemic made this means of advertisement far less useful. We found that sending recruitment emails out to professors and having these professors post our recruitment query on a course web page or forwarding our email allowed us to find a large pool of participants.

For the KRK and STRUDEL studies, we emailed between one and three faculty members, twice a week, asking for them to share the study announcement with their class. Faculty typically taught one or two courses with 20–100 students. We do not have data on how they shared it with their classes. We believe that some forwarded the email and that some posted the request as an announcement into a course management system. The number of participants recruited per class appeared to vary between 0 and 5 per class. If you have access to classes through their instructors or other methods such as announcements in class, this approach can replace or augment appearing in class and making a small presentation, which could have been done using Zoom. It is very useful when you cannot attend the physical classroom. In the case of the VITAMMINS study, one of the investigators gave a brief 5-minute presentation on the study during a nursing class using Zoom and then followed it up with an email.

We also recruited by email asking student groups (e.g., an honors program and a student-oriented church), as well as posting recruitment details on social media. These seemed to yield a higher number of participants; they were also larger groups, but there were not as many of these groups available as there were classes. We recommend this approach when it is available.
For the VITAMMINS study, we recruited in the relevant nursing class. Metaxas in his study used email and social media to recruit participants.

Participants were required to have access to a laptop or desktop computer, with a microphone and camera, and in one study, a piece of paper and a writing implement. Some software required a particular operating system, which was noted on recruitment materials.

We were able to compensate participants remotely in two ways. We could pay into their bursar's account, and we could provide gift cards by email. Class credit could be easily handled as well. It is useful to provide a cash equivalent in case participants do not wish to support the vendor of the gift card. In the studies offering both choices, the ratio was approximately 70% gift cards and 30% payments into bursar account.

Overall, through using emails to classes and groups, and perhaps a few recruited through posters, we were able to recruit over 200 participants for these studies during a variety of lockdown conditions, including when there was no lockdown and studies could be run in person. Emails, professors, and social media work as ways to recruit participants; however, we do not know if this method recruits different kinds of participants than previous approaches or can we quantify just how efficient it is.

Section Summary

- Emailing interested participants can be a useful form of contact and can even allow for reminders leading up to the time of participation.
- There are ways to pay participants online.
- Online class sites can be useful for advertising the study.

Materials

We had to make our materials run online or remotely. In the KRK, Strudel, and VITAMMINS studies, we had participants download a zipped folder containing the software needed. The results file was stored in the directory of the application, which was compressed at the end of the session and sent back to the experimenter using Zoom chat or email. The effect of the initial download time was decreased by starting it and then having the participants use a browser-based tutor. Additional online tutors were used for training materials and declarative memory quizzes. The tutor data was self-contained on our server. In the Metaxas study, he ran participants using a standalone online tutor.

Thus, the materials have to be able to run on participants' machines. You can either do this by using a web site or sending software to them via the communication media to be downloaded (e.g., Zoom) or have them download it from a web site. We chose to have them download it from the communication media so that they could not return to the system.

If you think running the software materials may be attractive to participants, you should take steps to limit reuse. If you have a website-based set of materials, you might wish to create a unique set of paths, one for each participant, and then remove the links afterward to avoid the materials being shared. Software remaining on remote machines remains a concern. In our examples, each participant had a unique account in the online tutor.

You need to keep in mind the size of the files that will be shared and make adjustments or find a way to share large files. An additional study that was run (not described here, Oury, 2022) used Zoom's remote-control function to provide access to experiment materials hosted from the experimenter's computer. Recording the data via remote control on your own machine also can alleviate privacy risks associated with some types of data collection (i.e., keystroke logs) by ensuring the recording software is never on the subjects' personal computers (Kim & Ritter, 2007), but on the experimenter's computer.

You will want to be cautious about updating your online study materials while running a study. During an in-person study, you can generally avoid modifying your materials, particularly software ("freeze" it); the hardware can be stable and can often be disconnected from the Internet to avoid automatic software updates.

Web-based systems can be more difficult to avoid modifying. This is particularly true if someone else is administering the machine your software is on. There are routine updates to web service software (e.g., PHP, Ruby, and SQL), and some universities require applying these updates, without regard to the risk to a running study. These updates can break some web-based systems. Our tutors, for example, could be and have been broken by a change in a library that was updated.

The risk of software updates is particularly a risk the longer your study runs. Some of our studies have had relatively long running periods because of the size, complexity, and special participant population.

If your study will have a long run and uses a web site to provide the stimuli, we recommend moving your software to a server that can be easily rebuilt and that is set up to not accept automatic updates. You might also like to have a backup site. Sometimes this means moving your website off campus to a fixed server that does not have to submit to the updates for security reasons (but has a concurrent risk of being broken by the lack of updates). In any case, be mindful of automatic software updates to your web-based stimuli.

Section Summary

- When running a study online, you must be deliberate about data collection and storage. This can include the use of a web browser that stores data or having a participant send a data file to the experimenter.
- A number of tools exist for running studies on the participants machine, for letting participants run on your machine, and for observation of participants.
- When passing large files, practice and check upload and download speeds locally.
- Using a server to support a study that is not under your control should only be done for short-term studies because the server can change due to software updates.

Running participant sessions

A number of tools have been developed to allow the collection of data from participants to be conducted remotely (e.g., Mechanical Turk; Pavlovia: https://pavlovia.org/; Prolific: https://www.prolific.co/; EPrime: https://pstnet.com/). While Mechanical Turk and Pavlovia are open source tools, Prolific and EPrime are not. Some universities provide site-license-based access to software such as EPrime, but this is not always the case for individuals looking to conduct online research. These systems allow any individual who meets a study's criteria to participate in it in their own space and on their own schedule or to create portable study software. While these tools have been useful, our tasks were too complex to fit easily into these systems.

Experimenters were each able to run about eight one-hour sessions per week, the number depending on participant and experimenter schedules. In addition to the participant not having to travel, the experimenter did not have to travel to the lab and thus both could slightly extend the range of their running times to include early evening. When participants were late, which occurred, they were emailed at 5 and 10 minutes past the starting time of their session. Because they were not traveling across campus to the experimental session, they could quickly attend the study by just attending the Zoom room.

Our studies used Zoom Professional to video chat with participants, both because our university provides Zoom and because Zoom allows file sharing of large files that are necessary for our studies. Other tools exist, such as Adobe Connect, which contain similar features. In many cases, these tools will be equivalent.

When running online, the participants' environment may be more distracting than a study room, and your presence via video chat is less than would be available in a running room. You will also have to be concerned about task complexity and the ability to direct attention. In our running scripts, we have to ask our participants for more help than during an in-person study. We instruct them to close doors, turn off music, and mute their phone.
Your ability to direct attention and keep participants focused on the task may be reduced via video chat compared to an in-person study. So, you should pilot the protocol to examine how to use the video interface, how to notice problems in the participants' space, and how to ask the participants to ameliorate any distractions. However, we did not have to drop a session for this problem across these studies.

Section Summary

- You can use commercial or homemade software when running experimental sessions online.
- Video-sharing tools contain useful features for running studies, and these tools should be used to one's advantage when designing an online study that is run in real time.
- Include in your protocol steps to reduce interruptions and distractions in the participants space, such as asking for cell phones to be turned off and doors closed.

Design and Procedure

The procedure of these studies was held as closely as possible to what would be expected for an in-person study. The KRK, Strudel, and Vitammins studies were based on in-person protocols. Extra directions were included for ensuring the participant was in a quiet space and for downloading the necessary software. The studies we have run included within- , between-participant, and mixed designs. This approach appears to support a wide range of designs.

We were able to train our research assistants more easily than in-person studies because they could easily come to watch participants being run in a session. The presence of an additional experimenter in a Zoom room seemed less intrusive than being in a physical room with the participant.

We found out early on that we had to have participants clear their autofill entries for assessments that were run in a browser. If we did not do this, the browser would automatically provide answers during assessments! Our work around for this issue was using the private browsing feature available in most Internet browsers (i.e., incognito mode in Google Chrome); using private browsing to access experiment websites should likely be the default choice when possible. We did not have any difficulties running our study based on different browsers or machines; our software and web interfaces functioned well across all machines. This will not be the case for all software and studies.

Section Summary

- Many types of studies can be run online (within-participant, between-participant, and mixed designs).
- An experimenter should be familiar with any program being used in the study, and the settings of such programs should always be checked for autofill.
- Training research assistants can be easily and unobtrusively done in a video conferencing tool.

Data Collection

Because our study was already designed to be run on a computer, the programs and assessments contained logs of participant behavior. These logs were either saved on the website server being used or were recorded in the program files and sent back to the experimenter after the session was run. The data recorded were thus the same as an in-person study.

Additionally, where we would have collected a participant's drawing, we took a picture using the web camera and screenshot to record participant drawings. This sometimes resulted in low-quality images. The use of a cell phone camera to record the image could fix this problem.

In these studies, we used cloud storage systems to support file sharing across experimenters, and as a way to back up data. We also backed up data in some cases to hard drives. We used Qualtrics, an online survey tool to implement informed consent forms and to gather demographic and other survey data. Qualtrics has a free online tool that allows individuals to create surveys that can be used to collect and analyze data (www.qualtrics.com). We have learned to look at the Qualtrics default parameters, and to turn off geolocation, which we did not need to record.

Section Summary

- Screenshot features of video conferencing tools can be useful for collecting records of physical creations (i.e., drawings). An experimenter should be familiar with any program being used in the study, and the settings of such programs should always be checked for autofill.
- Online storage systems can be useful for data storage in any type of study.
- Online survey tools exist to support data collection; you need not develop your own.

Conclusion

We found through these four studies that we could run studies online with minimal impediments and in some ways better and faster. The three ongoing projects are proceeding toward publication of the experimental data. We were able to get all these studies approved by our local IRB with no more difficulty than previous studies.

All the arguments for remote learning can apply to remotely running studies. We have discovered the benefit of reaching populations that are further from the university in time or space, which can allow a more diverse subject pool. This benefit has also been observed in remote nursing studies (Cantrell & Lupinacci, 2007). Additionally, we have been able to continue running through typical school breaks (i.e., summer and spring break), which is beneficial when participants must return for multiple sessions. Table 1 summarizes the general benefits and limitations of our online method.

Table 1 Advantages and disadvantages of remote running.

	A	B
1	***Advantages of Remote Running***	***Disadvantages of Remote Running***
2	*Not limited by physical space* *Greater geographic range of potential participants, e.g., multiple Penn State campuses*	Not able to share or use physical things like button boxes or props as easily Susceptible to Internet interruptions
3	*Allows sessions to occur when students are traveling, e.g., university breaks* *Can run more easily at early evening or early morning times* *Less transit time (uncompensated) by participants* *No greater transit time, perhaps less, for experimenters* *Less concerns about experimenter or participant physical safety because they don't have to travel or physically interact with each other*	At home environment, less control of audio, noise, and interruptions Participants might have to have the right kind of computer May have to install software on participant's computers Computers and displays may vary between participants Have to pull data back from participant's machine

There are limitations to online approaches. This approach will not be applicable to all situations. Running remotely will be difficult to perform when special hardware is required, such as a hardware implementation of the high-fidelity system. Similarly, this approach will have difficulty when a button box is required to get millisecond-accurate timing, although a button box could be delivered or shipped to users and returned by hand or post.

The apparatus software has to be slightly robust to different machines. The software to be run cannot be too large to download in a reasonable time. In the next online study in our lab, the software is larger than in these studies, so we are having the participants remotely access a desktop where the software is running.

However, running studies remotely using a video chat application to interact appears to support a wide range of studies, including HCI, computer-based human factors, usability studies where you can download or access the materials, and cognitive psychology where timing is not as important or the software to do timing can be downloaded. This approach opens up the ability to run studies during pandemics but also allows the ability to recruit participants from around the world and can save both the experimenter and the participant time. The next study in our lab was designed to be online, not just because of the pandemic, but also because it appears easier to recruit and to run.

Discussion Questions

1. What are some limitations to running studies online?
2. What types of studies can and cannot be moved online?
3. What are some risks to study success for studies run online?
4. What factors can lead to online studies getting different results than an in-person study?
5. Will on-line studies recruit different types of participants?

Multiple Choice Quiz Questions

1. What is the easiest way for data collected on a participant's personal computer be stored for analysis?
    a. By saving data on a password-protected website.
    b. By saving data in an application and sending the application to the experimenter through the video conference tool. CORRECT
    c. Having the participant email a file from a subdirectory.
2. Which is a benefit of running behavioral studies remotely?
    a. More variance in the running environment.
    b. Greater geographic range of participants. CORRECT
    c. Susceptible to Internet interruptions.
3. Which tool can be used to run participants in real time?
    a. Pavlovia.
    b. Mechanical Turk.
    c. Zoom. CORRECT
4. Which of the following is a disadvantage of running a study online?
    a. Allows less flexibility of times that study can be run.
    b. Can be challenging to run if special hardware is needed. CORRECT
    c. Limits the geographical access to potential participants.

Metadata

Acknowledgments

The work was supported by ONR (N00014-15-1-2275). Much of the work was done with Charles River Analytics supported by DHP (W81XWH-17-C-0002) and ONR (N00014-18-C-7015). The simulation software used in the studies was created by and with Charles River Analytics and are used under license. Comments from Marc Burns, Chris Garrison, Jong Kim, Elizabeth Meschia, Jacob Oury, Connor Reidy, Clare Robson, Martin Yeh, and Steve Zimmerman, and discussions with Ashley McDermott helped improve this report. Benjamin Bauchwitz, Mat Brener, Chad Chae, Na Li, Ashley McDermott, James Niehaus, William Norsworthy, Jacob Oury, Raphael Rodriguez, Farnaz Tehranchi, Shan Wang, Peter Weyhrauch, and Martin Yeh helped create and run the KRK study. Arunima Maheshwari, Ashley McDermott, James Niehaus, William Norsworthy, Sarah Stager, Peter Weyhrauch, and Martin Yeh helped create and run the Strudel study. Benjamin Bauchwitz, Alexis Fenstermacher, Chris Garrison, Nicholas Phillips, James Niehaus, and Jacob Oury helped create and run the VITAMMINS study.

Frank E. Ritter is required by Pennsylvania State University [sic] to include this paragraph: "Frank E. Ritter, the co-author of this paper, has financial interest with Charles River Analytics Inc.; a company in which Frank E. Ritter provides consulting services and could potentially benefit from by the results of this research. The interest has been reviewed and is being monitored by the Pennsylvania State University in accordance with its individual Conflict of Interest policy, for the purpose of maintaining the objectivity of research at the Pennsylvania State University." [sic]

Further Reading

Morgan, J. H., Cheng, C.-Y., Pike, C., & Ritter, F. E. (2013). A design, tests, and considerations for improving keystroke and mouse loggers. *Interacting with Computers*, *25*(3), 242–258.

Nussenbaum, K., Scheuplein, M., Phaneuf, C. V., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: Comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra Psychology*. https://doi.org/10.1525/collabra.17213

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgement and Decision Making*, *5*(5), 411–419.

Ritter, F. E., Kim, J. W., Morgan, J. H., & Carlson, R. A. (2012). Running behavioral studies with human participants: A practical guide (1st ed.). SAGE Publications, Inc.

Scott, K., Chu, J., & Schulz, L. (2017). Lookit (part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*, *1*(1), 15–29.

Web Resources

http://frankritter.com/rbs/

https://www.thechildlab.com/

https://nyu.databrary.org/

References

Cantrell, M. A., & Lupinacci, P. (2007). Methodological issues in online data collection. *Journal of Advanced Nursing*, *60*(5), 544–549. https://doi.org/10.1111/j.1365-2648.2007.04448.x

Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, *56*(5), 815–822. https://doi.org/10.1037//0022-3514.56.5.815

Csíkszentmihályi, M. (1990). *Flow: The psychology of optimal experience*. Harper and Row.

Emergency Nurses Association. (2014). *TNCC Trauma nursing core course provider manual* (7th ed.). Emergency Nurses Association.

Garrison, C. M., Ritter, F. E., Bauchwitz, B. R., Niehaus, J., & Weyhrauch, P. W. (2020). A computer-based tutor to teach nursing trauma care that works as an adjunct to high-fidelity simulation. *Computers, Informatics, Nursing*, *39*(2), 63–68. https://doi.org/10.1097/CIN.0000000000000637

Ghafurian, M., Reitter, D., & Ritter, F. E. (2020). Countdown timer speed: A trade-off between delay duration perception and recall. *ACM Transactions on Computer-Human Interaction*, *27*(2).

Kim, J., & Ritter, F. E. (2007). Automatically recording keystrokes in public clusters with RUI: Issues and sample answers. In *Proceedings of the 29th Annual Conference of the Cognitive Science Society, 1787*. Cognitive Science Society.

Kim, J. W., Ritter, F. E., & Koubek, R. J. (2013). An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science*, *14*(1), 22–37.

Metaxas, L. R. (2018). *Impacts of user sentiment on information recall, intrinsic motivation, and engagement in the context of intelligent tutoring systems*. Unpublished MS thesis. https://acs.ist.psu.edu/papers/metaxas18.pdf

Oury, J. D. (2022). *Dismal interrupted – how interruptions affect the learning and retention of complex procedural tasks* [Unpublished PhD thesis]. College of Information Sciences and Technology, The Pennsylvania State University.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgement and Decision Making*, *5*(5), 411–419.

Ritter, F. E., Kim, J. W., Morgan, J. H., & Carlson, R. A. (2012). *Running Behavioral Studies with Human Participants: A practical guide* (1st ed.). SAGE Publications, Inc.

Ritter, F. E., & McDermott, A. F. (2020). The effect of task fidelity on learning curves. In *Proceedings of the 18th International Conference on Cognitive Modeling (ICCM 2020)* (pp. 229–235).

Ritter, F. E., Ricupero, S., Yeh, M. K., Workman, D., Oury, J. D., & Stager, S. J, et al. (2022). *Testing a learning and retention theory using a troubleshooting task* [Technical Report no.ACS 2022-1]. Applied Cognitive Science Lab, Penn State.

Ritter, F. E., Tehranchi, F., Brener, M., & Wang, S. (2019). Testing a complex training task. In *Proceedings of the 17th International Conference on Cognitive Modeling (ICCM 2019)* (pp. 184–185).

Ritter, F. E., Yeh, K.-C., McDermott, A. F., & Weyhrauch, P. W. (in press). The effect of task fidelity on learning curves: A synthetic analysis. *International Journal of Human-Computer Interaction*.