Everyday Locations as Cues to Smoke:

Personalized Environments in Virtual Reality to Elicit Smoking Cravings

Alkis Antoniades

TU Delft

Everyday Locations as Cues to Smoke:
Personalized Environments in Virtual Reality to Elicit Smoking Cravings

by

Alkis Antoniades

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday April 5, 2022 at 13:30.

Student number:	4748147	
Project duration:	November 30, 2020 – April 5, 2022	
Thesis committee:	Dr. ir. W. P. Brinkman,	TU Delft, supervisor
	Dr. O. E. Scharenborg,	TU Delft
	Ir. N. Albers,	TU Delft, daily supervisor

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Abstract

Smoking is a leading risk factor negatively impacting the health of people, not only those partaking in it first-hand, but also to those around them. Different methods are available to assist people with quitting smoking, with various degrees of effectiveness. Researchers developing smoking cessation approaches would like to have controlled environments to test how effective they are before offering them as viable options. Virtual reality has been demonstrated to be an efficacious tool for facilitating the presentation of cues aimed at eliciting smoking cravings in the lab. Addiction, however, is a complex matter involving different parts of the brain, and how or when conditioned responses causing smoking cravings are activated depend on the individual in question. There are strong indications that personalization of smoking cues, or in other words using elements relevant in the addiction model of an individual, may elicit higher or at least consistent smoking cravings. The general aim is that developing smoking cessation approaches targeting those most relevant elements may have higher ecological validity and therefore be more effective in assisting people both with quitting smoking and maintaining that behavior change. To assist researchers with testing their smoking cessation approaches using a virtual coach, we have created a system enabling the presentation of virtual environments and facilitating communication between researchers and participants while the latter are viewing them. We have evaluated the effectiveness of our system in eliciting a familiar experience, which we posit is a major component tied to smoking cravings. Our results showed that personalized virtual environments elicited a more familiar experience than non-personalized ones. We also examined the usability of our developed user interface, as well as the sense of presence elicited by our system, both of which received positive scores.

Contents

1 Introduction

1.1. Problem Definition

Smoking is the second-highest leading risk factor globally to people's health and well-being, behind high blood pressure [69]. While there has been a decrease in smoking worldwide, it still is one of the most significant contributors to health problems in the world [93]. In 2015 in the United States, 68% of adult smokers said that they wanted to quit, in 2018 55.1% said that they had made a quit attempt in the past year, and 7.5% of adult smokers stated that they had successfully quit in the past year [89]. Both quit attempts and successes have seen an average rising trend globally [93], however, many people fail to quit successfully or relapse after some time. Therefore, it is important to develop and offer a wide range of smoking cessation methods for people to have at their disposal, so that they may choose one which suits them and their lifestyle. Much research has been done and is ongoing regarding smoking addiction and how people can be assisted with quitting. Researchers, however, would greatly benefit from having controlled environments in which to test whether their proposed methods are effective. Virtual reality approaches have demonstrated efficacy in eliciting smoking cravings in the lab [64], and we want to investigate how the technology can be used to assist researchers in their work.

1.1.1. Smoking Cessation Research

Evidence shows that behavioral counseling interventions are more effective in helping people quit smoking compared with no treatment or self-help material [89]. However, the same source reports that more than two-thirds of adult cigarette smokers who attempted quitting in 2019 did not use evidence-based treatment. This not only points to there being a need to develop and promote evidence-based treatments such as behavioral counseling, but to also make them easy to access and use, enabling people to easily take advantage of these approaches.

Researchers are continually investigating methods to not only help people with quitting smoking but also to develop effective behavior maintenance approaches to prevent them from relapsing. Either by improving upon current methods, by combining methods or by applying different principles of psychology with demonstrated efficacy elsewhere, to test whether they may also apply to combating smoking or other substance addictions [9, 23, 29, 37, 57, 61].

Smoking addiction research can typically be classified into either of two main categories. The first involves investigating cravings themselves, including cue-presentation methods, or the effects of variables such as demographic information or specific conditions, e.g. level of nicotine addiction. The second category involves using craving-elicitation methods with demonstrated efficacy, to test proposed smoking cessation approaches. Essentially, research in the second category usually builds upon findings of studies from the first one. Therefore, it is important to develop and demonstrate the reliability of methods to elicit smoking cravings, which others can employ in their own work relevant to either of the two aforementioned categories.

1.1.2. Smoking-Craving Elicitation

Cue-reactivity research involves eliciting cravings via exposing people to different types of cues [64]. These include:

1. Proximal cues, which are specific objects that are often associated with substance intake, such as cigarettes, ashtrays, lighters, in the case of smoking.
2. Contextual cues, such as different situations in which substances may be used, e.g. walking out of a building may trigger smoking cravings because it is the first opportunity to do so after being in a smoking-restrictive situation.
3. Complex cues, which relate to combinations of proximal and contextual cues, for instance a party, which would be a social event context where people are interacting and engaging in potentially triggering situations, such as drinking alcohol, smoking, offering / being offered cigarettes, etc.

In research using conventional (non-VR) means, these can be presented via:

1. Imagery scripts, where participants read or listen to a narration of various scenarios and are asked to imagine different things.
2. Photographs / images / videos depicting various objects, such as smoking paraphernalia, and/or potentially triggering situations.
3. In vivo exposure, where participants look at or interact with various objects such smoking paraphernalia, and/or are put in different potentially triggering situations, e.g. watch people smoke.

Combinations of cues and presentation methods may elicit cravings associated with different aspects of addiction, such as triggering memories, activating learned responses. [20, 31]. However, not all cue types are easy, or even particularly feasible for many research projects to utilize using traditional means, such as in vivo presentation. This is primarily because some cue types require an increasingly complex operational setup, have additionally imposed physical constraints, and consequently would require a project to have more funding. Even proximal cues which may be rather uncomplicated to present, necessarily require participants to be physically attendant in the lab. Therefore, alternative approaches have been investigated to perhaps help overcome this inherent complication, by utilizing virtual presentation methods to expose participants to various types of cues.

Furthermore, apart from the aforementioned inherent complexity in presenting a variety of cues in the lab, there is also the risk of having unaccounted-for variables skewing experiment results, e.g. environmental factors. And as such, it is important to have controlled environments in which to test not only the effectiveness of proposed approaches, but also various hypotheses aiming to investigate specific elements of addiction and how researchers may address them.

1.1.3. Virtual Reality in Cue-Reactivity

The use of Virtual Reality (VR) in cue presentation aimed at eliciting smoking cravings has been receiving increasing attention, and arguably for good reason. Virtual reality is a computer-based technology where users can be immersed in virtual 3D worlds of varying realism, offering different levels of navigability and interactability. It has been employed to assist in the assessment and even the treatment of various psychological conditions, such as post-traumatic stress disorder, different types of phobias, anxiety, eating disorders, addictive behaviors, etc. [42, 49, 64, 73]. The efficacy of using VR in eliciting not only smoking, but also other substance cravings such as drugs and alcohol via presenting different types of cues, has been demonstrated in many studies [10, 12, 30, 36, 42, 44, 48, 51, 60, 64]. The technology offers the distinct advantage of allowing people to have a variety of experiences through different virtual environments (VEs), perhaps similar to the real-world, and to be exposed to complex situations such as social interactions, with high ecological validity [11, 64].

Some studies have been conducted to identify complex-cue environments which are perhaps more effective than others in invoking smoking cravings in participants [36, 60, 87], perhaps pointing towards a trend in attempting to standardize nicotine craving research [11]. And indeed, some form of standardization for the overall methods used to elicit cravings may benefit research, by providing demonstrably and reliably effective methods for presenting cues leading to craving elicitation, to help in developing approaches aimed at helping people.

Finally, controlled environments inherently allow for the manipulation of different variables, to help disentangle them so that researchers can study specific potentially influencing factors. This is something VEs excel in due to their inherent flexibility in being malleable allowing for their complete control.

Therefore, the utility of employing virtual technologies becomes evident, especially when considering the complexity of experimental design imposed when wanting to perform analogous tests in the real world.

1.1.4. Research Limitations

There are, however, certain subjects which research has yet to address. Firstly, while moving towards standardizing which environments and settings to use in eliciting cravings may help simplify certain aspects of the research, addiction has many complex and individualized aspects [34, 38, 53, 85]. Therefore, we cannot always assume that cravings elicited by standardized generic scenes are necessarily as high as we can obtain, which would allow researchers to accurately test their proposed approaches. This may be due to the absence of certain cues individuals find important, or simply that those individuals are more susceptible to specific combinations of cues presented. This is demonstrated when considering that some cues may be more prevalent in some people's lives or more strongly associated with smoking. For instance, more heavily addicted people are more prone to smoking in the first few hours after they wake up, while those who are not are less likely to exhibit that behavior [63]. Or, for instance, young adults may differ from other smokers in terms of cue responses [86]. Therefore, precisely because of this complex nature of addiction which is not completely understood [20], it may be the case that personalizing VEs to the addiction model of each person could maximize their potential for craving elicitation. This is something that could contribute to field of cue-reactivity and help researchers with their own work, both in determining what cues are most effective, and in developing effective treatments.

Therefore, it becomes more apparent that it may be difficult or even infeasible to create comprehensive ”one-size-fits-all” solutions. And that while certain aspects of VEs could be standardized to fit most cases, others may require adapting them to people's individual addiction models to enhance the potential that VR can offer in eliciting smoking cravings.

1.2. Research Questions

Some work has already been conducted to investigate the personalization of certain types of craving cues [28, 29, 82], with findings suggesting that the impact it can have on cue reactivity may strongly depend not only on the method of cue presentation [31], but also on the type of addict [20]. There has not, however, been much research into explicitly comparing the effect that personalization has on cues presented in VR, which we envision may hold an advantage over generic environments that most studies have used.

That is, therefore, the main question this research project aims to address:

How can personalized virtual reality environments be created to help with nicotine cue-reactivity assessment?

This can be broken down further to formulate sub-questions:

1. What are the requirements for building virtual environments to elicit smoking cravings?
2. How do we design personalized virtual environments to elicit smoking cravings, for the purpose of enabling researchers to test their conversational agent approaches?
3. How do different virtual environments compare with each other in terms of the magnitude of self-reported cravings elicited in each, and how do users evaluate the usability of a system developed to present these virtual environments?

1.3. General Approach

To address these questions, we started in Chapter 2 by first exploring the relevant fields of psychology and addiction research to examine how others have created VEs to elicit smoking cravings. Investigating their findings also helped us determine and compile the different requirements which have informed our design decisions. These include functional, operational, and technical requirements, which were formulated by also examining research in the fields of games development and virtual reality, as well as by consulting researchers working on *Perfect Fit* [62], which this project is a part of. The aim here was to obtain information regarding standardized tobacco-craving and neutral (no-craving) scenes that have been used in cue-reactivity assessment research, as well as to determine methods by which to implement them in a way that fits our operational demands, which is discussed further in Chapter 3.

Finally, we designed and conducted an experiment to test the effectiveness of standard and personalized smoking environments, as well as neutral environments of both aforementioned types, presented in Chapter 4. This was done by examining how familiar the experience that was elicited by each, which we posit either has a causal or at least positive correlational relationship with smoking cravings. We additionally had participants in our experiment rate the usability of the system we developed to facilitate presenting these virtual environments, including the user interface implemented to allow researchers to test smoking cessation methods using a virtual agent.

2 Foundation

Having briefly touched upon the general subject matter, we now wanted to delve deeper into already-conducted research so that we could gain a better understanding of the domain and how we could contribute to it. This was done in an effort to answer our first research sub-question, namely:

What are the requirements for building virtual environments to elicit smoking cravings?

To undertake answering this question, we initially needed to examine the relevant research literature to gain sufficient insights. Firstly, however, we found it prudent to also have an idea of what information we wanted to extract from the domain to formulate a sufficient answer. Therefore, we started by breaking down our first research sub-question into:

1. What are the different types of cues which can feasibly be presented in a virtual environment?
2. How can different types of cues be presented in a virtual environment?
3. How do we decide which cues to present?
4. What are some elements which can impact the effectiveness of cue-reactivity in virtual environments?

This helped us construct a basic framework for what to look for in the literature on a theoretical level. To answer these, we first looked at research relevant to eliciting smoking cravings and investigated different types of cues, as well as how others have presented them in virtual environments. We then explored the idea of personalization for the purpose of eliciting smoking cravings in virtual environments, which we had previously identified as being a gap in research. To conclude our literature review, we investigated the concept of immersion and sense of presence, which along with the subject of personalization, seemed important to examine so that we could determine potential additional factors that can impact the effectiveness of VEs in eliciting smoking cravings. This investigation of relevant work helped us determine what the basic requirements would be for designing such VEs, therefore partially answering the questions posed above.

These requirements were then further refined after consulting with psychologists contributing to the *Perfect Fit* project and presenting them with scenarios illustrating potential solutions for each aspect of our proposed approach, which primarily focuses on determining how to create personalized VEs (PVEs). This helped us finish answering the above questions.

2.1. Cue-Reactivity Research

Much research is being performed into investigating various aspects of addiction and how it manifests in different people. Examining the reactivity of diverse types of cues is important in determining what causes cravings in different people, as well as how that can be simulated in controlled environments. First we wanted to examine what are the different types of cues that can be presented in a VE, how they can be presented, and which are actually feasible to do so. We then further investigated the element of personalization, which was previously identified as a research gap, to determine what has already been done and whether we can build upon it. Finally, we examined immersion and sense of presence, which are widely reported to be elements potentially affecting

2.1.1. Cue-Delivery Methods

Supplementary to the general types of cues mentioned previously (proximal, contextual, complex), there are more specific types of cues relevant to the method of delivery to a person [42, 73]:

1. *Visual*: Looking at images/video containing craving-cues, or in-vivo exposure.
2. *Auditory*: Listening to either explicitly suggestive audio, such as someone asking if you would like to smoke, or listening to ambient sounds from an environment which may be triggering, e.g. a bar.
3. *Haptic*: Interacting with objects that may trigger craving, e.g. using a lighter, unwrapping a pack of cigarettes, holding a cigarette, etc.
4. *Olfactory*: Smelling scents associated with smoking, e.g. tobacco smoke, coffee, alcohol, etc.
5. *Gustatory*: Tastes associated with smoking, similar to olfactory.
6. *Imagery*: Reading or listening to scripts which are designed to elicit emotions/cravings by having the subject imagine presented situations.

The first five types are experienced by physically being exposed to them, while the latter involves the person imagining those situations described and in a way causing themselves to have cravings. These cue types can also be combined to perhaps amplify their effectiveness. For instance, being presented with a video of a bar may be less effective than also being exposed to the audio one would expect to hear, and scents one would expect to smell [42, 73].

In the case of virtual reality, cues from five of the above types can be generated using different methods, with some being somewhat trickier to accomplish their presentation adequately. Gustatory cues are excluded from this since they necessarily require a more complex setup outside VR. Visual and auditory cues are rather straightforward to generate by simply presenting relevant material, such as photos or 3D objects, and audio. In the case of haptic and olfactory cues, however, more specialized equipment is required, such as haptic feedback gloves or suits, and scent-generation devices. That said, one could alternatively present haptic and olfactory cues using methods more akin to mixed-reality, such as having users hold a physical lighter or cigarette while they are immersed in a virtual world, or using candles and sprays to present olfactory cues. Although depending on the experimental setup, using these alternative approaches may cause unwanted distractions. For instance, if one is wearing a head-mounted display (HMD), being asked to manipulate a physical object may take away from the experience because they may not also be able to look at the object they are manipulating, or there may be inconsistencies with the interaction between the physical object and its virtual representative. And in the case of presenting olfactory cues, using conventional approaches does not provide the scent-on-demand experience that specialized devices can offer, unless researchers go to huge lengths in planning and executing complex scenarios. Such scenarios would likely require great coordination to present the correct olfactory cues with appropriate timing based on what the user is experiencing in the virtual world. And if not using an HMD, this may be distracting to the user and even more difficult to accomplish in a realistic manner in the lab.

In determining what combinations of cue types and delivery methods to utilize in virtual environments, arguably the biggest considerations we need to make are whether something adds significant enough value, and whether it is feasible to include. In other words, there needs to be a cost-benefit analysis. Naturally, it would be interesting to develop a system capable of presenting every type of cue using every delivery method available. However, it is difficult to predict whether the increased complexity that such a solution would entail, as well as the additional costs required, can actually pay off proportionately in terms of the experimental value added. That said, visual and auditory cues are the most feasible to present using even standard equipment, such as a computer with monitor and speakers. They are also arguably what virtual environments are currently best suited to present, i.e. visuals and audio.

2.1.2. Presenting Cues in Virtual Environments

Much work has been done in smoking addiction research to investigate potential causes for people to have cravings. As discussed previously, different cue types exist that can be presented via various methods. Cue-reactivity in each person also depends on whether cravings are caused by cues triggering some explicitly-or-implicitly-activated response. For instance, watching someone smoke, looking at a pack of cigarettes, being in a party with people smoking are cues typically evoking explicitly-activated responses (causing cravings). On the other hand, for instance, walking out of a building may act as an implicit cue. This is because the activity of exiting a confined space where smoking is typically not permitted can potentially signal an opportunity to smoke in some people [10]. Another example would be walking down the beach triggering some memory or emotions associated with smoking, which may also cause cravings. Different studies have presented various combinations of cues to investigate their efficacy in eliciting cravings.

Most research has implemented methods presenting:

1. Explicit visual proximal cues i.e., presenting smoking-related paraphernalia [1, 12, 13].

2. Implicit visual proximal cues shared by large amounts of people e.g., objects associated with alcohol or coffee, which many people relate with smoking [11, 13].

3. Explicit/implicit visual/auditory complex cues, such as a bar setting with people smoking and partaking in other activities associated with smoking e.g., drinking alcohol [9, 11, 13, 19, 29, 30, 32, 35, 51, 60, 88].

4. Explicit auditory complex cues e.g., being asked if one wants to smoke [11, 13, 88, 32, 50].

Some research has also included explicit and implicit olfactory cues, such as tobacco, coffee, alcohol, or food scents in an attempt to elicit cravings. However, the overwhelming majority did not explicitly examine the effects that including this type of cues has on cravings [19, 88, 1, 13], while one study found that the addition of scent did not significantly increase attention to cues or thoughts about smoking [87].

After investigating cue-reactivity research, we see commonalities in the VEs designed and used in experiments to test cue-reactivity. This may somewhat be due to focus groups or surveys being conducted prior to the actual VEs being designed, where participants are asked to rate which cues they think would elicit more smoking cravings, such as in [51]. Additionally, there has already been much research conducted in the field of addiction prior to the emergence of virtual reality and its establishment as a reliable tool in research, which has also informed as to cues which may likely be effective. Unfortunately, we have not seen much research explicitly comparing the effectiveness of different cues in eliciting smoking cravings in VR. Instead, most research involves presenting a multitude of cues and assessing the cravings self-reported by participants, as well as other variables, such as physiological data, emotions, demographics information, etc., to investigate potential effects. In general, research consulted has shown that combinations of above cues are effective in eliciting smoking cravings.

VEs employed in cue-reactivity research which have been frequently used include:

1. Rooms containing explicit and/or implicit visual cues, such as smoking paraphernalia and/or associated objects, such as a cup with coffee, or alcohol, etc. [1, 11, 12, 13, 37].

2. Scenes depicting a home setting, with the scenario being that the viewer and others in the environment have finished eating breakfast/lunch/dinner. After eating a meal has been shown to be a common occasion for people to smoke, however, this is also dependent on each person's addiction model. For instance, people with lower nicotine dependence have been shown to not crave smoking in the first hours after they wake up [32, 33, 36, 63].

3. Scenes depicting the viewer and others having a meal at a restaurant. Similar in some ways to the home setting mentioned above [32].

4. Scenes depicting a bar or party setting where people are drinking, smoking, dancing, talking, and in some cases offering the viewer a drink/cigarette. Consumption of alcohol has been associated

with smoking in some people. This is especially observed in social smokers who only smoke when in those types of situations involving public places, other people, and usually alcohol [9, 11, 13, 19, 32, 36, 44, 51, 61, 88].

5. Scenes depicting the viewer and others having coffee at a café. Similarly with alcohol, coffee consumption has been associated with smoking in some people [32, 36].

6. Scenes depicting a scenario where the viewer is waiting in the street for public transportation to arrive. Having to wait for something may trigger cravings to smoke in some people, or one may also see/hear/smell others smoke at those public places e.g., bus stop [30, 32].

Other research has also investigated more elaborate settings, such as creating multi-scene VEs. For instance [9, 10] used a virtual world of eleven interconnected environments, including an apartment building, restaurant with a bar, a cinema complex, a grocery store, where these were connected with streets allowing the participant to navigate between them.

There have not been studies explicitly comparing the effectiveness of different cue combinations to elicit smoking cravings e.g., VEs with only proximal cues vs VEs with only contextual cues vs VEs with complex cues. However, some studies have shown that VEs with complex cues (party with social interactions), were able to elicit similar self-reported craving levels as VEs containing only proximal cues (smoking paraphernalia room) [11, 12]. In general, the research investigated has reported that the environments listed above are effective in eliciting cravings when compared to neutral-cue environments.

Additionally, not many studies have directly compared the effectiveness of different complex-cue environments on craving. One, however, compared seven VEs similar to ones mentioned in the list above, to examine which were reported to elicit the strongest cravings [36]. Results indicated that in their experiment, VEs depicting 1. being in a pub, 2. having coffee at a café, as well as 3. having lunch at home elicited the highest cravings among their participants, in that order.

The role of context in causing cravings has also been demonstrated by [60], where they compared two identical convenience stores, one with explicit proximal cues such as a shelf behind the counter displaying tobacco products for sale, and the other where no smoking cues were present. The smoking-cue VE elicited significantly higher self-reported cravings, however, the neutral-cue VE also elicited higher cravings than an unrelated neutral-cue environment. The authors postulated that this perhaps points towards some participants having associated convenience stores with the sale of tobacco products, and therefore even in the absence of proximal cues some craving was elicited.

To summarize, the general concept of designing VEs aimed at eliciting smoking cravings involves depicting objects and places, and simulating scenarios which may be triggering to people. This includes not only explicitly presenting individuals with smoking-related objects and activities, but also presenting them with implicitly-associated cues. Arguably the biggest consideration of what to present and how to present it relies on operational requirements, such as whether something is feasible, or what kind of hardware we want to use. On the other hand, if certain cues are not implicitly associated with smoking in some people, they may not activate a response which would cause them to have cravings. Therefore, we posit that it may be valuable to also investigate whether we can obtain information on which cues may implicitly activate cravings in different individuals, that may enhance the efficacy of a virtual environment in eliciting cravings.

2.1.3. Virtual Environment Personalization

We have already briefly mentioned that we identified a research gap regarding the subject of VE personalization. This gap has been justified to some extent by the inherent complexity involved and the increased magnitude of work required to simulate experiences in VR that sufficiently correspond to the addiction model of each individual. We now look into how personalization has been employed in past research, as well as the reasoning behind why we believe it may lead to eliciting higher, or at least consistently high cravings.

Many VR studies have used some element of personalization, for instance having packs of cigarettes, posters, advertisements, matching each participant's preferred brand [1, 10]. One study even included the creation and customization of entire VEs in the Second Life online gaming program for each participant's self-reported triggers [29, 52]. However, none of these explicitly compared the effects of personalized with non-personalized environments. Conklin et al. performed research into the effect of personalization on elicited smoking cravings, first by means of an earlier study using imagery scripts [28], and later by presenting participants with photos on a computer monitor [26]. In their study using imagery scripts they compared personalized, standard, and personalized craving-cue imagery scripts of other participants. Their results showed that personalization did not increase cravings in imagery scripts, while using personalized imagery scripts of other participants actually elicited lower cravings. This may be because the cues contained in personalized scripts by others did not activate some implicit response. The authors also posited that there may be a limit on how much craving may be elicited by imagery scripts, which is perhaps why personalization seemed to not be more effective than standard approaches, something which may also potentially be a factor for other types of cues and presentation methods. In their later study, they had participants record photos from locations in which they smoke and do not smoke most frequently, and contrasted them with photos the researchers took from locations shown in an earlier study they did to determine which locations participants considered as being the most and least triggering to smoke [27]. Their results showed that personalized smoking-cue images elicited higher smoking cravings than standard and non-smoking-cue ones.

Therefore, we posit that there is a rather strong argument for hypothesizing that cravings elicited when presenting VEs personalized to the addiction model of a person may be higher, or at least the highest that can be feasibly elicited. For instance, implicit cues are dependent on what each person has associated with smoking, e.g. someone who doesn't drink coffee or alcohol has not associated smoking with them and so they will probably cause no cravings in that individual. Or for example, convenience stores may not be an effective contextual cue in people who have not associated them with seeing or buying cigarettes. Additionally, the time of day and level of nicotine dependence may also be a factor affecting smoking cravings, such as social smokers or less nicotine-dependent individuals may not smoke in the first few hours of the day, e.g. after breakfast. While this may be the opposite for more habitual smokers or more nicotine-dependent individuals. Finally, biological factors such as age may have an impact on the effectiveness of cues in causing smoking cravings [86].

All of these, while not necessarily easy to take into account when developing VEs, may be worthwhile to explore for potentially maximizing their effectiveness to elicit smoking cravings. This is because researchers attempting to test their own hypotheses, either regarding cue-reactivity or attempting to develop smoking cessation methods, may benefit from being able to elicit high or at least consistent craving levels from participants, so that they can more effectively assist people with overcoming them.

2.1.4. Immersion and Sense of Presence

Two prominent concepts concerning psychology research using virtual reality, is determining the role that immersion and sense of presence (SoP) have in the effectiveness of methods intended to be employed. While closely related, these do not correspond to the same thing. According to [75, 76, 77, 78, 81], *immersion* relates to what the technology offers in terms of allowing for sensory modalities comparable to their real-world equivalents. In other words, offering the tools which can enable the mimicking of what could be experienced in the real world. Immersion is also objectively measurable, for instance an HMD is a more immersive technology than a computer monitor, or an HMD capable of displaying content in higher fidelity is more immersive than one without that capability. *Sense of presence*, on the other hand, is the reaction of people to immersion, i.e. whether one actually feels they are present in the virtual world being experienced. This is something purely subjective and difficult to accurately measure [78], with Slater [75] presenting immersion as being being analogous to the wavelength of a color, which can be objectively measured, and SoP being analogous to how one perceives that color, which is an entirely subjective experience. There is, however, a positive correlation between immersion and SoP, and there is some indication that the two can be concomitantly increased [8, 16, 81, 76].

Since the SoP cannot necessarily be directly improved, increasing immersion has been shown to perhaps help improve it [76, 81]. Increasing immersion has to do with both the hardware and the software involved in presenting VEs. Firstly in terms of hardware, the more a technology offers an escape from the real world by isolating the senses and having virtual experiences perceived as being real, the more immersive it is. Sufficient results can be accomplished by for instance using an HMD and surround headphones, an approach prevalent in most relevant research projects. Better results may be achieved by using top of the line HMDs with very high resolution lenses, high quality surround audio, haptic feedback suits, and scent-generation machines, since those are more immersive technologies. This, however, is not necessarily feasible by most research projects, due to the increased costs and additional experimental complications the technology may introduce.

In terms of software, higher fidelity environments offering well-rounded combinations of cue types may increase immersion, and perhaps subsequently the SoP. However, similarly with the aforementioned hardware considerations, there is arguably an optimum point for the complexity involved and effort/resources required, relative to the experimental value gained. Furthermore, higher immersion hardware technologies have been shown to suffer proportionately more from lower-fidelity software approaches, as opposed to lower immersion hardware [16]. Therefore, the choice of software/hardware approach combinations arguably depends on what is feasible to implement, with a good balance of each perhaps being preferable.

The matter of immersion and SoP is relevant in psychology research, because many approaches rely on people's emotional states, which have also been shown to be affected by feeling as if you are physically present in a VE [15, 16, 59, 84]. For instance, when implementing VR-based approaches to help people overcome conditions related to phobias, anxiety, PTSD, etc., it has been shown that the more the client feels as if they are actually present in that VE, the more effective treatment approaches may be. In some cases, reflexive reactions which may be elicited can suffice to give psychologists important information regarding the mental and/or emotional state of a client. For instance, one is aware that they are viewing a VE while physically being present in the lab, however, being presented with a scenario where they are suddenly falling can still trick them into momentarily thinking that they are actually experiencing that in the real world. That is because reflexes are capable of reacting before the brain can reason about the situation and realize that it is not actually happening [83].

Similarly in smoking addiction research, there are indications that immersion and SoP may also be important to take into account. There has not been much investigation specifically into their effects, and results from what research has been conducted offer somewhat conflicting reports on their effects on craving. Firstly, results of [32, 33] suggest that smoking cravings may increase with more immersive systems causing a higher SoP. On the other hand, [30] report that people who recently quit smoking were having higher cravings in the high-immersive condition of their experiment, while in current smokers there was higher craving in the low-immersion condition. The two projects used fundamentally different types of environments, however, with the former using interactive 3D VEs and measuring self-reported craving during the experiment and SoP after VR exposure. As mentioned before, the SoP is entirely subjective and therefore difficult to accurately capture, however, using questionnaires such as ones proposed by [79] may give an approximate indication. Which is what the researchers did in this case and reported that the SoP was strongly related with self-reported craving levels. In the case of [30], they specifically tested immersion by having participants watch the same 360° video environment on a held smartphone (low immersion), or on the same smartphone placed in a VR headset (high immersion). They reported that their unexpected results may have also been influenced by different factors. 1. That the HMD VR experience may have been distracting enough that current smokers did not get comparatively high cravings as opposed to recent quitters, 2. that current smokers may have recently smoked therefore naturally having less cravings, as well as 3. positing that environments having been passively experienced, i.e. watching them as a video as opposed to being interactive, may have also influenced their results.

There is, therefore, some indication that striving to improve immersion, at least to some degree, may be a worthwhile endeavor in improving the effectiveness of cues presented to cause smoking cravings. Despite that, we do not see enough research focus placed on the matter to warrant overly concentrating our efforts there, but we should still perhaps consider more immersive approaches where feasible.

2.1.5. Basic Requirements

A system aiming at presenting cues in VEs to elicit smoking cravings, just like any other type of system, should fulfill certain requirements, which are split into functional, technical, and operational.

At this early stage, after performing our background research we identified a set of basic high-level requirements, listed in Table 2.1.

Type	Requirement
Functional	Create virtual environments containing smoking cues of various types, to elicit smoking cravings.
	Facilitate the measurement of elicited cravings by the virtual environments.
Technical	Determine software to use in developing virtual environments.
	Determine what hardware to use in presenting virtual environments.
	Adhere to good software engineering practices in developing the system.
Operational	Determine overview of how personalization of virtual environments may take place.
	Determine which cues are feasible to present based on other operational requirements.
	Determine how immersive virtual environments should be.
	Determine requirements of the *Perfect Fit* project.

Table 2.1: Basic System Requirements

These were rather general and some had not been entirely defined yet, primarily because we first needed to consult with experts contributing in *Perfect Fit*, which this project is part of.

2.2. Expert Consultation

Perfect Fit is a project undertaken by a multi-discipline collaboration of various universities and organizations in the Netherlands. It ultimately aims to produce a personalized-medicine approach in assisting people with adopting a healthier lifestyle, by means of an application employing a virtual coach. We first had a brief consultation with psychologists contributing to *Perfect Fit* to elicit some basic requirements, which we then used as a basis to design scenarios illustrating the options identified for different aspects of the proposed system. We subsequently presented these scenarios to a team of experts contributing to *Perfect Fit* and gained valuable feedback which helped us refine our requirements and update our general vision for the project.

2.2.1. Background and Initial Requirements Elicitation

While the specific methods for achieving the goals that *Perfect Fit* has and will set, including the type intervention or coaching approach will be determined later by those aiming to undertake the relevant parts of the project, there is a need to develop reliable controlled environments for them to test some of their hypotheses. That is where this project comes in, with the aim of designing and developing virtual environments which can be used to elicit smoking cravings in smokers who view them.

We had an initial consultation with two health psychologists experienced with smoking cessation approaches and treatment, contributing to the *Perfect Fit* project. We discussed what they generally think may be worthwhile matters to consider in creating virtual environments and the type of system that could assist them, as well as what basic requirements they have in terms of how they envision our solution will be used later on as part of the wider project.

These are supplementary to our basic requirements presented in section 2.1.5, and can be listed as:

1. Researchers contributing to Perfect Fit want to have virtual environments capable of effectively eliciting smoking cravings in participants.

As mentioned previously, researchers developing methods that aim in helping people quit smoking want controlled environments in which to test their hypotheses. So the requirement here involves providing them with virtual environments which can effectively elicit smoking cravings in participants, so that they can perform their various intended experiments, e.g. testing whether their proposed intervention approaches have the desired effects.

2. Researchers contributing to Perfect Fit want to have a means of communication with participants in the aforementioned virtual environments, to simulate interactions with a virtual coach.

In terms of communication, researchers would like the ability to exchange messages with participants during the experiment. In the application that *Perfect Fit* aims to develop, a virtual coach will be communicating with users. However, prior to its development researchers would like to simulate the intended front-end functionality by means of a ”Wizard of Oz” approach. This entails a human operator interacting with users using similar principles to how a virtual agent would e.g., a structured protocol with pre-set responses.

2.2.2. Scenarios and Discussion

Based on our literature review and this first expert consultation, we set out to distinguish the various aspects that our proposed system would have, and then determined different options that we envisioned feasible for each. Since we were aiming to not only construct VEs to elicit smoking cravings, but to also allow the researchers later on to use our methods to create VEs for their own experiments, we essentially have two types of stakeholders that we need to base our scenarios on. Therefore, we have differentiated these into the participant (client) and researcher perspectives.

In terms of attempting to satisfy the first requirement received regarding creating effective VEs to elicit smoking cravings, we used relevant aspects identified in our literature review to inform our decisions. As such, we conceptualized approaches using some form of personalization in the proposed solutions, since there are strong indications that adapting the presented cues to the addiction model of individuals may be more effective in eliciting a high level of cravings.

We structured our scenarios based on a set of propositions that we formulated after the initial expert consultation, which helped us in separating the different aspects of our proposed solution identified:

1. The client is able to provide relevant material required by researchers to create personalized virtual environments.
2. The researchers know what to personalize in virtual environments for each client.
3. The researchers find the concepts of immersion and sense of presence important in presenting virtual environments to elicit smoking cravings.
4. The researchers are willing to spend X time on preparing virtual environments for each client.
5. A limited library of 3D assets is sufficient to create suitably-personalized 3D virtual environments.
6. The client is willing and able to physically visit the lab to participate in the experiment.
7. The communication solution we envisioned sufficiently fulfills researcher requirements.

The following scenarios were presented to the experts consulted and a discussion took place to determine truth values for the above propositions. This was done by analyzing not only how the experts responded to each option for every scenario, but also the reasoning provided. A summary of the discussion regarding the different scenarios and the insights we found can be seen in Table A.1 of Appendix A.

2.2.2.1 Participant Perspective: Required Materials

This first scenario attempts to address our first proposition, ”The client is able to provide relevant material required by researchers to create personalized virtual environments.” The initial consideration regarding how and what to personalize to design VEs, involved examining what information or material we would require from participants prior to the experiment being conducted. We had two main contending options here:

1. Participants submit answers to a questionnaire, which the researchers would formulate and provide, attempting to elicit relevant information that can be used to inform decisions regarding how to personalize virtual environments.

This would include information related to their addiction model, such as their demographics information, the average number of cigarettes they smoke daily and their preferred brand, locations where they smoke most frequently, objects present in certain locations, etc.

2. Participants submit photographs and audio from the locations in which they smoke most frequently.

The idea was that similar information to what questionnaire answers could provide may be extracted by analyzing the photos, e.g. where do participants smoke most frequently, what objects may be triggering, etc. Additionally or alternatively, the photos themselves could be incorporated somehow in the VEs, along with the audio provided, which would act as an auditory cue.

Both options were viewed generally positively, however, the second was favored more because it could provide information that participants may not be consciously aware of, which could factor into their addiction model. Additionally, participants could also perhaps be asked to submit some information regarding their smoking history or demographics information, to be used where appropriate.

An important consideration had to do with the individuals who *Perfect Fit* may recruit into being experiment participants. There was discussion regarding wanting to involve people of low literacy or socio-economic status (LSES), due to there being evidence suggesting that individuals belonging to such groups may also be less health-literate [56] and therefore a vulnerable group that the project wants to help as well. In this case, we needed to consider potential differences in how LSES people would be providing the necessary material to the researchers, both in terms of technical proficiency and in time investment. In the case of photos, the experts suggested that they may already be taking photos in their everyday lives, so they may already have the skills required. In the case of questionnaires, the experts seemed uncertain as to whether LSES participants would be willing to make the effort required to provide adequate responses i.e., reading all the text and then responding with relevant information. In the case of time investment, there appeared to be more leeway, since the experts suggested that the process of preparing the required material should not take more than two hours. Which in our estimate is sufficient time for either option.

2.2.2.2 Participant Perspective: Viewing Virtual Environments

This scenario attempts to partially address the second and third propositions, ”The researchers know what to personalize in virtual environments for each client.”, and ”The researchers find the concepts of immersion and sense of presence important in presenting virtual environments to elicit smoking cravings.”

In terms of what personalized virtual environments (PVEs) to present, we had two primary options along with a third using combined elements from each. Participants would:

1. Experience 3D virtual environments containing personalized elements determined by researchers and stock sound corresponding to the visuals.
2. Experience virtual environments containing panoramic photos and audio submitted by themselves, of the locations in which they smoke most frequently.
3. Experience virtual environments containing 3D assets, personalized 3D elements determined by the researchers, photos submitted by themselves, and stock/submitted audio.

The main difference from a practical perspective, is that 3D environments are more feasibly made navigable, may contain interactible elements, and may present animated content. While on the other hand, these aspects are more difficult or even unfeasible to seamlessly incorporate in a VE presenting panoramic photos, and may potentially venture into uncanny valley territory, since there will necessarily be a difference in the graphical representation of elements shown, for instance adding 3D objects in a photo, and may likely affect immersion. The experts consulted did not view navigation inside the VEs as particularly important in eliciting cravings, but were not entirely certain as to what effects interaction with objects in the VE would have.

The concept of immersion was viewed as being moderately important, something consistent with research consulted which reported that environments offering higher immersion may be more effecting in eliciting smoking cravings. One study reported that environments composed of 360° photos generated similar levels of presence as 3D environments, even with differing levels of realism [15], which arguably makes all three of our options here viable in that regard.

The idea to utilize photos and audio taken by participants was viewed more favorably by the researchers consulted, primarily because it presents participants with objects or locations that they have already most likely associated with smoking, and therefore increasing the likelihood that they will have a craving response activation. More specifically, the idea of presenting participants with panoramic photos of the locations in which they smoke most frequently seems promising, because it would somewhat inherently offer a form of realism in terms of simulating elements in people's addiction model. Additionally, this reduces the effort required in having to determine exactly what to personalize for each person in every environment. Furthermore, this is useful in the case of *Perfect Fit*, since the cravings which participants may be helped to overcome will be elicited by VEs displaying some of the same cues they frequently get exposed to and have formed conditioned responses to. Something which may arguably increase the real-world transference of the planned intervention approaches.

2.2.2.3 Researcher Perspective: Designing Virtual Environments

Continuing from the previous scenario, we attempted to finish addressing our second proposition, "The researchers know what to personalize in virtual environments for each client.". Additionally, the scenario here addresses the fourth and fifth propositions, "The researchers are willing to spend X time on preparing virtual environments for each client.", and "A limited library of 3D assets is sufficient to create suitably-personalized virtual environments."

As mentioned previously in 2.1.3, personalization of VEs is not very well defined in literature, and can potentially be unfeasibly complicated and/or time-consuming. We have already presented some of our ideas for PVE presentation methods in the previous section, and here we examine what would designing these can entail.

The options considered were:

1. Design a number of VEs with 3D assets, based on what the literature has indicated to be more effective in eliciting cravings, e.g. a bar, a café, a living/dining room at home, etc. Allow for the personalization of certain elements in each scene, which are to be determined by researchers. Additionally/alternatively, personalization may take place in the form of which type of environments are presented, rather than the only of the elements personalized within each environment, e.g. showing a bar to someone who doesn't go to bars may not elicit cravings.

2. Design personalized VEs from scratch with 3D assets, similar to ones presented in literature. Allow the freedom of basing everything in every environment on what researchers deem relevant to the addiction model of each participant.

3. Utilize participant-submitted photos and display them in some way, such as part of VEs with 3D assets, or having participants view VEs entirely composed of photos.

The primary issue with the first two option above is that the researchers would need to know what to personalize, and also what to ask participants to obtain the necessary information that would allow them to sufficiently do so. Which as mentioned previously, may require an extensive elicitation process, something not necessarily easily undertaken, especially when also considering LSES people. Additionally, a limited library of 3D assets would have to be enough for sufficiently personalizing VEs to fit each participant's addiction model, therefore recreating for instance people's houses or other environments would not be feasible. Finally, a potentially moderate to relatively long period of time would be required to create PVEs for each participant, something primarily based on how proficient one is with the methods used, as well as on the amount of work required to obtain a satisfactory result.

On the other hand, incorporating participant-submitted photos and audio can be a streamlined and uncomplicated process, requiring as little effort as for instance copying files in the specified folder and editing a document with certain properties of the files added.

Presenting panoramic photos was generally viewed as the more favorable option by the researchers consulted, as was also discussed in the previous scenario, which made the choice here rather straightforward. Although in terms of why they also preferred this from the researcher perspective, the experts consulted were mostly concerned with the time requirement, as well as how complex the process would be. They indicated that they would not want to spend more than a maximum of 2-3 hours on personalizing VEs for each participant, and that the process should be uncomplicated enough for someone without programming expertise to easily undertake with some instructions.

2.2.2.4 Participant Perspective: Location and Hardware Interaction

The options here had to do with where participants would physically be located to take part in the experiment, and consequently, what hardware would be used. These aimed to finish addressing our third proposition, ”The researchers find the concepts of immersion and sense of presence important in presenting virtual environments to elicit smoking.”, as well as address our sixth proposition, ”The client is willing and able to physically visit the lab to participate in the experiment.”

The options presented were:

1. The participant takes part in the experiment at home.
   (a) The participant views the VEs on their computer monitor and uses a keyboard/mouse for interactions.
2. The participant goes to the lab to take part in the experiment.
   (a) The participant views the VEs through an HMD and uses a gamepad controller for interactions.
   (b) The participant views the VEs on a computer monitor / projector and uses a keyboard/mouse for interactions.

As presented above, if the participants were to take part in the experiment while at home, the only hardware interactions we can guarantee are feasible for mostly everyone, are to use a computer monitor and keyboard/mouse. We did not consider presenting VEs on a smartphone screen since that may break immersion too much and also additionally complicate interactions.

Alternatively, if participants were to visit the lab to take part in the experiment, there is a higher degree of freedom in what hardware we can have them use. They can of course be asked to use the same as mentioned above, or they can use an HMD and gamepad controller to view/interact with VEs, something which can also increase immersion and potentially improve the magnitude of elicited cravings.

Having participants go to the lab and use an HMD were the more favorably-viewed options, primarily due to this combination offering the most experimental control. After all, we have already argued for VR advantages in allowing for controlled environments, while on the other hand, having participants physically be in environments we cannot control somewhat defeats the purpose and potentially harms experiment results. This is primarily because we may not be able to guarantee that smoking cravings elicited are necessarily because of the VEs presented, rather than because of any other cues present in the participant’s physical environments. Furthermore, there may be additional distractions present which could not only reduce the effectiveness of the presented VEs in eliciting smoking cravings, but also potentially diminish any effects that immersion or sense of presence may have.

Finally, while this was not explicitly brought up by the experts consulted, if we are to also consider use cases involving LSES people as well, we may not be able to guarantee that everyone may have the necessary hardware at home i.e., a computer, or sufficient proficiency to use the provided application without extensive remote guidance. And even then, it may simply be compounding additional unnecessary complications that can potentially affect experiment results, such as environmental factors or distractions, which a controlled environment aims to avoid.

2.2.2.5 Participant/Researcher Perspective: Communication Method

Options in this aspect of the proposed system have to do with how we can facilitate communication between the researcher and participant during the experiment to partially address our final proposition, ”The communication solution we envisioned sufficiently fulfills researcher requirements.”. We broke this proposition in two parts; the first involving whether the communication solution has sufficient functionality, which is addressed here. There are two main options we identified, with the choice in the researcher’s perspective influencing which option in the participant’s perspective to take, and vice versa.

The identified options were:

1. Interact via free-form text input, similar to a typical messaging application.

2. Interact via the researcher messages being accompanied by corresponding responses among which the participant can select the one most closely matching what they want to say.

As discussed with the experts consulted, the options most closely resembling how the interface would appear in the final *Perfect Fit* application is preferable. This may increase the real-world transference of results obtained during *Perfect Fit* experiments, e.g. attempting to help participants overcome elicited cravings. That said, however, the functionality offered by the interface in the planned *Perfect Fit* application was somewhat unclear at the time that this project was underway, although offering a selection of pre-determined responses would more likely be the case.

Furthermore from the researcher's perspective, the envisioned solution for our project was to offer a selection of message templates, each having pre-determined responses. The researcher would be able to select one, and be allowed to edit both the message and available responses before sending them to participants. Again, this is to be consistent with how the *Perfect Fit* virtual coach will most likely operate, where it may have structured responses, rather than being similar to offering free-form communication. Additionally, the experts mentioned that the messaging style being uniform may be a good idea, which is in alignment with having pre-set templates, rather than each researcher free-form typing messages in their own style, that can impart the perception that different people are talking with the participant.

2.2.2.6 Participant Perspective: Communication Solution Appearance

This final scenario involved how the communication solution user interface (UI) would appear on the screen so that participants can interact with it. This attempted to finalize addressing the last of our propositions i.e., "The communication solution we envisioned sufficiently fulfills researcher requirements.". As mentioned in the previous scenario, we broke this into two parts, with the first involving whether the communication solution has sufficient functionality to meet researcher requirements, which we already examined. The second part involves whether the communication solution is designed in a sufficient manner for its intended purpose, which is examined here.

The two options identified were:

1. Having the UI overlaid on a 3D object, e.g. a smartphone.

2. Having the UI appear in a typical manner, i.e. flat against the screen.

There was some discussion about this, primarily because how the 3D model presented arguably was prone to breaking immersion. More specifically, its size, location, rotation, and that it remained stable on the screen even as the camera was rotating, were not realistic enough and so took away from the overall experience. Therefore, from our discussion we understood that they wanted the communication solution to not break immersion if it was to be implemented as part of a 3D model in the environment.

On the other hand, presenting the UI in typical fashion flat on the screen seemed to have a more favorable reaction, because its purpose was more apparent, and it looked separate from the environment. Therefore, purposefully breaking immersion in a way via a user interface appearing was preferable to immersion breaking accidentally because of the object design. That said, however, we agreed that most of the criticism regarding overlaying the UI on a 3D object had to do with how said object appeared. This was something we subsequently addressed by modifying the size, rotation, position of the object and made it remain centered in the field of view even as the camera looked around. This elicited a more favorable opinion from the researchers in a later consultation, and we agreed that this was the preferred option.

2.3. Requirements Refinement

After obtaining feedback on the scenarios presented, we now had enough information to refine and expand our basic requirements previously listed in section 2.1.5. We list our updated requirements in Table 2.2. This process helped us update our vision for the solution we planned to develop as part of this project.

Type	Requirement	Source
Functional	Create a system that allows for personalizing virtual environments using participant-submitted panoramic photos and audio.	Literature / Experts
	Facilitate the measurement of elicited cravings by the virtual environments.	Literature
	Facilitate communication between researcher and participant by means of a messaging solution.	Experts
Technical	Employ appropriate tools in designing virtual environments, and in developing the system.	Literature
	Adhere to good software engineering practices in developing the system.	Literature
Operational	Utilize appropriate assets and presentation methods that do not break immersion.	Literature / Experts
	Present environments which are more likely to contain cues relevant to each person's addiction model.	Literature / Experts
	Develop system to be accessible to people of low literacy/socio-economic status.	Experts
	Ensure that material preparation does not take participants longer than 2 hours.	Experts
	Ensure that virtual environment preparation / personalization does not take researchers longer than 2-3 hours for each participant.	Experts
	Allow researchers to select message templates with pre-set responses to send to participants.	Experts
	Develop participant communication interface that is conceptually similar to the virtual agent solution currently envisioned by *Perfect Fit* researchers.	Experts

Table 2.2: Updated System Requirements

2.4. Vision

As a summary of our findings, discussions, and refined requirements, we present our updated vision for the solution we aimed to develop as part of this research project. Both in terms of contributing to the overall *Perfect Fit* project, as well as to the field of VR-assisted addiction research.

The goal was to develop an application allowing the incorporation of panoramic photos and audio (material) to compose virtual environments, and the presentation of these environments by means of a head-mounted display. We aimed to instruct participants to provide us with the material which our application could then utilize and present for the purpose of eliciting smoking cravings in those participants. Additionally, because our system was intended to be used by researchers testing the efficacy of their proposed smoking-cessation methods via a conversational agent, we also implemented a solution to allow for communication with participants while they are experiencing said virtual environments.

3

Design and Implementation

Having refined our identified requirements for building virtual environments (VEs) to elicit smoking cravings, we now want to formally address our second research sub-question, namely:

How do we design virtual environments to elicit smoking cravings, for the purpose of enabling researchers to test their virtual coach approaches?

Reflected by our formulated requirements and the updated vision for our project, we expanded our scope from the general concept of designing standalone VEs, to also creating a system that can be used in a laboratory setting to assist researchers with testing their proposed smoking cessation approaches. The aim was for this system to have the capabilities of incorporating photos and audio to design VEs, presenting these to participants, and facilitating communication between researcher and participant while the latter are viewing them. The goal of presenting these VEs was eliciting the desired effects, such as smoking cravings in smoker participants, for the purpose of allowing researchers to test the efficacy of their proposed smoking cessation methods using a conversational agent. This chapter explores how we approached fulfilling our requirements listed in Table 2.2.

We first present a theoretical overview of how our proposed system aims to satisfy the identified requirements, in Section 3.1. We begin by presenting our solution approaches for each scenario discussed in Section 2.2.2. We briefly summarize each scenario along with the preferred option after taking into account input by *Perfect Fit* researchers, and we use that as the proposed solution to satisfy our formulated requirements. While Finally, we provide a high-level description of how we implemented the different aspects presented in the aforementioned design, in Section 3.2.

3.1. Design

First, we look at a theoretical overview of our proposed solution by discussing the design decisions we took for the different aspects that our system should have, based on feedback from our scenarios and the resulting formulated requirements. Succinctly, our system design is composed of three major components. These include the virtual environment presentation method, the material sourcing and preparation that preceded their incorporation into virtual environments, as well as the solution that facilitates communication between researchers and participants while the latter are viewing virtual environments.

3.1.1. Virtual Environment Presentation

The first aspect of our proposed solution we needed to address was how these virtual environments would look, and why we would want them to look that way. Relevant to our choices here were scenarios regarding the viewing of virtual environments (section 2.2.2.2), designing them (section 2.2.2.3), as well as obtaining the necessary material to do so (section 2.2.2.1). Here we aimed to fulfill three of our formulated requirements; one functional and two operational ones. These were 1. the creation of a system allowing the personalization of virtual environments using participant-submitted material, 2. the presentation of environments which are more likely to contain cues relevant to each person's addiction model, and 3. the use of methods that do not cause the system to break the viewer's immersion, if we could avoid it.

We chose to have our system present virtual environments comprised of panoramic photos and audio provided by participants, rather than have 3D environments containing personalized elements. We reached a consensus with *Perfect Fit* researchers that this would be the more suitable approach, since presenting visuals from the actual environments in which participants smoke most often may elicit stronger smoking cravings. This may be due to higher cue-reactivity caused by contextual cues that may otherwise be unaccounted-for if we were to use only certain elements of those locations. Especially since we cannot assume that even participants themselves would necessarily be aware of what item, circumstance, or their combination may act as a cue that activates a smoking craving response. Furthermore, having environments being composed of both 3D assets and real-life photos may case breaks in the sense of presence due to inconsistencies in the level of immersion elicited, if this composition is not realistic enough. This is something that may harm cue-reactivity, as opposed to having environments fully composed of either type of material. Similarly with audio, having environments contain sounds from the actual locations may contribute to the activation of contextual cues, rather than having stock audio from similar type of environments, or no sound at all, which is likely to harm the sense of presence elicited.

3.1.2. Material Sourcing

The second aspect of our proposed solution involves obtaining the material we want to use in designing the proposed VEs, as presented in the scenario discussed in section 2.2.2.1. This contributed to the fulfillment of the same requirements as the section above, with the addition of two operational requirements stating that 1. material preparation should not take participants longer than 2 hours, and 2. that the process should be accessible to LSES people.

Since we require audiovisual elements from physical locations which participants visit, we can request that they record and submit this material themselves. This can increase the likelihood for contextual cues being present, as well as improve the immersion of resulting environments. Certain factors influencing this are discussed below.

Firstly, participants should record this material from a position in which they most often spend time when there. For instance, consider the case where one is sending us material from their living room, where when usually there they spend the majority of their time sitting on the couch. We posit that recording material from that position may make the resulting environment more effective, rather than if they recorded the material from a different position. This is because if the goal is to have them imagine they are actually there when viewing the virtual environment, then presenting them with their usual viewpoint of those locations may increase the likelihood that contextual cues are present.

Secondly and similarly with the above, participants should record this material at a time of day at which they usually spend time there. For instance, if they most often smoke on their balcony in the morning, recording material there should also occur in similar conditions, so that the resulting VEs are more likely to contain relevant contextual cues, in terms of both visual and auditory elements.

In terms of obtaining audio, we deemed that having participants record it using an application of their choice was sufficient enough for our purposes. This is because audio is arguably easy to manipulate and produce a seamless result, so we could amend any volume or minor quality discrepancies with minimal to no issues. Instructing participants to record slightly longer audio clips than needed can help with adjusting for potential issues such as unexpected noises or the person putting their phone on a surface, which may harm immersion.

Taking our two additional requirements regarding material collection into account, we judged that having participants take photos or record video and audio given instructions and a contact point to receive support if necessary, was a good enough balance between receiving the material we needed, while not over-burdening participants. *Perfect Fit* researchers pointed out that mostly everyone nowadays owns a smartphone with the minimum capabilities required, and it is quite likely participants may use the skills necessary to complete the task in their every day lives. We also judged that material collection would take a reasonable amount of time per environment, with the visual element for each requiring 20-30 seconds, and the audio recording depending on the intended duration of virtual environment exposure, or some other criteria that researchers may have.

3.1.3. Material Preparation

As specified by our requirements, we wanted to make the virtual environment design process easy enough for researchers and make it achievable within a few hours per participant, as presented in the scenario discussed in section 2.2.2.3. This involves the processing of participant-submitted material before it can be incorporated in our proposed system and presented as VEs. For the visual element of our environments, we envisioned that researchers would employ relevant software to manipulate the raw material provided, and transform it into a format that could be used by our system, if necessary. The exact process was not yet concretized at this design stage because we needed to first test and finalize the relevant material sourcing process, however, the general idea was that preparation of visual material could involve panoramic photo stitching and / or photo editing.

Regarding the audio preparation process, we specified that researchers would be required to listen to the clip and judge whether it can be listened to at a natural volume, and whether there were noises present that could harm immersion, such as audio glitches or the phone being set on a surface. If not, they would be required to edit the file and adjust the audio clip accordingly. This, of course, relies on researcher capability to judge what a natural volume for each environment would be, and whether a sound would be immersion-breaking. However, we determined that to be the preferable option, rather than complicating the process from the other side by having participants normalize the audio themselves before submitting it.

3.1.4. Communication Approach

The last part of our system design relates to how we can facilitate communication between researchers and participants while the latter are experiencing VEs in VR, as presented by our scenarios discussed in sections 2.2.2.5 and 2.2.2.6. This aimed to fulfill one functional and four operational requirements; namely, 1. enabling communication between researchers and participants by means of a messaging solution, 2. using assets and presentation methods that do not break immersion, 3. developing a system accessible by LSES people, 4. allowing researchers to select message templates with pre-set responses to send participants, as well as 5. developing a participant communication interface that conceptually resembles the currently-envisioned *Perfect Fit* virtual agent solution.

Essentially, this communication solution was used to simulate interaction with a conversational agent aimed at supporting people to quit smoking and adopting a healthier lifestyle. This virtual agent, however, had not yet been developed or formally defined. Therefore, this communication system was designed as a *Wizard of Oz* solution, where researchers can interact with participants in similar fashion to how they would have their virtual agent interact with them. So the goal here was to make participants perhaps be under the impression that they were interacting with a virtual agent, rather than a person.

Succinctly, communication in our solution is facilitated through a messaging system, in which the researcher sends a pre-written message which has a set of pre-determined responses, and participants in turn select the response most accurately reflecting how they want to reply. The researcher and participant perspectives were separated and are discussed in the following sections.

3.1.4.1 Researcher Perspective

Starting with the interface available to researchers, it was designed with the goal of allowing for pre-determined messages and their available responses that participants can select from. The idea here is that we attempted to design something with similar functionality to what a typical virtual agent would be capable of. This includes having specific messages at its disposal which it can send depending on the situation, along with pre-defined responses that it can in-turn process and react to accordingly. The difference here is that researchers are responsible for identifying the situation and sending the message they believe appropriate. The reasoning for creating a design with similar functionality to a typical virtual agent is that because our system is aimed at assisting researchers with testing their proposed approaches in a controlled environment, which in this case include the use of a virtual agent solution as a behavior change support tool. Therefore, the testing environment should resemble the intended real-world use case, hence adding value to making participants be under the impression that they are interacting with a virtual agent rather than a real person, since this can increase the ecological validity of results. An abstract design of how this interface would appear is illustrated in Figure 1.

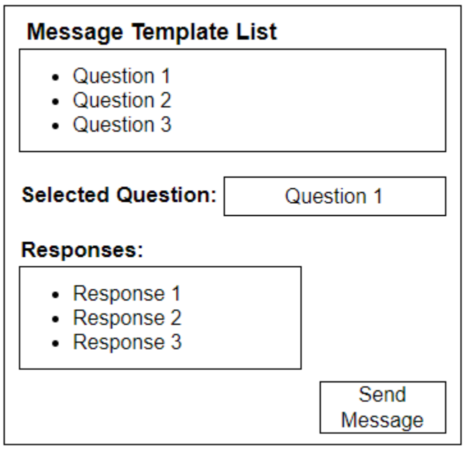


Figure 1: Researcher Communication UI Abstract Design

3.1.4.2 Participant Perspective

Regarding the interface that participants interact with, we also wanted this to appear in similar fashion to what they would expect to see in a typical messaging system employed by a virtual agent. This also helps us in satisfying our requirement regarding making the system accessible to LSES people by making the interface appear familiar to other messaging applications they may have used. The design here involves presenting participants with a message sent by the researcher alongside relevant responses they can choose from. Additionally, we wanted to increase immersion and intuitiveness of the interface itself by making it appear as a virtual smartphone with a typical chat interface. This is aimed to match the general use case of the intended virtual agent solution currently envisioned by *Perfect Fit* researchers, which involves participants interacting with a virtual agent on their phone. An abstract design of how this interface would appear is illustrated in Figure 2.

Figure 2: Participant Communication UI Abstract Design

3.2. Implementation

Next, we look at how we implemented these different aspects discussed in our design. Succinctly, our system is composed of a VR panorama viewer housed in a host-client networked solution that allows for exchanging messages via different types of interfaces depending on the user type. We start by presenting our decisions regarding the tools we chose to use in different aspects of our solution in Section 3.2.1, followed by a high-level summary of how each relevant major component was implemented. The virtual environment viewer is described in Section 3.2.2, and networking in Section 3.2.3. The overall system design, which is presented in the following sections, is illustrated by Figure 3. A higher-level illustration of using our system may be viewed in Figure fig: System Use Illustration. Figure 6 in Appendix D illustrates a lower-level representation of the processes involved in using our system.

The source code for the system we developed may be viewed and downloaded by visiting our Github repository [4].

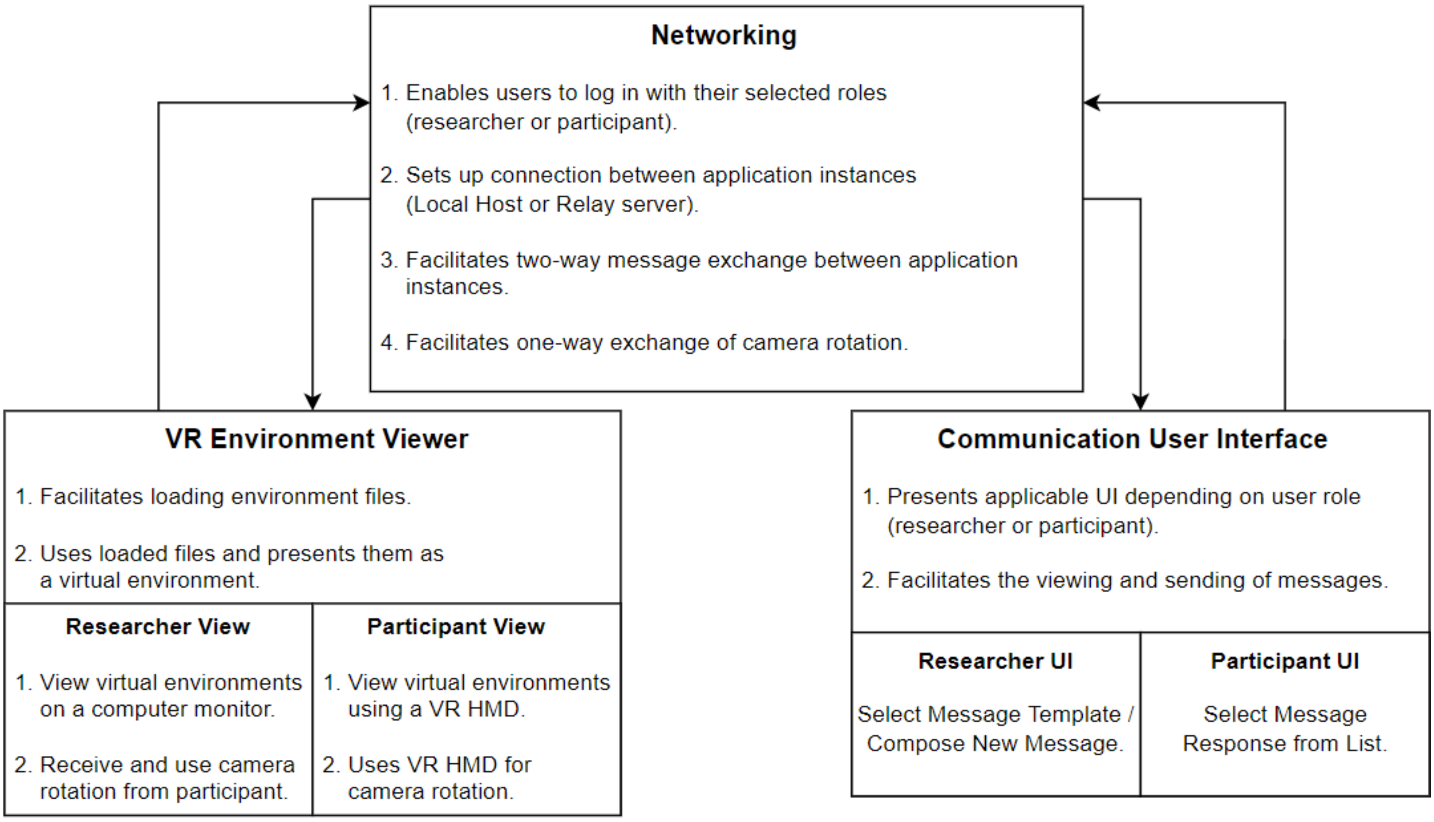


Figure 3: Overall System Design

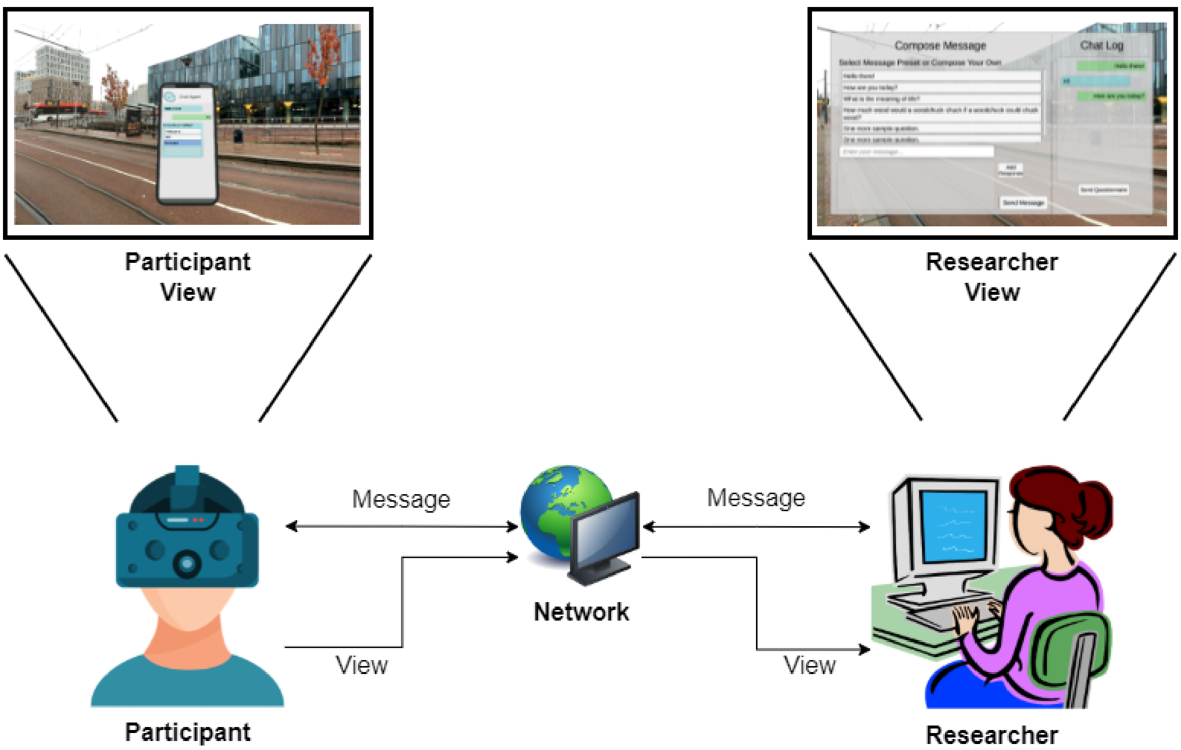


Figure 4: System Use Illustration

3.2.1. Tools Used

We first needed to decide which tools we would use for our system implementation, that would satisfy the following requirements; namely, 1. employing appropriate tools in both designing virtual environments and in developing the system, and 2. making sure that material preparation does not take researchers longer than a few hours per participant. We needed tools to process participant-provided material to be used in designing virtual environments, and a platform to use in implementing our proposed system. We also needed to determine how the material would actually be utilized by our system, and how the pre-determined researcher messages and their relevant responses would be imported. Each of these is discussed in the following sections.

3.2.1.1 Participant Material Processing

Regarding the actual material we would use to design virtual environments, while we initially envisioned participants directly recording and submitting panoramic photos, after further research we found that to not be realistically feasible. We then investigated the options of having them send us series of overlapping photos for each intended VE, or video.

This led us to search for a suitable tool that we could use process this material and obtain panoramic photos. The main conditions were that the tool be powerful enough to suit our purposes while not having a steep learning curve, so that it would also be accessible to *Perfect Fit* researchers. From the tools considered, *Microsoft Research - Image Composite Editor (ICE)* [67] was arguably the most powerful freely-available software capable of producing high quality panoramic images, while also being relatively easy to use.

After conducting a pilot run of our system, we determined that having participants submit videos was the preferable solution for ICE to produce visually cohesive panoramic photos. This allows for a higher potential allowance in user error, where since video is by definition a set of continuous images, we could avoid issues arising from stitching failures due to e.g., discrepancies with photo overlap.

One issue the overall solution had was that because the textures are panoramic photos taken using stock smartphone camera apps, the viewpoint while spanning between $200°$ to $360°$ horizontally, did not do so vertically. This meant that if looking up or down beyond the vertical edges of the texture, the viewer could see the image borders, which was immersion-breaking. We somewhat alleviated this by stretching the top and bottom 1% of pixels in each texture, so that at least the color scheme would be consistent when looking outside the normal vertical boundaries of the image. We deemed that the *GIMP image manipulation software* [68] was a suitable and freely-available solution that we could use to achieve both this and the image resizing step mentioned above, while being relatively uncomplicated to use for these purposes.

Regarding audio, we determined that the *Audacity* software [66] was an appropriate tool both for adjusting the volume of submitted files via modifying the amplitude of each clip, and for converting them to the appropriate format that our system can use. We judged that the tool, although powerful and offering a multitude of capabilities, made this an uncomplicated process even for people not familiar with either it or audio processing in general, which fulfilled the aforementioned requirements we set.

3.2.1.2 Development Platform

Next, we investigated different development platforms to determine which was suitable for the purposes of our project. We had three main requirements for the development platform: 1. has packages officially supported by HMD manufacturers, since we want to create a VR-capable system. 2. Can offer us the desired networking capabilities, so that we can facilitate communication between researchers and participants while the latter are viewing virtual environments. And 3. does not have a steep learning curve and has sufficient learning tools available, so that we can save development and maintenance time. We were provided with an *Oculus Rift CV1* HMD with an *XBox One* gamepad controller for development purposes, so package-wise, the platform we would use needed to be compatible. The two main contenders fulfilling our requirements were *Unity3D* and *Unreal Engine 4*. While the latter is arguably more powerful in some use cases, the former is more accessible for development and has recently released native support for larger texture size limits than the latter. Therefore, for the purposes of this project it was deemed the appropriate solution to use.

3.2.1.3 Importing Data

After deciding on the development platform, we determined how we would import the intended external data into the proposed system. This data included virtual environment material (photos / audio), as well as messages and their predefined responses that researchers could send participants. It was vitally important that the solution here would be accessible to people not familiar with software development, such as psychology researchers developing methods in the field of addiction, since they were the target audience for utilizing it.

Regarding virtual environment data, the relevant photo and audio files were to be placed in the *streamingassets* folder, which gets loaded into the application at runtime, so that there would not be a need to rebuild the application to add or remove environments. Researchers had to ensure that they followed the given naming convention so that photo and audio files could be matched with each other, and we determined this to be a sufficiently uncomplicated process.

For researcher messages, we chose to use JSON files imported using the *JSON.NET* package for *Unity3D*, because it offered us the functionality of importing the intended material in an uncomplicated manner, both from the development side, and from the end-user (researcher) side. Using JSON files meant that we could streamline the file import process from the researcher side, by ensuring that the file had an intuitive internal structure accompanied by instructions. Researchers were instructed to edit this file to add messages and their responses by copying the provided structure and populating it accordingly.

3.2.2. Virtual Environment Viewer

For viewing virtual environments, we used a solution involving utilizing a sphere 3D object. This sphere was implemented to use a shader that allowed the rendering of textures on the inside, rather than on its outside surface. We then positioned the camera, which users would view the scene from, inside this sphere so that when looking around they would be viewing the rendered texture.

3.2.2.1 Texture Considerations

As mentioned previously, we manipulated photos by stretching the top and bottom 1% of pixels. This was done in an attempt to reduce the immersion-breaking effects caused by realism limitations due to the nature of using the specific material. The entire height of each photo was visible in the neutral camera orientation, and therefore looking up or down meant that one could see the texture repeating outside the vertical bounds, as seen in Figure 6a of Appendix B. We initially conceptualized using dark-colored planes to block viewing outside them to perhaps reduce the negative impact on immersion, as shown in Figure 6b of Appendix B. However, due to this also sufficiently breaking immersion, we opted to pre-process the photos themselves before importing them and stretch segments from the top and bottom. This was primarily aimed at continuing the color pattern and preserving some level of immersion, and can be seen in Figure 6c. While neither seamless nor completely immersive, this was considered an acceptable limitation of our system, since obtaining the necessary equipment to collect fully-panoramic $360°$ photos was not feasible within the scope of this project.

3.2.3. Networking

In terms of networking, we needed a solution that could facilitate two functionalities; 1. communication between researchers and participants while the latter are viewing virtual environments, and 2. allowing researchers to see what the participants are currently seeing in the virtual environments. We opted to use the newly-released *Unity Netcode* package which is officially supported and offers the desired capabilities. For enabling communication using different computers and over the internet, we also chose to use the *Unity Relay* package and services, which allow for setting up a relay server that can facilitate users connecting to it and making use of the provided functionalities. Additionally, we enabled optionally setting up a server on the local host network. However, this only allows for two instances of the application to communicate on the same computer. Communicating over a local network or the internet would require more advanced approaches such as *port-forwarding* or *NAT punching*. This would not necessarily be feasible for researchers not familiar with such processes, such as psychologists working to develop smoking cessation methods. Additionally, this would not necessarily be something that can feasibly or securely be implemented in enterprise-type network solutions that are typically used by universities or other research institutes.

3.2.3.1 Initialization

The general login and setup process seen in Figure 6 of Appendix D involves users selecting their intended role i.e., researchers or participants. Selecting the former role has the system set up a server, and connect as host. Selecting the latter has the user connected as client to the server created by the former. In the case where a Relay server is used, a unique key is generated by the host which is required for clients to connect to the server, so this is given by researchers to enter at login. After successfully resolving login information, the applicable camera for the specific user type is activated, the corresponding UI is instantiated, and the VE specified is presented after loading the necessary files.

3.2.3.2 Communication Solution

For communication between researcher and participant, as mentioned in our design, we offered a specific interface depending on the user type. This is because of the inherent difference in the intended functionality between user types, where researchers are essentially employing a *Wizard of Oz* approach to select messages based on the situation, and participants are responding to those messages by choosing any of the given responses. Commonalities between the interfaces can be seen in how the chat history appears, where from the participant perspective the chat makes up the entire interface, while from the researcher side the chat is located next to the message template selection list. Researchers cannot directly interact with the chat interface other than scrolling through the history, while participants can select the response most accurately reflecting what they want to say via buttons located under the researcher message in the chat. Figure 5 illustrates the message exchange process.

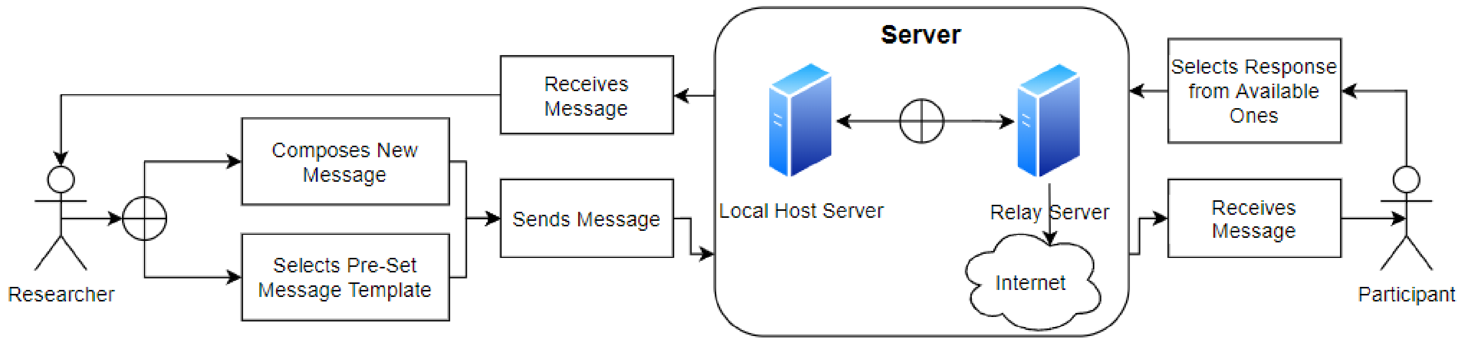


Figure 5: Researcher-Participant Message Exchange Process

The researcher interface takes messages and pre-set responses via a JSON file and loads them at run-time. This makes them available for the researcher to choose at any time through a list, where selecting a message presents its available responses that participants will be able to choose from. Researchers are enabled to edit both message and responses, or even compose their own from scratch. The researcher interface can be seen in Figure 6.

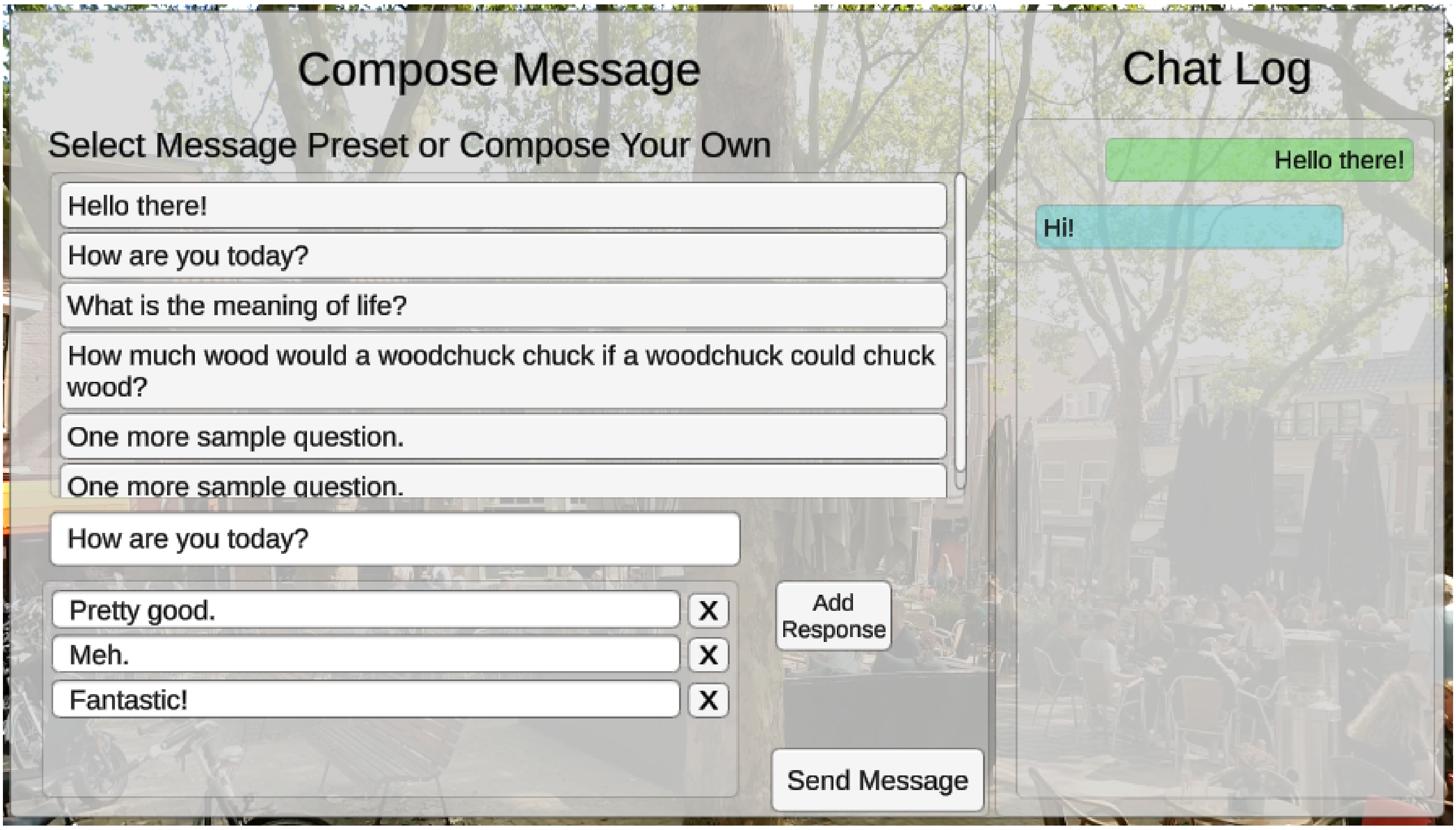


Figure 6: Researcher User Interface

The participant interface, as mentioned above, is composed of a chat panel allowing for the viewing of messages sent by researchers, and for the selection of an appropriate response. To increase immersion and a sense of purpose for the interface, this was overlaid on a virtual smartphone 3D object visible in the environment itself. This was placed in a central location, and would appear after an audible notification sound was played and a brief rumble on the gamepad controller to simulate smartphone vibration. Participants would then use the gamepad controller to select the response most closely matching how they want to reply to the researcher message, at which point they could press the relevant button to toggle visibility of the smartphone. The participant interface can be seen in Figure 7.

3.2.3.3 Camera View Sharing

To enable researchers to view what participants were currently viewing, we used two different cameras; one for each type of user. The participant camera was controlled by looking around using an HMD, while the researcher camera was updated with the latter's position. We have a short buffer for when the researcher camera is updated, so as to not overly tax the network connection, something which may potentially cause issues e.g., overloading the connection causing it to disconnect. When the participant camera rotation changes by 0.1 units, the updated rotation is sent over the network for the researcher camera to be updated. This process is illustrated by Figure 8.

Figure 7: Participant User Interface

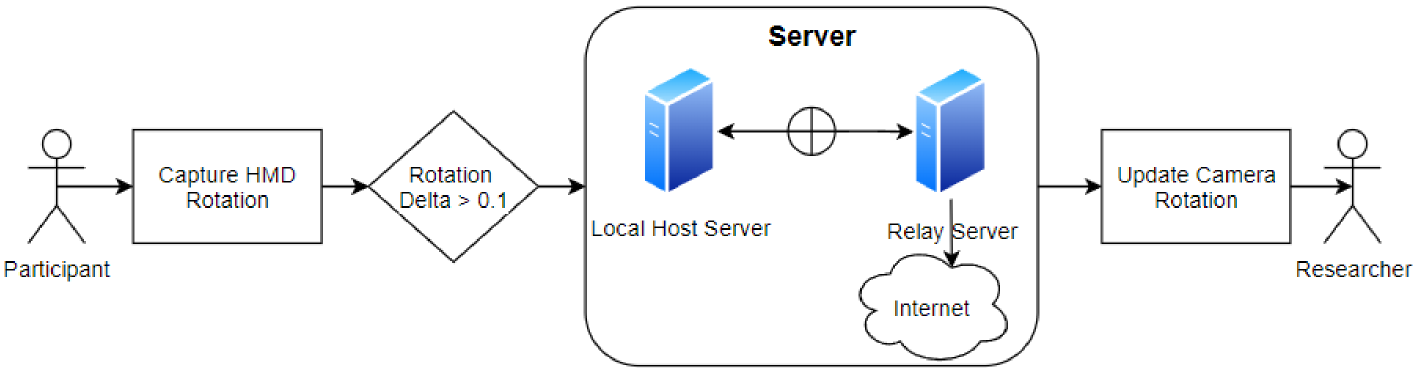


Figure 8: Researcher Camera Update Process

3.3. Summary

We have presented a description of our system design and implementation, along with some of the different considerations we had to make that resulted in our finalized solution. To summarize, we have designed a networking-capable system aimed at enabling researchers to test their smoking cessation methods with a virtual agent. This system allows for researchers and participants using their own instances of the application on different computers to interact with each other via the former sending messages with pre-determined responses, and the latter replying via selecting one of those responses. This communication occurs while participants are viewing VEs using a VR head-mounted display, and researchers are able to see what participants are seeing via camera rotation parameters being passed from the participant to the researcher application instance over the network connection. These VEs are composed of processed audiovisual material provided by participants. Next, we present the evaluation of our system that we performed.

4
Evaluation

Having presented a high-level design and implementation of our system aimed at eliciting smoking cravings in participants, we then wanted to evaluate how effective it was in doing so and also in eliciting a familiar experience. Our human research ethics application with ID number 1860 received approval from the TU Delft ethics committee, and we pre-registered our study in the Open Science Framework platform (OSF) [5]. Our study period was between December 2nd 2021 until January 27th 2022. This evaluation was done in an effort to answer our third and last research sub-question; namely:

How do different virtual environments compare with each other in terms of the magnitude of self-reported cravings and experience familiarity elicited in each, and how do users evaluate the usability of our system?

As mentioned in Section 2.4, there is a distinction between the overall system developed as part of this project, and evaluating its efficacy in eliciting effects. This is because the overall system aimed at assisting researchers with investigating the efficacy of their smoking cessation approaches includes a communication solution for the purpose of testing virtual agent strategies. While on the other hand, the evaluation that we wanted to perform was instead concerned with comparing the effectiveness of personalized and non-personalized virtual environments, as presented by our system, in eliciting the desired effects. We are hypothesizing that personalized VEs (PVEs) would evoke a more familiar experience than non-personalized ones, and we are basing on experts consulted that experience familiarity may be positively correlated with smoking cravings. Therefore, we are here examining the experience familiarity that personalized virtual environments presented by our system elicit. We additionally wanted to examine the usability of the user interface participants are asked to interact with in our main system, as well as explore whether our system is effective in evoking a sense of presence in participants.

To that end, we have designed an experiment based on previous work by Conklin et. al. [25, 26, 27], and created a modified version of our system to facilitate conducting it. This chapter first describes how we designed and set up our experiment as well as the process for running it, and then presents and discusses the results obtained.

4.1. Methods

4.1.1. Experimental Design

Through our experiment we wanted to study the impact that environment personalization has on the elicited experience familiarity. Personalization here is in the form of virtual environments composed of audiovisual material from locations in which people frequently spend time. A full list of location criteria is provided in Section 4.1.3.1. Experience familiarity refers to how familiar to their experiences in the real world participants think each virtual environment being presented feels.
The main hypothesis we wanted to test is as follows:

H1: Personalized virtual environments lead to a more familiar experience than non-personalized virtual environments.

We are additionally:

Q1. Exploring the usability of a user interface that participants are asked to interact with while viewing virtual environments.

Q2. Examining the ease of material collection by participants.

To test our hypotheses we used a single-factor within-subjects repeated measures design with personalization of the environment as our factor. Personalized and non-personalized virtual environments were presented in a counterbalanced order, which can be seen in Figure 5 of Appendix C, and is discussed further in Section 4.1.3.2. The conceptual design of our experiment is illustrated by Figure 9.

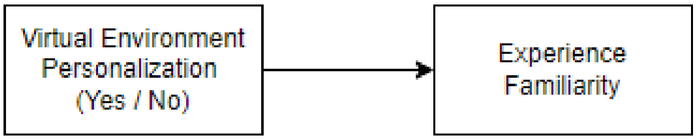


Figure 9: Conceptual Design

4.1.1.1 Exploration of Smoking Cravings

While here we are primarily focusing on whether personalized virtual environments presented by our system elicit a higher experience familiarity than non-personalized virtual environments, we are additionally interested in exploring whether personalized virtual environments can also directly elicit higher smoking cravings than non-personalized ones. To that end, we created an additional design that we used for eligible smoker participants, as described in Section 4.1.2. This was a 2x2 within-subjects repeated measures design with personalization and cue-type of the environment being our two factors. Essentially, 1. personalized smoking-cue, 2. personalized non-smoking-cue, 3. non-personalized smoking-cue, and 4. non-personalized non-smoking-cue environments were presented in a counterbalanced order, which can be seen in Figure 4 of Appendix C, and is discussed further in Section 4.1.3.2. The conceptual design of our experiment regarding smoking cravings is illustrated by Figure 10.

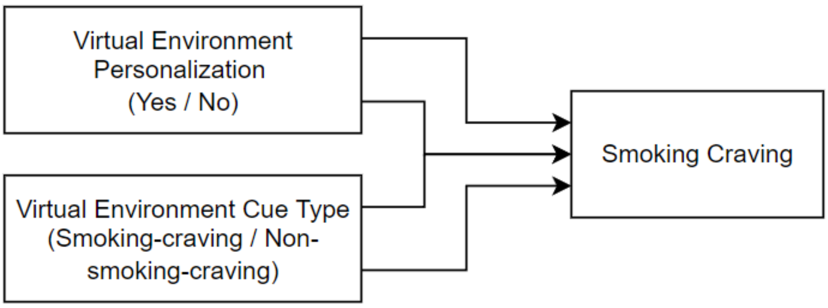


Figure 10: Smoking Craving Exploration Conceptual Design

4.1.2. Participants

We recruited participants using posts on social media, flyers, as well as in-person recruitment. Eligible participants were those over the age of 18, and in the case where they also fit the Center for Disease Control (CDC) criteria for ”current smokers” [21], they were placed in our exploratory design experiment. Additionally, participants needed to own or at least have access to a smartphone capable of recording high-quality video (at least 1080p resolution) and audio, so that they could collect the necessary audiovisual material.

Participants who completed the experiment were awarded with a choice of either a 32GB USB stick or a 5 Euro gift certificate for a popular online store in the Netherlands.

For our main design (single-factor) the planned sample size was 16. We used the software program G*Power to conduct a power analysis. Our goal was to obtain .95 power to detect a medium effect size of .3 according to Cohen [24], at the standard .05 alpha error probability. The resulting sample size was 20, however, since we are using Bayesian methods which require no more samples than frequentist methods and for which sample size is less relevant [22], and due to time constraints we aimed for a sample size of 16. For our exploratory analysis (2x2 design) the resulting sample size was 32. However, for the same reasons as above we aimed for a sample size of 24, although we primarily focused on getting 16 people overall for our main analysis. In both cases, the sample sizes were selected because they also fulfilled our counterbalancing requirement of assigning an equal number of participants to each of our environment presentation sequences, described in Section 4.1.3.2. In the case of smokers we needed a multiple of 8 as sample size, while for non-smokers we needed a multiple of 4 as sample size. So if a smoker was assigned to counterbalanced sequence 6, this would correspond to non-smoking sequence 2, therefore making the two sequence types compatible for experimental design and data analysis purposes.

One participant developed symptoms associated with cybersickness during the experiment after viewing three virtual environments. They were knowingly prone to having motion sickness but still wanted to participate. They attempted to push through and after viewing three more environments they could not continue. For this participant we used the partial data obtained for the virtual environments presentation. They were, however, able to fill-in the sense of presence and UI usability questionnaires. Our final sample size was 18 non-smokers and 1 smoker. These were 11 men and 8 women of various nationalities, with all of them holding higher education degrees, while most of them ($> 75\%$) were PhD students at various faculties and research groups at TU Delft.

4.1.3. Materials

4.1.3.1 Virtual Environments

To test the effect of personalization, we presented participants with both personalized and non-personalized environments to detect differences in the effects measured. As described in Sections 3.1.2 and 3.2.1.1, VEs were composed of panoramic photos and audio clips. These resulted from processing video and audio recordings, referred to as virtual environment material (VEM). Standard (non-personalized) environments were created using researcher-sourced VEM, while participants were requested to submit VEM which were then used to create personalized VEs (PVEs) to be presented to them.

Standard VEs. Standard virtual environments (SVEs) were composed of VEM sourced from those locations reported by Conklin et al. [26] which participants in their studies have reported to either smoke or not smoke most often. These smoking scenes are 1. bar, 2. having lunch at home, 3. bus stop, and the non-smoking scenes are 1. church, 2. gym, 3. shower. We visited corresponding locations and sourced that material, which we then used to create these SVEs that we presented participants with. The image for the SVE depicting a church can be viewed in Figure 7, and the one depicting a bus / tram stop can be viewed in Figure 8 of Appendix B. It is important to note that the smoking locations were specifically chosen so that people are able to smoke there. For instance, in an earlier Conklin study [27] one of the environments most rated to cause smoking cravings was a restaurant. After smoking was banned in most indoor locations, however, this environment was switched to be a scene depicting having lunch at home. This was to avoid potential issues arising from the unavailability to smoke at certain locations, such as people lacking those associations in their addiction model that would cause the activation of relevant conditional responses, in this case, the craving to smoke. The material we used in our experiment for the standard environments can be accessed in the ”Accompanying Material” folder in our Github repository [4].

Personalized VEs. PVEs were composed of VEM sourced by participants. We used similar specifications as in the study conducted by Conklin et al. [26] to provide a set of guidelines to help participants choose which locations to collect material from. These specifications were more detailed for smoker participants due to the nature of the intended environments, while we asked non-smoker participants to collect material simply from locations in which they most frequently spend time during a typical week. The specifications for collecting material from smoking and non-smoking locations that we gave smoker participants are listed in Table 4.1, and the full set of instructions given can be seen in Appendix F.

Smoking VEM	Non-Smoking VEM
Spend time at least once a week.	
Smoked at least 7 out of 10 times when there.	Smoked 3 or less times out of 10 times when there.
Found the difficulty of not smoking when there to be at least 5 on a 0-10 scale.	The difficulty of refraining from smoking when there was 5 or less on a 0-10 scale.
May feel at least somewhat negatively if they were not able to smoke there.	The strength of thinking about smoking when there was 5 or less on a 0-10 scale.

Table 4.1: Specification for Smokers to Collect VEM

Participants were given instructions as to how they could record video that would allow us to create $360°$ (or $> 180°$ if they only sit in that location) panoramic photos, and that they had to provide us with a sound recording lasting 3 minutes and 20 seconds in every location. Each VE was presented for 3 minutes, and the extra time was used as a buffer in case there were issues with parts of the audio (glitches, noises from handling the phone, etc.). Participants were additionally instructed to collect the requested material at a time of day when they usually spend time there, and from their usual point of view when in that location. This was to increase the chances that contextual smoking cues were present in the material. We requested that participants submit VEM from 8 locations (4 smoking and 4 non-smoking in the case of smokers), and we used 6 (3 each for smokers). The extra 2 VEM were requested in case there we were unable to use the material from one location, so that we could have a replacement.

4.1.3.2 Counterbalancing

In an effort to potentially avoid order, sequence, and carryover effects in our measured variables we presented VEs in a counterbalanced order. This was first created with the exploratory smoking-cravings design in mind using two factors (environment cue-type, and personalization), and was then simplified for our main single factor design by removing the environment cue-type factor and only considering environment personalization. We go in further detail about this process in Section C.1 of Appendix C, including diagrams illustrating the counterbalanced virtual environment sequences in Figures 4 and 5 of Appendix C.

4.1.3.3 Attention-Focusing Tasks

During a pilot run of our experiment, participants generally expressed that the VE exposure time period was rather long, and that it was prone to getting boring since there was nothing to do in the environments. This could have potentially left a negative impression of the system and VEs, causing them to have a worse experience than they would normally have, thus perhaps affecting their questionnaire responses. And especially for our exploratory (smoking craving) design, boredom could have also increased the smoking cravings elicited [91]. Therefore to combat this, while participants were viewing virtual environments, they were sent a notification on the virtual smartphone at 45, 90, and 135 seconds, instructing them to focus more on certain aspects of the environment, such as objects closer or further away. This was in an attempt to give participants something to do, while not distracting them from what we wanted them to do, that is, to experience the VEs.

4.1.3.4 Modified System

To facilitate conducting our experiment we created a modified version of our system. As mentioned at the start of this chapter, the evaluation had to do with testing the efficacy of PVEs in eliciting the desired effects (experience familiarity and smoking cravings). While on the other hand, the main system we designed primarily focuses at facilitating communication between researchers and participants while the latter are viewing VE. We therefore created a version able to run using a single instance of the application, as opposed to requiring two instances and host-client connection in the case of our main solution. Consequently, we also omitted the communication component while retaining the participant user interface, on which we presented questionnaires aimed at evaluating the effects elicited. Lastly, we implemented an automated solution to assist in presenting participants with different environments and collecting their responses to questionnaires, to minimize interaction with researchers during the experiment.

Virtual Smartphone. While used in our main solution for participants to respond to researcher messages, for the purposes of our experiment we used the virtual smartphone to present a questionnaire at the end of each virtual environment presentation. This was done primarily so that we could assess its perceived usability, while also giving us a convenient way to present the intended questionnaires without having to take participants outside the virtual environments, or present a separate user interface. How this modified participant user interface appears is illustrated in Figure 11.

Figure 11: Modified Participant User Interface

4.1.4. Measures

The primary variable relevant to our main hypothesis is experience familiarity. We additionally wanted to examine the usability of the user interface participants were asked to interact with (digital smartphone), the sense of presence elicited by the virtual environments, as well as how difficult participants found the material preparation to be.

4.1.4.1 Experience Familiarity

We measured how familiar the experience in each virtual environment felt to participants, compared to experiences they have had in the real world. To do so we used a single-point question; namely: ”How familiar to your experiences in the real world does your experience in this current virtual environment feel?”. We measured this on a scale of -5 (completely unfamiliar) to 5 (completely familiar), with the mid-point (0) labeled ”Neutral.”

4.1.4.2 User Interface Usability

We examined how participants perceived the virtual smartphone they were asked to interact with, and so we had them respond to post-trial questionnaires. We used two questionnaires to assess this; the Component-Based Usability Questionnaire (CBUQ) [14] which collects responses regarding perceived ease-of-use, and an additional set of questions aimed at examining specific aspects of the user interface.

These latter questions asked participants to evaluate their experience with the virtual smartphone on a scale of (-5,5) in terms of its size, position and orientation on the screen, text legibility, UI movement preference, their opinion on the notification sound that was played, as well as how they found the duration of each virtual environment viewing. A full list of these additional usability questions can be seen in Table E.1 of Appendix E. Furthermore, after each of these questions participants were asked to provide a free-text response for why they gave that specific score.

4.1.4.3 Sense of Presence

To assess the perceived sense of presence elicited by the virtual environments we presented participants with, we used the Igroup Presence Questionnaire (IPQ) [43]. This has been constructed using questions contained in questionnaires of other studies [18, 41, 80, 92] along with questions validated by the Igroup [65, 70, 71, 72]. This questionnaire is composed of 14 questions on a scale of -3 to 3, and assesses sense of presence in four dimensions; general presence, spatial presence, involvement, and expected realism.

4.1.4.4 Ease of Material Collection

Finally, we asked participants to give us an assessment of how difficult they found material collection to be. This was on a scale of -5 (very difficult) to 5 (very easy), with a mid-point (0) labeled ”Neutral.” Our operational requirements state that the process needed to be accessible to LSES people and not take longer than two hours, and therefore we aimed to assess how difficult participants perceived it.

4.1.4.5 Smoking Cravings

To measure smoking cravings in smoker participants, we used the short-form of the Tobacco Craving Questionnaire evaluated by Heishman et al. [39, 40]. We chose this because it was reported to assess smoking cravings in four dimensions (emotionality, expectancy, compulsivity, and purposefulness). Heishman et al. additionally reported that results of their study are congruent with relevant underlying theories exploring addiction, conditioning and cognition, therefore adding value to the measures employed. It is also composed of a relatively short number of questions (12), which would not overburden participants to complete after every environment presentation.

4.1.4.6 Metadata

Binary metadata was also collected from VEs relevant to our exploratory analysis. 1. Whether there were proximal smoking cues present in the environment i.e., cigarettes, rolling tobacco and assorted paraphernalia, ashtrays, cigarette smoke. 2. Whether there were popular contextual smoking cues present, that is, coffee, alcohol, food in the environment. 3. Whether there were people present in the environment.

4.1.5. Procedure

We scheduled appointments for participants to come to the TU Delft campus to participate in our experiment, and they were instructed to prepare and submit the necessary material a few days in advance to their appointment, for us to create PVEs.

Before the experiment began, participants were asked to complete a short instructional exercise, having them look around in a virtual environment and locate specific items. This was aimed at making them aware of the spatial rotation capabilities of VR. The VE was composed of an image of the Milky Way Galaxy, which was selected for the likelihood of its neutrality as pertaining to causing smoking cravings (relevant to our exploratory analysis). Participants were then presented with the virtual smartphone which was used in our communication solution, and were instructed as to how they could use the gamepad controller to interact with the user interface for the purpose of answering and submitting the questions we would present them during the experiment.

A high-level overview of the experiment process is illustrated by Figure 5 in Appendix D, and here we provide a more detailed description. The experiment itself started by having participants put on the HMD, presenting them with an empty dark-gray-colored virtual environment, and having them respond to the relevant questionnaires so that we could collect baseline levels of the variables measured. This was important for smoker participants so that we could gauge how high their initial smoking cravings were at the experiment start, however, we also took an experience familiarity measurement for the general process to be consistent for everyone. Participants were then presented with VEs based on the counterbalanced sequence they were assigned to, for a time period of three minutes each. Shiffman et. al. (2013 [74]) reported as three minutes seemingly being the period of time in which smoking cravings peak when exposed to relevant cues, which is why we chose it as our virtual environment presentation time.

After the three minutes had elapsed, participants were presented with the experience familiarity question and smokers were additionally presented with the smoking craving questionnaire on the virtual smartphone. Once they had submitted their answers they were presented with a transitional environment (photo of the Milky Way galaxy) and were instructed to remove the HMD and relax for 30 seconds. This relaxation time was aimed at both combating fatigue that could arise from viewing virtual environments, and to potentially avoid participants developing symptoms associated with cybersickness. When those 30 seconds had elapsed, participants were instructed to again put on the HMD and were then presented with the next environment in their sequence. This cycle ended when they had submitted questionnaire responses for the last environment presented (12 in total).

Lastly after removing the HMD, participants were asked to complete an online questionnaire to collect responses regarding the sense of presence elicited by our system, the usability of the user interface and virtual smartphone they were asked to interact with, as well as how difficult they found the material collection process to be.

4.1.6. Data Pre-Processing and Analysis Strategies

To analyze our obtained data, we employed Bayesian statistics and created different models to infer how likely our hypotheses were to be true. Before doing so, however, we pre-processed the obtained data and computed indices where appropriate.

4.1.6.1 Data Pre-Processing

For experience familiarity scores the only pre-processing that took place was simply the combination of responses from every user into a single dataframe.

For data obtained via the CBUQ, we calculated Chronbach's alpha (0.95, N = 19), which is above the threshold of 0.7 ($\pm$ 0.1) mentioned by Lowenthal [54]. We then computed an index by calculating the mean scores of each user, to obtain a single value representing the interface usability.

For our additional usability questions we also reversed the score of one question, and we additionally transformed the scores of two questions from a (-5, 5) to a (0, 5) scale. This was because the middle value was the one with positive sentiment regarding our user interface, while the two extremes were negative. For instance, in the question regarding the size of the virtual smartphone, the two extremes were labeled "Too small" and "Too big", with the middle (0) value being labeled "Just Right." For the purposes of our analysis, the two extremes were functionally identical and therefore we found it appropriate to transform the scale to reflect the middle value being the one reflecting positive sentiment.

For sense of presence, we first computed Chronbach's alpha and obtained a result of (0.88, N = 19), which is consistent with the reliability of 0.85/0.87 reported by the Igroup [43]. Then, according to the Igroup guidelines, even though the questions were on a scale of (-3, 3), they needed to be transformed to a (0,6) scale and additionally, three questions with negative wording (SP2, INV3, REAL1) needed to have their score reversed. We then computed an index for each subscale of the questionnaire; namely 1. General Presence, 2. Spatial Presence, 3. Involvement, and 4. Experienced Realism, by calculating the mean scores obtained.

4.1.6.2 H1: Impact of Virtual Environment Personalization on Experience Familiarity

We employed a multi-level Bayesian approach on the impact of environment personalization on experience familiarity score. Our independent variable was the environment personalization (0 or 1), and our dependent variable here was the experience familiarity score.

For our predictors we considered a general mean, a random intercept for each participant, and the fixed effect of whether the environments were personalized or not. We fit models that consecutively contained more effects, meaning that our first model only contained the general mean and random intercepts, and the last model contained the general mean, the random intercepts, as well as our fixed main effect.

We fit a t-distribution and used diffuse priors according to McElreath [55] when estimating our models, and have also conducted a prior sensitivity analysis to assess the impact of different settings for the priors. We saw no change in the estimated posterior probability that our hypothesis is true with our sensitivity analysis. We compared our estimated models based on the Watanabe-Akaike Information Criterion (WAIC) [90], which gave us an estimate of prediction error and the quality of each model relative to every other model, and we reported 95% highest posterior density for estimators.

For the experience familiarity score we sampled from the posterior of an estimated model. Based on these samples, we computed the posterior probability of H1 being true. We evaluated this posterior probability based on the guidelines listed by Chechile (2020, pp.238) [22]. We additionally extended the guidelines by Chechile to also account for posterior probabilities less than 0.5 based on Andraszewicz et al. (2015) [3].

Furthermore, we used the Region of Practical Equivalence (ROPE) as a secondary method to evaluate the results as described by Kruschke (2018) [46]. ROPE lets us know whether our results are significantly different from a negligible magnitude. The size of the ROPE was based on the fallback guidelines described by Kruschke, based on a small Cohen's effect size [24]. We computed 95% Highest Posterior Density Intervals (HPDI), which are the narrowest interval containing the specified probability mass to help indicate which are the most credible points in a distribution, in combination with ROPE [45, 47] to make a definitive decision regarding our hypothesis. We tested whether our hypothesis could be rejected, confirmed, or a decision should be withheld depending on the location of the HPDI relative to the ROPE, according to Kruschke's guidelines.

4.1.6.3 Q1: User Interface Usability, Q2: Ease of Material Collection

For each additional user interface usability question, as described in Section 4.1.4.2, as well as for the perceived ease of material collection by participants, we proceeded analogously to our H1, but only fit a Bayesian model with a general mean. We also did not use ROPE, and we estimated the posterior probability that the mean score was greater than 0 (or 2.5 in the case of 0-5 scale) for each question.

For the data obtained from the CBUQ we proceeded analogously to the above by fitting a Bayesian model on the index obtained. We estimated the posterior probability that the mean score was greater than the norm value of 5.29, above which we can argue that our system may be perceived as easy to use [14].

We additionally conducted a prior sensitivity analysis to assess the impact of different settings for the priors of our models. We found no difference in the posterior probabilities calculated.

4.1.6.4 Exploratory Analysis

As an exploratory analysis we examined the effect that sense of presence had on elicited experience familiarity. Similarly to previous sections, we fit a Bayesian model for each sense of presence sub-scale with a general mean, random intercept for each participant, and also a model using all subscales. We calculated the posterior probability that the hypothesis "this subscale has as positive effect on elicited experience familiarity" holds based on the guidelines listed by Chechile (2020, pp.238) [22], for each model. We additionally conducted a prior sensitivity analysis to assess the impact of different settings for the priors of our models, and found minimal (<0.017) to no difference in the posterior probabilities calculated.

Furthermore, we examined the sense of presence elicited by our system overall by fitting a Bayesian model with a general mean and random intercept for each participant, and compared these scores to reported norms provided by the Igroup [43] in a database of 542 research cases. We additionally conducted a prior sensitivity analysis to assess the impact of different settings for the priors of our models, and found minimal (<0.007) to no difference in the posterior probabilities calculated.

Additionally, we examined the free-text responses to our additional user interface usability questions and participants' perceived ease of material collection to investigate the rationale they had for the given scores.

Lastly, we also examined the feasibility of conducting a study such as ours, pertaining primarily to participant recruitment.

4.2. Results

This section contains results obtained by our analysis using data collected via our experiment. We start by presenting results for our main hypothesis regarding experience familiarity, and then move on to results for usability questions regarding our user interface as well as ease of material collection. Finally, we present results of our exploratory analysis regarding sense of presence. Our dataset and analysis code may be obtained at the 4TU ResearchData repository [6].

4.2.1. H1: Impact of Virtual Environment Personalization on Experience Familiarity

We first created a boxplot of the data to visualize how experience familiarity (ExpFam) elicited was impacted virtual environment personalization, which is illustrated by Figure 12. We can see that while ExpFam scores for standard virtual environments (PVEs) are distributed across the entire range of available values (-5,5), this does not appear to be the case when considering personalized VEs (PVEs). PVEs overall score higher in ExpFam than non-personalized ones with a median value of 4 as opposed to 2 in the case of SVEs. The results of our analysis can be seen in Table 4.2.

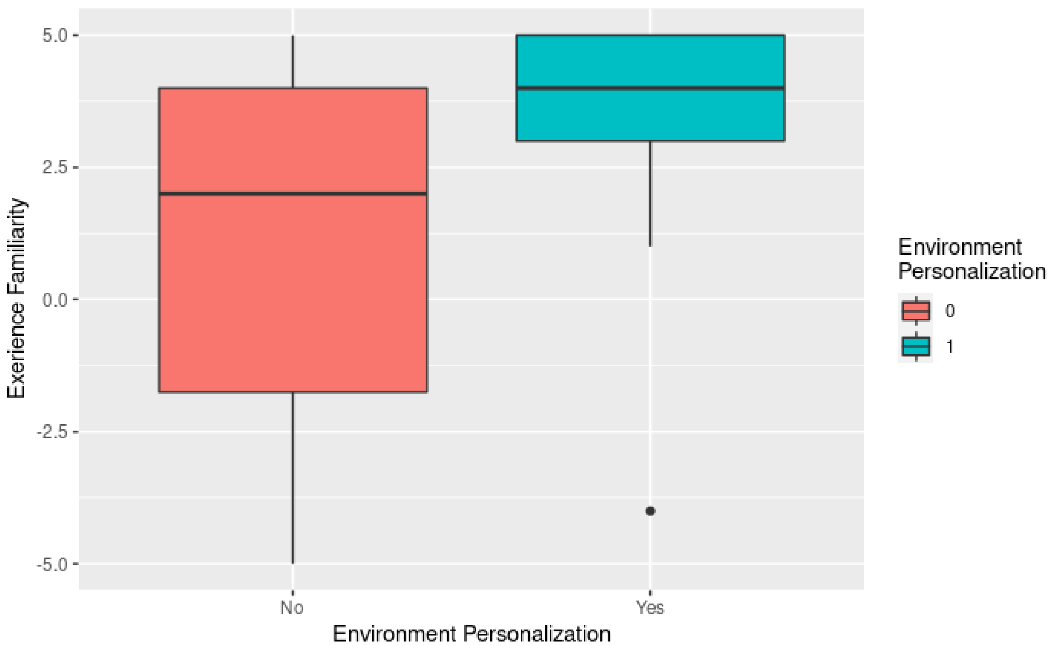


Figure 12: Experience Familiarity Across Environment Personalization

	Mean	STD	2.5% HPDI	97.5% HPDI
ExpFam	1.36	0.23	0.93	1.83

Table 4.2: Personalization Effect on Experience Familiarity Based on Model Estimators.

We sampled from the posterior of our second model (m1), which uses a general mean, random intercept for each participant and the personalization fixed effect, and computed the posterior probability that our hypothesis is true. This resulted in a probability of > 0.9995. By consulting the guidelines listed by Chechile [22], this leads us to conclude that our hypothesis is nearing certainty.

We also computed the HPDI for our predictor and calculated the ROPE to observe the location of the former relative to the latter. The ROPE calculated was in the range of (-0.1, 0.1), while our 95% computed HPDI was (0.92, 1.82). Since the ROPE lies outside and is lower than the HPDI, we can reject the null hypothesis defined as there being negative or no relation between virtual environment personalization and experience familiarity. A comparison between our calculated HPDI and ROPE is illustrated by Figure 13.

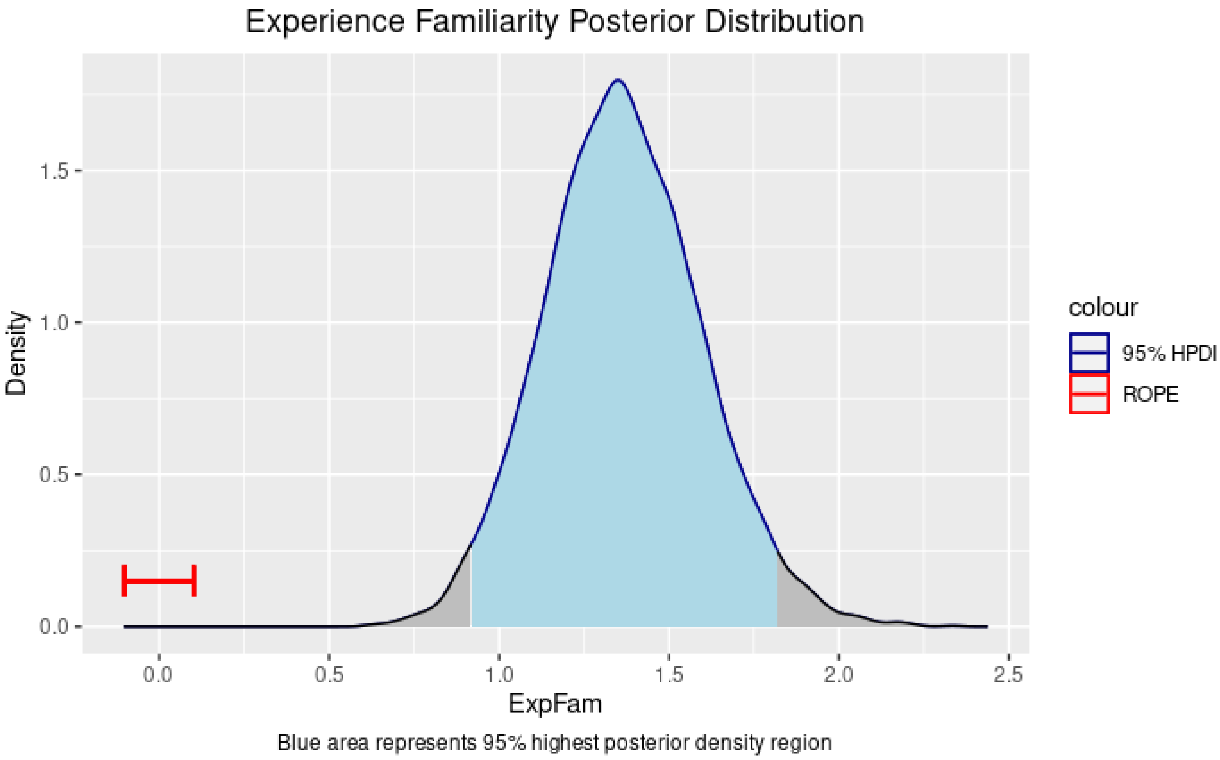


Figure 13: Comparison of Posterior Highest Density Intervals - Region of Practical Equivalence.

Finally, we computed the WAIC and compared our two models, the results of which can be seen in Table 4.3. We can see that the model also containing our fixed effect (m1) scored lower in the WAIC (better), and has a lower standard error (SE) than the model only containing a general mean and random intercept (m0).

	WAIC	SE	Weight
h1_m1	975.33	40.79	1
h1_m0	1016.71	43.21	0

Table 4.3: WAIC Comparison Between the Constructed Two Models.

4.2.2. Q1: User Interface Usability, Q2: Ease of Material Collection

The results of our analysis for the component-based usability questions (CBUQ), our additional usability questions (AUQ), as well as the ease of material collection (EoMC) can be seen in Table 4.4. We list the mean, standard deviation, 95% HPDI, as well as the scale in which each was measured. We provide the latter to put the obtained results in context, since the CBUQ was measured using a different scale than our additional questions, and two questions in the latter had their scale transformed for analysis purposes.

	Mean	STD	2.5% HPDI	97.5% HPDI	Scale Measured
CBUQ	6.22	0.21	5.81	6.61	0,7
AUQ1	3.04	0.38	2.23	3.71	-5,5
AUQ2	5.0	0	5	5	0,5
AUQ3	2.9	0.54	1.85	3.91	-5,5
AUQ4	2.52	0.49	1.55	3.46	-5,5
AUQ5	2.55	0.46	1.64	3.44	-5,5
AUQ6	2.8	0.7	1.39	4.13	-5,5
AUQ7	2.33	0.39	1.59	3.09	0,5
EoMC	0.21	0.61	-0.99	1.47	-5,5

Table 4.4: Usability and Ease of Material Collection Questions Results
Abbreviations: STD: Standard Deviation, HPDI: Highest Posterior Density Intervals, CBUQ: Component-Based Usability Questionnaire, AUQ: Additional Usability Questions, EoMC: Ease of Material Collection

Firstly, for the CBUQ we can compare the obtained HPDI to the 5.29 norm mean reported by Briman et al. [14]. Since both the mean score and the HPDI lower limit of our user interface are higher than the norm mean, we can argue that our user interface would be easy to use.

For our AUQ, the mean scores for all questions other than 7 are larger than the neutral value (0 or 2.5 in the case of our transformed scales), and the HPDI do not include the neutral value. AUQ7 mean score is slightly lower than the neutral value of 2.5, and the HPDI contains values lower than the neutral value.

Finally, for the EoMC, the mean score is slightly larger than the neutral value of 0, but the HPDI contains negative values.

4.2.3. Exploratory Analysis

4.2.3.1 One Smoker Case Study

One participant was a smoker, and therefore was given the corresponding material collection instructions and was placed in our 2x2 experimental setup consisting of presenting them with personalized or non-personalized, smoking or non-smoking VEs. They were rather vocal about feeling high smoking cravings when presented with personalized smoking environments, as well as when presented with certain non-personalized smoking environments which they were familiar with, something which is discussed further in the next section. An overview of the data recorded showed higher cravings elicited by personalized smoking environments. This hinted rather positively for the overall effectiveness of our system, however, a more thorough evaluation would be necessary before we can make a data-driven conclusion on the capabilities of our system to elicit high smoking cravings, or whether personalized smoking-cue virtual environments would elicit higher smoking cravings than non-personalized ones.

4.2.3.2 Free-Text Responses and Discussion

Below we provide a general summary of responses received from our free-text questions after every AUQ and EoMC question posed. We additionally provide a brief discussion regarding our findings for each question or group of questions. Specific participant responses may be referenced at certain points, which can be accessed in our dataset uploaded to the 4TU ResearchData repository [6].

AUQ1: Text Legibility. In general, most text seemed to be sufficiently legible. Questions, however, were answered via labeled sliders and the font of those labels which was smaller than the rest was harder for some people (P2, 7, 9, 16, 17, 18) to read because of general blurriness experienced by them.

> ”Some of the smaller text was a bit blurry, but overall it was quite clear.” - P2

AUQ2: Virtual Smartphone Size. Most participants reported that the virtual smartphone felt like the usual size of a real smartphone and did not take too much screen space. Some reported that it blocked view of the environment enough to give it focus, but not so much that they could not see where they were in the virtual environment (P2, 18). A few reported that a larger screen might have been easier to read (P7, 19), and some others reported that perhaps making it smaller would have been nice (P1, 15).

AUQ3: Location of Virtual Smartphone in Field of View. Most participants found where the virtual smartphone was located in their field of view to be in a natural position comparable to where they usually hold their phone in the real world. P3 additionally mentioned that its position helped with seeing it and responding to what was intended immediately.

> ”It was close to where I actually hold my phone (in my field of view).” - P15

AUQ4: Orientation/Tilt of Virtual Smartphone. Similarly to the above, the orientation / tilt of the virtual smartphone mostly felt natural, however, P11 mentioned that when looking in extreme angles (far up or down), this detracted from the realism.

> ”Felt like when using a smartphone in real world.” - P7

AUQ5: Notification Sound and Virtual Smartphone Appearing. Participants generally reported that the notification sound followed by the virtual smartphone appearing provided a good indication of what needed to be done next, and that it was consistent with a real smartphone. P3 reported being startled by the notification sound and accompanied gamepad controller rumble at first but was soon accustomed to it and even came to expect it. Further, P14 mentioned that it may have been more startling if the virtual smartphone appeared without warning, and the notification sound acted like a calm warning.

AUQ6: Virtual Smartphone Moving with Camera vs Remaining Stable. Opinions for this question were a little more mixed. Some participants liked that the virtual smartphone moved with them looking around, since reading the notifications or responding to the questionnaire were the current objectives, and that they would not forget or have to look for it (P2, 3, 8, 14, 19). P3 and 10 also liked that they could toggle the smartphone visibility once they have read the message presented. On the other hand, some participants reported that they would have preferred if the virtual smartphone remained stable because in their opinion it would have been more akin to the real world where people do not necessarily move their smartphones with their head movement (P7, 15), and that this caused view blocking. Additionally, P12 felt that since the rest of the environment was static, the virtual smartphone moving felt more intrusive.

> ”I feel like its more intrusive to move with you as the rest of the environment is completely static.” - P12

While it would be difficult to judge without having a comparison of interaction with the interface if it remained stable in the environment as opposed to always being centered in their field of view, some people seemed positive to that idea. It is also worth noting that some participants seemed to have forgotten that they could make the interface appear and disappear at will by pressing the corresponding button on the gamepad controller. We believe that is the case based on observations during the experiment and from some comments received. We hypothesize that if they had not forgotten, certain comments in this and previous questions regarding the blocking of views, would perhaps have not been made.

AUQ7: Virtual Environment Viewing Duration. The comments here had generally somewhat negative sentiment. The viewing itself did not seem to be negative towards the environments themselves, but the duration seemed to be rather long for most participants. The general sentiment was that it only took a short amount of time to view their entire surrounding and did not know where to look after a while. P8 reported that they started to daydream, while P3 reported that at a certain point they were just waiting for the notification sound to signify the end. P7 hypothesized that the viewing duration may have felt longer due to the number of environments they were asked to view.

> ”Maybe because of the high number of environments.” - P7 (In regards to giving a score reflecting feeling that the VE viewing duration was too long.)

P12 reported that unfamiliar environments encouraged them to explore their surroundings, as opposed to familiar environments, however, due to the lack of high image fidelity and their inability to move, it became frustrating when they could not further investigate something that caught their interest.

> ”Most of the environments are static and familiar, which makes them uninteresting. Stranger environments encouraged me to investigate more, but the lack of resolution or inability to move became frustrating because if something does catch my interest, I can’t do anything about it (like look further or move around).” - P12

Ease of Material Collection. The score was rather neutral but sentiment was somewhat negative. This was especially because a Covid-19-related lockdown which either gave people (P12, 15) less locations they could record in during their routine, or made certain locations outright inaccessible, such as gyms. Most participants reported that the process itself was not difficult, but it was rather time consuming, especially if they had to go out of their way to visit certain locations to record the necessary material. P2 and 14 reported having difficulty deciding where to record, P2 and 13 reported being uncomfortable recording in locations with other people present, and P3 even mentioned that they kept forgetting to record in familiar locations.

> ”It’s hard to get reminded of recording this. It’s particular hard to remember when you are at your familiar spots, as they are that, familiar, not triggering things to you (such as a need to record it). Then, it took me some time to understand it would be used directly in the VR environment. Recording the sound for that long is bothersome.” - P3

4.2.3.3 Sense of Presence

Results for the sense of presence questionnaire can be seen in Table 4.5. These are contrasted with the mean scores in the Igroup database of 542 research cases to examine how our system compared with a set of norm values for the questionnaire.

	Mean	STD	2.5% HPDI	97.5% HPDI	Igroup DB Mean
General Presence	4.26	0.3	3.67	4.82	3.39
Spatial Presence	3.89	0.25	3.4	4.38	3.75
Involvement	3.43	0.3	2.81	3.99	2.99
Experienced Realism	2.74	0.21	2.34	3.14	1.74

Table 4.5: Sense of Presence Questionnaire Results Compared with Research-Reported Norms

Overall the range of obtained results for each subscale are in line or in some cases better than the mean scores contained in the dataset of 542 cases provided by the Igroup [43], therefore reflecting positively on the overall sense of presence elicited by our system.
To investigate whether experience familiarity was impacted by sense of presence, we constructed models for each subscale. Results can be seen in Table 4.6. The posterior probability calculated here refers to how likely it is for the hypothesis ”This subscale has a positive on the experience familiarity elicited” to hold.

	Mean	STD	2.5% HPDI	97.5% HPDI	Posterior Probability	Chechile Interpretation
General Presence	0.09	0.14	-0.18	0.38	0.76	Only a casual bet on h.
Spatial Presence	0.03	0.15	-0.3	0.31	0.57	Not worth betting on h.
Involvement	0.19	0.14	-0.08	0.45	0.92	A promising but risky bet for h.
Experienced Realism	0.04	0.17	-0.3	0.36	0.59	Not worth betting on h.

Table 4.6: Sense of Presence Effect on Experience Familiarity in Separate Models

Comparing the calculated posterior probability with the guidelines listed by Chechile [22], we can see that for some scales it is not worth betting on our hypothesis (Posterior Probability in range of $[0.5, 0.75)$), and for others it is only a casual bet on our hypothesis (Posterior Probability in range of $[0.75, 0.9]$) or a promising but risky bet (Posterior Probability in range of $[0.9, 0.95]$). This leads us to conclude that there is some indication that experience familiarity may somewhat be positively affected by the sense of presence elicited by our system.

4.2.3.4 Study Feasibility

Lastly, we wanted to examine how feasible we found conducting our study to be. We made compromises in terms of the technologies used due to operational requirements, however, the biggest issue impacting the feasibility of our study related to difficulties in recruiting participants. Initially, the aim was examine smoking cravings elicited by personalized virtual environments, and so we set out to recruit smokers to participate in our experiment. However, due to the Covid-19 pandemic and relevant lockdowns, our university campus having become smoke-free making it difficult to locate people, as well as our limited budget, we were unable to recruit the necessary amount of smokers to conduct the study.

We created posts on social media and put up flyers, but even after more than fifty people viewed our informed consent form, none signed it. We additionally hired a teaching assistant to help us recruit smokers, and out of the fifteen people who supposedly showed some interest in participating, only one scheduled an appointment and later had to drop out. The one smoker who participated was through word of mouth from an acquaintance.

In hindsight, what was required of participants may have been a factor perhaps discouraging people from wanting to participate. People were asked to visit physical locations and record audiovisual material, which was reportedly uncomfortable for some non-smokers who did participate. They were also asked to physically visit the lab to participate in our experiment, something which also is likely to decrease participant retention rates. Coupled with our limited budget and therefore a comparatively small reward, we believe that these factors were detrimental in recruiting smokers.

4.3. Results Discussion

In this section we discuss and interpret our results given the analysis performed on the obtained data. We first begin with our main hypothesis (H1), then proceed with our user interface usability scores, ease of material collection, and free-text responses (Q1, Q2). Finally, we discuss our exploratory analysis regarding sense of presence and the one-smoker case study.

4.3.1. Experience Familiarity (H1)

The results obtained demonstrate that in our experiment, whether a virtual environment is personalized or not can have an impact on the magnitude of the experience familiarity elicited. Therefore, we have confirmed that in our experiment, virtual environment personalization can be viewed as a positive predictor of how familiar the experience in a virtual environment feels, assuming the validity of the question posed.

4.3.2. User Interface Usability and Ease of Material Collection (Q1, Q2)

For our AUQ apart from question 7 have all received scores higher than the neutral value (0 or 2.5 in the case of our two transformed scaled), indicating generally positive user feedback regarding aspects of the user interface usability we wanted to examine.

AUQ 7 on the other hand, received a slightly negative overall score, and the EoMC received a relatively neutral score. This is consistent with our previous impression from feedback we have received, that indeed participants found the virtual environment viewing time to be slightly longer than desired, and material collection to be a somewhat tedious process.

4.3.3. AUQ Free-Text Responses

In general, sentiment and comments reflected positively on the user interface. It is worth noting that a number of people did not sufficiently adjust the HMD on both axes (via a physical slider on the device and by shifting it vertically on their face), even after initial instructions to do so. Therefore, some issues concerning image fidelity or text readability may have been avoided otherwise. Additionally, given the varying opinions regarding the size, orientation, and location of the virtual smartphone, for future directions it may be worthwhile to create presets and run a survey on what most people prefer by having a direct comparison, before a specific configuration can be posited as being preferable. Alternatively, customization options could be offered for each user to tailor the interface to their liking, so that interaction with it feels more natural to them.

4.3.4. Limitations

While our results regarding PVEs eliciting higher experience familiarity than SVEs look promising, there are certain limitations which may hinder their wider generalizability, pointing towards more research being warranted. These limitations primarily relate to our sample and to our material.

Firstly, while we would have preferred to directly test the effects of virtual environment personalization on smoking cravings, we were unable to recruit a sufficient number of smoker participants. We therefore resorted to only testing experience familiarity which we were originally positing it could be a factor in activating smoking craving responses in smokers. So while our results indicate that virtual environment personalization has a positive effect on experience familiarity elicited, additional testing specifically on smoking cravings would be required before we can definitively say that personalized virtual environments indeed cause higher smoking cravings.

Secondly, we cannot exclude the possibility that participants may have responded in a certain way to be in compliance with experimental demand [58] i.e., made a guess as to how we would have wanted them to respond so that our hypothesis is accepted.

Additionally, there may have been some variance in the interpretation of the question and what we were asking participants to rate regarding experience familiarity. We had participants rate how familiar to their experiences in the real world did the experience in each virtual environment feel. This could have been interpreted as us asking about whether when exposed to a virtual environment they felt similarly to how they feel in the real world, or it could alternatively be interpreted more literally as us asking whether an environment looked familiar to them. A location looking familiar may mean that you have been there before or that you have seen something similar. While on the other hand, a virtual environment feeling familiar may mean that you feel as if you would feel in the real world, implying part of what may be encompassed by a sense of presence.

Finally, material for non-personalized environments were collected in Delft, which is also the city where our university is located. Therefore, since most of our participants live in or near Delft, some of those locations may have been familiar to them as well. This may be reflective of our non-personalized environment experience familiarity results being distributed in the entire range of values available.

5

Discussion and Conclusion

We have performed research leading to the design of virtual environments (VEs) and the development of a system to present them. This system makes use of VR technologies for the purpose of providing a controlled environment for researchers to test their smoking cessation approaches using a conversational agent (CA). We then evaluated how effective our system is in eliciting a familiar experience in users who are presented with different VEs.

This chapter first presents a summary of our answers to the posed research sub-questions, and discusses our findings. We then present our contributions, limitations we have identified in our research project, and explore future work that may overcome them as well as improve other aspects in the field of virtual reality.

5.1. Findings

We had posed the research question:

How can personalized virtual environments be created to help with nicotine cue-reactivity assessment?

We broke this down into three sub-questions to address different aspects related to answering our main question, with answers to these succinctly presented below.

Our first sub-question was defined as:

What are the requirements for building virtual environments to elicit smoking cravings?

To answer it we performed research into the field of addiction research and virtual reality, as presented in Chapter 2, to examine which qualities a virtual environment must have to be effective in eliciting smoking cravings. Succinctly, the answer our research produced is that most importantly a virtual environment must contain smoking cues i.e., elements that can cause a smoker to have smoking cravings. Cues can be presented and perceived via visual, auditory, olfactory, gustatory, or tactile sensory systems. Virtual environments by themselves, without external specialized hardware such as scent machines, can only present visual cues. Therefore, we considered what can be presented using a visual medium to elicit smoking cravings. These include 1. proximal cues, which refer to objects directly related to smoking e.g., cigarettes. 2. Contextual cues, such as objects or situations that have been associated with smoking e.g., coffee or alcohol. And 3. complex cues, which are a combination of the above e.g., a bar scene with people smoking and drinking alcohol. Past research has already demonstrated the efficacy of presenting these types of cues using virtual reality to elicit smoking cravings in smokers, and we therefore used it as a basis for building our own virtual environments. Hence, the basic requirements for building virtual environments to elicit smoking cravings is for them to contain cues associated with smoking behavior. In the case of this research project, the cues we could feasibly present were visual and auditory.

Our second sub-question was defined as:

How do we design personalized virtual environments to elicit smoking cravings, for the purpose of enabling researchers to test their conversational agent approaches?

In this project we investigated the element of personalization, i.e., whether the locations in which people smoke most often would act as effective cues to elicit smoking cravings in those people. To answer this sub-question, as presented in Chapter 3, we framed our solution around also developing a system capable of presenting these VEs we would design. This system also facilitates communication between researchers and participants for the purpose of testing CA strategies. Based on the requirements discussed in answering the previous sub-question, we needed to present visual and auditory cues to elicit smoking cravings. Since we wanted to explore whether being in the locations people usually smoke would cause them to have smoking cravings, we considered what they would see and hear when in those locations as being potential smoking cues. Therefore, we decided to use audio-visual material from locations in which people have associated most with smoking, from their point of view when usually there. We then processed and incorporated this material into a $360°$ VR viewer we developed so that we could present it to participants in a manner aiming to feasibly simulate them being there.

We additionally developed user interfaces that would offer specific functionality depending on whether the user was a researcher or a participant, for the purpose of simulating interaction with a CA. The intent here was to enable researchers to test the smoking cessation assistance strategies they would develop. For researchers we offered a UI that was capable of selecting message templates or composing messages and their intended responses and sending them to users through a network connection. While for participants we offered a UI that appeared as a virtual smartphone, on which researcher messages were received. The UI enabled them to reply by selecting from a list of pre-determined responses the one that most accurately reflected what they wanted to say, which in-turn was sent back to the researcher through the network.

Our third sub-questions was defined as:

How do different virtual environments compare with each other in terms of the magnitude of self-reported cravings elicited in each, and how do users evaluate the usability a system developed to present these virtual environments?

Finally, to answer this we conducted an experiment to examine the likelihood that personalized VEs would cause higher smoking cravings than non-personalized ones, and we additionally investigated the perceived usability of our system, as presented in Chapter 4. We presented VEs composed of material recorded by us in locations mentioned by previous work as popularly triggering smoking cravings, such as a bus stop, and compared the elicited effects against VEs composed of participant-provided material.

Experience familiarity elicited was higher in personalized VEs, and this is a promising first step in examining the impact that they could have in also eliciting higher smoking cravings. This is because according to experts consulted, an element of familiarity with locations in which people smoke most often could have created an association between the location and the subsequent smoking behavior. Therefore, the implication would be that the more familiar a smoking location would be, the more likely that people have formed conditioned responses which would be activated when present there. And by corollary, this could also imply that they would have an increase in their craving to smoke when presented with a virtual environment representing that location.

We were unable to recruit the necessary amount of smokers to perform formal analysis on elicited smoking cravings, and therefore we instead opted to only test how familiar was the experience elicited by the VEs presented by our system, which originally was our second measure. Our one smoker participant was rather vocal about feeling higher smoking cravings in personalized and some non-personalized smoking locations, and an overview of the data recorded showed higher cravings elicited by personalized smoking-cue virtual environments. This hinted rather positively for the overall effectiveness of our system, however, a more thorough evaluation would be necessary before we can make a data-driven conclusion on smoking cravings elicited.

The participants in our experiment generally had a positive opinion about our user interface, both reflected by our data analysis and their free-text responses. The size, position, and orientation of the UI was seen generally as appropriate, with some participants reporting a preference for slight changes, perhaps hinting toward there being a need to offer customization options. Additionally, the sense of presence elicited was higher than the norm values calculated using a database of 542 research cases provided and maintained by the questionnaire designers [43], and our analysis found that sense of presence may have some positive effect on experience familiarity. This would imply that sense of presence is likely to be a factor worth considering when designing systems to elicit a familiar experience, and also perhaps smoking cravings.

5.2. Contributions

Overall, we have developed solutions to provide researchers with a controlled environment to test their smoking cessation approaches using a CA. Where on the other hand, conducting similar tests involving everyday locations and situations people find themselves in, would not be presently feasible to do so in vivo. To that end, we have also provided a communication solution to facilitate interaction between researchers and participants while the latter are being presented with virtual environments, and thus provided a way for researchers to test their CA approaches. We have concluded that the VEs designed using participant-provided material cause a higher experience familiarity than standard ones, which we assert is a promising first step in determining whether they would also cause higher smoking cravings. Having demonstrated that our solution can reliably elicit experience familiarity to some extent, can be used as a foundation to posit whether additional effects, such as smoking cravings, may be elicited using similar means. This can be used as a starting point to examine the effects of other types of cues on the same measured variables, examine what other effects may be elicited such as emotions, investigate the impact that other methods of visual cue presentation can have, as well as explicitly testing whether smoking cravings are correlated with experience familiarity.

5.3. Limitations

The first identified limitation involves the equipment available to record material for the purpose of designing virtual environments, and the consequent generalizability of our findings. We opted to use panoramic photos and audio to design virtual environments because of the type of material we could feasibly request participants to submit. This has affected the resulting virtual environments and potentially how effective they were in eliciting desired effects, in our case, experience familiarity, smoking cravings, system usability. Ideally we would have preferred to use an omnidirectional camera to record seamless $360°$ sphere photos or even video, however due to our operational requirements we needed to develop a solution that would be feasibly accessible both for use by the amount of people in our target sample size, and for the purposes of the wider *Perfect Fit* project. Even so, we cannot extrapolate the findings from using our still photos to claim that using video would necessarily offer an improvement in the effects elicited just because this would be a more immersive experience. The goal of providing a more immersive experience would be to enhance the sense of presence elicited. And even though we did find some indication for sense of presence to have effect on the resulting experience familiarity, there may be additional variables which we would need to first investigate and account for. For instance, De Bruijn et al. [30] mentioned that in their study, an on-rails $360°$ video experience may have been novel and distracting enough to not cause the hypothesized increase in smoking cravings they were investigating. Therefore, we argue that such approaches would first need to be studied further so that we do not inadvertently add confounding variables.

Secondly, we cannot assert that our findings are necessarily generalizable in the wider population of smokers. This is because while we did observe a higher elicited experience familiarity in non-smokers when presented with personalized virtual environments, we cannot conclude that this would also be the case in smokers, or that in their case, smoking cravings themselves would necessarily be higher. As mentioned before, our one smoker case study had promising results, however, further testing would be required using a broader sample of smokers belonging to diverse relevant categories of e.g. ,gender, smoking frequency, literacy, socio-economic status, technology perception, nationality.

Lastly, regarding the communication solution we developed, we implemented a *Wizard of Oz* approach allowing the simulation of users interacting with a CA using pre-defined responses for each researcher message. While the participant UI design was based on typical messaging apps, it may not accurately represent the interface that researchers would like it to simulate. Nor would it have the capabilities that researchers may want to offer in their own CA approach, such as a more realistic chat app potentially also allowing for free-text participant responses, or an app having more functionalities than only chat.

5.4. Future Work

Reflecting on the aforementioned limitations, it would be interesting to see the improvement that more immersive methods in presenting virtual environments could have. For instance, one could take advantage of advances in the field of image or video analysis and terrain generation for the purpose of constructing navigable and interactible 3D worlds based on real locations e.g., [7, 17]. For instance, one could record photos, video, and even lidar or radar or sonar inputs from their living room, which would then be used to generate a 3D representation of that location, potentially also including an accurate 3D representation of each item. Subsequently, image recognition techniques could be employed to detect objects and apply not only the correct textures over them, if not already performed by previous steps, but potentially also identify their function/purpose, so that a set of interactions could be enabled. Overall, especially since based on our analysis we have found that sense of presence can impact effects elicited, we believe that the enhanced realism which could result from using such methods may be of great benefit to not only the field of addiction research, but to virtual reality in general.

Additionally, as mentioned in our limitations, we implemented a *Wizard of Oz* approach to simulate interaction with a CA. Alternatively, in the case where one wants to connect an external CA to our solution, an adapter could be implemented to allow the two systems to interact with one another, as opposed to only simulating that interaction. This could include for instance, the CA system exporting its utterances to a file, which our system would import, present to users, and then export their responses which would in-turn be imported by the CA system for analysis.

Furthermore, additional cue types could be presented, which were not feasible for us due to the limited scope of our project. While we were able to present visual and auditory cues, we did not have the necessary equipment to allow the presentation of olfactory or tactile cues, which could have potentially elicited a more realistic experience and perhaps increase the likelihood that conditioned responses would be activated. Such cues could be presented using scent machines and wearable equipment e.g., tactile gloves. This could enable researchers to present cues such as tobacco smell or associated ones such as coffee, as well as potentially allow for simulating the feeling of unwrapping a pack of cigarettes or even using a lighter or matches to light a cigarette.

Finally, simulating interactions with a virtual smartphone while viewing virtual environments could be a more realistic experience. We used a traditional gamepad controller because of our limited scope. However, solutions ranging from one-handed VR controllers, to using tactile gloves could vastly improve the feeling of holding and interacting with a virtual smartphone. Additionally, head movement could be enabled while keeping the virtual smartphone at a specific position depending on where the user expects to be holding it e.g., at a position relative to the actual position of the hardware interface.

5.5. Final Remarks

Virtual reality is a tool which can be used to immerse users into virtual worlds, allowing the elicitation a variety of responses. We can create controlled environments not only to test the efficacy of various approaches in different fields prior to their implementation in the real world, but to also solve real problems by facing them in a different way. We can also interact with one another using embodied communication methods, train skills, or even create immersive experiences for entertainment purposes. This technology shows promise in what it may be capable of, including but not only in the field of addiction or psychology research, which warrants further research into advancing what may be possible to achieve.

References

[1] Acker, J., and MacKillop, J. Behavioral economic analysis of cue-elicited craving for tobacco: A virtual reality study. *Nicotine and Tobacco Research 15* (2013), 1409–1416.

[2] Allen, M. Counterbalancing. *The SAGE Encyclopedia of Communication Research Methods* (2017).

[3] Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., and Wagenmakers, E.-J. An introduction to bayesian hypothesis testing for management research. *Journal of Management 41*, 2 (2014), 521–543.

[4] Antoniades, A. VR system to present environments composed of 360° images and audio, and enabling interaction with a virtual smartphone. URL: `https://github.com/PerfectFit-project/CravingCuesInVR`.

[5] Antoniades, A., and Albers, N. Experience familiarity of personalized environments and virtual interface usability in virtual reality. URL: `https://osf.io/9vkfz`.

[6] Antoniades, A., Albers, N., and Brinkman, W.-P. Everyday environments as cues to smoke: Personalized environments in virtual reality to elicit smoking cravings - data and analysis code. URL: `https://doi.org/10.4121/19433867`.

[7] Arenas, R. T., Delmas, P. J., and Strozzi, A. G. Development of a virtual environment based image generation tool for neural network training. *2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ)* (2020).

[8] Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., and Rey, B. Immersion and emotion: Their impact on the sense of presence. *Cyberpsychology and Behavior 7* (2004), 734–741.

[9] Baumann, S., Neff, C., Fetzick, S., Stangl, G., Basler, L., Vereneck, R., and Schneider, W. A virtual reality system for neurobehavioral and functional mri studies. *Cyberpsychology, Behavior, and Social Networking 6*, 3 (2003), 259–266.

[10] Baumann, S. B., and Sayette, M. A. Smoking cues in a virtual world provoke craving in cigarette smokers. *Psychology of Addictive Behaviors 20* (2006), 484–489.

[11] Bordnick, P. S., Graap, K. M., Copp, H., Brooks, J., Ferrer, M., and Logue, B. Utilizing virtual reality to standardize nicotine craving research: A pilot study. *Addictive Behaviors 29* (2004), 1889–1894.

[12] Bordnick, P. S., Graap, K. M., Copp, H. L., Brooks, J., and Ferrer, M. Virtual reality cue reactivity assessment in cigarette smokers. *Cyberpsychology and Behavior 8* (2005), 487–492.

[13] Bordnick, P. S., Yoon, J. H., Kaganoff, E., and Carter, B. Virtual reality cue reactivity assessment: A comparison of treatment- vs. nontreatment-seeking smokers. *Research on Social Work Practice 23* (2013), 419–425.

[14] Brinkman, W.-P., Haakma, R., and Bouwhuis, D. G. The theoretical foundation and validity of a component-based usability questionnaire. *Behaviour & Information Technology 28*, 2 (2009), 121–137.

[15] Brivio, E., Serino, S., Cousa, E. N., Zini, A., Riva, G., and Leo, G. D. Virtual reality and 360° panorama technology: a media comparison to study changes in sense of presence, anxiety, and positive emotions. *Virtual Reality* (2020).

[16] Cadet, L. B., and Chainay, H. Memory of virtual experiences: Role of immersion, emotion and sense of presence. *International Journal of Human Computer Studies 144* (2020).

[17] Cannavo, A., D'Alessandro, A., Maglione, D., Marullo, G., Zhang, C., and Lamberti, F. Automatic generation of affective 3d virtual environments from 2d images. *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (2020).

[18] Carlin, A. S., Hoffman, H. G., and Weghorst, S. Virtual reality and tactile augmentation in the treatment of spider phobia: a case report. *Behaviour Research and Therapy 35*, 2 (1997), 153–158.

[19] Carter, B. L., Bordnick, P., Traylor, A., Day, S. X., and Paris, M. Location and longing: The nicotine craving experience in virtual reality. *Drug and Alcohol Dependence 95* (2008), 73–80.

[20] Carter, B. L., and Tiffany, S. T. Meta-analysis of cue-reactivity in addiction research. *Addiction 94*, 3 (1999), 327–340.

[21] Center for Disease Control and Prevention. NHIS - Adult Tobacco Use - Glossary. `https://www.cdc.gov/nchs/nhis/tobacco/tobacco_glossary.htm`, Aug 2017.

[22] Chechile, R. A. *Bayesian statistics for experimental scientists*. MIT Press, 2020.

[23] Choi, J. S., Park, S., Lee, J. Y., Jung, H. Y., Lee, H. W., Jin, C. H., and Kang, D. H. The effect of repeated virtual nicotine cue exposure therapy on the psychophysiological responses: A preliminary study. *Psychiatry Investigation 8* (2011), 155–160.

[24] Cohen, J. A power primer. *Psychological Bulletin 112*, 1 (1992), 155–159.

[25] Conklin, C. A. Environments as cues to smoke: Implications for human extinction-based research and treatment. *Experimental and Clinical Psychopharmacology 14*, 1 (2006), 12–19.

[26] Conklin, C. A., Perkins, K. A., Robin, N., McClernon, F. J., and Salkeld, R. P. Bringing the real world into the laboratory: Personal smoking and nonsmoking environments. *Drug and Alcohol Dependence 111*, 1-2 (2010), 58–63.

[27] Conklin, C. A., Robin, N., Perkins, K. A., Salkeld, R. P., and McClernon, F. J. Proximal versus distal cues to smoke: The effects of environments on smokers' cue-reactivity. *Experimental and Clinical Psychopharmacology 16*, 3 (2008), 207–214.

[28] Conklin, C. A., and Tiffany, S. T. The impact of imagining personalized versus standardized urge scenarios on cigarette craving and autonomic reactivity. *Experimental and Clinical Psychopharmacology 9* (2001), 399–408.

[29] Culbertson, C. S., Shulenberger, S., Garza, R. D. L., Newton, T. F., and Brody, A. L. Virtual reality cue exposure therapy for the treatment of tobacco dependence. *Journal of cyber therapy and rehabilitation 5* (2012), 57–64.

[30] de Bruijn, G. J., de Vries, J., Bolman, C., and Wiers, R. (no) escape from reality? cigarette craving in virtual smoking environments. *Journal of Behavioral Medicine* (2020).

[31] Drobes, D. J., and Tiffany, S. T. Induction of smoking urge through imaginal and in vivo procedures: Physiological and self-report manifestations. *Journal of Abnormal Psychology 106* (1997), 15–25.

[32] Ferrer, M., García-Rodríguez, O., Maldonado, J. G., Pericot-Valverde, I., and Villa, R. S. Efficacy of virtual reality for triggering smoking craving: relation with level of presence and nicotine dependence. *Studies in health technology and informatics* (2010).

[33] Ferrer-García, M., García-Rodríguez, O., Pericot-Valverde, I., Yoon, J. H., Secades-Villa, R., and Gutiérrez-Maldonado, J. Predictors of smoking craving during virtual reality exposure. *Presence: Teleoperators and Virtual Environments 22* (2013), 423–434.

[34] Field, M., and Cox, W. M. Attentional bias in addictive behaviors: A review of its development, causes, and consequences. *Drug and Alcohol Dependence 97* (2008), 1–20.

[35] Gamito, P., Oliveira, J., Baptista, A., Morais, D., Lopes, P., Rosa, P., Santos, N., and Brito, R. Eliciting nicotine craving with virtual smoking cues. *Cyberpsychology, Behavior, and Social Networking 17* (2014), 556–561.

[36] García-Rodríguez, O., Pericot-Valverde, I., Gutiérrez-Maldonado, J., Ferrer-García, M., and Secades-Villa, R. Validation of smoking-related virtual environments for cue exposure therapy. *Addictive Behaviors 37* (2012), 703–708.

[37] Goldenhersch, E., Thrul, J., Ungaretti, J., Rosencovich, N., Waitman, C., and Ceberio, M. R. Virtual reality smartphone-based intervention for smoking cessation: Pilot randomized controlled trial on initial clinical efficacy and adherence. *Journal of Medical Internet Research 22* (2020), 1–15.

[38] Goldstein, R. Z., and Volkow, N. D. Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex. *American Journal of Psychiatry 159*, 10 (2002), 1642–1652.

[39] Heishman, S., Singleton, E., and Moolchan, E. Tobacco craving questionnaire: Reliability and validity of a new multifactorial instrument. *Nicotine and Tobacco Research 5*, 5 (2003), 645–654.

[40] Heishman, S. J., Singleton, E. G., and Pickworth, W. B. Reliability and validity of a short form of the tobacco craving questionnaire. *Nicotine and Tobacco Research 10*, 4 (2008), 643–651.

[41] Hendrix, C. M. Exploratory studies on the sense of presence in virtual environments as a function of visual and auditory display parameters.

[42] Hone-Blanchet, A., Wensing, T., and Fecteau, S. The use of virtual reality in craving assessment and cue-exposure therapy in substance use disorders. *Frontiers in Human Neuroscience 8* (2014), 1–15.

[43] IGroup. Presence questionnaire. Retrieved from `http://www.igroup.org/pq/ipq/index.php`.

[44] Kaganoff, E., Bordnick, P. S., and Carter, B. L. Feasibility of using virtual reality to assess nicotine cue reactivity during treatment. *Research on Social Work Practice 22* (2012), 159–165.

[45] Kruschke, J. *Doing Bayesian Data Analysis*, 2 ed. Elsevier Science & Technology, 2014.

[46] Kruschke, J. K. Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods and Practices in Psychological Science 1*, 2 (2018), 270–280.

[47] Kruschke, J. K., and Liddell, T. M. The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review 25*, 1 (2017), 178–206.

[48] Kuntze, M. F., Stoermer, R., Mager, R., Mueller-Spahn, F., Bullinger, A. H., and Roessler, A. Immersive virtual environments in cue exposure. *Cyberpsychology and Behavior 4* (2001), 497–501.

[49] Ledoux, T., Nguyen, A. S., Bakos-Block, C., and Bordnick, P. Using virtual reality to study food cravings. *Appetite 71* (2013), 396–402.

[50] Lee, J., Lim, Y., Graham, S. J., Kim, G., Wiederhold, B. K., Wiederhold, M. D., Kim, I. Y., and Kim, S. I. Nicotine craving and cue exposure therapy by using virtual environments. *Cyberpsychology and Behavior 7* (2004), 705–713.

[51] Lee, J. H., Ku, J., Kim, K., Kim, B., Kim, I. Y., Yang, B. H., Kim, S. H., Wiederhold, B. K., Wiederhold, M. D., Park, D. W., Lim, Y., and Kim, S. I. Experimental application of virtual reality for nicotine craving through cue exposure. *Cyberpsychology and Behavior 6* (2003), 275–280.

[52] Linden Lab. Second life, 2003.

[53] Loewenstein, G. Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes 65*, 3 (1996), 272–292.

[54] Loewenthal, K. An introduction to psychological tests and scales. *Psyccritiques 42* (01 2001).

[55] McElreath, R. *Statistical Rethinking*, 2 ed. CRC Press, 2020.

[56] Meijer, E., Gebhardt, W. A., Van Laar, C., Kawous, R., and Beijk, S. C. Socio-economic status in relation to smoking: The role of (expected and desired) social support and quitter identity. *Social Science and Medicine 162* (2016), 41–49.

[57] Moon, J., and Lee, J. H. Cue exposure treatment in a virtual environment to reduce nicotine craving: A functional mri study. *Cyberpsychology and Behavior 12* (2009), 43–45.

[58] Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist 17*, 11 (1962), 776–783.

[59] Pallavicini, F., Pepe, A., Ferrari, A., Garcea, G., Zanacchi, A., and Mantovani, F. What is the relationship among positive emotions, sense of presence, and ease of interaction in virtual reality systems? an on-site evaluation of a commercial virtual experience. *Presence: Teleoperators and Virtual Environments 27* (2020), 183–201.

[60] Paris, M. M., Carter, B. L., Traylor, A. C., Bordnick, P. S., Day, S. X., Armsworth, M. W., and Cinciripini, P. M. Cue reactivity in virtual reality: The role of context. *Addictive Behaviors 36* (2011), 696–699.

[61] Park, C. B., Choi, J. S., Park, S. M., Lee, J. Y., Jung, H. Y., Seol, J. M., Hwang, J. Y., Gwak, A. R., and Kwon, J. S. Comparison of the effectiveness of virtual cue exposure therapy and cognitive behavioral therapy for nicotine dependence. *Cyberpsychology, Behavior, and Social Networking 17*, 4 (2014), 262–267.

[62] Perfect Fit: Targeting key risk factors for cardiovascular disease in at-risk individuals using a personalized and adaptive approach. Retrieved from `https://commit2data.nl/projecten/perfect-fit`.

[63] Pericot-Valverde, I., García-Rodríguez, O., Gutierrez-Maldonado, J., Ferrer-García, M., and Secades-Villa, R. Evolution of smoking urge during exposure through virtual reality. *Studies in health technology and informatics 167* (2011), 74–79.

[64] Pericot-Valverde, I., Germeroth, L. J., and Tiffany, S. T. The use of virtual reality in the production of cue-specific craving for cigarettes: A meta-analysis. *Nicotine and Tobacco Research 18* (2016), 538–546.

[65] Regenbrecht, H., and Schubert, T. Real and illusory interactions enhance presence in virtual environments. *Presence: Teleoperators and Virtual Environments 11*, 4 (2002), 425–434.

[66] Audacity Team. Audacity, 2021.

[67] Microsoft Research. Image composite editor, 2015.

[68] The GIMP Development Team. GIMP, 2021.

[69] Ritchie, H., and Roser, M. Smoking. *Our World in Data* (2013). https://ourworldindata.org/smoking.

[70] Schubert, T., Friedmann, F., and Regenbrecht, H. Embodied presence in virtual environments. *Visual Representations and Interpretations* (1999), 269–278.

[71] Schubert, T., Friedmann, F., and Regenbrecht, H. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments 10*, 3 (2001), 266–281.

[72] Schubert, T. W. The sense of presence in virtual environments:. *Zeitschrift fur Medienpsychologie 15*, 2 (2003), 69–71.

[73] Segawa, T., Baudry, T., Bourla, A., Blanc, J. V., Peretti, C. S., Mouchabac, S., and Ferreri, F. Virtual reality (vr) in assessment and treatment of addictive disorders: A systematic review. *Frontiers in Neuroscience 13* (2020).

[74] Shiffman, S., Dunbar, M., Kirchner, T., Li, X., Tindle, H., Anderson, S., and Scholl, S. Smoker reactivity to cues: Effects on craving and on smoking behavior. *Journal of Abnormal Psychology 122*, 1 (2013), 264–280.

[75] Slater, M. A note on presence terminology. *Presence Connect 3* (01 2003).

[76] Slater, M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences 364* (2009), 3549–3557.

[77] Slater, M. Immersion and the illusion of presence in virtual reality. *British Journal of Psychology 109* (2018), 431–433.

[78] Slater, M., Lotto, B., Arnold, M. M., and Sanchez-Vives, M. V. How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicologia 40* (2009), 193–210.

[79] Slater, M., Steed, A., McCarthy, J., and Maringelli, F. The influence of body movement on subjective presence in virtual environments. *Human Factors 40* (1998), 469–477.

[80] Slater, M., and Usoh, M. Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments 2*, 3 (1993), 221–233.

[81] Slater, M., and Wilbur, S. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments 6* (1997), 603–616.

[82] Staiger, P. K., and White, J. M. Cue reactivity in alcohol abusers: stimulus specificity and extinction of the responses. *Addictive Behaviors 16*, 5 (1991), 211–221.

[83] Steuer, J. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication 42* (1992), 73–93.

[84] Tieri, G., Morone, G., Paolucci, S., and Iosa, M. Virtual reality in cognitive and motor rehabilitation: facts, fiction and fallacies. *Expert Review of Medical Devices 15* (2018), 107–117.

[85] Tiffany, S. T. A cognitive model of drug urges and drug-use behavior: Role of automatic and nonautomatic processes. *Psychological Review 97* (1990), 147–168.

[86] Traylor, A. C., Bordnick, P. S., and Carter, B. L. Assessing craving in young adult smokers using virtual reality. *American Journal on Addictions 17* (2008), 436–440.

[87] Traylor, A. C., Bordnick, P. S., and Carter, B. L. Using virtual reality to assess young adult smokers' attention to cues. *Cyberpsychology, Behavior, and Social Networking 12*, 4 (2009), 373–378.

[88] Traylor, A. C., Parrish, D. E., Copp, H. L., and Bordnick, P. S. Using virtual reality to investigate complex and contextual cue reactivity in nicotine dependent problem drinkers. *Addictive Behaviors 36* (2011), 1068–1075.

[89] U.S. Department of Health and Human Services. *Smoking Cessation: A Report of the Surgeon General - Executive Summary*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, GA, 2020.

[90] Watanabe, S. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res. 11* (dec 2010), 3571–3594.

[91] WEST, R., and SCHNEIDERS, N. Craving for cigarettes. *Addiction 82*, 4 (1987), 407–415.

[92] Witmer, B. G., and Singer, M. J. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments 7*, 3 (1998), 225–240.

[93] World Health Organization. *WHO global report on trends in prevalence of tobacco use 2000-2025 third edition WHO global report on trends in prevalence of tobacco use 2000-2025, third edition*. 2019.

Appendices

A

Scenarios Summary

Here we present a summary of scenarios which we discussed with psychologists contributing to the *Perfect Fit* project, to refine our requirements and base our system design, in Table A.1. A longer discussion regarding the scenarios and the resulting requirements can be found in Section 2.2.2, and Table 2.2.

Table A.1: Scenarios Summary

Scenarios	Options	Insights
Participant Perspective: Required Materials	Participants submit answers to a questionnaire to inform how to personalize VEs.	Second option more favorable because it can include elements not consciously considered as affecting an addiction model. Process needs to be accessible to LSES people, so asking them to take photos may be easier than having them answer questionnaires.
	Participants submit photographs and audio from the locations in which they smoke most frequently.	
Participant Perspective: Viewing Virtual Environments	View 3D VEs containing personalized elements determined by researchers and stock sound.	3D environments are more flexible and offer navigability and interactability, but impact on elicited smoking cravings is uncertain. Immersion was viewed as important. Photos may be more realistic are more likely to contain elements implicitly associated with smoking behavior, while also easier to obtain and incorporate in creating VEs. They may also offer higher ecological validity.
	View VEs containing panoramic photos and audio submitted by themselves.	
	View VEs containing a combination of elements from the two options above.	
Researcher Perspective: Designing Virtual Environments	Design VEs with 3D assets, based on literature. Allow for personalization of certain elements and/or choice in environment type.	Not easy to elicit the information that would be required to personalize VEs. Since preferable case in the previous scenario was to have participants submit photos, environments entirely composed of them was seen as more favorable.
	Design personalized VEs from scratch with 3D assets. Base everything on what researchers deem relevant to each addiction model.	
	Utilize participant-submitted photos and display them e.g., as part of VEs with 3D assets, or having VEs entirely composed of photos.	
Participant Perspective: Location and Hardware Interaction	Participants take part in the experiment at home, view VEs on their computer monitor, and use a keyboard/mouse for interactions.	Lab experiment using an HMD was viewed most favorably because this option offered the most experimental control and would make the process more accessible to LSES people.
	Participants go to the lab, view VEs using an HMD and use a gamepad controller for interactions.	
	Participants go to the lab, view VEs on a computer monitor / projector, and use a keyboard/mouse for interactions.	
Participant/Researcher Perspective: Communication Method	Free-form text input, as in a typical messaging application.	Envisioned solution for the *Perfect Fit* application is to have a conversational agent offering pre-determined responses to its utterances. Therefore, the corresponding option was viewed more favorably since it would increase the real-world transference of results.
	Researcher messages accompanied by corresponding responses that can be chosen.	
Participant Perspective: Communication Solution Appearance	UI overlaid on a 3D object e.g., a smartphone.	Initial 3D object option was viewed unfavorably because it broke immersion too much, while a typical UI would be more purposeful. 3D object size / position / rotation was refined to increase level of immersion and purposefulness, and it was then viewed as the preferred option.
	UI appears in a typical manner i.e. flat on the screen.	

B

Virtual Environments

B.1. Image Vertical Bounds

The first two Figures in this section illustrate the vertical image borders of our virtual environments due to how the audiovisual material was sourced i.e., recording a $360°$ video in portrait orientation using a typical smartphone camera. The third Figure illustrates how we aimed for this to be less immersion-breaking, by stretching the top and bottom of the image.

(a) Vertically-Repeating Texture

(b) Using Blocking Planes

(c) Texture Stretching

Figure 6: Configurations for Dealing with Image Borders.

B.2. Standard Virtual Environments

Here we have two of the standard environments we used, by processing material sourced by us based on what past work has reported are locations more likely to either cause or not cause smoking cravings. The images appear similarly to the files ready to imported by our system i.e., the horizontal boundaries form a $360°$ rotation and the vertical boundaries are stretched to avoid users seeing outside the image.

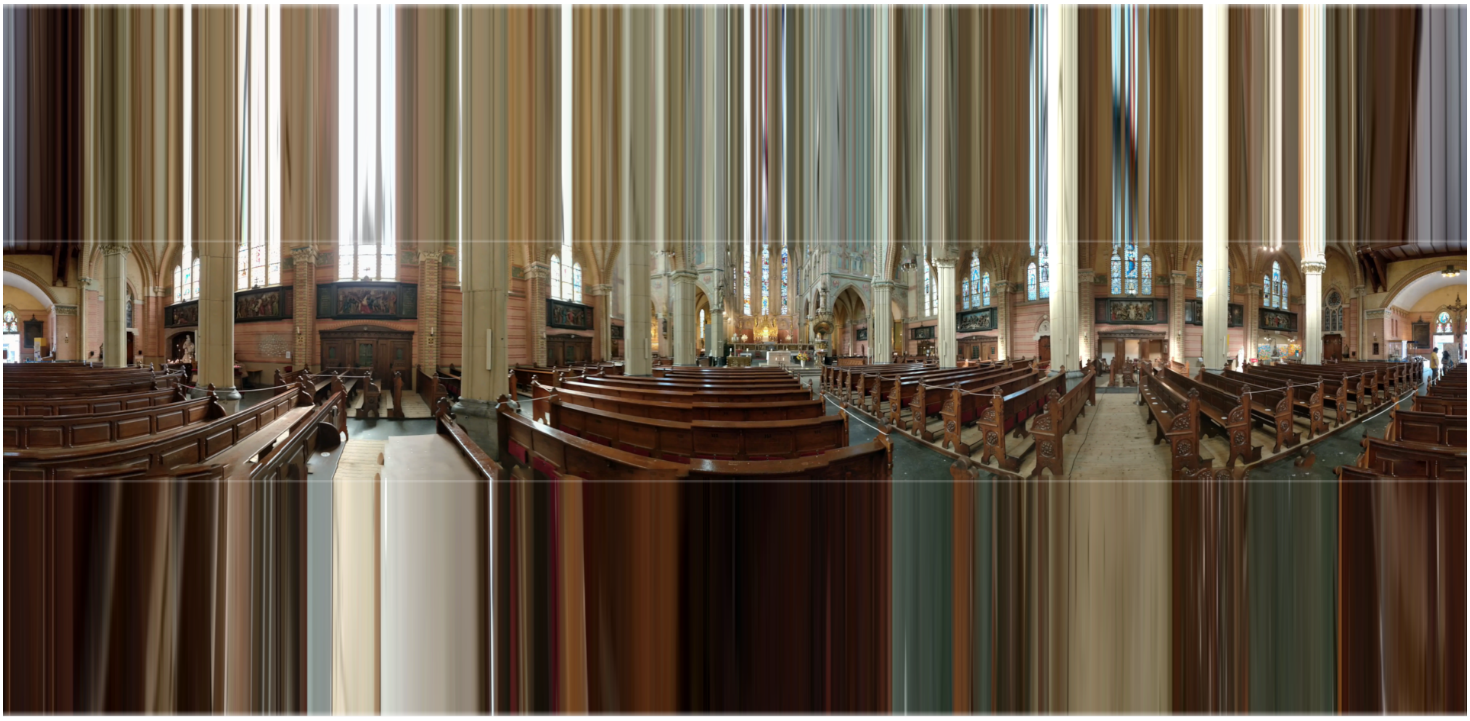


Figure 7: Non-Personalized Non-Smoking Environment (Church).

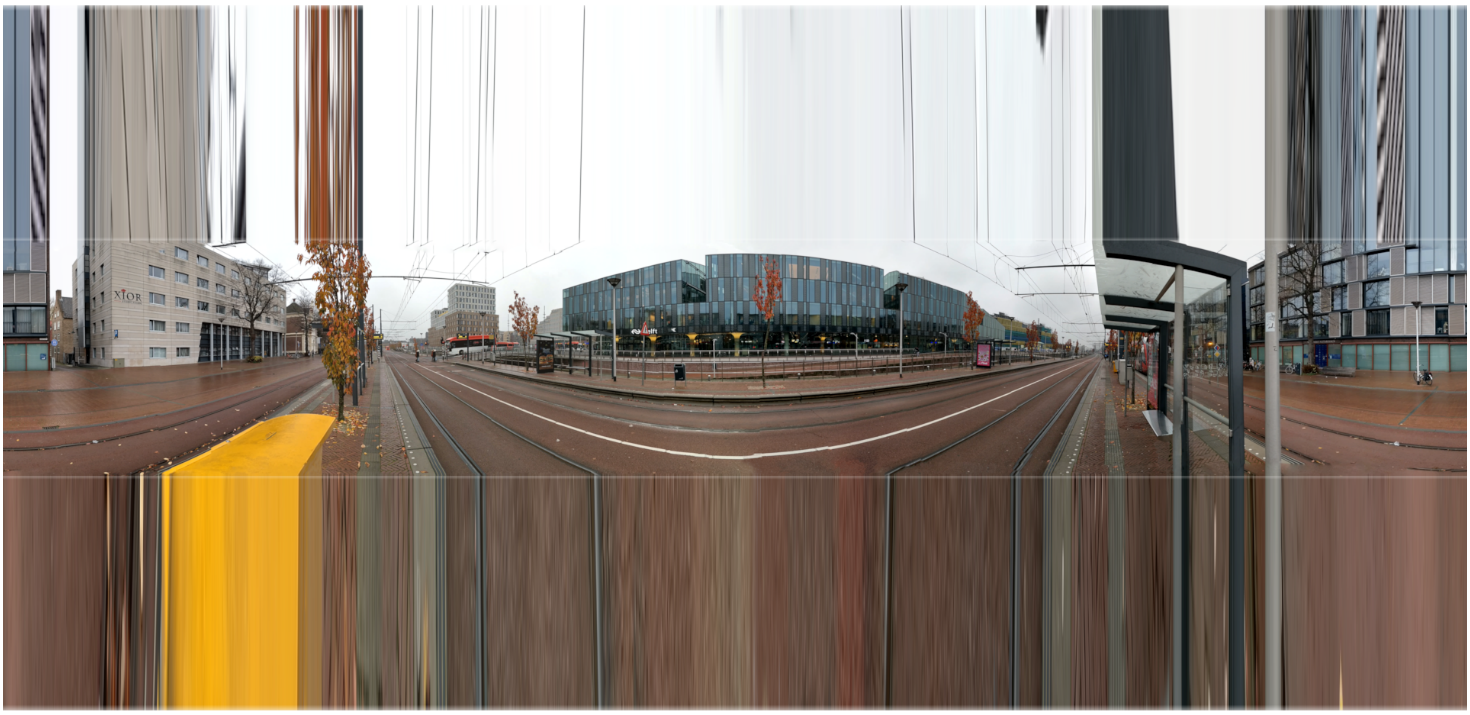


Figure 8: Non-Personalized Smoking Environment (Bus/Tram Stop).

C Experimental Setup Additional Information

Here we provide additional information regarding our experimental setup. We start by presenting further details about how we created counterbalanced sequences, and then briefly describe the automation implemented to run our experiment without requiring direct researcher input.

C.1. Counterbalancing Details

For smokers, we presented 12 environments; 3 personalized smoking-cue, 3 personalized non-smoking-cue, 3 non-personalized smoking-cue, and 3 non-personalized non-smoking-cue. We created all combinations of presenting 12 environments of said types, and followed guidelines similar to Conklin et al. [26], as well other principles of counterbalancing to achieve a modified Latin square design [2]. This resulted in 8 distinct combinations which are illustrated by Figure 4. The specific requirements used to reach our resulting 8 combinations can be seen in Table C.1. This resulted in every sequence containing two instances of adjacent personalized-personalized environment pairs, and two instances of adjacent non-personalized-non-personalized environments pairs.

In the case of non-smokers, we adapted the counterbalancing to only keep the personalization factor, leading to four unique sequences (since only one factor is used we have one unique non-smoking sequence per two unique smoking sequences). This is illustrated by Figure 5. Thus, if a non-smoker participant is to be assigned a sequence, we pick a smoker randomization sequence and then find the applicable one in the non-smoker sequence. For instance, if a non-smoker is assigned to smoker sequence 6, they are actually assigned to non-smoker sequence 2 (because they are identical if not considering the environment cue-type factor).

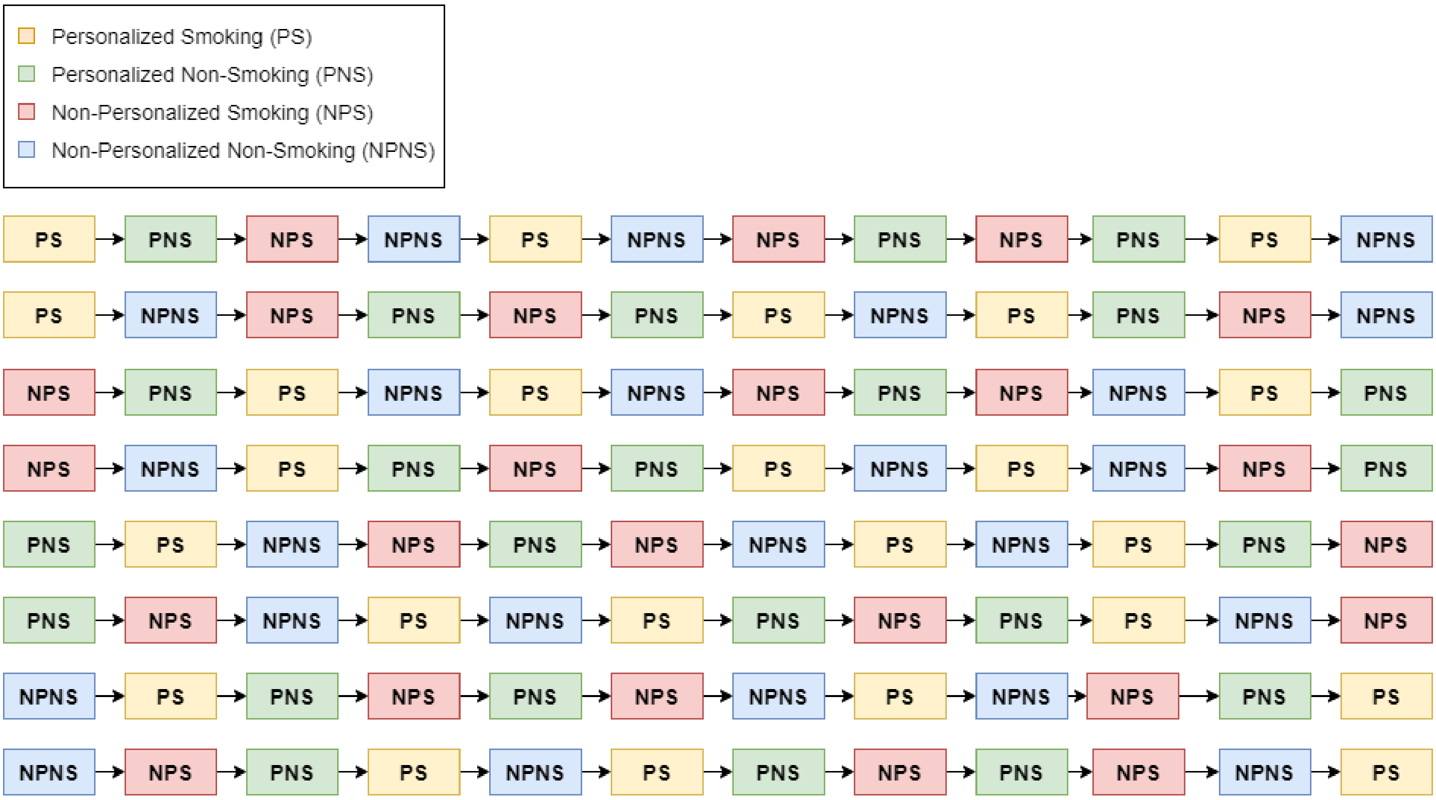


Figure 4: 8 Counterbalancing Sequences - Smoking

1. No adjacent smoking-cue pairs.
2. 3 smoking and 3 non-smoking-cue environments to occur in the first and last six trials
3. 3 personalized and 3 non-personalized environments to occur in the first and last six trials.
4. No more than two consecutive personalized or non-personalized environments.
5. Each environment type to occur an equal number of times in each position across randomizations.
6. No consecutive-pair pattern repeating more than two times.
7. No more than two pairs of adjacent same-factor personalization pairs.
8. Equal least number of adjacent same-factor personalization pairs. (In an attempt to limit and balance the amount of adjacent same-factor personalization, i.e. two consecutive personalized or non-personalized environment pairs in each sequence)

Table C.1: Smoking Counterbalancing Sequence Criteria

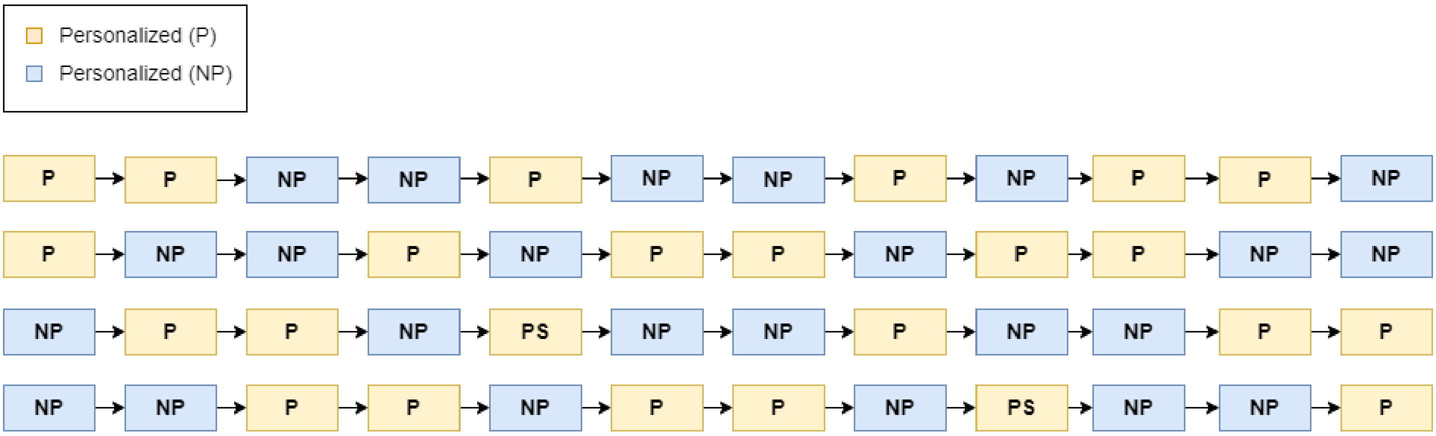


Figure 5: 4 Counterbalancing Sequences - Non-Smoking

C.2. Experiment Run Automation

To run our experiment without requiring researcher input we implemented a finite state machine (FSM) solution. This was to automate the process and minimize participant interaction with researchers, which could have impacted environment-elicited effects. Our FSM is illustrated by Figure 6, where each box represents a state and its associated actions, arrows represent state transitions and their triggers, and the filled-in circle represents program termination.

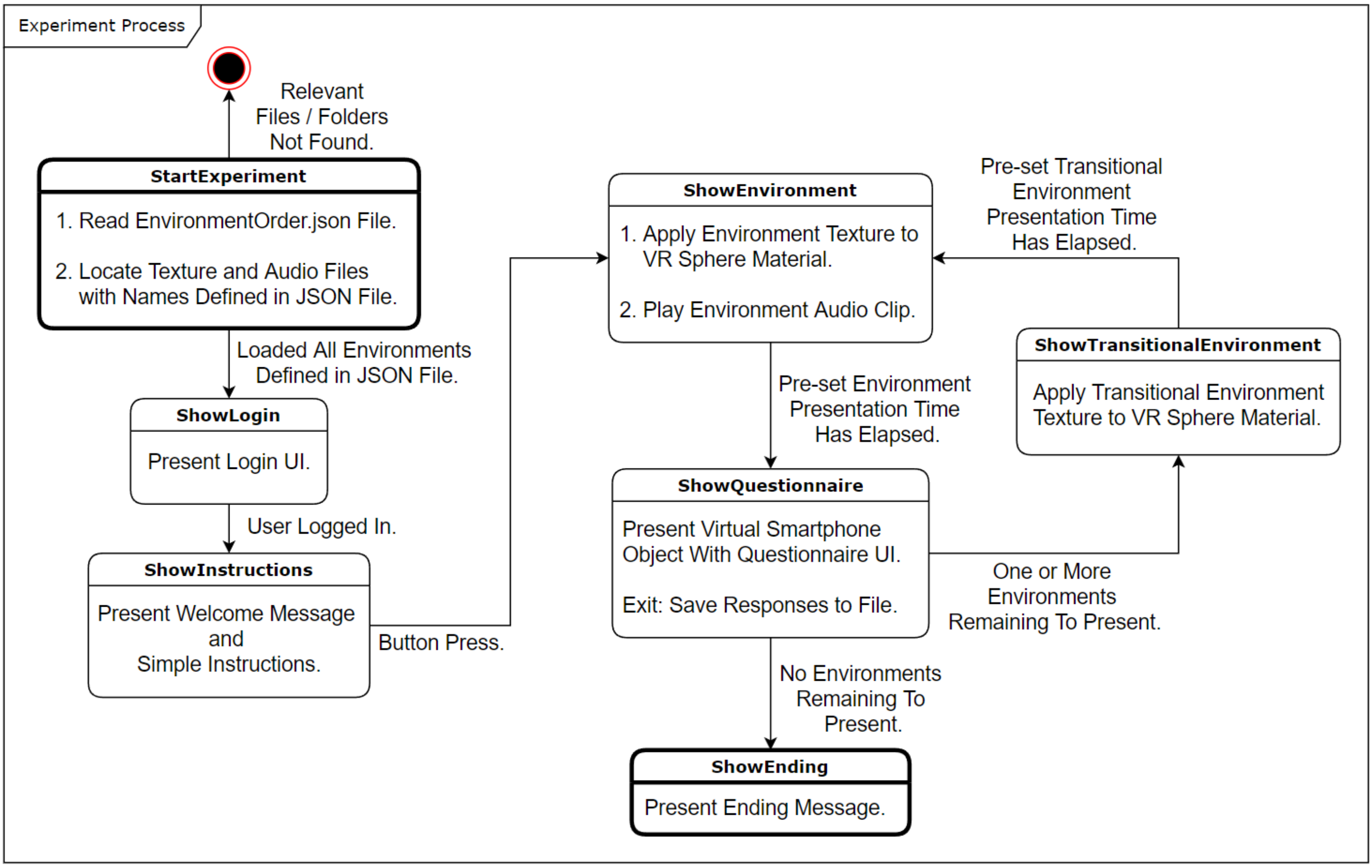


Figure 6: Experiment State Diagram

D

Additional Diagrams

Here we provide additional diagrams regarding our system implementation and experiment design.

This diagram illustrates a high-level design of our experiment. The process is succinctly defined as presenting different virtual environments, taking measurements on elicited effects, taking a short break, and then presenting the next environment until there are no more left to present.

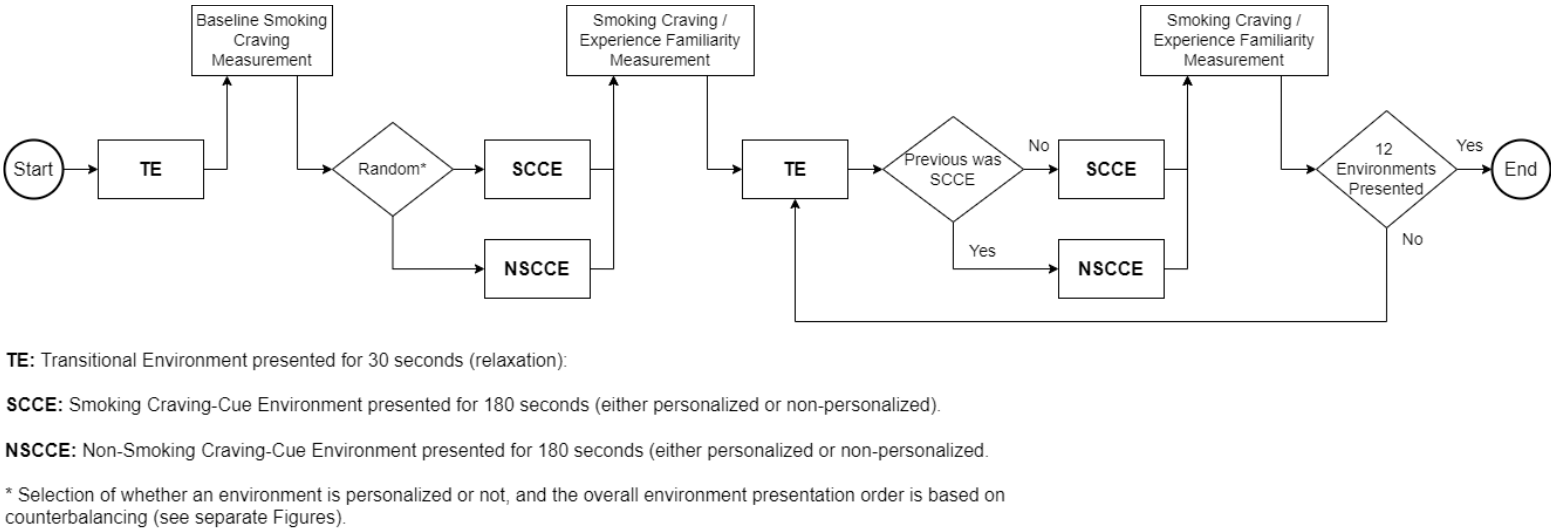


Figure 5: High-Level Experiment Process Diagram

This diagram illustrates a high level overview of our networked system use described in Chapter 3. Succinctly, this process is defined as the user logging in with the appropriate role (researcher or participant), with the system in-turn instantiating the corresponding UI, presenting a loading screen, loading environment files, and presenting the appropriate virtual environment.

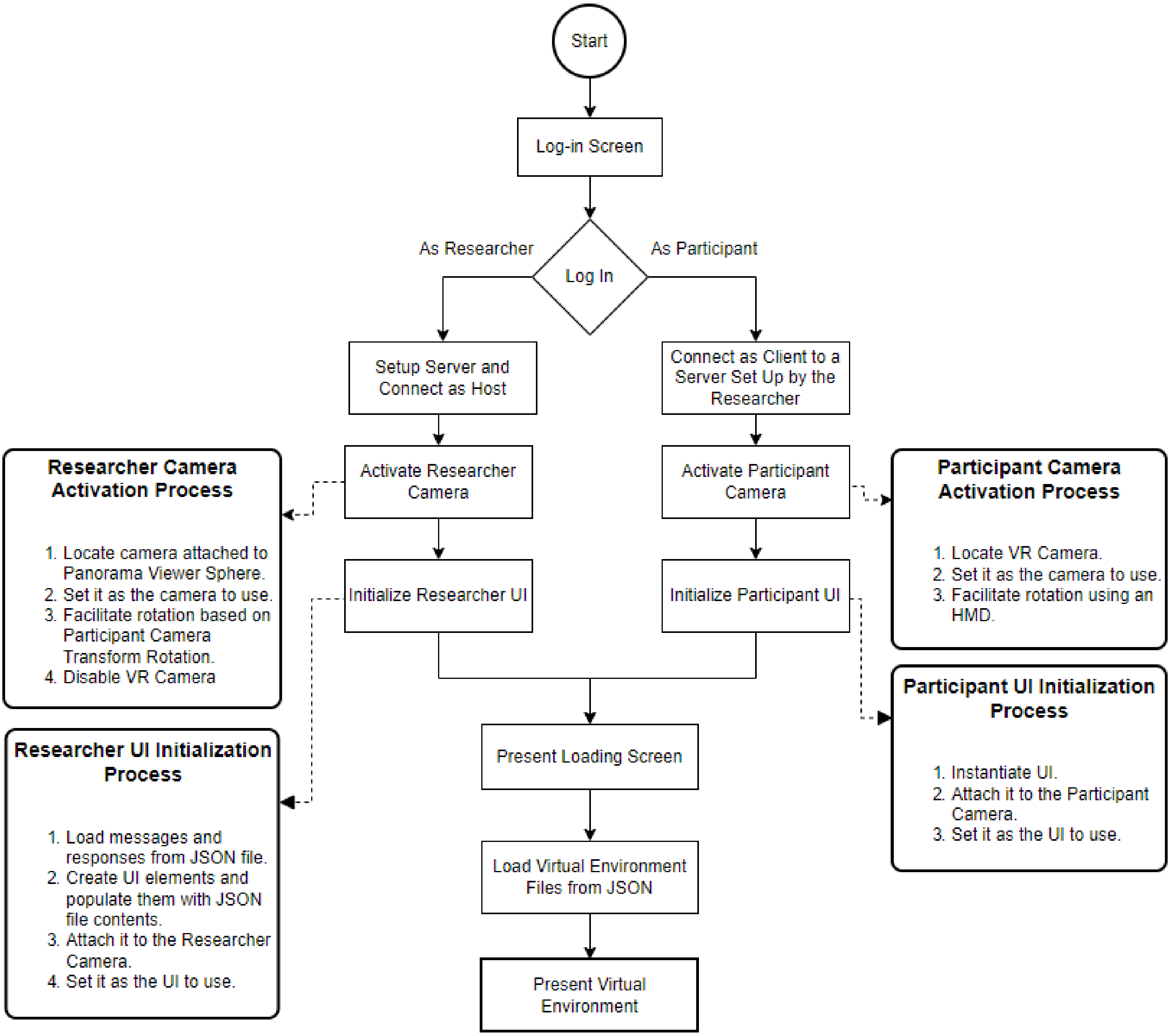


Figure 6: General System Startup and Use Processes.

E

Questionnaires

1. How easy or difficult was it to read the text displayed on the virtual smartphone screen?	-5 (Very Difficult), 0 (Neutral), 5 (Very Easy)
2. The size of the virtual smartphone was:	-5 (Too Big), 0 (Just Right), 5 (Too Small)
3. The location of the virtual smartphone in your field of view was:	-5 (Very Inappropriate), 0 (Neutral), 5 (Very Appropriate)
4. The orientation (rotation / tilt) of the virtual smartphone was:	-5 (Very Inappropriate), 0 (Neutral), 5 (Very Appropriate)
5. The notification sound and the virtual smartphone appearing was:	-5 (Very Disturbing), 0 (Neutral), 5 (Very Intuitive)
6. I would prefer the virtual smartphone to: **A. Move** with me when I look around in the virtual environments and remain stable at the **same position in my field of view**. **B. Not move** with me when I look around in the virtual environment and remain stable at the **same location in the environment**.	-5 (Completely prefer **A**), 0 (Neutral), 5 (Completely prefer **B**)
7. I found the viewing duration of virtual environments to be:	-5 (Too Short), 0 (Just Right), 5 (Too Long)

Table E.1: Additional Usability Questionnaire

F

Audiovisual Material Collection Instructions

Here we provide the instructions given to instruct participants in how to collect and submit the audio-visual material required to create virtual environments. These specific instructions were for smoker participants, however, the only differences can be found in the ”Location Selection” Section, where smoking-specific instructions are given. Non-smoker participants were simply instructed to ”Consider the top 8 locations in which you spend time at least once a week.”

Instructions for submitting material:

You are requested to submit video and audio from 8 locations. 4 locations in which you most frequently smoke, and 4 locations in which you most frequently do not smoke during a typical week. In the case where you have less than 4 locations for each, please contact the researcher responsible at a.antoniades@student.tudelft.nl to perhaps reach an alternative solution.

Collecting material will take a minimum of 30 minutes.

Below you will find instructions as to:

1. Which locations to choose.
2. How to record video:
    a. If you most frequently stand in that location.
    b. If you most frequently sit in that location.
3. How to record audio.
4. How to prepare and submit the material.

1. Location Selection:

Try and think of places in which you most frequently smoke during a typical week, as well as places in which you do not smoke.

Examples are places such as your office/workspace, your living room, the bus/tram stop / train station that you usually go to, places where you have lunch / dinner / desert / coffee, the gym / supermarket you usually go to, the street in front of your house, etc.

For places where you smoke most frequently, consider the top 4 places in which you:

1. Spend time at least once a week.
2. Smoked at least 7 out of 10 times when there.
3. Found the difficulty of not smoking when there to be at least a 5 on a 0-10 scale.
4. May feel at least somewhat negatively if you were not able to smoke there.

For places where you do **not** frequently smoke, consider the top 4 places where:

1. You spend time at least once a week.
2. You smoked only 3 or less times out of 10 times when in that place.
3. The strength of thinking about smoking when there was 5 or less on a 0-10 scale.
4. The difficulty of refraining from smoking when there was 5 or less on a 0-10 scale.

For each of those locations, please record and submit video and audio using the instructions below, at a time of day when you usually spend time there. It might be easier and less time-consuming if you record material when you are usually at those places, rather than specifically going there to record.

If possible, please try to make sure that videos do not contain material directly related to smoking, e.g. cigarettes, cigarette packs, ashtrays, smoke.

It is important that you record the material from your usual point of view in that location, and at a time of day when you usually spend time there.

2. Video:

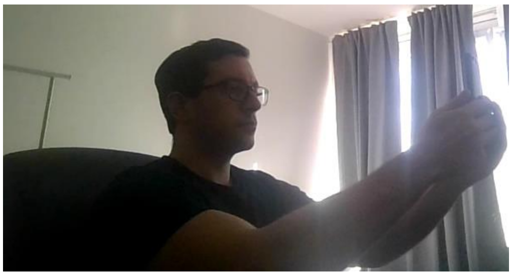
1. Set your smartphone's camera to take the highest quality video it can (at least 1080p). Please do a quick internet search for how you can achieve this on your specific smartphone model.
2. Stand or sit in your chosen location at the place where you most often stand/sit there.
    a. In places where you both stand and sit, only take a sitting video if you most frequently sit when there.
3. Hold your smartphone vertically at an eye-level, i.e., for the camera to "see" the environment from the same height as you can see, at about the half-way distance between your face and the furthest you can stretch your arms out.

2. a. If standing or sitting on a rotating chair:

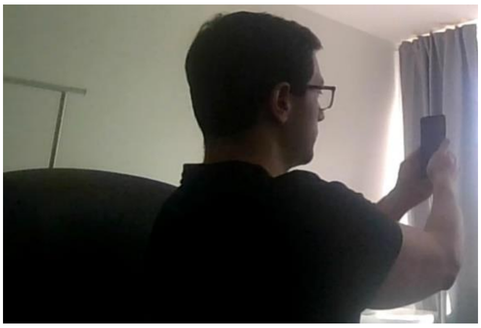
a. Hold the smartphone at that height/distance and begin recording video.
b. Holding the smartphone stable, rotate your entire body to the right while continuing to record video.
c. Try to make sure that you are holding the smartphone at a stable height/distance and that you are not tilting it on any axis. Rotate your body/chair, not the smartphone.
d. Continue recording video while rotating to the right, until you complete at least a 360-degree rotation.
e. Try to end a little further to the right than your starting orientation.
f. Aim to rotate at an even pace, and for the video to last at least 30 seconds.

2. b. If sitting somewhere you cannot naturally rotate 360° (e.g. car):

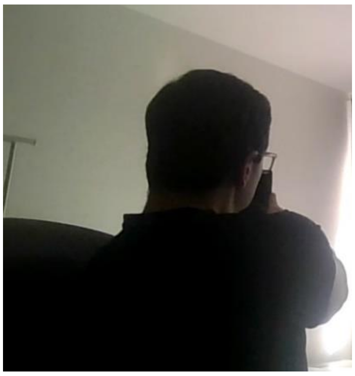
a. Hold the smartphone stable at your eye-level in front of you but do not start recording video yet.



b. Rotate your waist as far as you can on the left without rotating the chair or arms.

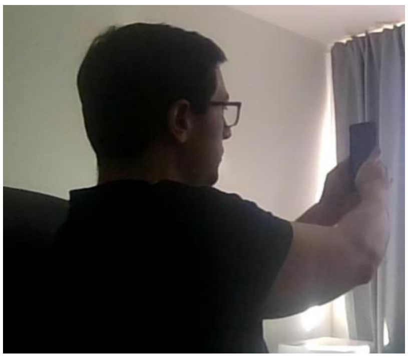


c. Then rotate your arms a little further while keeping the same rotation angle.

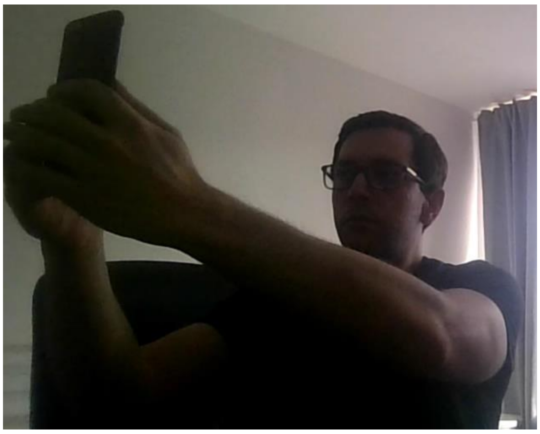


d. Begin recording video.

e. Holding the smartphone at that height/distance, rotate your arms to the right until the smartphone reaches the middle of your face.



f. Holding the smartphone stable at that height/distance, rotate your waist to the right.
g. Continue rotating your waist to the right until you reach the furthest point you can rotate your waist, without rotating the chair or arms.



h. At that point, rotate only your arms a little further to the right while keeping the same rotation angle.

i. End video recording.
j. Try to make sure that you are holding the smartphone at a stable height and distance from your face, and that you are not tilting it on any axis.
k. Aim to rotate at an even pace from start (d.) until finish (i.), and for the video to last at least 20 seconds.

3. Audio:

1. Using your chosen audio recording app on your smartphone, please record audio for **3 minutes and 20 seconds** at the same location in which you have taken video using the method above.

4. Submitting Material:

Please name the files accordingly, using the following naming convention:

“least_smoking_video_x”, “least_smoking_audio_x”, “most_smoking_video_x”, “most_smoking_audio_x”, where “x” refers to the number of the file with video and audio numbers matching.

For instance, a set of files for the same least smoking location would be: “least_smoking_video_3.mp4”, “least_smoking_audio_3.wav”.

Please **do not** convert the files to a different format before submitting them.

Compress the material to an archive using your preferred application, e.g. WinZip, WinRAR, 7zip, etc. If using an iPhone, you can achieve this by using the compress function.

1. Visit https://wetransfer.com/
2. You do not need to set up an account, but you may need to verify your email address.
3. Upload the archive you have created.
4. In the “Email to” field, please enter a.antoniades@student.tudelft.nl
5. In “Your email” field, please enter the email address you entered in the informed consent form and time slot booking.
6. Enter a suitable title and message.
7. Press the “Transfer” button to initiate transferring the material.

You can email a.antoniades@student.tudelft.nl with any questions regarding the material recording process, or renaming / compressing / sending the files.