COMPARING VARIOUS MEASURES TO PREVENT COVID-19 SPREAD IN OFFICES THROUGH SIMULATION MODELLING

COMPARING VARIOUS MEASURES TO PREVENT COVID-19 SPREAD IN OFFICES THROUGH SIMULATION MODELLING

MASTER THESIS SUBMITTED TO DELFT UNIVERSITY OF TECHNOLOGY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN **COMPLEX SYSTEMS ENGINEERING AND MANAGEMENT**

FACULTY OF TECHNOLOGY, POLICY AND MANAGEMENT

BY

EMMA DEIJKERS

STUDENT NUMBER: 4456475

TO BE DEFENDED IN PUBLIC ON SEPTEMBER 20 2022

Graduation committee

First Supervisor	: Dr. Y. Huang, Systems Engineering and Simulation
Second Supervisor	: Dr. ir. C. van Daalen, Policy Analysis
External Supervisor	: Dr. ir. D.C. Duives, Faculty of Civil Engineering, TU Delft

Executive summary

COVID-19 and the containment strategies that followed it have had a great impact worldwide. One of the effects was that many gathering places, such as offices, had to be closed. Consequently, the closing of offices has put a mental strain on many people, in addition to causing the negative economic effects. Therefore, it is important that a safe way to keep offices open amid the pandemic is found, which also ensures people do not get infected and experience negative effects on their health. Non-Pharmaceutical Interventions (NPIs), such as wearing masks and social distancing, are an important way to do this. However, to pick the right NPIs for one's office, knowledge is needed on what the effects of the different NPIs within the given context are. However, as of now, there are not many models that can provide a detailed estimation of COVID-19 spread and quantify the impact of NPIs in office spaces. Based on the state-of-the-art of models for this purpose, various knowledge gaps are found:

- Most models in current research only allow for the implementation of one or two NPIs, if any at all.
- In addition, the results of these models still have relatively high uncertainty for two reasons.
    - The limited extent of movement modelling they incorporate.
    - The models are not explicitly designed for measuring infection risk in real-life office scenarios or even in office spaces in general.

Therefore, a model is developed in this thesis research to contribute to these knowledge gaps. Using this model, more information can consequently be found on the effects of the various NPIs in offices. Hence, this thesis research answers the following main research question:

What is the effect of Non-Pharmaceutical Interventions on COVID-19 spread in office spaces?

The existing model of the PedestrianDynamics - Virus Spread (PeDViS) of Duives et al. (2022) is chosen as the base model to be extended into a model that allows the comparison of the effects of NPIs in an office space. In this thesis research, a simulation modelling approach is used, consisting of four phases: (1) model design, (2) model implementation, (3) experimentation and (4) model assessment of the impact of the different NPIs. Data on office layouts and behavior was collected via observation, supplemented with desk research, to serve as input for the first two phases. The resulting output of the conceptualization and model implementation phases is an adaption of the PeDViS model. This adaption focuses on simulating the spread of COVID-19 in an office environment. The simulated office consists of personal offices, a coffee area, restrooms, meeting rooms and hallways. As the existing version of the PeDViS model is tailored to a restaurant situation, the activity scheduling of the office employees is not based on existing code but created in this thesis in order to be able to focus on office activities. The following office activities are included in the model: having meetings in meeting rooms, having small meetings in offices, working in one's office, getting coffee and using the restroom. The NPIs of which the effects are explored in this office scenario are as follows:

- All employees wear a mask while walking
- Increasing the rate of cleaning
- Increasing the ventilation rate from standard rate to the rate advised by government officials
- Allowing only 50 % of the employees to come into the office
- Allowing a maximum number of people in the meeting rooms, equal to 50 % of the capacity of the meeting rooms

Next, in the experimentation phase, a sensitivity analysis is performed to learn more about the impact of some of the variables related to the meetings held in the office. The meeting variables whose values are

varied are: (1) the number of meeting room meetings per day, (2) the number of office meetings per day and (3) the group size of meeting room meetings. For this sensitivity analysis, set-ups differing the values for these three variables were constructed. In the analysis, the average number of infections per workday was gathered for each set-up. The three meeting-related variables were shown to be able to partly explain the number of infections. However, other factors are expected to also play a big role in the infection numbers.

Of the three meeting variables, the number of infections is found to be most sensitive to the number of meeting room meetings held on a day. However, based on the set-ups, the group size of the meeting room meeting is found to have the biggest overall impact, due to it having a broader value range (up to 14 people in a meeting versus up to 8 meetings on a day). The frequency of office meetings was found to also impact the infection numbers, but have a lower impact than the other two variables. This could potentially be explained by the smaller group size of 3 people in the office meetings. This smaller size could have two effects. First, it leads to fewer direct infections during the meeting itself due to the fewer other people a person comes into contact with during the meeting. Second, it leads to less build-up of virus in the office meeting areas, as there is a smaller probability the infected person had a meeting in the same area earlier. Besides the effect of the smaller group size, office meetings also have a shorter duration than meetings in meeting rooms, which also makes them less risky.

In the final phase of model assessment, the effectiveness of the NPIs in containing the spread of COVID-19 in the office is tested. To test the impact of the NPIs, all possible combinations of NPIs are defined as different policies. The policies thus differ in the NPIs that are implemented. The method of policy analysis is used to compare the impact of the NPIs and NPI policies on containing the COVID-19 spread. Based on the number of infections found per set-up during the sensitivity analysis, two set-ups are chosen to use in the model to test the impact of the NPIs. The first one is the set-up with the number of infections that was closest to the overall average number of infections found based on all set-ups. The second set-ups is the one with the highest overall number of infections.

Based on the policy analysis, various conclusions are drawn. However, there is some uncertainty that should be kept in mind for these conclusions. This uncertainty is due to the high variability in the number of infections within an implemented NPI policy. The cause of the variability is expected to be the difference in which employee is the infected one per model run. Therefore, there is some uncertainty in stating the effectiveness of the policies and no guarantees can be given about the effectiveness of any set of NPIs. This should be kept in mind when interpreting the conclusions on the the effectiveness of the NPIs.

Per NPI, information was found on their individual effectiveness, before diving into the policies combining the NPIs. First of all, having all employees wear a mask while walking seems very effective in reducing the number of infections, even more so than found in other literature. A reason for the increased effectiveness of masks could be the long time people spend in the office, leading to building-up of the amount of virus in the building. Wearing a mask while walking would then reduce the amount of virus particles employees inhales while walking through or standing in different areas in which there is a high virus build-up.

Second, it is found that the NPI of restricting the size of the meetings held in meeting rooms is often effective in reducing infections. This is in line with the finding of the sensitivity analysis that the size of these meetings plays a large role in the number of infections. The cause for this decrease in infections is expected to be a combination of a lower density of people during meetings and less people being in close contact with other people for the prolonged time it takes to have a meeting. Based on this, it seems like restricting the size of meetings, especially preventing very large meetings, would be a good way for offices to lower virus transmission.

Next, increasing the ventilation was shown to often bring down the number of infections. This effect is limited, however, when other effective NPIs are present or when the initial situation of the office does not have a particularly high infection risk. In both cases, this means there are less virus particles present in the air. Therefore the increased ventilation can bring about a lower absolute decrease in virus aerosols present in the office and thus in the amount of virus people are exposed to.

The fourth NPI, increasing the frequency in which all surfaces in the office are cleaned, was shown to have limited effect in these offices. This effectiveness is in line with findings in literature, as virus transmission via virus particles on surfaces and hands does not seem to have much impact on the number of infections. By extent, cleaning the surfaces more often cannot be expected to have much impact on the virus transmission.

Finally, it was found that limiting the number of people allowed to come to the offices can have inconsistent effectiveness. This inconsistent effect is expected to be due to the increased number of meetings the employees present in the offices have when this NPI is present. This increase in the number of meetings per employee is based on the assumption that when people can only come to the office on certain days, they plan more of their meetings to be on the days they are at the office. However, this higher number of meetings in turn causes people to come into contact with more other people during the day, increasing the infection risk. This also explains the interaction effect found, that when more other NPIs are present, the NPI of limiting the number of people in the office is more effective. In this case, even though there is a higher number of meetings, there is a lower risk of people getting infected during these meetings. This effect is then combined with the fact that there are still less people present in the office than can get infected, leading to a lower number of infections.

In other research, limiting the number of people in a building did consistently reduce the number of infections. In these cases, this was due to the lowering of the density of people, which was not the case in the office situation of this thesis. Therefore, it is advised that when this NPI is implemented, employees get encouraged to do their meetings online and when at home, in addition to setting a limit to the size and frequency of meeting at the office. Also due to the interaction effect of the NPI with other NPIs, it is advised that this NPI is implemented in combination with other NPIs, to reduce the virus spread during meetings.

The effectiveness of combining the NPIs into policies was reviewed next. In addition, it was reviewed which combination of NPI allowed for an R-value below 1, which is a guideline the Dutch government often follows when judging the situation regarding the pandemic. The lowest number of infections were found when all NPIs were present. When infections need to be kept down at all costs with unlimited resources, it is therefore advised to implemented each NPI. If the goal is to keep the R-value in the office below 1, it was found at least one of the following NPIs needs to be present:

– all employees wear masks while walking

or

– the maximum capacity of the meeting room is set to 50 %, which has to be combined with the NPI of increasing the ventilation to the government recommended levels in contexts that give reason for a higher initial infection risk

To keep the R-value below 1 even in the worst-case scenario, the mask-wearing NPI always needs to be present in any combination of NPIs making up a policy.

Various limitations of this research were identified and recommendations for future work were given for part of them. The most important recommendations for future work are:

- Extending the model to include shared work spaces and the activity of having lunch.
- Testing the impact of more aspects relating to activity scheduling, such as the duration of activities or the order in which they are performed.

- Gather data on people’s behavior when the number of days they can come to the office is limited, to learn more about the inconsistent effect of this NPI. Additional simulations to test the effect of this NPI while keeping the number of meetings per person the same is also recommended.
- Test the impact of the NPIs in a context in which the initial infection risk is low.

Based on the knowledge gathered on the effectiveness of different NPI policies and starting points for future work, this thesis provides several insights for policy-makers regarding COVID-19 prevention policies. These are insights into the usefulness of the different (combinations of) NPIs, advice for the use and implementation of the NPIs and recommendations for government funding research. Besides these insights, this thesis has made various contributions, both scientific and societal, contributing to the knowledge gaps.

First of all, a virus transmission model focusing on an office context has been developed, of which not many exist in current literature, especially not with the extent of movement modelling of this model. Valuable knowledge has been found for both more extensive movement modelling of an office context and the influence of various aspects relating to meetings on the number of infections. Another scientific contribution the model provides is that the effects of more than one NPI are included in this model, with many existing models only containing one or two NPIs, if any at all. In addition, the thesis also provides a potential strategy that can be used to compare the effectiveness of these NPIs, consisting of a policy analysis that includes various statistical tests and considerations regarding the acceptable numerical values required for an NPI to provide significant or acceptable prevention of virus transmission risk. Finally, the filling in of these research gaps allows the final scientific contribution, which regards the knowledge gathered on the effectiveness of the NPIs for certain office contexts. The thesis also provides societal contributions. First of all, the knowledge this thesis provides can contribute to reducing the number of people affected by the COVID-19 virus, through either sickness or even death, after getting infected in the office. In addition, the model created and knowledge gathered from it provides offices with a way to choose the right NPIs to safely stay open, which has both economic and psychological advantages.

Preface

This thesis report, *Comparing various measures to prevent COVID-19 spread in offices through simulation modelling*, concludes my time as a master student Complex System Engineering and Management (CoSEM) at the Delft University of Technology. This research covers a topic everyone has some experience with: COVID-19. When I started writing my thesis proposal, in February 2022, we were still in lockdown and my subject certainly felt very relevant. Currently, in August of 2022, COVID-19 is luckily not as present in our daily lives anymore. However, winter is coming, and I am hoping my work could contribute to the handling of the virus in the future. During my research process, I had great supervision, advice and support from many people, who I would like to thank.

First of all, I want to thank Yilin, who during our biweekly meetings could always provide me with an answer to my many questions, in addition to coming up with points I had not even thought of yet and very concrete advice on how to further improve my work. Second, I would like to thank Martijn, who was able to provide a solution to each error or problem I had using the Nomad model by either responding to an email of mine right away or being willing to meet with me to explain the model further. I also want to thank Els, for always providing me with very detailed feedback, new insights and important practical information when needed and also for helping me get started by letting me do my observations in front of her office. Also, I want to thank Dorine for actively thinking of important aspects I should think of or include in my work each time we met and for encouraging me to better explain and think about the choices I made in my thesis. I want to thank Xinyi, for graciously being willing to meet with me, sharing her old thesis work and providing her knowledge as a fellow former thesis student working with the PedestrianDynamics - Virus Spread (PeDViS) model. Also, I want to thank Büsra for always providing me with detailed answers to my many questions about the model and the virus spread part.

Next, I would like to thank my friends and family. First, I want to thank Saskia for reading my thesis and providing useful (and uplifting) feedback to me, in addition through living to the high and low times of my thesis with me. Of course, my parents also deserve some praise for raising me to be disciplined and hardworking and have perseverance, while also fully supporting each undertaking I ever wanted to take, among which are my studies and my thesis. I also want to thank my friends for studying with me and making the many hours working on the thesis more bearable, especially Esther and Joost, who spend many, many hours/afternoons studying with me. Finally, I want to thank Ilja for being my biggest supporter while writing this thesis, who helped me both by providing well-thought-out and to-the-point advice and also reading my whole thesis, but also supporting me by giving me confidence in my own abilities and cheering me on and helping me when needed.

Emma Deijkers

Table of contents

Abbreviations

ODD:	Overview, Design Concepts and Details
NPI:	Non-Pharmaceutical Interventions
PeDViS:	PedestrianDynamics - Virus Spread
CoSEM:	Complex System Engineering and Management
TPM faculty:	Technology, Policy and Management faculty, of the TU Delft
R-value :	Reproduction value
SSO:	SamenSlimOpen, a research project

1 Introduction

Since the discovery of the COVID-19 virus in December 2019, it has had a great impact worldwide (WHO, 2021). First of all, there have been important health effects, with already 21332 deceased since the start of the pandemic in the Netherlands alone as of September 2022 (Dutch central government, 2022a) and many recovered patients experiencing long term health problems (Hofman, 2022). In an effort to slow the spread of COVID-19, the Dutch government has at various instances decided to close restaurants, offices, public events and other gathering places in an effort to slow down the spread of COVID-19 (Dutch National Institute for Health and Environment, 2020). These measures, however, also have some significant adverse effects. The economy even dropped by 8.4 % during the first lockdown (Central Bureau for Statistics, 2022). In addition, the lockdown measures were shown to take a severe mental toll on people for various reasons (NOS & I&O Research, 2020). One of the reasons for this mental strain is that many people had to work from home during significant parts of the pandemic, due to the Dutch Government's COVID-19 guidelines (Dutch central government, 2022b). Emphasizing the gravity of this topic, the probability of epidemics occurring is only becoming higher over time, as the world is more connected than ever with large scale movements of people and goods (Watkins, 2020). This highly connected structure of society facilitates rapid disease spread between distanced areas, causing seemingly spontaneous outbreaks (Watts & Strogatz, 1998). For all of these reasons, it is very important to be able to choose the right strategies to predict and contain the spread of viruses such as COVID-19. Choosing the right strategies entails knowing what the effects of the different measures within these strategies are.

1.1 Importance of a safe operation of offices using Non-Pharmaceutical Interventions

During the largest part of 2022, offices in the Netherlands were allowed to be open again (Ministry of General Affairs, 2022a). However, caution is still warranted. It has already been shown that despite a high vaccination rate, the COVID-19 virus can rampantly spread (Dutch central government, 2022b; Dutch central government, 2022c). In addition, the COVID-19 virus constantly mutates (National Institute for Public Health and the Environment, 2022), and the next variant could possibly cause graver health effects, bringing about the need for a new lockdown. Therefore, to keep society open in a safe way and thus decrease the adverse economic and mental health effects experienced, adequate use of Non-Pharmaceutical Interventions (NPIs) to limit virus spread is very important. Examples of NPIs are mask-wearing, social distancing, limiting the number of people allowed in a certain space and hygiene and cleaning measures (CDC, 2020; CDC, 2021; Fouda et al., 2021; National Institute for Public Health and the Environment, 2021; WHO, 2021). NPIs can also be used for the safe functioning of office spaces during a pandemic. Before an adequate approach can be chosen for keeping offices safely open using NPIs, an improvement is needed in office management's ability to choose the right prevention measures to operate their offices safely with low infection risk. However, as of now, there are not many models that can provide a detailed estimation of COVID-19 spread and quantify the impact of NPIs in office spaces. Therefore, the objective of this thesis research is as follows:

To help offices safely stay open with sufficient certainty of low infection risk by providing a model that allows for the comparison of the effects of various NPIs to contain COVID-19 in real-life office settings.

1.2 Outline

The remainder of the thesis is outlined as follows. In chapter 2, the key concepts of this research are defined, the search process to find the literature is explained, a literature review is performed, the knowledge gap is stated and the main research question is formulated. In chapter 3, the existing model used to answer the main research question is described and the choice for this model is explained. In chapter 4, first, the research approach of this thesis is created based on literature and this approach is then elaborated upon. The sub-questions and the methods used to answer the sub-questions are also stated in this chapter. In chapter 5, information of the data collection process and the conceptualization of the model can be found. In chapter 6, the model implementation is covered. Experimentation using the model is then described in chapter 7 and model assessment is performed in chapter 8. Chapter 9 contains the discussion, which includes the literature implications of the results, the limitations found and the recommendations based on them, insights for policy-makers, the contributions of this thesis and the thesis' suitability with the CoSEM master program. Finally, chapter 10 gives the conclusions of this thesis.

2 Literature review

In this chapter, a literature review is performed to identify the research gap and to gather knowledge on the start-of-the-art of models allowing the comparison of NPIs. In order to do this, first of all, the key concepts are defined. Second, the search process for finding the literature is described in section 2.2. Next, the literature is reviewed in section 2.3. Based on this review, the knowledge gap is identified and the main research question is formulated in section 2.4.

2.1 Core concepts

To find the literature needed for the review, a search plan was set up. First, the most important concepts were defined. As mentioned in section 1.2, the objective of this proposed thesis research is *to allow for the comparison of the effect of different NPIs and combinations of them, to help offices to stay open with sufficient certainty of low infection risk.* To accurately measure the effects of NPIs, different types of models need to be combined. First of all, a COVID-19 transmission model, which incorporates the different modes of transmission, is key. However, only using a COVID-19 transmission model is insufficient, as contextual values and the behavior and movement of people also play an important role in the spread of the virus (Cartenì, Di Francesco & Martino, 2021; Raza, Ali & Hussain, 2021*).* The movement of pedestrians and their interactions with the environment can be represented by individual movement modelling (Hoogendoorn & Bovy, 2004). It can thus be concluded that to reach the objective, there is a need for a model that allows the comparison of the effects of ***NPIs,*** on the ***covid spread*** in ***indoor spaces*** (offices), using a model that combines ***individual movement modelling*** and virus ***transmission models.*** These bolded terms form the key concepts used in this research, and are thus the parts on which the literature review is focused.

2.2 Search process

To find literature on the state-of-the-art of models predicting virus spread, a search plan was set up. The key concepts defined in section 2.1 were used as search terms in various configurations, which can be seen in Table 1. All articles were found through the Google Scholar Database. A total of 16 articles, relating to the estimation of COVID-19 spread in indoor spaces, which also consider behavioral and contextual aspects, were collected, read and analyzed. The following search terms were used to find the literature, finally settling on the one in the last row as giving the best results.

Table 1: The search process

Search term	Rationale
Covid AND spread AND "individual movement modelling " AND "indoor spaces" AND "transmission models" AND NPI	To find information on research that covers all of the key concepts.
Covid AND spread AND "*pedestrian dynamics"* AND "indoor spaces" AND Model AND NPI	The term pedestrian dynamics was found to give better results on the movement of individuals, and the term transmission did not add much more information, with the term spread already used.
covid AND spread AND pedestrian AND "indoor spaces" AND NPI AND model	Using only the term pedestrian proved to be enough to show *pedestrian movement modelling of some form.*
pedestrian AND covid AND indoor	It was discovered additional sources could be found using even fewer search terms, as this allowed for more results.

In addition, both forward and backwards snowballing was used, using the articles already found through the database. A time constraint was set to filter the articles, as COVID-19 is a relatively new virus, and articles can only include the COVID-19 virus if they have been published after 2018. Due to this constraint, all resulting articles were published very recently and therefore show the state-of-the-art research in this field. As there was much literature available on COVID-19, there was no need to extend the search area to other types of viruses.

2.3 Discussion of literature

To learn more about the state-of-the-art, a literature review is performed, focusing on the core concepts as discussed in Section 2.1. However, first some more explanation is given on the concept of movement modelling, because it was discovered during the review that movement modelling entails many different aspects. To bring some structure into this concept, the categorization of Hoogendoorn and Bovy (2004) is used. This categorization organizes the movement of pedestrians into three choice levels: strategic (activity, activity location and departure time choice), tactical (mode and route choice) and operational choices (how pedestrians react to changes in the environment).

An overview of the literature review can be seen in Table 2. The review is structured around the key concepts, which are represented as columns in Table 2, with the movement modelling thus defined by three levels: strategic, tactical and operational choice modelling. The columns of Table 2 from left to right are: the authors and the year the research was published, the presence of aspects related to / the model used to represent strategic and tactical choices modelling, operational choice modelling, the type of indoor space it simulated and finally whether the model also allowed the measurement of the effect of one or more NPIs, and if so which NPIs.

2.3.1 Limited research available within the level of detail needed

A large part of the literature on COVID-19 spread and containment focuses on high scale levels, such as on national or local level (Dzau et al., 2020; Zhao et al., 2020; Tomar & Gupta, 2020) or subnational level (Tuite et al., 2020; Prem et al., 2020). Even less of the literature focuses on indoor spread, despite the probability of infection being higher indoors than outdoors (EenVandaag, 2020). Analyzing the state-of-the-art further, it was concluded much of the research that focused on indoor COVID-19 spread did not contain any pedestrian dynamics. Examples of this are Dai & Zhao (2020) and Huang et al. (2020), in which the pedestrians do not move and no specific type of indoor space is given. To conclude, the state-of-the-art of indoor COVID-19 spread models containing pedestrian dynamics is fairly limited as is. Based on this more limited amount of research available, a review is performed.

Table 2: Overview of literature

Author & Year	Strategic and tactical choice modelling	Operational choice modelling	Type of space	NPIs present, of which effect can be measured
Alvarez Castro & Ford (2021)	Personal characteristics influencing behavior, activity scheduling and layout of campus used	Speed pedestrians assigned per area, take shortest path to location of activity and move randomly within that area till next task	Different campus areas: faculty, supermarket, leisure and other buildings	Masks, early lockdown, self- isolation

Cuevas (2020)	Activity scheduling, personal characterization	Maintain behaviors that are characterized by simple rules	Finite 2-dimensional space	Hygiene protocols and restricted mobility
Cui et al. (2021)	Mean dwell time individuals	Social force model	Room of given size without obstacles, with entry and exit.	None
Delval et al. (2021)	Map given	Walking speed initially universal, but depends on environment, evacuation behavior towards given staircase.	School building	Ventilation strategies
D'Orazio et al. (2021)	Very simplified layout without obstacles (like walls), behavior based on area pedestrian is at and the activity schedule (classes and breaks)	Focused on crowd density instead of on movement	University buildings	Types of masks
Duives et al. (2022)	Demand generator, activity schedular, person characterization (walking speed and compliance), layout partly defined by user input	Nomad	Restaurant	Among others: masks, payment at table, coats at table, rate of air change, number of table cleaning
Fang et al. (2020)	Activity scheduling, layout of the ship used	Movement rules of a discrete evacuation model of Fang et al., (2010)	Cruise ship [1]	Self-protection percentage
Harweg, Bachmann & Weichert (2020)	Layout of supermarket used, individuals get attributed a set of destinations specified on the map.	Force-model	Supermarket	Social distancing, allowed number of people
Islam et al. (2021)	Activity scheduling, demand generation, personal characterization, classroom layout used	Social force model	Classroom-type indoor spaces (e.g., classroom, auditorium, food court, and meeting room)	Alternative policies (namely entrance and exit policy, seating policy, and room layout)
Li & Yin. (2021)	Demand generation, individual having	Cellular Automata models	College canteen	Various combinations of not eating in canteen,

[1] As many of the activities defined in the model are indoors, this study is for purposes of simplicity considered as an indoor transmission model

	preferences, layout of canteen used, activity scheduling, rule-based choices for destinations			increasing ventilation rate, controlling student inflow rate, improving self-protection rate and quarantine
Romero, Stone & Ford (2020)	One-way and two-way pedestrian traffic within a one-dimensional corridor for the case of two student	Walking speed, distance to other pedestrians	Academic building	Physical distancing
Ronchi & Lovreglio (2020)	Different crowd movement scenarios can be defined, assumptions of crowd behavior can be set (circulation paths, group behavior, building occupant interactions, etc.), different sophistication levels of crowd movement and behavior possible in relation to crowd model in use	Movement based on crows model selected, microscopic model (including walking direction and distance to others)	Confined spaces	None
Salmenjoki et al (2021)	Layout used, activity scheduling.	Move to destinations while only considering avoiding obstacles	Supermarket and bar	Opening bars vs supermarkets, number of people allowed and ventilation
Sijjadi et al. (2021)	None	Social force model	Rooms of given size	Social distancing
Xiao et al. (2021)	Universal relaxation time, roughly defined layout	Social force model with universal desired speed and personal space	Public indoor space with defined size and no obstacles	Reduce non-essential trips
Xu & Chraibi (2020)	Individual shopping time, defined layout of supermarket, behavioral rule: pedestrians want the counter with the fewest customers	A collision-free velocity model, composed of a direction sub-model and a speed sub-model	Existing supermarket	Maximum number of customers, obligated shopping cart, social distance kept

2.3.2. Limited options for comparing the effect of NPIs

First of all, it can be seen that in much of the literature in Table 2 the effects of only one or two NPIs, if any at all, are measured. In addition, the models presented do not focus on allowing a comparison of what NPIs or combinations of NPIs would be most effective for containing COVID-19 spread. A model allowing the user to explore the usefulness of different NPIs and combinations of them would be a useful tool for office managers, as they often have multiple measures available to contain the COVID-19 spread. Examples of such measures are the obligation for people to wear a face mask, keeping 1.5 meter distance from one another or more frequent cleaning of surfaces in the offices.

2.3.3 Limited extent of movement modelling

In addition to the need to compare different NPIs, the results of such comparison need to be based on a model that incorporates sufficient movement modeling of individuals, as this is an important aspect of virus spread. Therefore incorporating more detailed movement modeling would allow for more certainty in the estimations of virus spread when implementing the NPIs. As can be deduced from the column *Strategic and tactical choice modelling* in Table 2, the extent of strategic and tactical choice modelling differs greatly among models. Only a few models (Duives et al., 2022; Islam et al., 2021; Li & Yin, 2021; Alvarez Castro & Ford, 2021) provided a great level of detail in terms of activity scheduling, personal characterization and demand, while also setting a life-like context. The other studies are all lacking in one or more of these aspects. Salmenjoki et al. (2021) and Fang et al. (2020) both only have the detailed activity scheduling and a layout of the situation, while Cuevas (2020) also includes personal characterization aspects but does not have a very specific context set. Some models make use of simple behavioral rules to route pedestrians, in combination with having a defined layout (Xu & Chraibi, 2020; Harweg, Bachmann & Weichert, 2020). Different approaches to determine pedestrian behavior, such as crowd movement scenarios (Ronchi & Lovreglio, 2020) or basing it on the area in which the pedestrians find themselves (D'Orazio et al., 2021), can also be seen. Other studies barely use any information relating to movement modeling, such as Delval et al. (2020), Romero, Stone & Ford (2020) and Sijjadi et al. (2021). Or studies just include one personal characteristic and roughly set a context (Cui et al., 2021; Xiao et al., 2021). It can thus be concluded that there is a limited collection of models that use an extensive amount of movement modelling to model the spread of COVID-19 in a confined space, implying a certain level of uncertainty in the results.

2.3.4. Gap regarding models focusing on office spaces

Finally, it is concluded that there are barely any studies specifically focusing on real-life office/meeting spaces. What comes closest would be the meeting room scenario of Islam et al (2021). Next, the most suitable models to be used for office space would be the ones not made for a specific type of indoor space, such as Sijjadi et al. (2021) or Ronchi & Lovreglio (2020). While Cui et al. (2021), Cuevas (2020) and Xiao et al. (2021) also provide general indoor spaces, there are no obstacles present within these spaces (such as walls), making them less suitable. The model of Duives et al. (2022) is data-driven, thus allowing flexibility in the scenario to be simulated, but does not yet work for indoor spaces other than restaurants. Other types of indoor spaces that are often modelled are campus/academic spaces (Delval et al., 2021; D'Orazio et al., 2021; Castro & Ford, 2021; Li & Yin, 2021; Romero, Stone & Ford, 2020) and supermarkets (Salmenjoki et al., 2021; Harweg et al., 2020; Xu & Chraibi., 2020). Additional indoor spaces that have been modelled are bars (Salmenjoki et al., 2021) and a cruise ship (Fang et al., 2020). The limited suitability of the models for real-life office spaces also implies uncertainty in the results.

2.4. Research gap and main question

For offices to apply the right set of NPIs to keep their facilities open with sufficient certainty of low infection risk, it needs to be possible for office managers to compare the implementation of various (combinations of) NPIs in their office. However, based on the literature review, it can be concluded that, first of all, most of the current research in this area only allows for the implementation of one or two NPIs, if any at all. In addition, the results of these models still have relatively high uncertainty for two reasons. First, due to the limited extent of movement modelling they incorporate. Second, due to the models not being explicitly designed for measuring infection risk in real-life office scenarios or even in office spaces in general.

The current state of the art is thus not allowing the comparison of many NPIs and does not provide a high certainty of low infection risk. In contrast, there are often many NPIs available that can be implemented in offices, and to make an informed decision for implementing the right set of NPIs, comparison between all of them needs to be possible. In addition, a high certainty of sufficient protection against COVID-19 infection is desired, as too high an infection risk at offices could lead to them being closed again. Therefore, a model that allows the comparison of the effect of multiple NPIs in office spaces with sufficient certainty still needs to be developed. Consequently, the main research question is as follows:

What is the effect of Non-Pharmaceutical Interventions on COVID-19 spread in office spaces?

3 Choice for the Pedestrian Dynamics – Virus Spread model

In this chapter, more information is given on the existing model that simulates the spread of COVID-19, to build upon in this research and the reasoning behind this choice. In section 3.1, the use of a base model and the choice for the PedestrianDynamics – Virus Spread (PeDViS) model as base model is explained. In 3.2, a detailed description of the PeDViS model is presented. Finally, in section 3.3, the existing versions of the PeDViS model are elaborated upon.

3.1 Base model

Based on the literature in chapter 2, it is concluded that *a model that allows the comparison of the effect of multiple NPIs in office spaces with sufficient certainty* still needs to be developed. However, in this literature review it could also be seen that there are many models that, although not perfect for the objective of this thesis, have very useful aspects and elements. Using an existing model would also allow for more time to focus on other aspects in more detail, such as the office movement modelling. Therefore, it was decided that it would be wiser to use one of these models as a base model that would be adapted, than to start from scratch. The choice of model is based on the following characteristics of the model that were compared in the review:

- the extent of movement modelling implemented
- the type of indoor space for which it was designed
- whether it allows for the comparison of the effect of NPIs, and if so, how many

Ultimately, the PedestrianDynamics – Virus Spread (PeDViS) model of Duives et al. (2022) is chosen as most suitable to be used as a base model to further extend towards comparing multiple NPIs in an office setting for several reasons.

First of all, the PeDViS model allows for detailed modelling of pedestrian movement, containing all three levels of the earlier defined pedestrian movement modelling: strategical, tactical and operational choices. In that respect, the PeDViS model was one of the few models to provide a great level of detail in the activity scheduling part of movement modelling, and in personal characterization and demand, while also providing a life-like context.

In addition, in accordance with the objective of this thesis, a model is needed that allows the users to compare the effects of various NPIs. The PeDViS model was one of the models with the highest number of NPIs incorporated. The model also allows for easy comparison of the effect of these NPIs, as the model can be run for various combinations of NPIs present, for which outputs regarding the virus transmission are then given. By running various scenarios, which differ in terms of NPIs used, the user can see which combination of NPIs is most effective in containing COVID-19 spread. The PeDViS model already has defined some NPIs that could also be useful in an office setting, such as wearing masks or adjusting the frequency of cleaning surfaces. In addition, it is also possible to implement new NPIs into the model, either through the input data or by adjusting some of the underlying code.

3.2 Background on the PeDViS model

The PeDViS model has already briefly been described in Chapter 2, but some more extensive information is given in this section. The PeDViS model is developed by Duives et al. (2022) to simulate the spread of COVID-19, which is currently tailored to a restaurant scenario. The model couples a data-driven activity assignment model (Duives et al., 2022), the individual movement dynamics of Nomad (Hoogendoorn & Bovy, 2004; Campanella, 2016), the virus spread model of QVEmod (Duives et al., 2022) and a risk identification model based on the dose-response model (Duives et al., 2022; Nicas & Sun, 2006). Figure 1 shows the coupling of these different elements. The activity scheduling part serves as input for the operational dynamics part, which in the case of PeDViS is the Nomad model. The Nomad model then simulates the individual routing and the movement dynamics. The QVEmod model, an epidemiological model for Quantifying Viruses in Environments, then takes the output of Nomad, a set of trajectories, as input, combined with information about the environment and individuals, and simulates the virus spread and transmission. This simulation consists of both simulating the virus particles an infected individual emits at different locations based on their trajectory and the amount of virus the other pedestrians are exposed to, based on the amount of virus present in their trajectories. The risk identification model finally calculates each individual's probability of infection based on the pedestrians' exposure to COVID-19 through three exposure routes: aerosols, droplets and fomites. A more detailed description of the different parts is given in the following sections.

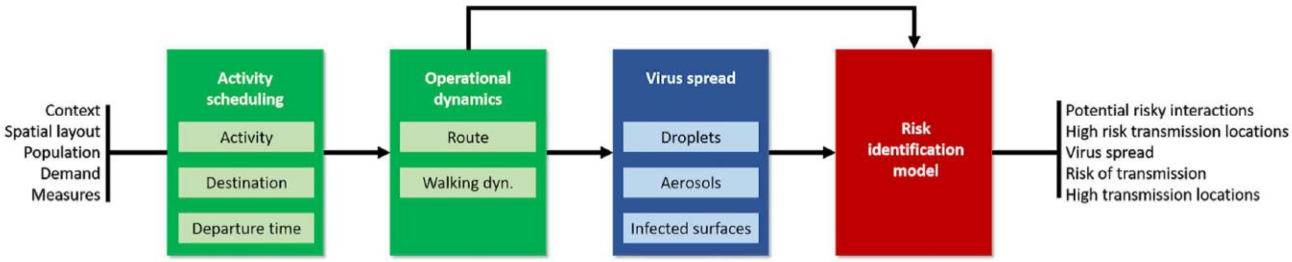

Figure 1: Elements of the PeDViS model (Duives et al., 2022)

3.2.1 Activity scheduling and operational dynamics

The two elements in green in Figure 1 represent the modeling of the behavior of the pedestrians in PeDViS. To better explain these parts of the model, the categorization of movement modelling from Hoogendoorn & Bovy (2004), which was introduced in the literature review of Chapter 2, is described here in more detail. According to Hoogendoorn & Bovy, three levels of pedestrian movement choices can be distinguished, namely strategic, tactical and operational choices. The strategic level entails choices relating to the activity of an individual, the activity location and their departure time. Tactical level entails choices regarding the activity schedule, functional area (the area where a certain activity is performed) and the mode and routing of the individuals towards these functional areas. The operational level is the "actual walking behavior" of the individuals (Hoogendoorn & Bovy, 2004) and explains pedestrians' movement choices in the presence of objects and other pedestrians.

The activity scheduling block in Figure 1 represents the strategical level of the pedestrian behavior and serves as input for the pedestrian dynamics element, which in the case of PeDViS is the Nomad model. The activity scheduling thus determines the strategical choices: activities, destinations and departure time. The activity scheduling element in PeDViS is a data-driven model, meaning data that serves as input for the model are the context, layout, population, demand and measures (NPIs). This entails that the user enters scenario-specific data to start the activity scheduling, which is also called the activity assignment model. The activity assignment model consists of four parts: generation of groups, visit characteristics, person characterization and activity scheduler. These four parts and their interactions are shown in Figure 2

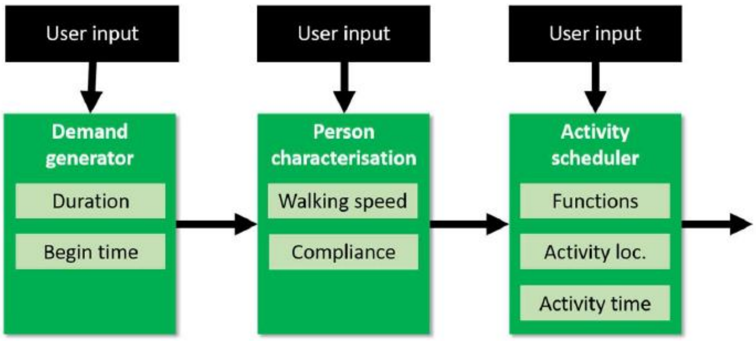

Figure 2: Elements of activity assignment model

The visit characteristics are based on user input regarding the demand and average visit characteristics, such as opening time of the venue, average duration of visits and estimated demand per hour. Based on this information, the demand generator generates the inflow of pedestrians into the model. Next, the personal characteristics of each individual have to be defined. In the restaurant scenario of Duives et al. (2022) only walking speed is varied, but some of the other possible characteristics to vary are reaction time and the size of the individuals. In addition, the activities that can be performed should be defined, along with their location and their duration. Finally, the activity patterns people follow, defining which activity they do at what time, have to be defined. For the restaurant version, an activity scheduler has been designed, where the user has to input some information relating to the demand and personal characteristics, such as how long people stay in the restaurant, how many people come into the restaurant and at what times. After this data is inputted, the model generates each person's activity schedule. This data-driven nature of the activity scheduling allows users to enter data into the model that is context-specific. This allows for real-life situation virus transmission modeling, making the choice for the best (combination of) NPIs office-specific, and therefore more accurate and useful. This way, after some adaption, office managers could even use the tool to input specifics of their office and then come to the ideal set of NPIs to safely manage their offices. In addition, the data-driven nature provides allows us a good alternative to deal with the limited work on activity choice modelling in offices.

The operational dynamics element in Figure 1 represents the tactical and operational choices, using the agent-based Nomad model. At these levels, another classification is important to recognize, namely microscopic and macroscopic models. Macroscopic models focus on studying characteristics on a system level (Di Francesco & Rosini, 2015), while microscopic models represent each pedestrian as an individual (Bellomo et al., 2012) and have a separation between individuals and the environment (Bandini et al., 2014; Bellomo & Dogbe, 2011). The Nomad model is a microscopic model on both the tactical and operational levels. Because of microscopic models' characteristic of treating each person as an individual with different properties (Wang, 2021), microscopic models are suitable to use in the context of this research: the movement of individual pedestrians in an office setting.

On a tactical level, there are different types of microscopic models on pedestrian behavior to distinguish: Rule-, force-, goal- and utility-based models (Campanella, 2016). The Nomad model is a utility-based model (Hoogendoorn et al., 2004). With utility-based models, agents make a trade-off between an infinite number of different utilities, while trying to maximize their overall utility. Utilities are the different functions making up the state of the agent (Campanella, 2016). Its principle of least effort (Zipf, 1949) includes both the walking and the activities performed in between (Campanella, 2016), thus also focusing on the activities used in activity scheduling (choice modelling). Besides Nomad, there are various other types of utility-based walker models, such as discrete choice models (Antonini et al., 2004) and fuzzy logical models (Nasir et al., 2012).

There are six distinct types of models that implement the operational level movement choices; namely i) Graph-based models, ii) Continuous models, iii) Cellular Automata models, iv) Force-base models,

v) Velocity-based models and vi) discrete choice models (Duives et al., 2022). The first two types are macroscopic, and the last four are microscopic models. The operational level of Nomad used in the PeDViS model can be defined as a Force-based model (Duives et al., 2022). Force-based models assume that various (social) forces influence a pedestrian's behavior, either repulsing or attracting them. These forces can come from, for example, obstacles or other individuals in the model (Helbing & Molnar, 1995). Force-based models allow for the inclusion of tactical choices, allowing the presence of the two levels of movement choice in the Nomad model.

3.2.2 Virus spread and risk identification model

The third and fourth element in Figure 2 are the virus spread and risk identification model. The virus spread is simulated through a transmission model. Many different transmission models exist that are suitable to simulate virus spread in indoor spaces. The model that is used in the PeDViS model is the QVEmod model (Duives et al., 2022), an epidemiological model for Quantifying Viruses in Environments. QVEmod is a spatially explicit model incorporating three routes of infection: virus-laden aerosols, virus-laden droplets (larger respiratory particles that do not remain airborne) and fomites (objects contaminated by droplets). Including all three routes allows for more accurate estimations of virus transmission by better inclusion of spatial and temporal dynamics (Duives et al., 2022).

The model simulates virus emission of individuals into the air as virus-laden aerosols and onto surfaces as virus-laden droplets. These aerosols and droplets travel and diffuse throughout the environment and become less infectious over time. Susceptible individuals in the environment can in turn be infected with the virus by inhaling these aerosols or touching surfaces contaminated with virus-laden droplets (fomites) (Duives et al., 2022). It can be noted that there are two main types of transmission models: agent-based and equation-based. Equation-based entails that mathematical equations lie at their foundation and the simulation consists of evaluating these equations (Paranuk et al., 1998). QVEmod is considered an agent-based model. This allows it to take into account the dynamics of human interaction and the choices of individuals pedestrian as defined in Nomad. In addition, the QVEmod model is a microscopic movement model, same as the Nomad model, lending itself well to the spatial variation and dynamic interaction between individuals, provided by the input from the Nomad model.

The information on the spread of the virus on surfaces and in the air produced by the QVEmod model is taken as input for the risk identification model. This information is then used to estimate the risk each individual has of being infected by the COVID-19 virus, based on the amount of virus they came into contact with during their time in the environment (Duives et al., 2022). By extension, the total number of infections resulting from the event can be estimated. The relationship between exposure to the virus and infection risk is built on the dose-response relationship by Nicas (1996) and is calculated through all three exposure routes.

3.3 Existing versions

Currently, there are only two types of contextual settings the PeDViS tool has been tailored to: the restaurant setting (Duives et al., 2022) mentioned earlier and an outdoor event setting (Wang, 2021). The original restaurant version was developed through the SamenSlimOpen research project between the TU Delft, the University of Wageningen and the Erasmus MC hospital. The resulting model focuses on an indoor restaurant setting, in which various groups visit a restaurant and restaurant employees walk between tables.

The restaurant version was designed to be a tool for restaurants to help them get insight into the possible spread of the COVID-19 virus in their restaurants and what the effect would be of various measures (Duives, 2022). For this reason, an open-access web-based simulation environment was designed, called the SamenSlimOpen app. In this app, users can input specifics of their restaurant setting, such as the size of the indoor space, the locations, material and size of the objects (chairs, tables, bar), among other things. Various restaurant-specific objects and their function have been predefined in the webtool for the user to choose from, such as a register where people pay and a coatrack where people hang their jacket and pick it up at the end. In addition, it contains more general objects such as an entrance, exit, and toilet. Some behavior aspects the user can fill in are the number of people that come in per hour, when people come in and how long different activities (such as going to the restroom) take. The number of employees visiting the tables and what tables they visit can also be chosen, although this function is not active in the latest version (Duives, 2022). Finally, the user can also decide what measures they want to implement per model run. The measures to choose from are:

- cleaning frequency of surfaces
- ventilation rate
- whether or not face masks are obligated while walking
- whether or not tables can have multiple shifts per day/simulation

After setting all these configurations, simulations can be run, after which the webtool also allows the user to compare the outcome of different simulation runs. Duives et al. (2022) also performed a case study using their own model given a specific restaurant setting.

The second version of the PeDViS model that exists, is the adaption made to study the COVID-19 transmission risk at large outdoor event settings, made by Wang (2021). This version consists of a case study application based on the Amsterdam Open Air music festival, which took place on June 1st and 2nd, 2019. For this adoption, the Gaasperpark, in which the festival took place, was modeled as the environment of the model. The behavior of the pedestrians was based on movement data obtained from the event. As the behavior of people at a festival is significantly different from the behavior of people in a restaurant, the restaurant scheduler from the version of Duives et al. (2022) could not be used. Therefore, new code was to be written to generate personalized schedules for the different functional areas that each person visited and at what time.

However, these existing versions of the PeDViS model are not suitable for simulating an office setting. First of all, there are various differences between a restaurant setting and an office setting. One difference is that people often are in the office for a much longer amount of time, increasing the infection risk. Second, people know more people within the office room than they would in a restaurant, and thus interact with more people. Finally, there is just one group of people working at the office instead of the two groups in the restaurant setting (staff and visitors). The adaption for outdoor events is also unsuitable for the office setting, among other things because it is mostly based on outdoor virus transmission, which has a very different virus transmission from indoor virus transmission.

4 Research Design

In this chapter, the research approach of the thesis is determined and elaborated upon. In section 4.1, the modelling approach is chosen and described and the sub-questions are formed based on this approach. In 4.2, the methods used per sub-question are explained.

4.1 Choice for an approach

As stated in the previous chapter: an existing model, functioning as a base model, needs to be adjusted to create a model that enables office managers to compare various NPIs based on their the effect on containing COVID-19 transmission in office spaces. This type of model would improve the decision-making process for office managers regarding which NPIs should be implemented in their offices. It is widely acknowledged that integrated assessment models enhance the effectiveness of decision-making (Bland, 1999; Carnevale et al., 2012: Voinov & Bousquet, 2010) and the objective of the research is the design of and simulation using a model. Therefore, it is decided a simulation modelling approach is followed. There are multiple approaches to building a model (Malleson, 2018). The simulation modelling approach followed in this thesis is based on the general modeling cycle and shaped after the examples of the modeling cycle by Malleson (2018) and the modelling steps of Brugnach et al. (2005). The resulting modeling approach that is followed can be seen in Figure 3. Based on this approach, the research in this thesis is structured in four main phases: forming a model design, consisting of data collection and conceptualization (phase 1), model implementation (phase 2), experimentation (phase 3) and model assessment (phase 4).

The data collection in the first phase is needed for multiple reasons. First, it serves as input for the conceptualization and is used to explore what elements are present in real life and which are the core elements that need to be included in the model. This is done because no conceptual model on the behavior of employees in an office setting, fitting for this thesis, exists as of the start of this research. Second, the data collection also serves as input for the model implementation phase, where data is needed on what values need to be assigned to the variables defined in the conceptualization.

After the data collection is completed, the phase of conceptualization can start. Padilla et al. (2019) have the following definition: "Conceptualization (of a model) represents a collection of concepts and relationships." This step is needed to essentially scope and frame the model. The next phase of the model implementation focuses on *how* the model works and the implementation of the conceptual model into a computerized model. The model implementation results in a working model, the dynamics of which are then explored in the experimentation phase. Finally, the model can be used to help answer the research question by running simulations and assessing the results.

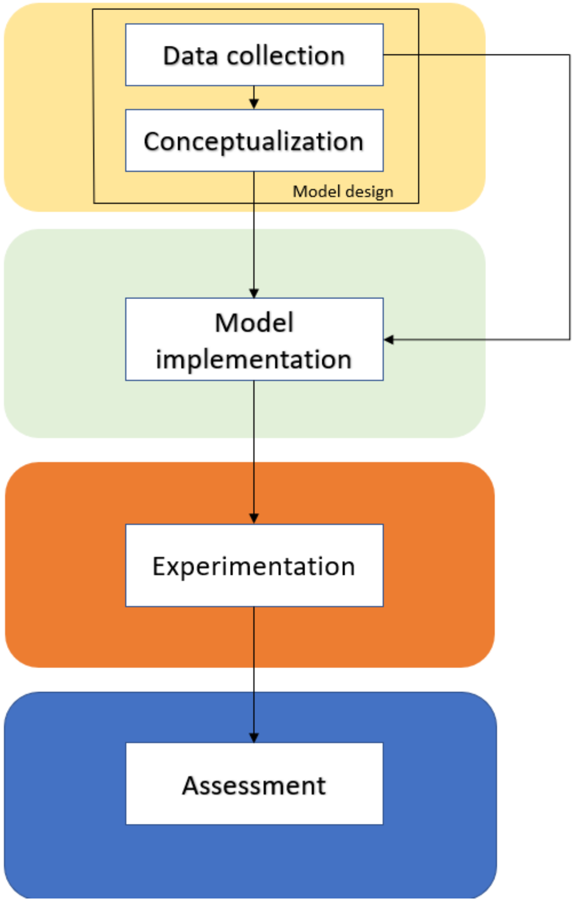

Figure 3 : The Simulation Modelling approach

This structuring of the thesis into 4 phases is implemented by having each phase represented by one sub-question. The sub-questions are:

1. What is the typical setup of office spaces and the behavior of the people working there? (Phase 1)
2. How to quantitatively model the behavior of people working in an office and the resulting virus spread accordingly? (Phase 2)
3. What is the influence of the parameter values related to the part of activity scheduling focusing on meetings on the virus spread? (Phase 3)
4. What is the effect of the NPIs on the office space scenarios being simulated? (Phase 4)

These sub-questions consequently lead to an answer to the main research question.

4.2 Methods used

Per phase, a different method is used to answer the associated sub-question. In the first phase, the first step is data-gathering. The main protocol used here was observation of office employees, supplemented by desk research. Observations were held at the TU Delft employee areas for this purpose. Having the data, the model can be described in the conceptual model phase and model implementation phase. There are many different ways a model can be described, depending on, for example, the model paradigm used (in this case agent-based) or the subject of the model (in this case COVID-19 spread based on pedestrian movement). A widely used technique is the use of a modelling protocol, as mentioned in Padilla et al. (2019), Banks (1998) and Dai et al. (2020). A modelling protocol provides a more standardized way to capture the model descriptions, often having more clear rules to follow. Agent-based Modelling researchers believe a specific protocol should be followed at the beginning of model building to facilitate communication (Moore et al., 2007). A widely used example of this technique in agent-based modelling research is the Overview, Design Concepts and Details (ODD) protocol (Dai et al., 2020). This ODD protocol (Grimm et al., 2006; Grimm, 2010) is used in this thesis to both explain the PeDViS model and conceptualize the adaption of the model and consists of the following agent-based modelling description elements:

1.) Purpose
2.) Entities, state variables, and scales
3.) Process overview and scheduling
4.) Design concepts
5.) Initialization
6.) Input data
7.) Sub models

An overview of the steps, grouped into the three main blocks (Overview, Design Concepts and Details), can be seen in Figure 4.

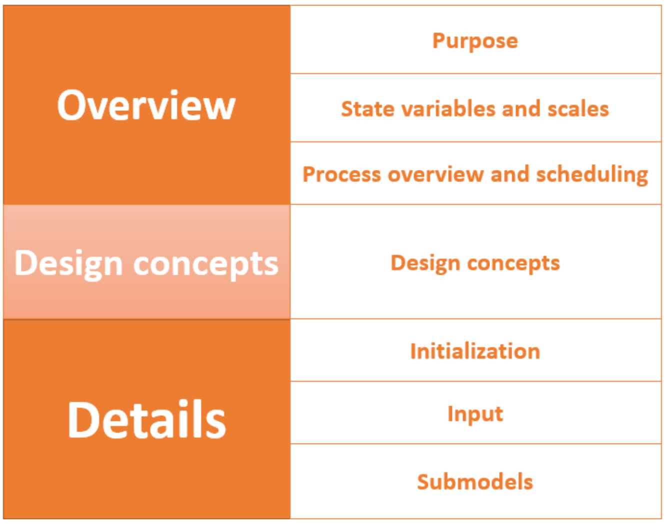

Figure 4: The seven steps of the ODD protocol, grouped into three main blocks (Source: Grimm, 2004)

The first four steps of the ODD protocol are covered in the conceptualization phase as these more high-level steps provide a good abstract description, needed to produce a conceptual model. In the model implementation phase, the last three steps, focusing on the details, are covered, as the model implementation focuses more on the exact input and initialization of the model.

After the model is finalized at the end of the model implementation phase, the effect of varying parameter values related to the part of activity scheduling focusing on meetings is explored in the experimentation phase. As these values relating to the activity scheduling in the office, including meetings, are based on the data collection and modelling of this thesis, not much is known yet about their influence on virus transmission. Therefore, it would be interesting to learn more about their influence on the number of infections. The focus on the meeting part of activity scheduling is chosen as meetings are expected to have an important influence on virus transmission, due to the group gathering taking place in meetings. The exploration of the influence of these parameters is done by a sensitivity analysis. Outputs are generated for different set-ups relating to the meetings held, getting an overview of the infection risk per set-up. Having performed the sensitivity analysis, it can then be concluded whether the meetings held in an office have a significant impact on the number of infections and if so, what relationships exist between those variables. It also allows for a set-up in which the NPIs can be tested in the next phases, now having information on the infections for different set-ups relating to the meetings held in an office.

In the final phase, the model assessment, the results of the experimentation phase are used to provide different office set-ups on which the impact of various NPIs can be tested. To test the impact of the NPIs, all possible combinations of NPIs are defined as different policies. The policies thus differ in the NPIs that are implemented. Using these policies, policy analysis is used to compare the impact of the NPI combinations on containing COVID-19 spread. Based on their impact, it can then be concluded what the overall effect is of each NPI and which NPIs are most suitable to apply in the simulated office space.

The flow diagram in Figure 5 shows the structure of the research. Per chapter, the main activities are shown, along with the methods used. The flow of data between the various phase can also be seen in the arrows.

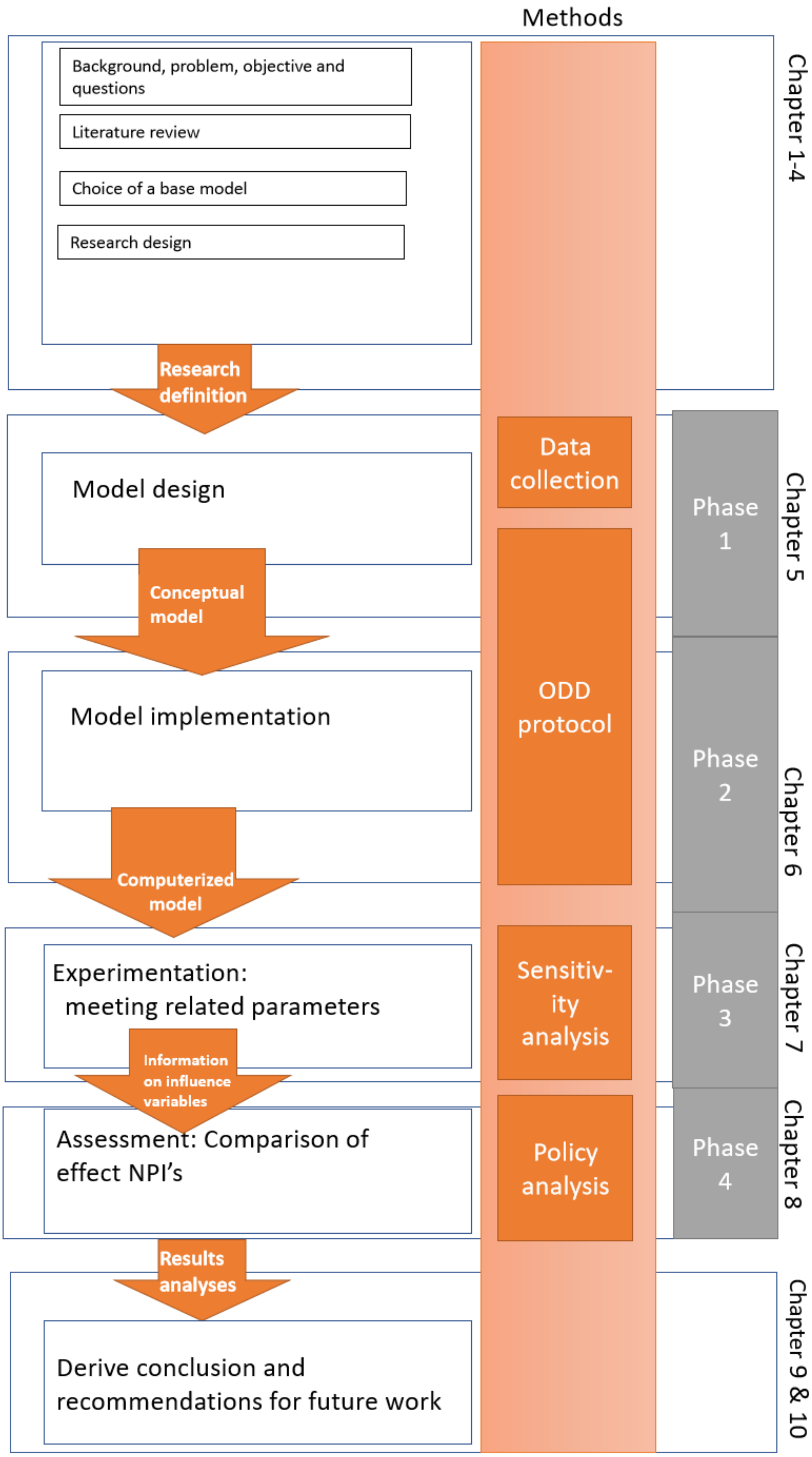

Figure 5: Flow Diagram

5 Model Design

This chapter represents the conceptualization phase of the model. First, the data collection performed ahead of the conceptualization is described. In the next section, 5.2, the first sub-question of the thesis is answered by following steps 1-4 of the ODD protocol.

5.1 Data collection

Ahead of designing the conceptualization, it was necessary to collect some initial data to get insight into the typical layout and behavior of office employees that need to be modelled and to better define the scope of the model. This data was collected through a combination of the method of observation, desk research and analyzing employees' work schedules, as existing conceptual models on employee behavior in offices were not yet available. In this section, the observation protocol that was followed is described, along with the other data gathering methods used and types of data found.

In the period from the 14th of March 2022 until the 25th of March 2022, a total of six days of observations were performed, at various locations at the TU Delft employee areas. The employee office areas at the campus of the Technical University (TU) Delft were chosen as the general location to perform these observations because the author could easily access them without much prior arrangements. Within this location, three spots were chosen to observe the layout of the offices and the behavior of its employees. The first location was on the third floor of the faculty of Civil Engineering, where its employee meeting rooms are located. The second observation spot was the third floor of the Technology, Policy and Management (TPM) faculty, which contains the employees' personal offices, a coffee area and two bathrooms. The third observation spot was on the first floor of TPM, where employee meetings could be observed through a glass wall. An overview of the different observation locations, the time spent observing at each of them and the type of behavior observation possible per spot can be seen in Table 3.

Table 3: Overview of the observation performed, regarding locations, dates, observed areas and types of observations made

Observation spot	**3rd floor Civil Engineer**	**3rd floor TPM**	**1st floor TPM**
Observation time	one whole day, one afternoon and one morning	Two whole days	One morning
Visible areas	1 Coffee area	1 Coffee area, 1 printing area	1 Coffee area and meeting room
Movement between areas and the entering and leaving of the functional areas that is visible	Between the 4 meeting rooms, 1 coffee area and 2 (gender separated) restrooms	Between the roughly 10 personal offices, 1 coffee area, 2 (gender separated) restrooms and 1 printing area	Between 1 Coffee area and 1 meeting room
Duration within a functional area timed	4 Meeting rooms, 1 coffee area and 2 restrooms	Coffee area, restrooms, printing and meetings in personal office	1 Meeting room

Observing these various employee areas at the TU Delft and the employees' use of the space provided various types of information. First of all, it provided some insight into the general layout of office space and what design of the space would be efficient. For example, coffee areas were observed to often be situated close to meeting rooms, and to often be visited prior to and during a meeting. Therefore, it was concluded that it would also be fitting to have the coffee area and meeting rooms located close to one another in the model.

In addition to information on the layout, the behavior of the employees was studied in more detail. These observations were done with the use of timestamps. First of all, the time at which people entered and left the different areas of the office, such as a coffee area or a meeting room, were timestamped. If it was possible to look inside an area, such as through a glass wall of a meeting room, the employees' behavior within that area was observed and the activities they performed within the area were timestamped. Each translocation between the different areas was also timestamped. The timestamps went to the level of detail of 10 seconds when possible and needed but often were on the level of minutes. In addition to the time stamps, often additional notes were written down to provide some context, to help keep an overview of the observations or to include some additional information. The time spent observing each area and the type of observations made there can be seen in Table 3. An example of how the observations were registered can be seen in Figure 6, which shows the observation sheet of a meeting room for a small period of time. Similar protocols for writing down the observations were used for each of the functional areas.

15-3-2022	meeting room 2.66	duration			number of people in the meetingroom	number of people entering or leaving at timestamp	
:43-9:11	empty						
:11-12:05	meeting	114.00		09:11	1.00	"+1"	person opens meeting room, gets coffee in between
				09:16	2.00	"+1"	2nd person enters
				09:17	2.00	0	2nd person gets coffee
				09:26	3.00	"+1"	someone enters
				09:56	5.00	"+2"	2 new people enter
				09:58	6.00	"+1"	one new person enters. Two people get coffee together and one person visits the bathroom
				10:01	7.00	"+1"	1 new persons enters, who gets coffee first
				10:02	8.00	"+1"	1 new person enters, who quickly gets coffee in between
				10:03	8.00	0	meeting stars again and everybody is inside
				10:06-10:09	8.00	0	someone leaves for a little bit and come back later
				break			
				11:02	0.00	0	4 people enter shortly after one another

Figure 6: Example of part of a filled-in observation sheet for a meeting room

The data retrieved from the observations was supplemented by data gathered from both work schedules and desk research. The work schedules were provided by two employees. Both work schedules spanned a two-week period and contained information on their daily schedule, such as the time and location of meetings, the number of participants in their meetings and the number of meetings they attended per day. Finally, some additional desk research was done to learn more about the frequency in which the different locations in the office are visited.

Based on the information that can be gathered from the observations, desk research and the work schedules obtained, various insights into the behavior within an office could be deduced. These behaviors include the time spent within an area, the frequency and the order in which the functional areas were visited, the number of people attending each meeting and the number of people present in a functional area. The details of the values found and data analysis performed can be found in Appendix A.

5.2 The conceptualization

The conceptualization of the model is described in this thesis using the first 4 steps of the ODD protocol, covering the overview and design concepts blocks. The resulting conceptualization of the model leads to answering the first sub-question:

What is the typical setup of office spaces and the behavior of people working there?

5.2.1 Purpose

The first step of the model is the purpose of the model. This has already been touched upon in the chapter 1, but is reiterated here for the sake of the ODD protocol:

The objective/purpose of this thesis research is to help offices safely stay open, by providing a model that allows for the comparison of the effects of various NPIs to contain COVID-19 in real-life office settings.

The model does this by simulating the virus spread in an office through three routes while taking into account the office layout and pedestrian movement within.

5.2.2 Entities, variables and scales

After having described the first step of the ODD protocol, the second step can be reported: Entities, variables and scales Grimm et al. (2010). The PeDViS model, which is used to make the final model, simulates pedestrian (the employee) movement and virus spread in indoor spaces. Due to the extensive nature of the indoor space environment, in addition to the extensive nature of the office employee's behavior, the choice was made to first describe the environment in a high level of detail in this step, and then focus on the entities, variables and scales not relating to the environment, such as the office employees and the virus.

Grid cells

Spatial units, in the form of grid cells, are present in the model for both the Nomad and the QVEmod part. The grids represent 0.1 m x 0.1 m rectangles inside the Nomad model on which pedestrians can walk. In QVEmod, the environment consists of grids of 0.5 m x 0.5 m and are proximate to the space one person occupies. An office consists of different functional areas, areas where different functions are performed, such as a meeting room for meetings, a personal office for working and a canteen for having breaks. The layout of the model needs to be divided into the different types of functional areas included in the model. Therefore, a selection of the types of functional areas included is made as a first step. Different possible functional areas for offices can be seen in the first column of Table 4. To better represent the behavior in office spaces, the choice was made to include several of these functional spaces in the model, instead of modelling just one type of area. Due to the limited scope of this thesis, the choice was made to include only four functional areas. In Table 4, it can be seen which functional areas were chosen and which were not and for what reasons.

Table 4: Potential functional areas in offices to include in the model

Functional areas in an office	Included?	Reason for inclusion/exclusion
Personal offices	Yes	Important areas, which office workers visit every workday
Group working spaces/shared offices	No	Individual office space is expected to be easier to model, and having both types of offices (personal and shared) in the model would provide less new information than adding other types of functional areas
Coffee machine area	Yes	Important areas, which office workers visit (almost) every workday
Breakroom	No	Relatively similar to the restaurant scenario, and therefore would have less added value to research than the other areas
Bathrooms	Yes	Important areas, which office workers visit every workday
Meeting rooms	Yes	Important areas, which office workers visit (almost) every workday
Wardrobe	No	Lesser frequency of use compared to the other areas
Printing area	No	Lesser frequency of use compared to the other areas

The grid cells in the model have to be divided into which types of functional areas they belong to. In addition, according to Grimm et al. (2010), the size of a spatially explicit model should also be specified. To have some more background on appropriate sizes for the offices and their functional areas, some initial information on office setups and behavior within them (used in the next steps of the ODD protocol) was gathered. To define the size of the space used for the office, first, a definition had to be made of the type of office to be modelled. The choice was made to use a small office setting for the model, for reasons of simplicity. According to the definition of the European Commission, a small office is defined as having less than 50 employees (European Commission, 2020). Therefore, it was decided the layout would need to be able to facilitate 40 office employees. Currently, many people also have the option to work from home, so it is very much possible there are less than 40 people in the office on a given day. The number of employees present is therefore variable, but the maximum number of employees at the office is set at 40, for which sufficient facilities need to be present. External people visiting the office, such as business partners or acquaintances, are outside of the scope of this model. Therefore, all pedestrians shall be employees working in this office.

The required sizes and quantities for each of the functional areas are determined, based on the laws regarding office spaces and the frequency of use of the different areas and the observations made in the data collection phase. An overview of the sizes and quantity per functional area can be seen in Table 5. Details on how these sizes and quantities were arrived upon can be found in Appendix B. The size of the coffee area was originally set to be 6 m^2 (2.0 m x 3.0) based on the data. However, the size has been slightly altered, reducing the measurements in both directions by 0.2 meters due to the need to also incorporate walls.

Table 5: Sizes and quantities of the different functional areas in the office

Functional area	Size	Number
Personal office	12.2 m^2 (2.8 x 4 m)	40
Meeting room	30 m^2 (5 m x 6 m)	2
Coffee area	5.04 m^2 (1.8 m x 2.8 m)	1
Bathroom	8 m^2 (2 m x 4 m)	1 area, consisting of 4 stalls and 4 sinks

Environmental agent entities

Next to the grid cell entities, there are also several agent entities related to the environment present in the model. These are the various input in the Nomad model (TU Delft, 2021) and QVEmod that make up the layout of the office. An overview of the entities and their state variables is given in Table 6.

Table 6: Environment agent entities and their state variables

Entities	Explanation	State variable relating to Nomad	State variables relating to QVEmod
Walkable areas	The area in which the pedestrian can potentially walk	ID, Coordinates	Coordinates, air change rate, amount of virus-loaded aerosols and droplets in the air
Obstacles	An area of the walk level, which pedestrians cannot access, and they try to avoid. Examples of this are walls, chairs, tables etc.	Type (shape), coordinates, ID, (radius, rotation), see-through	Coordinates, amount of fomites on surface, amount of the surface being touched, material
Destinations	Locations where a pedestrian goes to perform an activity, such as the coffee machine or the restroom.	Type(shape), coordinates, ID	Coordinates, amount of fomites on surface, amount of being touched, material
Sources	Location where pedestrians enter the model	Type (shape), ID, Coordinates, (direction vector)	Coordinates

Setup of the functional areas

According to the ODD protocol, “with spatially explicit models that include spatial heterogeneity, a figure representing the model area in a typical configuration can be useful”. Such figures are therefore produced in this section, using the earlier defined sizes and environment entities as input. In Table 7, Table 8, Table 9 and Table 10, the layouts of respectively the personal offices, the two meeting rooms, the coffee area and the restrooms can be seen.

One important aspect should be mentioned about the personal office layout. Based on the information provided by an employee alongside their work schedule, the knowledge was added that in general, meetings with less than four people were held in the meeting area of personal offices instead of in meeting rooms. As the meeting area in a personal office thus has a different main function than the rest of the personal office, it is regarded as a separate functional area from here on out. The personal offices are thus all split into two functional areas: the personal office, the place where employees work on their own and the meeting area. In the top half of the office in Figure 7, the two possible layouts for a personal office can be found, which both consist of a meeting area and a meeting area.

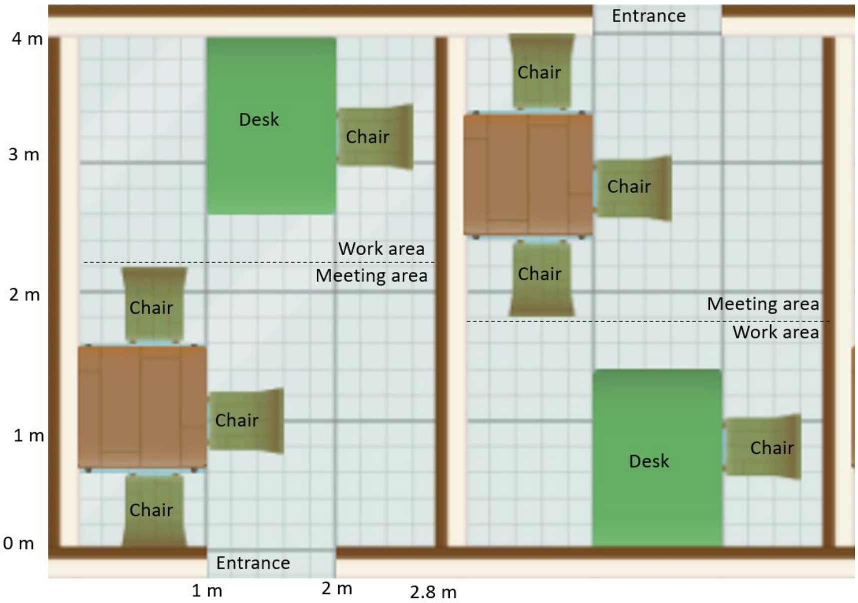

Figure 7: Layout of personal office with meeting area

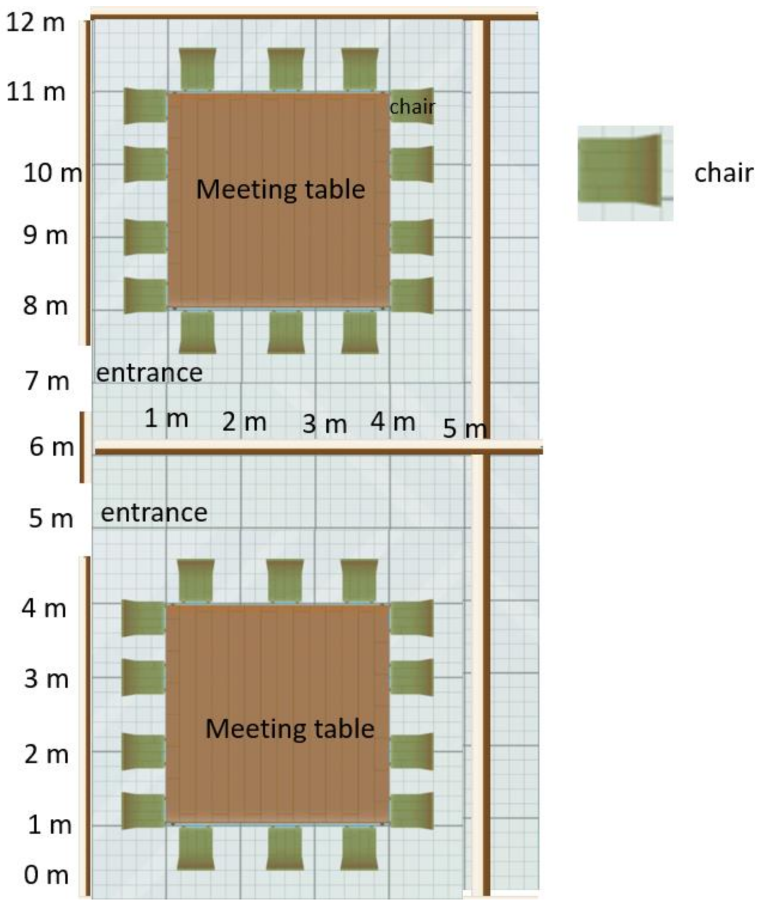

Figure 8: Meeting rooms layout

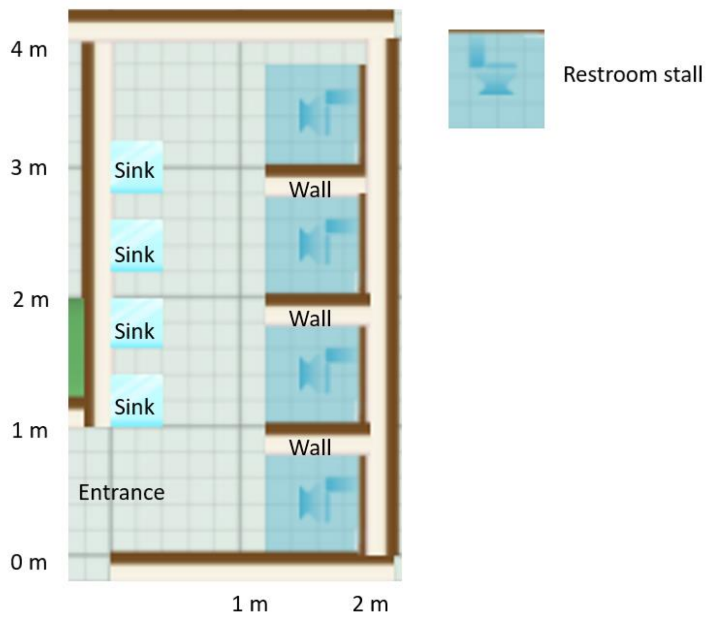

Figure 9: Layout of restrooms

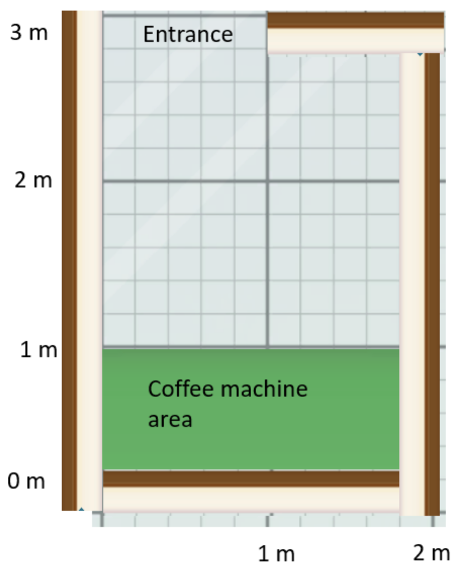

Figure 10: Layout of the coffee area

Setup of the entire office

After having defined the size and setup of each of the functional areas within the office, the overall setup can be designed. At first, it was considered to simulate each of the functional areas separately. However, first of all, a high amount of movement between the different areas was observed, which also has potential for virus spread and would be ignored if the functional areas would be modelled separately. In addition, in a real-life office, there exists a substantial amount of complexity in terms of movement between areas and the scheduling of people visiting these areas one or more times per day. This complexity is deemed to be important to also have represented in the model simulations. Therefore, it was decided to design one big layout for the model showing the whole office, that connects each of the different functional spaces. There are no doors in the office layout, as these also did not exist in the original restaurant version and they are not expected to play a large role in the virus transmission. In Figure 11, the setup of the entire office, based on the setups and sizes of the functional areas, can be found.

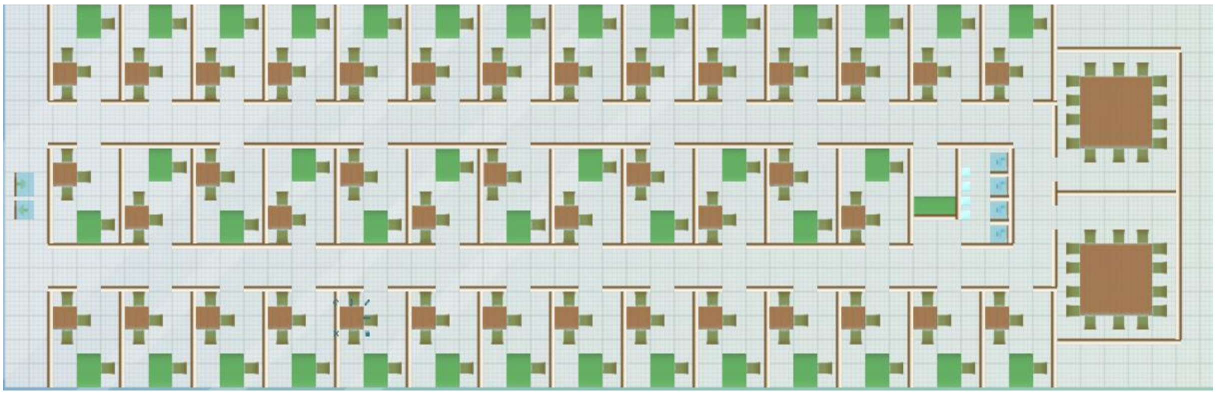

Figure 11: Setup of the office for the model

All walls in the office are 0.2 m thick in accordance with the standard sizes in the SamenSlimOpen webtool. The office has a rather long shape with two long hallways, something that was also often observed in the studied offices. There are three rows of offices, 14 offices on the first and third row and 12 in the middle row. At the end of the middle row of offices, a coffee area and the restrooms can be found. These are across from the two meeting room, which is convenient if people want to quickly get a coffee or use the restroom before, during or after a meeting. The side of the personal offices on which the entrances are located are alternated at the middle row, to keep the amount of traffic in both hallways equal. The main entrance and main exit can be seen on the far left side of the office and are used by all employees at the start and end of the workday respectively.

Non-environmental entities

After having defined the environmental part of step two of the ODD protocol, the entities, variables and scales not relating to the environment, such as the employees and the virus, are now described to finalize step 2. The PeDViS model consists of different models: the activity scheduling, the Nomad model, QVEmod and the risk identification model. Therefore, different hierarchical levels can be distinguished. The remaining entities are therefore ordered per hierarchical layer. The activity scheduling layer can be seen as the lowest layer, providing input to the Nomad model, with the Nomad model then providing input for the QVEmod layer. The risk identification layer is then built upon the QVEmod layer, thus being the top layer. An overview of the entities and their state variables is given in Table 7. The office employees are seen in all layers, as they have state variables relating to different layers. Some state variable values are assigned individually, while others have the same value assigned for each employee. Details on this are given in the model implementation chapter. Details on the state variables in the Nomad, QVEmod and risk identification layers can be found in Appendix C.

Table 7: Entities and state variables per hierarchical layer PeDViS

Hierarchical layer	Entities	State variable
Risk identification	Office employees	Virus exposure, being infected or not
QVEmod	Office employees	Emission rate of virus (if infected), respiratory activity currently performing, infectiousness, inhalation rate, level of fomites on hand, proportion of virus emitting through different routes, infectious dose Wearing a mask while walking, Compliance with mask-wearing and measures influencing the air change rate
	The virus	Decay rate in aerosol form, deposition rate in droplets, diffusion coefficient in aerosols, decay rate on different surfaces, proportion of virions reaching respiratory cells, transfer efficiency from hand to surface
Nomad	Office employees	Pedestrian parameters related to walking behavior, activities assigned
Activity scheduling	Office employees	Personal schedule, assigned personal office, Compliance with NPIs related to activity schedule

Finally, the time scale of the model should be addressed. All time steps have been chosen in accordance with the timesteps used in the original version of PeDViS of Duives et al. (2022). First of all, this entails that the timestep for the virus emission calculations are set to 0.5 minutes. NOMAD then generates pedestrian trajectories at each 0.1 second time step. The trajectories of the pedestrian, which are an output of Nomad, can have an accuracy of up to a 20 second time step. However, the timestep can also be bigger if the individual has not changed their location during the 20 second time step.

Each simulation is run for an entire working day, from 8:30 till 17:30, a duration of 9 hours, which is in accordance with the office times observed in the data collection phase. The choice is made not to simulate a whole week, because the difference between the crowdedness of different days (on a Tuesday often more employees being present than on Wednesday) can also be simulated by differing the number of employees between simulations.

5.2.3 Process overview and scheduling

The third step consists of describing the processes and scheduling in the model. This entails describing the processes the model entities, such as the employees and virus, go through and how their state variables are updated. First, some description of the processes is given. As the processes an employee goes through are called activities in Nomad, the processes are hereafter referred to as activities. The activities are again separately defined for the different models of which PeDViS consists, starting with the activities within the Nomad model. For the Nomad model, different activities can be defined. In Nomad, these are called activities and are performed by the pedestrians (the employees for the office adaption). For the office setting, office-related activities need to be defined. These have been selected, based on the observation at the TU Delft. In Table 8, all activities included in the Nomad part of the model have been defined in the column “Activity”. In the column “Functional area” it is stated in which functional area(s) this activity is performed.

Some activities that are observed in offices have been left out of the model. The first reason for this is that activities that do not take place in one of the functional areas incorporated in the model, such as having lunch as this is held in a lunch and/or break room, which is not included in the model. Other reasons for activities not being included are that the activity is not expected to provide a significant increase in infection risk and/or was expected to be very hard to incorporate into the activity scheduling part of the model. Some of the activities observed that were left out were: people stopping by a personal office to have a short talk from the door opening, people hanging their coat on a coatrack in a meeting room and people (un)packing their own bag or people quickly getting grabbing/putting away. The last activity could be seen for meeting rooms, personal offices and the coffee area. In the case of the meeting room, it was seen that people often entered, put their bag and coat away and then quickly got a coffee or visited the restroom before returning. As the short activity of putting away of bags of stuff is excluded, this is then just be seen as people entering the meeting room *after* using the restroom or coffee machine.

Table 8: Activities in Nomad, defined per functional area

Activities in Nomad	Functional area
Use toilet	Restroom
Wash hands	Restroom
Small meeting	Personal office
Meeting	Meeting room or office meeting area
Work at desk	Personal office work area

Besides the activities in the office that need to be incorporated through the Nomad part of the model, other activities need to be included through the QVEmod part. An overview of these activities can be seen in Table 9. These activities can be performed by either the employee or the virus. The activity of emitting virus by employees happens when an office employee performs respiratory activities. The amount of virus a person emits depends on both what type of activity they do (talk or sing) and whether or not they are wearing a mask (an employee state variable).

Table 9: Activities in QVEmod, defined per executer

Activities in QVEmod	Executor
Emit virus	Employee
Contaminate surfaces	Employee
Inhale virus	Employee
Self-inoculation	Employee
Decay	Virus
Diffuse	Virus
Fall	Virus

Infectious employees can contaminate the objects they touch with the COVID-19 virus, such as tables or chairs, or by the virus particles they emit landing on an object’s surface. Susceptible employees can then in turn self-inoculate by touching these contaminated surfaces. Contaminated surfaces can also be cleaned in the model, after which all of the surfaces in the model are free of the virus again. Virus particles decay over time and the air change rate (the ventilation) influences how fast this happens, with a higher air change rate causing faster decay of the virus. For QVEmod, there is a certain order in which the spread of and contamination with the virus iterates. The order in which the processes happen can be seen in Figure 12

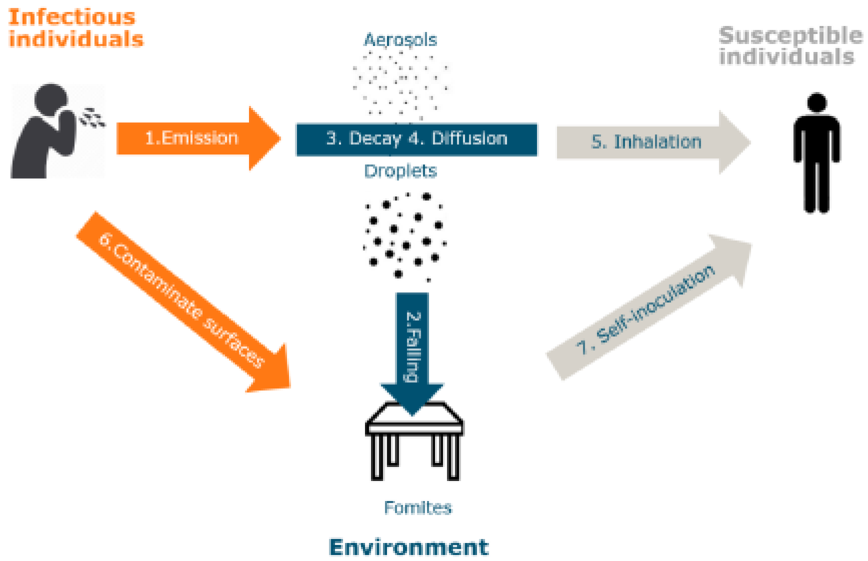

Figure 12: Order of virus spread and contamination (source: Duives et al. (2022))

Finally, the risk identification layer consists of only one process: the calculation of the infection risk of each individual. The risk of being infected depends on how much virus a person has been exposed to during the work day.

The calculation of the infection risk and the final determination of whether an individual is infected is applied to each office employee in the model. The probability of being infected depends on the amount of exposure to the virus the individual has accumulated, either through inhalation or touching contaminated surfaces. Whether an employee is exposed to the virus in these ways depends on what place in the office he is doing a certain activity and at what time, if he is at a certain location at the same time or just after an infected individual is also there or not. This in turn depends on the work schedule the employee has. The work schedules of employees are determined by the activity scheduling in the model and brings us to the scheduling part of this step. The activity scheduling in the model focuses on the scheduling of the activities of the Nomad part, due to the way the PeDViS model is set op. This sequence of activities, called activity schedule, differs per employee and is assigned to them the moment they enter the model. These are the schedules the employees have per workday. The order of these activities in the schedules is for a large part based on transition matrices, the values of which are based on information retrieved through the data gathering process.

However, using a transition matrix would not allow for the incorporation of scheduled meetings in the schedules, meaning that a group of people are in a meeting room at the same time and for the same period of time to hold a meeting, as it would be in real life. To be able to model such meetings and the consequent virus spread, it is decided to include the behavior related to meetings by starting the process of creating the activity patterns by planning such meetings. After these meetings are planned, the remaining gaps in the activity schedules can then be filled with the other activities, in an order based on the transition matrix. The number of meetings held on a day and the times and sizes of these meetings then need to be predefined, after which the right amount of employees are picked for these meetings. These employees then get the meetings added to their schedule, for the allocated time of the meeting.

The filling in of the activity schedules of the employees thus happens in multiple steps. Each employee starts with an empty schedule, a rough example of which can be seen in Table 10. The schedule allows for the assignment of each of the nomad activities to a certain time that day, for the corresponding duration of that activity. The timestep is specified in the next phase of model implementation.

Table 10: High-level example of an employee schedule

Timestep 1	Timestep 2	Timestep 3	….	Timestep x	Timestep x+ 1	Timestep x+ 2

In these empty schedules, first, the meeting room meetings are added. This is done for all of the employees, one at a time. This is followed by the assignment of office meetings, which again is done one employee at a time. After this is finished, each employee has a schedule filled with various meeting room meetings and office meetings at scheduled times. A rough example of such a schedule can be seen in Table 11.

Table 11: A rough example of an employee schedule after only the meetings have been assigned

Timestep 1	Timestep 2	Timestep 3	….	Timestep x	Timestep x+ 1	Timestep x+ 2
		Meeting in meeting room 1		Meeting in personal office 3		

After having the employees' schedules filled with meetings, the next step is to know for each employee what time they enter and leave the office that day, to know from and till what time other activities can be scheduled. When this is known for all employees, the time gaps between the meetings can be filled in with the other activities: working in their office, getting coffee or using the restroom. This scheduling of these activities is based on transition matrices, as inspired by the activity scheduling for PeDViS done by Wang (2020). One schedule at a time is taken and processed from start to end, while each time gap gets filled with activities. The choice of the activity is based on the transition matrices each time. The activity a person just came from, determines the probabilities of going to each of the other activities. The transition matrices provide these probabilities, with each row in the transition matrix representing one of the activities a person could just have come from. The columns represent the activities they can pick from to go next. A simplified version of a transition matrix can be seen in Table 12.

Table 12: Set-up of a transition matrix

	Personal office	Meeting room	Coffee	Toilet + sink	Office meeting
Personal office	X	%	%	%	%
Meeting room	%	X	%	%	%
Coffee	%	%	X	%	%
Toilet + sink	%	%	%	X	%
Office meeting	%	%	%	%	X

Then, the choice between the possible next activities to perform is based on the probabilities of each activity relating to that row. When an activity is picked, it is assigned to the schedule for the duration of that activity. Activities keep being picked till the time gap is filled. This eventually results in a fully filled-in schedule, a rough example of which can be seen in Table 13. The example only shows a high-level version of the schedule and does not show that activities can also be assigned for multiple time steps, matching their duration.

Table 13: Rough example of a fully filled- in schedule

Timestep 1	Timestep 2	Timestep 3	….	Timestep x	Timestep x+ 1	Timestep x+ 2	Timestep x +3
Not yet in office	Enter office	Meeting in meeting room 1	…	Meeting in personal office 3	Restroom	Leave office	Not in office

The whole process of the gradual filling in of the employee schedules described is visualized in Figure 13.

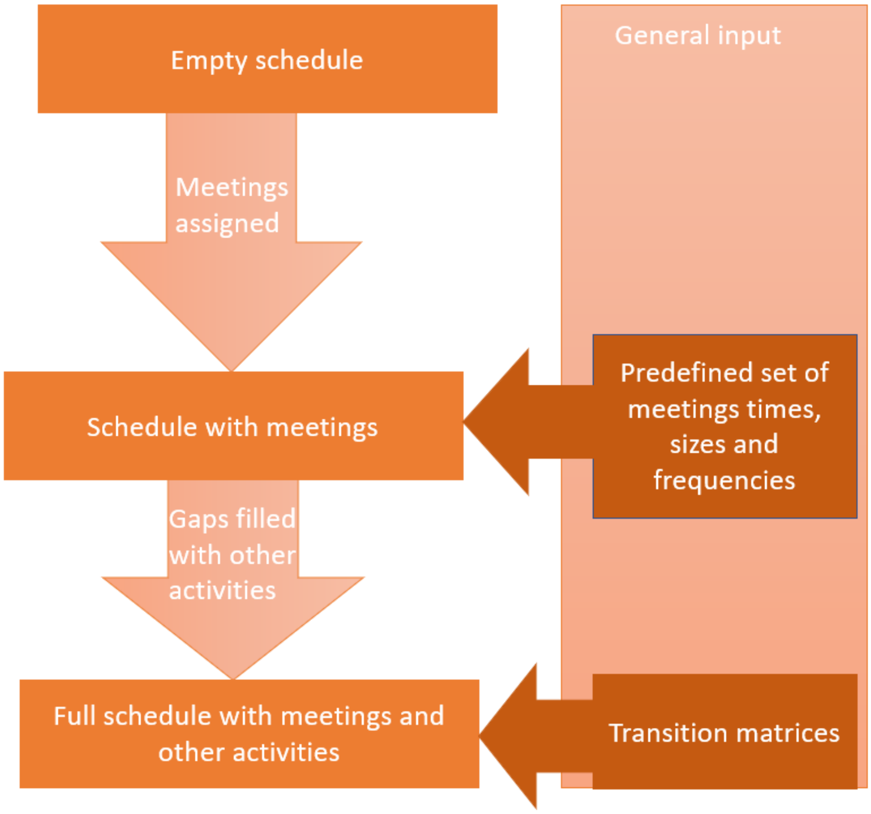

Figure 13: Flowchart showing the activity scheduling of filling in the work schedules

5.2.4 Design concepts

The fourth step consists of describing the design concepts, listed by the ODD protocol, that are present in the model. The design concepts that the ODD protocol lists are: Basic principles, Emergence, Adaptation, Objectives, Learning, Prediction, Sensing, Interaction, Stochasticity, Collectives, Observation. The design concepts present in the model are discussed in this section. The first design concept that should be discussed are the basic principles of the model. These are the modelling of pedestrian movement, the movement of virus spread and the modelling of infection risk, based on the former. Each of these are represented as sub models being part of the PeDViS model.

The second ODD design concept present is emergence. Emergence entails that system-level phenomena emerge from the individual traits of low-level entities (Grimm et al., 2006). In the case of this model, the number of infections overall is what emerges from the model, based on the individual traits of the model such as their initial health situations (infectious or not) and their individual schedule, dictating at what place they are at what time and the amount of virus emitted on each location by individuals. The next design concept present is sensing, which is about the internal agent's (employees in this model) state variables or environmental state variables individuals are assumed to sense/know, and consider in their adaptive decisions (Grimm et al., 2006). In the model, this is relevant in the Nomad part. Individuals have to sense the nature of entities close to them, (whether they are obstacles, pedestrians or destinations). Based on this nature, they can either be attracted or repulsed by them. In addition, they know their own schedule, leading them to the right location, their personal destinations, at the right time. To do this, they also have to know the location of each of the destinations.

Fourth, in the model, the design concept of interaction can be found. This concept focuses on the kinds of interactions that are assumed among individuals. The main implicit interaction of the model is one pedestrian infecting another with the COVID-19 virus. This can happen through various types of interaction, such as inhaling the contaminated aerosols emitted by another person or touching contaminated surfaces.

The design concept of stochasticity can also be found in various places in the model. First, this happens for employees' attendance of meetings, as for the predetermined meetings, the employees that attend are randomly drawn from the population. Second, stochasticity can be found in the order in which employees perform the other activities, as this order is based on the probabilities found in various transition matrices. The length of their activity schedule is also stochastic, as the times at which employees enter and leave the office building are randomly drawn. In the QVEmod part, there are two additional stochastic elements. First of all, the choice of which individual is the infected individual is chosen randomly. Second, whether a specific virus exposure results in infection is also stochastic, as this determination is made based on the comparison of their infection probability against a randomly drawn number. Due to the stochastic nature of the model, replications of the simulation need to be conducted to verify the output.

The final design concept from the checklist that is present is observation, which is about what data is collected for testing, understanding and analysis and how and when this data is collected (Grimm et al., 2010). The PeDViS model can produce different types of output after a whole simulation has run, such as the amount of virus each person inhales over time or how much virus they get on their hand. The amount of virus-laden aerosols, droplets or fomites can also be given for each location in the environment over time. However, in this thesis, the model analysis focuses on the risk of people getting infected. Therefore, only the average number of infections is collected for analysis to test the impact of the NPIs.

6 Model implementation

This chapter represents the model implementation phase, which covers the last three steps of the ODD protocol: the model initialization, the input data and the sub models. The initialization and input data are not covered in separate steps, but are explained together in section 6.1, as they are very intertwined in this model. Next, the sub models relating to the adapted model are provided in 6.2. These steps then lead to a function model, providing an answer to the second sub-question:

How to quantitatively model the behavior of people working in an office and the resulting virus spread accordingly?

6.1 Initialization and data input

The two steps of initialization and data input are jointly described in this section in four parts, roughly relating to the different elements of the PeDViS model consists of: (1) the Nomad part of the model, (2) the activity scheduling part, (3) the QVEmod part and finally (4) the NPIs that are chosen to be incorporated in the model (relating to the lowering of the infection risk).

6.1.1 Nomad model infrastructure and pedestrian parameters

The Nomad model requires three main input files that have to be constructed and initialized: the infrastructure, the pedestrian parameters and the scenario definition. The content of the first two files is discussed here. The third file, the scenario definition, contains the activity scheduling of the Nomad model and is, therefore, discussed in the next section of activity modelling.

First, the infrastructure for the Nomad model is created. The layout used is the same for each simulation, and consists of the environment sketched in section 5.2.1. Infrastructures in Nomad consist of the four types of elements explained in section 5.2.2: *walkable area, obstacles, destinations and sources.* The design for the layout of the office has to be translated into these types of elements to be able to correctly implement it in the Nomad model. Additionally, the exact location in terms of x and y coordinates of each object has to be determined. For this thesis, an infrastructure file for the office was generated using the SamenSlimOpen webtool (Samen Slim Open Project, 2022), which allows the user to design their own layout of an indoor space, which can then be generated into input files for the Nomad model. In the resulting file, adjustments were made by hand to alter some details that could not be defined in the webtool, which can be found in Appendix D. A graphical user interface (GUI) is used to check if the layout design is formatted correctly and whether pedestrians are able to reach all of the destinations. The modelled infrastructure that resulted from this process, and is used for all of the simulations, can be seen in Figure 14. The source, which in the entrance of the office, can be seen on the far left. The exit through which all employees leave at the end of the day can be seen right below it. The full XML file can be found in Appendix E.

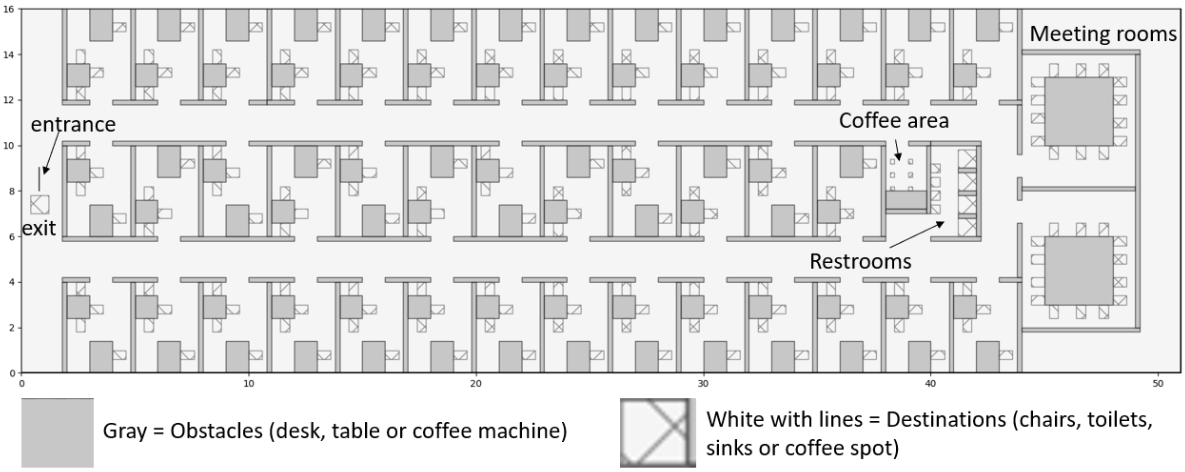

Figure 14: Office layout as Nomad input XML file

The second input file consists of the pedestrian parameter values, which are influenced by different pedestrian behavior assumptions. The values for the pedestrian parameters established for the restaurant version of Duives et al. (2022), which used the work of Campanella (2016), were used in this model. For these values, no 1.5 meter distancing between pedestrians is incorporated, as this is also not incorporated in the model. All pedestrians are modelled to have the same parameter values, except for their desired walking speed, which they draw from a normal Gaussian distribution. The employees get these variable values assigned to them the moment they enter the simulation. The full input file can be found through the link in Appendix E.

6.1.2 Activity scheduling for Nomad

The activity scheduling of the model has to be defined from scratch, due to the office setting of the model entailing different behavior than the restaurant setting of the base model. This activity scheduling is modelled in the scenario file, which is the third input file required to run Nomad. The scenario file states, as the name suggests, which 'scenario' is simulated. For the experimentation phase and model assessment phase, multiple 'scenarios' are used as input for the model. In this section, however, the focus lies on creating a base scenario file.

The first step of the activity scheduling is to define input data and initialization of the activities the pedestrians can perform in Nomad. These activities correspond to *the activities in Nomad* defined in 5.2.3. The scenario file required data on the name, location (corresponding to a location in the infrastructure file), and duration for each activity. First, the different names and locations are specified here in Table 14.

Table 14: Activities defined in the office space

Activity	Destination
Get Coffee	Coffee machine
Meeting in meeting room 1	One of the chairs in meeting room 1
Meeting in meeting room 2	One of the chairs in meeting room 2
Use toilet 1	Toilet 1
Use toilet 2	Toilet 2
Use sink 1	Sink 1
Use sink 2	Sink 2
Meeting in own office meeting area	One of the chairs in meeting area own office
Meeting in other office meeting area	One of the chairs in meeting area other office
Work in personal office	Chair behind desk in own office

The *destinations* are based on the locations specified in the conceptualization. Having further specified the activities for modelling, the activity scheduling can be modelled. The activity scheduling is modelled by writing Python code to generate these schedules. As explained in the concept chapter, each employee has an individual activity schedule in which first the meetings are added. Next, the time gaps between the meetings are filled with the activities that are not meetings, using transition matrices.

Meeting set-ups and adding meetings to the schedules

To incorporate meetings in the schedules, it first needs to be determined how many meetings people typically have and how large these meetings are. The goal is to ultimately provide useful set-ups, relating to the meetings, representing a typical office day to test the various NPIs on. However, to allow for the construction of schedules representing different types of office situations, relating to the frequency and size of meetings, different meeting set-ups are composed related to these values. In the next chapter of experimentation, several of these meeting set-ups are then chosen to serve as set-ups in which the NPIs can be implemented and compared.

To represent the meeting behavior, the two types of meetings defined in the activities have to be considered: meetings in the two meeting rooms and meetings in the meeting area of the personal offices. In the data collection phase, data was gathered on the frequency of these meetings and the number of people present per meeting, which are both important factors in simulating the virus spread. Therefore, to accurately represent the different possible behaviors with regard to meetings in the office spaces, different combinations of values for these factors have been composed. As to not make the scheduling of these meetings too complex, it was decided that set-ups relating to the meetings held in an office, presenting combinations of these values would be composed. To create these set-ups, the values of three variables were varied: the number of meeting room meetings held in an office on a day, the number of people present per meeting room meeting (constant during the day), and the number of meetings held in offices per day. An overview of these variables and the values they can take in the meeting set-ups are given in Table 15. A brief explanation of how these values were chosen is given here. The choice is made to use the maximum, minimum and sometimes average value observed for these variables, because this allows insight into the virus transmission for both a wide range of different office behavior and for a more average/representative type of office behavior.

In these set-ups, with regard to the meetings held in the meeting rooms, the number of meetings and the number of people attending can be either the maximum value, minimum value or average value found in the data collection phase. In the case of the meeting room meetings, the maximum size of the meetings is also equal to the number of chairs in the meeting room. If an average value was not a round number, this number is rounded to the nearest integer, as these variables need whole numbers as input.

For the meetings in personal offices, the number of people present per meeting was not varied in the set-ups. This is done, among other reasons due to the number of people in these meetings only being able to be one of two values: two or three (three is the maximum number of seats and less than 2 people sitting there would not be a meeting). To keep the number of set-ups down, the constant value of three people attending the meetings is chosen, as the average value of 2.5 people per meeting is considerably harder to implement. This way, the maximum amount of potential infection per office meeting based on the number of people joining can be considered. In addition, the number of meetings in personal offices per person per day has been shown to range from 1 being the smallest positive number to 6. However, based on data collection and to be able to fit all of the meetings into the schedule, the value of a maximum of 4 office meetings per day was chosen.

These values of the number of meetings per day per person have been translated to the number of these types of meetings in the whole office per day, which due to all of these meetings consisting of 3 people, translates to a maximum of 53,3 meetings held in personal offices per day, rounded down to 53 and a minimum of 13,33 meetings, which is rounded down to 13.

Table 15: Parameters chosen to be varied in the set-ups, and the values they can hold

	Number of people in a meeting room	Number of meeting room meetings held per day (rounded down to nearest integer)	Number of office meetings per day (rounded down to nearest integer)
Average value	7	3	Not included
Minimum	3	1	13
Maximum	14	8	53

The values in Table 15 are varied in the set-ups, with the addition of one set-up in which no meetings are held during the day. This set-up was added to allow for comparison with a base set-up with regard to the impact of meetings. The combinations formed in the set-ups have been reviewed as to how many meetings a day a person could be having at a maximum. As the highest number of meetings a person had in a day was observed to be 6, which consisted of both personal office meetings and meeting room meetings. Therefore, it was decided that the average amount of meetings an employee has on a day should not be more than 6. For this reason, the set-up with the combination of having the maximum number of people in the meeting room meetings with the maximum number of meeting room meetings per day was deleted, as this set-up would lead employees to have up to 7 meetings a day on average. This led to 18 set-ups ultimately being formed. The set-ups can be found in Table 16, showing per set-up how many of each type of meeting are held per day and how many people visit the meeting room per meeting.

Table 16: Meeting schedule set-ups

Set-up number	Number of people per meeting	Number of meeting room meetings per day	Number of personal office meetings held per day
1 minimum set-up	0	0	0
2	7	3	53
3	7	8	53
4	7	1	53
5	14	3	53
6	14	1	53
7	3	3	53
8	3	8	53
9	3	1	53
10	7	3	13
11	7	8	13
12	7	1	13
13	14	3	13
14	14	8	13
15	14	1	13
16	3	3	13
17	3	8	13
18	3	1	13

As can be seen in Table 16, per set-up all meetings have the same number of people attending. This choice is made to keep the number of set-ups down and to allow easy comparison between the resulting virus spread with different sizes of the meetings. The duration of the meetings is not a variable that is varied in these set-ups, both to allow for easier scheduling and due to limited time to study the effect of these differences. The highest number of meetings held at meeting rooms, 8, which translated to four meetings per meeting room per day, would still allow all meetings to be held on one day, with time left for other types of meetings (office meetings). Finally, the number of employees considered is in all cases set to 40, which is the number of employees working in the office modelled. This value is kept constant in these set-ups, as the effect of varying it is explored in chapter 4, as one of the NPIs.

The next step is to have the exact times of these meetings defined. With regard to the meetings held at the meeting rooms, it is assumed these are equally split between the 2 rooms. This entails that with the maximum amount of 8 meetings per day, 4 are held in each meeting room. In the case of 3 meetings a day, the 3 meetings are split between having one meeting room hosting 1 meeting and 1 meeting room having 2 meetings. In the case of there being only 1 meeting in a meeting room on a day, only one meeting room is used. For the personal offices, in the case of 53 office meetings being held per day, the meetings are equally distributed between the offices, resulting in one or two meetings held in each office per day (53 meetings / 40 offices = 1.3, so one or two). When 13 office meetings are held per day, some offices host one meeting per day while other offices host zero. However, to precisely plan these meetings, their duration also has to be determined. This duration per type of meeting is presented in Table 17. The *real-life duration* per activity is based on the data collection phase. The choice is made to scope the activity scheduling to be per minute, as this provided an ideal balance between accuracy and computing power. Therefore, the durations have been rounded to the nearest positive integer number of minutes. In addition, the *modelled duration* of the office meetings has also been rounded off or up to the nearest number of quarter hours, to allow for easier scheduling. The duration of the meetings in meeting rooms has been shortened into two steps. First, the modelled duration is reduced to 60 minutes, to allow for more meetings to take place within one day. Next, the meeting duration has been reduced even further to 56 minutes, to allow people some time between consecutive meetings, to do something else in between meetings, like get coffee. These deviations from the observed values are expected to not distort the model simulation too much, as the differences are a relatively small reduction in activity duration for activities that both have a long duration compared to the other activities in the model.

Table 17: Real-life durations and modelled duration for the different types of meetings

Activity	Real-life duration (seconds)	Modelled Duration (seconds)
Meeting in meeting room	68 * 60	56 * 60
Meeting in office meeting area	48.18 * 60	45 * 60

Based on the frequencies of meetings per room, the start times of the meetings are determined. In Table 18, Table 19, Table 20 and Table 21, the times at which the meetings take place are given, which differs based on how many meetings a day happen according to the chosen set-up. The columns refer to the specific meeting slot, for example, ‘meeting 4’ being the fourth time slot in which that type of meeting can take place. Only a maximum of 4 meeting slots is needed for both types of meetings. For the meeting room meetings, this is due to the maximum number of meetings being 8, and at each timeslot two meetings take place starting at the same time. For the office meetings, this is because there are at maximum 4 meetings held per person per day.

It is not common for multiple meetings within an office to end at the exact same time in an office. This would also lead to extreme crowdedness, as people often go get coffee together or go to the restroom right after meetings. Therefore, it was decided that both for the meeting room meetings and the office meetings, the meeting time should be set up in a way to not let all meetings of the same type and timeslot end at the same time. The approach taken to achieve this differs per type of meeting (in a meeting room or in the meeting area of a personal office). For the office meetings, this is done by making two shifts per timeslot of meetings. These shifts here take the shape of roughly half of the meetings starting and ending 10 minutes earlier than the second shift. This approach is taken, due to the even higher amount of people going into meetings at the same time slot. Both the start and end times needed to be adjusted to be able to keep the duration of the meetings constant and to be able to really compare the effect of other variables than the duration.

Table 18: Meeting times office meetings early shift

Office meeting times early	Meeting 1	Meeting 2	Meeting 3	Meeting 4
1 meeting per day per person	13:50-14:35	x	X	x
4 meetings per day per person	10:00-10:45	11:55-12:40	13:50-14:35	15:45:16:30

Table 19: Meeting times office meetings late shift

Office meeting times late	Meeting 1	Meeting 2	Meeting 3	Meeting 4
1 meeting per day per person	14:00-14:45	x	X	x
4 meetings per day per person	10:10-10:55	12:05-12:50	14:00-14:45	15:55:16:30

Table 20: Meeting times meeting room 1

Meeting room 1 times	Meeting 1	Meeting 2	Meeting 3	Meeting 4
4 meetings in meeting room	9:00 – 9:56	10:55-11:51	12:50-13:46	14:45-15:41
2 meetings in meeting room	9:00 – 9:56	10:55-11:51	x	x
1 meeting in meeting room	10:55-11:51	x	x	x

Table 21: Meeting times meeting room 2

Meeting room 2 times	Meeting 1	Meeting 2	Meeting 3	Meeting 4
4 meetings in meeting room	9:00 – 9:49	10:55-11:44	12:50-13:39	14:45-15:34
2 meetings in meeting room	9:00 – 9:49	10:55-11:44	x	x
1 meeting in meeting room	10:55-11:44	x	x	x

For the meetings in the meeting rooms, another approach is chosen, because there was not enough time left in a day to take the approach of the two shifts with a 10 minutes gap between them. It is decided that 1 of the meetings should end earlier than the other meeting in that time shift, to still prevent the crowdedness in the coffee and restroom area. Therefore, it is scheduled that the meetings in meeting room 2 end 7 minutes earlier, as based on the model simulations, this is just enough time for everybody from the first meeting to leave the coffee area and restrooms before the second meeting is done, therefore altering the duration of the meeting the minimum amount needed.

The earliest start time of a meeting is 9:00, as everybody is then assumed to be inside the office and the latest ending time is 16:30. All of these meeting times are designed to fit all possible combinations of the number of meetings given by the set-ups and thus to not have overlapping meeting times. Depending on the exact meeting times an employee gets assigned, the time he has between two consecutive meetings can either be just the one minute to walk from one place to another, or a longer break starting from either

5 till up to 15 minutes. This entails that sometimes employees just go straight from one meeting to the next, but often they go to the restroom, get coffee or stop by their office and maybe do some work in between, mirroring the behavior observed in offices.

Next, having defined the activities and the different set-ups regarding the meetings held, the code for generating the schedules could be written. The code is designed to let the user input the number of schedules desired and the set-up they want to implement. Based on this information, first, the requested amount of empty schedules are generated, one for each employee. The details on how these individual schedules are assigned to employees can be found in Appendix F. In this thesis, the standard value of 40 employees is assumed, the maximum number of employees that can be present in the office. Each schedule consists of 540 rows, each row representing one minute of the day, simulating the time from 8:30 to 17:30. The simulation set-up sets the model to run till 17:40. This 10 minute extension is implemented to ensure that all employees have indeed left the office when the simulation ends, as employees are shown to have an up to 8 minute delay on their schedule. The schedule consists of two columns, the first one representing the time, and the second one the activity the person does at that specific time. An example of the schedule can be seen in Table 22

Table 22: Example of empty activity schedule in the code

Time	Activity
1	Not yet in office
2	Entered office
3	Work in personal office
4	Work in personal office
……….	
536	Uses restroom
537	Leaves office, heads for exit
538	Has left office
539	Has left office
540	Has left office

Next, based on the set-up chosen, the meeting room meetings are added to the schedules, followed by the office meetings. This is done by looping through the applicable meeting times for that set-up, and for the corresponding times fill in the meeting for the matching duration. This is done for the number of people that are needed to fill the number and size of meetings selected, who are randomly selected. For the office meetings, one of the three people in the meeting is scheduled to be the employee to whom the office belongs. In the case of meeting room meetings, people randomly pick one of the chairs available in the meeting room. The first person to add this activity to their schedule is able to choose from all fourteen chairs, and each subsequent person has to choose from the chairs that are still available.

The intermediate output during the activity scheduling code is a schedule for each employee filled with various meeting room and office meetings, with a specific chair assigned, in all of the rows corresponding to all of the allocated timesteps.

Putting the remaining activities into the schedule

After having the employees' schedules filled with meetings, the next step taken is to assign each personal schedule with the time that an employee enters and leaves the office. The time they enter is randomly drawn from a uniform distribution between 8:30 and 9:00 and the time they leave is a randomly drawn time between 17:00 and 17:30, in accordance with the timing people usually enter and leave offices based on personal experience and what was seen during the observations. Then, the code fills in the time gaps between the meetings with the other activities of working in their office, getting coffee or using the restroom. First, the durations of these activities are also defined. The durations can be seen in Table 23

Table 23: Real-life duration and modelled durations of the activities, other than the meetings

Activity	Real-life duration (seconds)	Modelled Duration (seconds)
Get Coffee	100	2* 60
Use toilet 1	70	60
Use toilet 2	70	60
Use sink 1	40	60
Use sink 2	40	60
Work in personal office	No observations	Multiple options (1, 30, 60 or 120) * 60.

The real-life durations are again based on the data collection phase, and the modeled durations are the durations of the activities that were eventually incorporated into the model. The durations are again rounded off to exact minutes, due to the scope of the activity scheduling being per minute. In the data gathering process, no 'standard' duration of being in one's office was found due to people usually being in there as a base location until they headed to one of the other activities. Therefore, various durations can be assigned, with the choice of duration depending on the context during the scheduling process and time left till the next meeting. The duration of 1 minute working is assigned in a different way than the other durations of working in one's offices: when the activity of working in one's office is chosen but there are less than 30 minutes left till the next meetings, working in the office for one minute is assigned for each minute till an employee has to go to their meeting.

This scheduling of these activities is based on various transition matrices. The scheduling code in Python code, for each schedule, goes through the entire schedule and fills each time gap with activities and then moves to the next time gap in that schedule. Which activity gets filled in at each timestep, is based on how much time an employee has left before the next meeting and on which activity they just came from. The activity a person just came from determines the probabilities of going to each of the other activities. When an activity is selected, this activity is added to the schedule for the corresponding timesteps. Then the next activity is picked, using the last activity as input for the transition matrix, till the time gap is filled.

For the scheduling, four versions of the transition matrix were used. The full transition matrix showing all activities can be seen in Table 24. As can also be deduced from this table, when a person first enters the building, they always head to their own office first, which could be either to work or for a shorter time to drop off some stuff. When a person has limited time left before their next meeting, some activities are not an option anymore to do before that meeting. If an employee has for example only three minutes left, he does not go to work in his office but could get a cup of coffee or use the restroom. In this case, a reduced transition matrix, shown in Table 25, is used. Also, it should be considered that if an employee has just been to the restroom and then has gone to get coffee for two minutes, he does not go back to the restroom right after. The reverse is also true when going to the restroom after getting coffee. This has been incorporated into the scheduling, resulting in the employees basing their next activity on the adjusted transition matrices In Table 26 and Table 27, after having already been into the restroom or to get coffee in the last 5 minutes, respectively. Finally, if an employee has just finished an activity and has 3 minutes or less left till their assigned time to leave the office, it is assumed they leave the office already. This assumption is made because it allows for easier scheduling and it is not very realistic for employees to then get coffee or work for just 2 minutes, also needing one minute of walking to reach the exit. Both data on the frequencies of order in which pedestrians visit the functional areas and frequencies of visits to the different functional areas has been gathered, on which the probability values in the transition matrix are based. The details of the calculations can be found in Appendix G.

Table 24: Full transition matrix

	Work in office	Coffee	Toilet
Work in office	0	0.5	0.5
Meeting room meeting	0.7	0.15	0.15
Coffee	0.75	0	0.25
Restroom	0.67	0.33	0
Office meeting	0.5	0.45	0.05
Entrance	1	0	0

Table 25: Transition matrix for coffee or toilet choice

	Coffee	Toilet
Work in office	0.5	0.5
Meeting room meeting	0.5	0.5
Coffee	0	1
Restroom	1	0
Office meeting	0.9	0.1

Table 26: Transition matrix office or coffee choice

	Work in office	Coffee
Work in office	0	1
Meeting room meeting	0.82	0.18
Coffee	1	0
Restroom	0.67	0.33
Office meeting	0.5025	0.4525

Table 27: Transition matrix office or toilet choice

	Work in office	Toilet
Work in office	0	1
Meeting room meeting	0.82	0.18
Coffee	0.75	0.25
Restroom	1	0
Office meeting	0.725	0.275

The filling in of the activities in the schedules starts at the assigned time the employee enters the office and ends at the assigned time they leave the office. When employees come into the office, it is assumed they go straight to their office. This assumption is made to simplify the schedule and as because it is expected that this behavior is representative of real life, as people either start working in their office right away or want to drop off their coat and/ or bag before heading towards any of the other activities. In Figure 15, a flow chart is shown representing the progress of how it is determined what activity is chosen to fill *the current* time gap. An employee thus goes through the decision-making process of this flowchart for each time gap between meetings, each time starting at the yellow circle. This flowchart does not include the choice of skipping the activity of getting coffee when the line is too long.

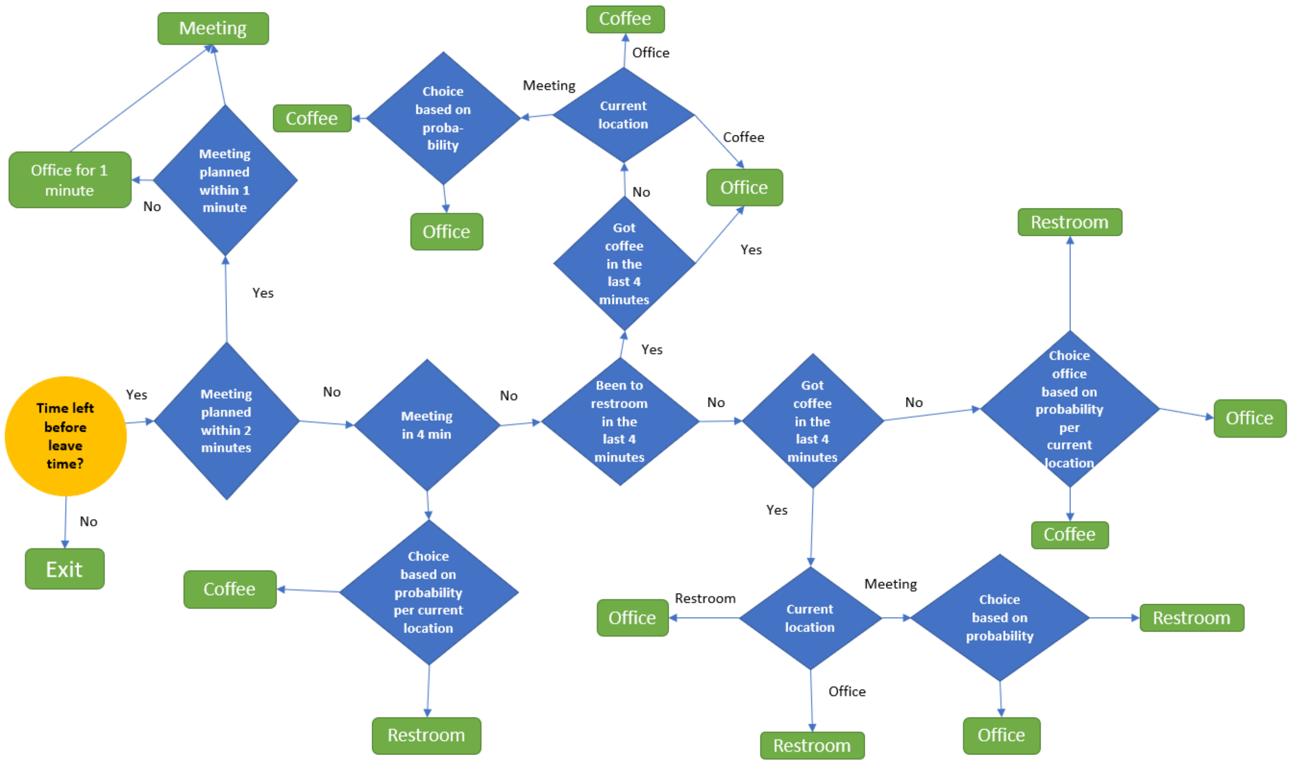

Figure 15: Flow chart of choosing the next activity to perform

During the run, the pedestrians have one personal office assigned, which is always the location where they perform the “work at office” activity. In addition, during an office meeting, this person will always be present if a meeting is held in their office. When going to the restroom, a pedestrian will first check which stalls are already occupied and will then randomly choose one of the stalls that are available.

When being assigned to work at their office, one of the four durations of office work defined for this activity is randomly chosen, based on how much time a person has left till their next meeting. For example, if a person only has 1,5 hours left, the option of working at the office for 2 hours is not taken into consideration. When getting coffee, a person will get right to the coffee machine if the coffee area is empty. However, if someone is already getting coffee, he will go stand next to him and possibly have small talk. The same happens for the next people entering the coffee area when two or more people are already getting coffee, and they will just join behind them. This process has been simplified, in that people then just spend the two minutes assigned to the activity of getting coffee in that one place, to allow for it to be more correctly scheduled. This can be assumed to be mirroring real-life behavior, as often when many people are present at the coffee machine at the same time, this is because they have come from the same meeting and then often one person makes the coffee for the other people and hands it to them. Also, the infection risk should be roughly the same as when everyone got their own coffee while still waiting for each other in the coffee area, as fomites only play a small role so touching the coffee machine is not the most important aspect.

It can also be seen in visualizations of the model, that the coffee area is notably busier right after one or more meetings have ended. The reason for this is that people often go by the coffee area after meetings, usually accompanied by one or more people from their meeting. Therefore, this behavior is also deemed to mirror real life well. In cases there are already 6 people in the coffee area, and one or more people are also planning to get coffee, it is scheduled that the next person then deems the line for coffee to long and chooses another activity (go to office for a certain duration or to the restroom). This is done not to let the coffee area get too crowded, as this would both causes people to stand on top of each other more as it would not be very realistic to have that many people wanting to wait for coffee. The activity schedules do not include the activity of walking from one location to another, both because activities are stationary in Nomad and because there is no defined time for it, as the duration depends on the distance between the locations of the first activity and the next and an employee's personal walking speed. If a location is close to the previous one, it can take only a few seconds, while some of the longest and busiest distances (from an office on the far left to a busy meeting room on the right) can take up to 75 seconds. As it is important that the times assigned in the schedules match the timing of an activity in Nomad, among other things to ensure all people are at the meeting at the same time or people can accurately check if a restroom is occupied at that minute, the time an employee spends walking from one place to another also has to be incorporated.

The walking time is no set value, as it depends on both the distance that needs to be crossed from the previous location to the next and the walking speed of the employee. For these reasons, there is also some randomness in the exact time people arrive at their meeting, with some being a bit early and others a bit late. This simulates the behavior of people arriving at a meeting more realistically than having them all arrive at the same second. As this randomness is present after each activity an employee performs on a day, this unaccounted randomness in an employee's timing can add up during the day, if his walking time is for example mostly very short or very long, instead of balancing itself out with both long and short walking times. However, the activity scheduling is set up to be per minute and the longest walking time from one location to be next was found to be one minute. Therefore the walking time is not implemented to be based on distance and speed, as these would lead to more detailed walking times per second needing to be implemented. For this reason, the walking time incorporated is finetuned based on the several simulations that could be visualized in the GUI. Trying different configurations, finally the walking time is included by incorporating one minute of walking time after some of the activities, and zero after other activities. The details of this can be found in Appendix F. This configuration led to the most fitting average walking time and thus the best results in terms of how well people followed their schedule time-wise, with people deviating up to 8 minutes from the schedule. All other configurations caused larger deviations from the schedules. It was also observed that despite this cumulative randomness, people still generally leave between 17:00 and 17:30 and therefore it can be concluded that overall the walking time of 1 minute balances itself out quite nicely during the day.

Finally, the complete schedules are incorporated into the Nomad model, by assigning them to the right employee once they enter the building, with the time they enter the office in Nomad being equal to the time they enter the office according to their schedule. The activity schedules thus differ per set-up chosen, but throughout the use of one set-up, the same activity schedule is used per employee to allow for comparability. The full activity scheduling code, some additional explanation in the form of comments and the input file for Nomad can be found through the GitHub link in Appendix E.

Verification of the activity scheduling

In this section, the verification of the activity modeling produced is described. Verification is only performed for this activity scheduling part, as the model is based on the existing PeDViS model, of which the Nomad, QVEmod and infection risk parts have already been calibrated by others, such as Duives et al. (2022). As the activity scheduling is the only part that is new in addition to the existing PeDViS model and has not been tested yet, verification is only applied to this part. Therefore, some verification is performed to check if the Nomad part of the model performs as intended using the activity scheduling input and thus corresponds to the conceptual model (Wilensky & Rand, 2015).

First, it was checked in the python code if each set-up produced the right number and size of the meetings. Also, some of the schedules, randomly chosen, were checked to see if they looked as expected: no empty spaces, no unexpected sequences, the activities being scheduled for the right amount of time and people always working in their own office. Finally, due to the adjustments in place to prevent the coffee area from getting too crowded, it was checked how often people initially wanted to get coffee, but then decide it is too crowded and go to another activity. It was also checked which people this happened to, and how often it happened per person. Next, a graphical user interface (GUI) showing the walking behavior of the pedestrians was used to check the employees' walking behavior in the Nomad model.

Using the GUI, several checks were performed:

- The general movement of a whole day in several set-ups was checked, to see if all walking behavior mirrored real-life behavior, such as people not standing on top of each other or in places they are not supposed to stand.
- Multiple times, one employee's movements and activities were followed and compared with their schedules, to check whether their activities and the timing of their activities match their schedule
- For multiple set-ups, it was checked if the meetings started and ended at the expected time, that were assigned in section 6.2, and whether these meetings indeed contained the right number of people. This was already checked in the code, but also again here in the GUI, to check if the implementation of the schedules in Nomad provided the expected result.
- It was checked if everyone entered and left the office in the designated timeframes, entering between 8.30 and 9:00 and leaving between 17:00 and 17:30.
- For each set-up, the general walking behavior of the whole day was globally checked, to make sure no expected situations happen, such as people getting stuck or getting extremely off of their schedule

6.1.3 QVEmod

For the QVEmod part of the model, there are two main elements that need to be defined: the infrastructure and the QVEmod parameters.

Infrastructure and touch frequency

First of all, in addition to the infrastructure file, that defines the office layout for the Nomad part of the model, a layout has to be made for the QVEmod part of the model. The SamenSlimOpen webtool is used to make this file, as doing it by hand again would be too time-consuming for this type of file. The input that is required consists of the objects in the layout that either can be contaminated or have an effect on the spread of aerosols/droplets. Therefore, some of the furniture (chairs, tables, the coffee machine) and the walls have to be defined in the model. For the furniture included in the model, the material of the furniture also has to be chosen when setting up the layout in the webtool. The four materials that can be chosen in the webtool are: wood, glass, plastic and cloth. The material of an object in turn affects several parameters related to the virus transmission. An overview of the elements put in the file and their chosen material can be seen in Table 28.

Table 28: Material each of the types of objects in the model consists of

Object	Material
Desk	Plastic (due to the keyboard being the most touched part)
Meeting table in personal offices	Wood
Meeting table in meeting rooms	Wood
Chairs	Wood
Coffee machine	Plastic

In addition, the frequency with which each object is touched needs to be defined. The touching frequencies were chosen based on literature or calculations based on assumptions on behavior. An overview of the touching frequencies per type of furniture can be found in Table 29.

Table 29: Level of touch per surface/object

Surface (object)	Touching frequency per hour
Personal desk (keyboard)	29 (Zhang & Li, 2018)
Meeting table in personal offices	15 (Lei et al., 2020)
Meeting table in meeting rooms	15 (Lei et al., 2020)
Chairs	3 (Zhang & Li, 2018)
Coffee machine	105

For the touch rate of the personal desks, a touch rate of 28.98 per hour of one's keyboard found by Zhang & Li (2018) is taken as a proxy, as this can be assumed to be the object office employees touch the most at their desk and is also assumed to be something present at each office, in accordance with what was observed during the data collection phase. However, as the touching frequency can only be defined as an integer in QVEmod, the touch rate has been rounded off to the nearest integer of 29.

The touch rate of meeting tables, however, is based on Lei et al. (2020), which focuses on a restaurant setting. This choice is made because the touch rate of 15 touches per hour suggested by Lei et al. (2020) seems more fitting than the lower touch rate for desk/tables suggested by Zhang & Li (2018), which assumes there are objects such as keyboards and a mouse present on the table, which are then the most touch objects, which is not the case for the meeting tables. The touching frequency of the chairs was determined to be 3.17, based on the work of Zhang & Li (2018). This value is again rounded off to the nearest integer of 3.

The touching frequency of the coffee machines is based on the information that the coffee machine has to be touched three times, twice clicking the order and once grabbing the coffee. Combined with the information that it takes 1.71 minutes to get coffee according to the data collection done, this is calculated towards a touch rate of 3 * (60/1.71) = 105.26 per hour. This is then rounded down to 105 touches per hour.

The restrooms are a part of the layout for the QVEmod part of the model, as the structural objects in the restroom area influence the transmission of virus-laden aerosols and droplets. However, the sinks and toilets are expected to play a limited role in the transmission of the virus through fomites, because people are assumed to wash their hands right after, so the additional amount of fomites they get on their hands from using the restroom will only be on there for a very short amount of time and can therefore be neglected. To incorporate this in a simplified way, the touch rate for the sinks and toilets is set to 0. Each of the values for both the material and touch-rate are set before simulations begin, and do not change during a run.

The touching frequencies for the meetings tables are the same as the values used in the PeDViS model of Duives et al., (2022). The other values have been implemented as is and have not be adjusted in the model based on additional sensitivity testing.

QVEmod parameter values

For several parameters in the QVEmod part of the model, that have been shown in 5.2.2, a value needs to be defined before simulation. For the original version of PeDViS focusing on a restaurant (Duives et al., 2022), these values have already been defined. Upon reviewing these parameters and their values, it is concluded that these values should not differ in an office setting, with a few exceptions, including the earlier defined touching frequency. Therefore, the same values have been used as in Duives et al (2022), in which the same values are assigned to each employee. These parameters, their values, explanations and their sources can be found in Appendix H. The two other parameters that are given different values are the mask filter efficiency and the air change rate, as these are influenced by the NPIs incorporated and the compliance of the employees. Therefore, these values are determined in the next section on NPIs. Each run with the model start with all of the parameters related to the amount of virus in the environment, the amount of exposure for all individual and the amount of virus on individuals' hands set to zero. Additionally, for each run one employee is randomly selected to be the infectious individual, being the source of the virus spread.

6.1.4 NPIs

Having the model more defined, the NPIs can now be specified. Based on desk research, different possible NPIs for office settings were considered and then reviewed with the help of a multi-criteria table. The details of this selection process can be found in Appendix I. The final selection of NPIs that are included as options in the model to put into practice are :

- Having all employees wear a mask while walking
- Increase the rate of cleaning of all contaminated surfaces, such as the tables, coffee machine and restroom sinks from once a day to twice a day
- Increase the ventilation rate from standard rate to the rate advised by government officials
- Limited office capacity: Allowing only 50 % of the employees to come into the office, others work online
- Allowing a maximum number of people in the meeting rooms, equal to 50 % of the capacity of the meeting room

Each NPI can be individually incorporated into the model at the start of a run, to be able to test the impact of each NPI on the virus transmission separately or in combination with other NPIs. Some of the NPIs are already included in the original PeDViS model, such as the ventilation rate, cleaning contaminated surfaces and wearing a mask. For these NPIs no modelling adjustments are performed to be described here, and the exact values to be chosen for them are discussed in the model assessment chapter. The number of people allowed at the office is something that can be easily simulated with the activity scheduling code made. A user can just fill in the Python code how many schedules need to be made, which is equal to the number of employees. The code then provides the right input for the Nomad model to run the correct simulation. The maximum capacity for meeting room meetings can be changed by setting the value for the maximum meeting room capacity NPI to be on in the Python code.

Compliance

However, perfect compliance with all of the NPIs cannot be assumed. For each NPI, different levels of compliance are incorporated, an overview of which can be seen in Table 30. First of all, the effectiveness of the rule to wear a mask depends on whether and how correctly people wear them. Second, The increase in the ventilation rate depends on whether people indeed open windows and doors. Third, the maximum number of people allowed depends on whether people come to the office despite being asked not to come that day. Therefore, for these three NPIs the expected level of compliance is incorporated by adjusting the effectiveness of the NPIs. For the cleaning rate, this is assumed not to be influenced by compliances, as this cleaning is performed by an external factor in the model, the cleaning crew, of which the compliance is not considered here and who are expected to follow the instructions they are paid to do. Also, full compliance with the maximum occupancy of meeting room meetings is expected, based on two reasons. First, it is hard to disobey this rule due to the monitored booking system of meeting rooms offices often have. Second, there is a social control of at least some of the people in the meeting wanting to adhere to the rules.

Table 30: Incorporation of compliance for the NPIs present in the model

NPI	Incorporating compliance by
Wearing a mask while walking	Cannot have some wear and some don't, so do less efficiency representing some people not wearing it correctly
Increased rate of cleaning	Not impacted by compliance
Increased ventilation rate	Typical level not altered, but higher level is.
Maximum number of people allowed	Value adjusted based on compliance
Maximum occupancy of meeting room meetings	Not impacted by compliance

Numerical values

For each NPI, there are two options: no extra measures or a prevention option, in which a measure is taken. This way, an NPI can be turned on or off in the model. The exact numerical values of the two options per NPI are shown in Table 31

Table 31: Numerical values modeled for the NPIs

NPI	No extra measures option	Prevention option
Wearing a mask	No mask wearing	Mask required while walking, with filter efficiency of mask as 31.6 %
Cleaning of contaminated surfaces	Only cleaned before the start of the day and at the end (every 9 hours)	Cleaning halfway during the day (every 4.5 hours)
Ventilation rate	3.0 air changes per hour	5.04 air changes per hour
Maximum number of people allowed at office	40	26
Maximum occupancy rate for meeting rooms	No maximum occupancy rate	Maximum occupancy rate of 50 % for meeting room meetings, meaning maximum of 7 people in meeting rooms. No adjustments for office meetings

A description of the values for the NPI options is given here. First of all, for the mask-wearing NPI, the no extra measures option is people not wearing a mask at any moment. This is compared with people wearing a mask while walking. The filter efficiency of the masks is assumed to be 40 %, based on the value used by Duives et al (2022). However, the compliance of mask-wearing was found to be 79 %, according to the most recent data (Dutch Central Government, 2021a). In the model, it is not possible to have only a part of the people wear a mask. Therefore, the compliance is incorporated in the value of the mask filter efficiency, which is consequently lowered to 31.6 %, to incorporate this impact. This value is incorporated from both the aerosol and droplet filtering, with the assumption that masks block most droplets and the filter efficiency for droplets is therefore ten times higher than that of aerosols (Asadi et al., 2020), based on which the filter efficiencies are assigned accordingly for aerosols and droplet filtering.

For the cleaning, the no measure option is set to comply with the guidelines set by the Dutch National Institute for Public Health and the Environment, of cleaning certain areas at least once a day (National Institute for Public Health and the Environment, 2019). At the start of a simulation, no virus is present on surfaces and the office can be considered "cleaned" before opening. As the model is designed to set a rate of per how many hours all surfaces get cleaned, this value is set to every 9 hours for the standard situation, meaning it only gets cleaned at the end of the days, and thus only once a day. For the improved cleaning situation, all surfaces also get cleaned at the middle of the day, setting the cleaning to once every 4.5 hours. No higher value of cleaning is chosen, as it entails all surfaces get cleaned, and thus having a higher frequency of cleaning would lead to a lot of interruption in real-life for people working at their desk that then has to be clean or being in a meeting while the table has to be cleaned etc. The 'normal' ventilation rate is set to 3 air changes per hour, in accordance with the ventilation rate often found at indoor spaces (CIRES, 2020). As this is a value based on measurements, instead of the value being based on advice or assuming perfect compliance, the effect of compliance does not need to be added to this value anymore. The higher ventilation is set to 5.04 air changes per hour. This value is derived from the indoor air change rate of 6, based on the recommendation and regulations set by the government (Ministry of Social Affairs and Employment, 2018; National Institute for Public Health and the Environment, 2020), but with incorporating the impact of compliance with this measure. The most recent government data found that 68 % of people comply with ensuring improved ventilation (Dutch Central Government,2022d). Incorporating this compliance with the cooperation needed to raise the air change rate per hour from 3 to 6, results in the air change rate of 5.04.

Next, the options for the number of people allowed at the office are determined. In the option with no interventions, all employees are allowed/requested to come to the office, meaning all 40 employees are present. For the intervention scenario, the previous government guideline of having people work at home half of the time is taken (Ministry of General Affairs, 2022b), which would entail 20 employees are present. However, the Dutch compliance with working from home was found to be 70 % (Dutch Central Government, 2022e). Therefore, the amount of people staying home is assumed smaller, and the number of people in the office is 26.

When this NPI is implemented, the meeting set-ups are not altered, meaning the same amount of meetings in both office and meeting rooms are held per day. The meetings also keep the same group size as much as possible. However, the largest meeting room meetings have to change to consisting of 13 people instead of 14, because otherwise it would not be possible to have two meetings at the same time with 26 employees. Keeping the set-ups largely the same, but having less employees present, leads to the employees present having roughly 50 % more meetings on a day than they would normally have, with the exact amount of meetings depending on the set-up. This increase the number of meetings an employee has per day is assumed to be similar to behavior in real life offices where limitations on the number of days employee are allowed at the office are in place, as it is expected people prefer to schedule their meetings to be mostly on the days they are at the office. Hybrid meetings and people in the office having meetings only are not considered in this NPI.

There is no maximum occupancy rate for the meeting rooms in the no measures option. The meetings are in that case as full as assigned by the set-up used. However, if the NPI of having less people in the office is present, due to only 26 employees being present, the maximum meeting size used is then 13 for both meeting rooms instead of 14, due to not enough people being available. In the option with measures, the 'office rule' is that there is a maximum occupancy rate of the meeting rooms. In this office layout, this entails that a maximum of 7 people are allowed at the meeting rooms. It is assumed that if meetings are bigger than 7, some people would have to join the meeting online, resulting in hybrid meetings. In the model, this is added by having people simply not join in the meeting room, which results in these people being in their office most of the time, due to the long length of this activity and the short duration of getting coffee or using the restroom. The office meetings are not included in this rule as this rule cannot easily be implemented and monitored in an office, in addition to being less risky due to the small size. As compliance is assumed not to influence this NPI, the value of people present in the meeting rooms can be set to a maximum of 7.

6.2 Sub models

The final step of the ODD protocol consists of describing the sub models, that represent the processes in this model (Grimm et al., 2006). An important aspect of this description is the parameterization of the model, which can be done by providing the mathematical skeleton of the model. Such a skeleton can consist of the model equations, rules and one or more tables presenting the model parameters. As the PeDViS model consists of multiple models (QVEmod, Nomad, infections and activity scheduling), the mathematical skeleton exists for each of these model. However, for all of these models except for the activity scheduling, the sub models in the form of mathematical equations made for the original restaurant version of the model from Duives et al. (2022) were used. As these equations were not adapted and are thus not a part of the work of this thesis, these are not described in this section. The mathematical frameworks for the Nomad, QVEmod and the infection risk calculation models can be found in Appendix C. Only the sub models of the activity scheduling are, therefore, presented in this chapter. This is done on a high level of abstraction, as the full code on the scheduling can be found on GitHub, through the link in Appendix E. To present the skeleton of the activity scheduling, first Table 32 is presented showing a selection of important parameters used in the activity scheduling.

Table 32: Parameters used in activity scheduling

Parameter	Explanation
Tick	Current timestep
time_left	Time (number of ticks) a person has left before they leave the office or have their next meeting start
set-up	The set-ups with regards to the meetings in an office, specified in this chapter
max_people_meetingroom	Maximum amount of people allowed in the meeting rooms
allowed_capacity_meetingroom	Percentage of maximum capacity allowed in meeting room
capacity_meetingroom	Number of chairs in meeting room, 14 in both meeting rooms in office layout
nr_of_employees	Number of employees present in office
mr_meetings_day	Number of meeting room meetings held in the office on a day
nr_people_meetingroom	Number of people present for each meeting room meeting
meeting_room	In which meeting room in the office meeting is held, meeting room 1 or 2
size_office_meetings	Number of people present for each office meeting
number_office_meetings_shift	Number of office meetings per office meeting shift
nr_people_officemeeting_day	Number of office meetings held in the office on a day
office_meetings_person_day	Number of office meetings a person has on a day
meeting_shift_meetingroom	Timeslots in which meeting room meetings are held
meeting_shift_office	Timeslots in which office meetings are held
begin_time_meetingroom_meeting	Tick at which an employee's next meeting room meeting starts
end_time_meetingroom_meeting	Tick at which an employee's next meeting room meeting ends
begin_time_office_meeting	Tick at which an employee's next office meeting starts
end_time_office_meeting	Tick at which an employee's next office meeting ends
activity_duration	Duration of the activities of going to the restroom, sink, getting coffee or doing office work
leave_time_employee	The tick at which an employee leave the office at the end of a workday
available_coffee_spots	Which of the 6 spots to get or wait for coffee are not occupied
closest_coffee_spot_available	The available coffee spot closest to the coffee machine, with 1 being the closest and 6 the furthest away. If all occupied is 'none'.
available_restroom_stalls	Restroom stalls not occupied at current tick
next_meeting	The next meeting an employee has, given the current timestep
available_activities	Activities an employee can choose from to have as their next activity
next_activity	The next activity an employee chooses to do at the next tick
current_activity	Activity the employee is doing in the current timestep
probabilities_of_activities	The probability of each activity being chosen, based on the current activity
chosen_chair_meetingroom	Which chair a person picks to sit on during a meeting in a meeting room
available_chairs_meetingroom	Chairs in a meeting room that are not picked to sit on by another person yet
location_office_meeting	Where office meetings is held, in own office or the office of someone else
chosen_chair_office_meeting	Which chair a person picks to sit on during a meeting in an office

Having an overview of the most important parameters, some of the most important rules and equations for the activity scheduling are given. The activity scheduling is not based on extensive mathematical equations, but on coding, with the exception of the equations used to assign people to meetings. Therefore, first some of the high-level rules used for the activity scheduling are given, after which the equations used to schedule meetings are provided.

Rules:

- Initial values of *mr_meetings_day*, *nr_people_meetingroom* and *office_meetings_person_day* based on chosen *set-up*
- *available_activities* = all activities complying with *activity_duration* <= *time_left,* exception for the activities of going to the restroom or getting coffee, if that activity has already been done in the last 5 minutes
- *probabilities_of_activities* = transition matrix value, with row of transition matrix based on the *current_activity*
- *next_activity* = random choice between the *available_activities*, based on *the probabilities_of_activities*
- *time_left*= minimum of (*leave_time_employee - tick*, *next_meeting - tick*)
- *begin_time_meetingroom_meeting* = choose right start time from python dictionary, given current *meeting_shift_meetingroom* and *meeting_room*
- *end_time_meetingroom_meeting* = choose right start time from python dictionary, given current *meeting_shift_meetingroom* and *meeting_room*
- *begin_time_office_meeting* = choose right start time from python dictionary, given current *meeting_shift_office*
- *end_time_office_meeting* = choose right start time from python dictionary, given current *meeting_shift_office*
- *available_chairs_meetingroom* = chairs in *meeting_room* – all *chosen_chair_meetingroom* already selected by other employees
- *chosen_chair_meetingroom* = randomly choose one of *available_chairs_meetingroom* → mark as occupied and remove chair from *available_chairs_meetingroom*
- *chosen_chair_office_meeting* = (when location_office_meeting = own office) → chair 1
- when *location_office_meeting* = someone else's office → randomly pick between chair 2 and chair 3
- *closest_coffee_spot_available* = minimum(*available_coffee_spots*).

When getting coffee:

- When *available_coffee_ spots* is a value between 1-6 → go to *closest_coffee_spot_available*
- When *available_coffee_ spots* is a value is 'none' →change next_activity from coffee to another one of the a*vailable_activities*

When going to restroom:

- Randomly choose one of *available_restroom_stalls*

Equations

- *max_people_meetingroom = allowed_capacity_meetingroom* * *capacity_meetingroom*
- *nr_people_meetingroom* = min(*max_people_meetingroom*, initially set *nr_people_meetingroom*)
- when (*nr_of_employees* / 2) < *nr_people_meetingroom*
  → *nr_people_meetingroom = nr_of_employees*/2
- *number_office_meetings_shift = nr_of_employees*// *size_office_meetings* (rounded down to integer value)
- *nr_people_officemeeting_day = number_office_meetings_shift * office_meetings_person_day*

7 Experimentation

In this chapter, a sensitivity analysis of the influence of the different meeting set-ups on virus transmission is performed, to answer the third sub-question:

What is the influence of the parameter values related to the part of activity scheduling focusing on meetings on the virus spread?

A sensitivity analysis is an exploration of how sensitive a model is to a particular initial condition, in this case, the set-up of the meetings, and gives a better understanding of the sensitivity in the model (Wilensky & Rand, 2015). This way, the impact of the differences between the set-ups can be explored, after which several set-ups are chosen upon which to base the policy analysis for the NPIs. The set-ups, containing different amounts and sizes of meetings, were composed in the previous chapter based on the data collection done in this thesis.

As the work schedules of the employees are based on the data collection and modelling of this thesis, not much is known yet about their influence on virus transmission. Therefore, it would be interesting to learn more about the influence of the activity scheduling on the number of infections, focusing on the meeting aspect. As meetings are the activity of the largest duration that are done in a group setting, in addition to having the most people in one space, it is expected to have a significant influence on virus transmission, which is the reason this part of activity scheduling is chosen to focus on. Finally, the choice to perform the sensitivity analysis only with the parameters relating to a part of the activity scheduling, is also made because the parameter values used for Nomad and QVEmod have already been tested in Duives et al. (2022). For these reasons, the average number of infections per set-up is retrieved through simulations, which are then compared in a sensitivity analysis. In 7.1, the number of simulations performed is given and explained. In section 7.2 the sensitivity analysis is performed and insights based on this analysis are described. Finally, in 7.3, two set-ups are chosen to perform the policy analysis in Chapter 8, based on the results found in this chapter.

7.1 Simulations and statistical tests approach

To get the average number of infections per set-up, multiple model simulations need to be performed per set-up, due to the stochastic nature of both the Nomad and QVEmod parts of the model. To keep the number of replications and the run time needed down, the choice was made to run each simulation with one run of the Nomad model, but with multiple runs of the QVEmod model. This entails that for each run, the activity scheduling and paths of the employees are the same, but that each run a different employee is the infected one. For all of these runs, the same random seed was used, as this way, per set-up the same employees were chosen to be the infectious individual as in the other set-ups, allowing for better comparison between set-ups. The employees to be infected each run are randomly selected, and it is checked whether no employee was picked twice. It is observed that the employees picked are quite diverse in terms of including both employees that came in early, late and in between, allowing a representative number of infections per set-up. The number of runs with the QVEmod part is set to fulfill the minimum amount of runs needed to tackle the stochasticity of the resulting output, by using the formula of Truong et al. (2015), which can be seen in Figure 16.

$$n = [\frac{z_{\alpha/2} S}{E}]^2$$

Figure 16: Formula for required number of runs (Truong et al. (2015)

- *N is the number of required runs*
- *S is the standard deviation of the sample*
- *E is the desired margin of error.*
  This value is calculated by the equation of the mean of the sample x the desired percentage error.
- *$Z_{\alpha/2}$ is the critical value of the normal distribution for α/2, with the desired confidence interval being 100 (1- α)*

A percentage error of 10 % is chosen, as this value provides a good balance between keeping the number of runs down to an amount that could be run within the time of this thesis research and focusing more efforts on the policy analysis, while still providing some certainty in the outcome of the results. Using this formula, each set-up is run at least as many times as the formula prescribed. This resulted in each policy being run 10 times, as this was roughly the number of runs needed per scenario. It was then decided to make the number of runs the same for each policy as only few additional runs were needed and because a universal number of runs ensures better equity of variance, one of the requirements for the statistical tests. All this results in various values for the number of infections per set-up, each representing a different employee being the infected one. Finally, per set-up the average number of infections could be collected, based on different values for the number of infections. As the resulting number of infections per set-up determines the choice of what set-ups are chosen to test the impact of NPIs, none of the NPIs are implemented in these simulations. To these two set-ups, the policies containing the different combinations of NPIs are applied, and the average number of infections resulting from them are retrieved.

Having collected the number of infections for each set-up, statistical tests can be performed on these values. Both independent t-tests and Ordinary Least Squares (OLS) Linear Regression analyses are performed for the sensitivity analysis. First of all, independent t-tests are performed comparing the average number of infections between each combination of two set-ups. This way, it can be tested if the values significantly differ from one another, and consequently what the effect is of changing one or more of the parameter values. The independent t-test was chosen as opposed to the dependent t-test, as the groups of employees representing each set-up are not the same, as in each set-up people have a different activity schedule that defines them and the set-ups do not influence one another but are created independently.

In addition to the t-tests, an OLS Linear Regression analysis is performed, to learn more about the effect each individual parameter relating to the meetings has on the number of infections. This way, the effect of each parameter on the number of infections can be analyzed, separate from the possible influence of the other parameter, for all set-ups (Te Grotenhuis & Van der Weegen, 2013). The coefficients resulting from this analysis show how much the number of infections changes if the value of that parameter changes. Finally, the Adjusted R-Squared value is retrieved, to learn about the proportion of the number of infections that is explained based on meeting parameters. The Adjusted R-Squared value is taken to compensate for the multitude of independent variables (Investopedia, 2021). More information on how the runs were set up in the model simulator, the compliance with the requirements to do these statistical tests and the numerical results of these analyses can be seen in Appendix J.

7.2 Sensitivity analysis

The average number of infections per set-up, found in section 7.1 can be seen in Table 33. Each run starts with one infectious individual, which is not counted into the resulting number of infections. Therefore, the average number of infections is equal to the Reproduction Value (R-Value) for these simulations. The R-value shows how fast a virus is spreading and how many people are infected by one person carrying the virus on average (Dutch Central Government, 2022f). This value is an important KPI used by the Dutch government to keep track of the state of the COVID-19 spread (Dutch National Institute for Health and Environment, 2022)

Table 33: Meeting schedule set-ups, with darker shades of orange signifying a higher number of infections

Set-up number	Number of people per meeting	Number of meeting room meetings per day	Number of personal office meetings held per day	Number of infections/R-value
1 minimum set-up	0	0	0	0.335
2	7	3	53	1.169
3	7	8	53	1.498
4	7	1	53	1.061
5	14	3	53	1.596
6	14	1	53	1.248
7	3	3	53	1.241
8	3	8	53	1.073
9	3	1	53	1.088
10	7	3	13	0.684
11	7	8	13	0.639
12	7	1	13	0.588
13	14	3	13	1.101
14	14	8	13	2.014
15	14	1	13	0.664
16	3	3	13	0.605
17	3	8	13	0.611
18	3	1	13	0.576

Based on Table 33, some preliminary observations can be made. Generally, the set-ups in which the maximum value for the number of office meetings is used, have a higher number of infections than the set-ups in which the minimum number of office meetings is used. There are two exceptions to this. First of all set-up 13 has a higher number of infections than 37.5 % of the set-ups with a maximum amount of office meetings. The second exception is set-up 14 having the highest number of infections overall. Both set-up 13 and 14 have the maximum amount of people attending the meeting room meetings, with the average or maximum value chosen for the frequency of the meeting room meetings. Another observation is that the two set-ups with the highest number of infections, set-up 14 and set-up 5, both have a meeting room meeting size of 14 people. This hints at the importance of the size of the meeting room meetings.

To get a better understanding of the relationship of the parameters on the infection risk, the results of the statistical tests are reviewed. First, the adjusted R-squared value is assessed to learn more about to what extent the meeting parameter explored in the analysis influence the number of infections. The adjusted R-squared value and its significance can be seen in Table 34. The adjusted R-squared value found indicates that 64% of the observed variance in the number of infections can be explained by the meeting variables, which can be stated with some certainty due to its significance (p-value below 0.05). However, as the adjusted r-squared value is not 100 %, not all of the variation in the dependent variable is explained by the meeting variables. Other factors than the meetings in an office are thus expected to still play a large role in the number of infections. This could, for example, be other aspects of the employees' work schedules that differ between the different set-ups, such as getting coffee or using the restroom.

Table 34: Adjusted R-squared value for all meetings parameters influencing the number of infections

Adjusted R-Squared	Significance
0.639	0.000559

To learn more about the individual impact of each meeting variable on the number of infections, the regression coefficients are also assessed. The statistically significant coefficient for each of the variables can be seen in Table 35.

Table 35: The coefficient, value range and change in the number of infections for each of the parameters

Variable	Coefficient	Value range	Coefficient * value range
Number of people in meeting room	0.0464	0-14	0-0.6496
Number of meeting room meetings	0.0488	0-8	0-0.3904
Number of office meetings	0.0110	0-53	0-0.583

The values for the coefficient can be interpreted the following way: for the parameter of the number of people in the meeting rooms, for each additional person being in a meeting room meeting, the average number of infections rises by 0.0464 infections. Even though the variables of the size and frequency of meeting room meetings seem to be the most important at first glance, due to their higher coefficients, the value ranges of the variables are also very important. The number of meeting room meetings has the highest coefficient and the infection risk could thus be indicated as more sensitive to this variable than to the other two variables. However, this value also has the smallest value range, as there is a limited amount of meeting room meetings that can be held on a day. Therefore, in the last column, the calculation of the coefficient over the value range is shown, showing the maximum impact the variable can have on the number of infections. In this column, it can be seen that the number of people in the meeting room can have the most impact on the number of infections, followed by the number of office meetings and lastly the number of meeting room meetings. The group size of the meeting room meetings having the largest impact could be partly explained by the fact that it has the broadest value range. In addition, each additional person per meeting can be multiplied by the number of meeting room meetings per day in terms of how many additional people attend these meetings during the entire workday. The relatively low impact of the frequency of the office meetings could potentially be explained by the small size of these meetings. This smaller group size could have two effects. First, it leads to fewer direct infections during the meeting itself due to the fewer other people a person comes into contact with during office meetings. Second, it leads to less build-up of virus in the office meeting areas, as there is less probability the infected person had a meeting in the same area earlier. Besides the effect of the group size, the office meetings also have a shorter duration, which also makes them less risk-prone than the meeting room meetings, which are 5 to 10 minutes longer.

Next, the results of the independent t-test are reviewed, which can be seen in Table 36. The (red) cells that contain ‘dif’ entail that the sample in the row and the sample in the column significantly differ from one another based on the t-test. The other (green) cells show the combinations that have not been proven to differ significantly. Within those cells, the non-significant p-value can be found.

Table 36: Results of independent T-tests for sensitivity analysis

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	x	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif
2	dif	x	0.06	0.39	dif	0.61	0.50	0.41	0.44	dif	dif	dif	0.72	dif	dif	dif	dif	dif
3	dif	0.06	x	dif	0.65	0.19	0.11	dif	dif	dif	dif	dif	0.09	dif	dif	dif	dif	dif
4	dif	0.39	dif	x	dif	0.23	0.10	0.92	0.79	dif	dif	dif	0.84	dif	dif	dif	dif	dif
5	dif	dif	0.65	dif	X	0.10	0.06	dif	dif	dif	dif	dif	dif	0.08	dif	dif	dif	dif
6	dif	0.61	0.19	0.23	0.10	x	0.96	0.24	0.26	dif	dif	dif	0.50	dif	dif	dif	dif	dif
7	dif	0.50	0.11	0.10	0.06	0.96	x	0.09	0.06	dif	dif	dif	0.45	dif	dif	dif	dif	dif
8	dif	0.41	dif	0.92	dif	0.24	0.09	x	0.87	dif	dif	dif	0.88	dif	dif	dif	dif	dif
9	dif	0.44	dif	0.79	dif	0.26	0.06	0.87	x	dif	dif	dif	0.95	dif	dif	dif	dif	dif
10	dif	dif	dif	dif	dif	dif	dif	dif	dif	x	0.66	0.26	dif	dif	0.87	0.35	0.39	0.26
11	dif	dif	dif	dif	dif	dif	dif	dif	dif	0.66	x	0.51	dif	dif	0.84	0.65	0.72	0.49
12	dif	dif	dif	dif	dif	dif	dif	dif	dif	0.26	0.51	x	dif	dif	0.46	0.72	0.66	0.86
13	dif	0.72	0.09	0.84	dif	0.50	0.45	0.88	0.95	Dif	dif	dif	x	dif	dif	dif	dif	dif
14	dif	dif	dif	dif	0.08	dif	dif	dif	dif	dif	dif	dif	dif	x	dif	dif	dif	dif
15	dif	dif	dif	dif	dif	dif	dif	dif	Dif	0.87	0.84	0.46	dif	dif	x	0.56	0.52	0.44
16	dif	dif	dif	dif	dif	dif	dif	dif	Dif	0.35	0.65	0.72	dif	dif	0.56	x	0.90	0.66
17	dif	dif	dif	dif	dif	dif	dif	dif	dif	0.39	0.72	0.66	dif	Dif	0.52	0.90	x	0.61
18	dif	dif	dif	dif	dif	dif	dif	dif	Dif	0.26	0.49	0.86	dif	dif	0.44	0.66	0.61	x

In Table 36, it can be seen that the average number of infections in set-up 1 differs significantly from each of the other set-ups. Set-up 1 is the only policy that does not contain any meetings, while all of the other policies contain meetings in both meeting rooms and offices. The meetings thus seem to have a significant impact on the number of infections, despite the conclusion based on the adjusted R-value that the number of infections cannot be described very well based on the meeting parameters. It could be possible that with other values, more data or better compliance with the assumptions of the regression analysis, a higher adjusted R-value could be found.

In addition, it was found that many of the set-ups significantly differ from one another in the average number of infections associated with them. Excluding set-up 1, the infection numbers of 63 % of the set-ups differ significantly from each other. This entails that often the changing of one or more of the parameter values has a significant impact on the number of infections. The next step is then to see what the set-ups that do not significantly differ from one another have in common, to learn more about the influence of the different parameter values.

Set-ups 10, 11, 12, 15, 16, 17 and 18 do not significantly differ, the number of infections between these set-ups ranging only between 0.576 and 0.68. Each of these blocks have the minimum number of office meetings per day, while the number of meeting room meetings and the size of the meeting room meetings both cover the whole range of possible values in these sets. This implies that the office meetings have an important influence on infection numbers. Set-up 13 and 14 are the only set-ups with the minimum number of office meetings per day, that have a significantly different number of infections than the other set-ups with the minimum number of office meetings. Both set-up 13 and 14 have the maximum number of people in a meeting room, while having either the average or maximum number of meeting room meetings, resulting in high infection values. Set-up 14 is even the set-up with the highest number of overall infections. This implies that the amount of office meetings has a significant influence, but if the number and size of the meeting room meetings are very large, this will also significantly influence the infection numbers.

The second block of matching set-ups can be found between set-ups 2, 4, 6, 7, 8 and 9, with the number of infections ranging between 1.073045 and 1.595525, thus having a higher average number of infections than the previous block. What these set-ups have in common is that they are all the set-ups with the maximum value for number of office meetings per day, excluding set-up 3 and 5 which also have this maximum value for number of office meetings. For the number of people in the meeting room they all have the minimum or average value, and for the number of meeting room meetings, they cover the full range of values. Set-up 3 and 5 significantly differ from this block, but not from each other, both having a higher number of infections than the other set-ups with the maximum number of office meetings value. These higher values can be attributed to both having more people attending meeting room meetings during the day than the other set-ups have.

The two blocks show similar patterns. The average number of infections seems to lie very close for set-ups that have the same value for the number of office meetings per day. The exception to this is when a set-up has a high number of people attending the meeting room meetings, and at least an average frequency of meeting room meetings. In this case, the highest overall infection risks are attained, making the set-ups differs significantly from other set-ups with the same frequency of office meetings, but smaller and less meeting room meetings. The size of meeting rooms therefore seems to be the most important indicator of the number of infections, which is in line with the coefficient values for the parameters found in Table 35.

It should be noted that the number of office meetings does only have a minimum and maximum value incorporated in the set-ups, highlighting the difference between the different values. This could be the reason that the difference between the number of infections based on the number of office meetings is clearer. Another important point to note is that the set-up of having a high number of office meetings and a large size and high frequency of meeting room meetings does not exist and could be expected to have the highest number of infections.

Based on the results of the analysis, some conclusions can be drawn. According to the adjusted R-value the meeting parameters' chosen values explain some of the variance found in the number of infections. However, as the meeting parameters cannot perfectly explain all the variance in the infection numbers, other factors are also expected to play a big role. Based on the number of infections per set-up and the results of the independent t-tests, it is found that having no meetings led to a significantly different number of infections in an office than in any office set-up in which meetings were present. This indicates that meetings do have an important impact on the number of infections. Based on the coefficients for the meeting parameters found through the regression analysis, the number of infections is the most sensitive to the number of meeting room meetings held. However, taking into consideration the plausible values changes of the different parameters in an office setting, the order of most influence seems to be: the size of meeting room meetings, frequency of office meetings and frequency of meeting room meetings. This order seems to be in line with the average number of infections found per set-up and the differences between them. The set-ups with the highest number of infections all had the largest size meeting room meetings, while a secondary split in the number of infections could be found based on the number of office meetings.

7.3 Choosing set-ups for the policy analysis

The set-ups that are to be used for the policy analysis on the impact of the NPIs are chosen based on:

- the average number of infections found per set-up
- the influence of the different variables relating to meeting frequency in size
- the differences between the various set-ups in terms of their initial values

Consequently, two set-ups are selected. The first is the set-up that is the most representative in terms of the number of infections relating to the amount and size of the meetings. The second is the set-up that causes the highest amount of infections. This set-up is chosen to learn what NPIs would best help a set-up with high infection risk. The set-up with the highest infection risk is set-up 14. To find the most representative set-up, first, the average number of infections using all of the set-ups is calculated, excluding set-up 1, as this set-up does not contain any meetings. Comparing the average number of infections of all set-ups to this overall average, it could be concluded that set-up 4 provided an average number of infections closest to the overall average. Therefore, this set-up is picked to the representative of the office situation in terms of meetings going forward. An overview of the chosen set-ups can be seen in Table 37.

Table 37: Set-ups chosen for further inspection

Chosen set-up	Type of set-up
Set-up 4	Most representative number of infections in relation to meetings
Set-up 14	Highest infection risk

These two set-ups together are also expected to give a good indication of the impact of NPIs for several other reasons. First of all, they have different values from one another for each of the meeting variables, which can be seen in Table 38, making it possible to see the effect of the NPIs on different variable values. As each of the variables has been shown to have some influence on the number of infections, this is important. Second, the number of infections of the set-ups are shown to be significantly different from another, thus explicitly showing the impact of the NPIs on different infection risk levels.

Table 38: The parameters value and the number of infections of the chosen set-ups

	Size of meetings held in meeting rooms	Number of meeting room meetings per day	Number of office meetings held per day	Number of infections
Set-up 4	7	1	53	1.061
Set-up 14	14	8	13	2.014

8 Model assessment

In this chapter, the impact of the NPIs on the number of infections in the model is assessed. This is systematically done using policy analysis. This leads to an answer to the final sub-question:

What is the effect of the NPIs on the office space scenarios being simulated?

In 8.1 the general approach for the policy analysis for the two set-ups is described. Next, in 8.2 and 8.3 policy analysis is performed for set-up 4 and set-up 14, respectively. Finally, the conclusion of the results is given in 8.4

8.1 Policy analysis approach

As mentioned earlier, the model contains five NPIs, which each can be turned on or off:

- Wearing a mask while walking (Yes or No)
- Increased rate of cleaning of contaminated surfaces, such as the tables, the coffee machine and restroom sinks (Once or twice a day)
- Increased ventilation rate (Standard rate or recommended rate)
- Limited office capacity: Maximum number of people allowed at the office (all employees or 50 %)
- Maximum occupancy rate meetings (no maximum or 50 % occupancy)

These NPIs are combined into all possible combinations to form the policies. To compare these different policies, policy analysis is then performed to compare their effectiveness in reducing the number of infections. Therefore, for each scenario, the number of infections in the office using these NPIs is found by running the model multiple times, due to the stochastic nature of the model. Similar to the model runs for the sensitivity analysis, the model is simulated with the same run of the Nomad model each time, but with multiple runs of the QVEmod model with the same random seed, meaning that a different employee is the infected individual each run.

The formula of Truong et al. (2015) is again used to determine the required number of runs. However, this time with a percentage error of 5 % instead of 10 %, as used in the sensitivity analysis in Chapter 7. The percentage error of 5 % is chosen as this is a fairly commonly used percentage error to provide insightful results. Using the formula of Truong et al., each policy is run at least as many times as the formula prescribed, resulting in between 15 and 20 runs performed per scenario. The exact number of runs per policy can be found in appendix K.

It should be noted that because the same random seed is used as input for each model run, the same employees would be chosen for each run. However, fewer employees are present when the 50 % office capacity NPI is implemented (26 instead of 40). Therefore, the employees chosen for each run when this NPI is present differ from the employees chosen in the runs when this NPI is not present. For example, if in the policy without this NPI employee 35 was chosen to be one of the infected individuals, this person cannot be picked to be one of the infected individuals when the office capacity is limited to 26 people as there is no employee 35 then. Instead, one of the 26 employees is chosen to be the infected individual. This change in the infected person could possibly influence the results, as this causes a person with another schedule visiting other places to spread the virus.

In chapter 7, two set-ups relating to the number and size of meetings held in the offices are chosen to perform the policy analysis on: set-up 4 and set-up 14. In Table 39, these set-ups and the number and size of the meetings belonging to them are presented, along with the average number of infections they produce in the model. Finally, the two chosen set-ups are run again for the number of runs, given an percentage error of 5 %. This is done because these set-ups now represent one of the policies, more specifically the policy in which no NPIs are present. To compare all policies based on the number of infections with an equal percentage error, these set-ups therefore had to be run again. As the number of infections per set-up were calculated with a percentage error of 10 % for the sensitivity analysis, the number of infections found slightly differs from the values presented in Chapter 7.

Table 39: The two chosen set-ups and the parameter values and number of infections belonging to them

Set-up	Number of people in meeting room meeting	Number of meeting room meetings per day	Number of office meetings per day	Average number of infections
4	7	1	53	1.125256
14	14	8	13	2.112

To these two set-ups, the policies containing the different combinations of NPIs are applied, and the average number of infections resulting from them are retrieved. However, for set-up 4 the NPI relating to a maximum capacity of 50 % allowed in the meeting rooms cannot be applied, as this set-up already has the number of people in the meeting rooms equal to this maximum capacity. Therefore, the impact of this NPI is only tested in set-up 14.

Having defined the policies and having chosen the two set-ups to apply them on, for each scenario, the number of infections can be retrieved and compared, first for set-up 4 and then for set-up 14. To learn more about the difference between the number of infections resulting from the policies, independent t-tests are performed. When a significant difference/decrease is shown between two policies, it can be concluded that the NPI that is added to one of the policies has a significant impact on altering the number of infections.

In addition to the t-tests, an Ordinary Least Squares (OLS) Linear Regression analysis is performed to learn more about the effect each individual NPI has on the number of infections. The coefficients resulting from this analysis show how much the number of infections goes down if the NPI is present within the given set-up. Finally, the Adjusted R-Squared value is retrieved, to learn about the proportion of the number of infections that is explained based on the NPIs present. The Adjusted R-Squared value is taken for both set-ups to compensate for the multitude of independent variables, the various NPIs. The compliance with the requirements to do these statistical tests and the numerical results of these analyses can be seen in Appendix K.

8.2 Policies for set-up 4

First, the results of the policy analysis tested on set-up 4 are assessed. In Table 40, the policies are ranked from most to least effective from top to bottom. Per scenario, whether each NPI is present, the average number of infections/R value, the coefficient of variation and both the absolute and percentual reduction of the average number of infections compared to the policy with no NPIs is given. For the NPI columns, a 'P' stands for that NPI being present and 'NP' stands for the NPI not being present in that policy. For the percentual decrease in infections, the cells are color-coded in various shades of green, with darker green meaning a higher decrease in infections. As could be expected, the policy with all NPIs present is most effective in reducing the number of infections. In addition, the average number of infections is highest if no NPIs are present.

Table 40: Number of infections per NPI policy in set-up 4, with darker green meaning a higher decrease in infections

Wear mask while walking	Increased Cleaning rate	Increased ventilation rate	Number of people allowed at the office at 50 %	Average number of infected/ R -value	Relative standard deviation of infections (%)	Absolute decrease in infection	Percentual decrease in infections (%)
P	P	P	P	0.089	29.2	1.037	92.1
P	NP	P	P	0.093	30.1	1.032	91.7
P	P	P	NP	0.094	25.5	1.031	91.6
P	NP	P	NP	0.102	26.5	1.023	90.9
P	P	NP	P	0.118	23.7	1.009	89.6
P	NP	NP	P	0.122	25.4	1.004	89.2
P	P	NP	NP	0.124	21.8	1.001	89.0
P	NP	NP	NP	0.131	22.1	0.994	88.4
NP	P	P	P	0.802	17.3	0.324	28.8
NP	NP	P	P	0.805	17.3	0.329	28.4
NP	P	P	NP	0.874	21.4	0.252	22.4
NP	NP	P	NP	0.881	21.3	0.244	21.7
NP	P	NP	P	1.038	16.5	0.087	7.7
NP	NP	NP	P	1.043	16.6	0.083	7.3
NP	P	NP	NP	1.118	19.3	0.008	0.7
NP	NP	NP	NP	1.125	19.4	x	0

Based on the results shown in Table 40, the t-tests and the OLS Linear Regression, an overview of the overall effect of each NPI is given in Table 41. In the first column the NPI is stated, while in the second column the *coefficient* of the NPI from the OLS Linear Regression analysis is given. This value states the overall influence of the NPI on the number of infections taken over all policies while excluding the possible interaction effect caused by other NPIs. The percentual decrease in infections each coefficient brings about, given the number of infections in this data set, is added in parentheses. The next column shows whether this coefficient is found to be significant and can thus be stated with some certainty. If the coefficient is not significant, the p-value of the coefficient is given to provide a complete view. The p-value here is the probability of having the given number of infections for the policies while in reality the averages for the number of infections of the two policies are the same.

Next, the column *highest decrease* shows the highest decrease of infection the NPI brought about between any two policies in which the NPI was added. For each NPI, this *highest decrease* occurred between policies where no other NPIs are present. In the fourth column, the *lowest decrease* in infections the NPI brought about between any two policies is given. In all cases, this lowest decrease is found in the two policies in which the most other effective NPIs are present. All NPIs can thus be said to have the highest percentual impact when no other NPIs with a significant impact are present, which could be due to the higher number of initial infections the NPI can then help to bring down.

Table 41: Overview of the effect per NPI in set-up 4

NPI	Coefficient (percentual decrease)	Significant coefficient	Highest Decrease	Lowest decrease	Significance decreases
Mask wearing	-0.8517 (75.7%)	Yes	88.4 %	63.4 %	Always
Cleaning	-0.0060 (0.5%)	No (0.861)	0.7 %	0.4 %	Never
Ventilation	-0.1346 (11.6 %)	Yes	21.7 %	2.5%	Always
Reduced number of people in office	-0.0425 (3.8 %)	No (0.232)	7.3%	0.5 %	Never

Based on the outcomes of these analyses, the least effective NPI seems to be the increased cleaning frequency, which does not have a significant effect when added to a scenario, while the percental changes are always below one percent. The second least effective NPI seems to be to limit the number of people in the office to 50 % of the number of people working there. While no significant decrease is proven in any scenario, it should be noted that in this set of policies, the policy including the 50 % of people NPIs does always show a slightly lower number of infections than the policy that does not have this NPI but does have the same other NPIs present. Therefore, it could be that if the t-tests were based on more extensive data, the differences would be significant, as more data available leads to more probability of significance (Verbeek, 2008). The limited effectiveness of restricting the number of people in the office could be due to the people being present having more meetings. While having less people in the office causes less people overall to potentially get infected, the people in the office having more meetings and thus more interaction with groups of people could in turn increase the risk of infection. This can be further explored in set-up 14, in which even more and larger meetings take place.

The NPIs of wearing a mask and increasing the ventilation rate both show a significant impact on reducing the number of infections in any scenario. The percentual impact of wearing a mask is much higher however than the impact of increasing the ventilation, with the coefficient for the mask-wearing NPI being more than 6 times the size of the coefficient for the increased ventilation. In addition, in case the mask-wearing while walking NPIs is already present, having either both the cleaning and limited office capacity or having only the ventilation NPI does not make a significant difference from one another.

Finally, the Adjusted R-Squared value, shown in Table 42, is assessed. The found value is very high, in addition to being significant. The adjusted R-squared value found indicates that roughly 98 % of the observed variance in the number of infections can be explained by the presence of NPIs. The presence of the NPIs can explain the number of infections quite well in this set-up.

Table 42: Adjusted R-squared value for NPIs in set-up 4

Adjusted R-Squared	Significance
0.978	9.78^{-10}

8.3 Policies for set-up 14

For set-up 4, the different policies, in this case 32 policies, are simulated again and the average number of infections per policy is found. In Table 43, the policies are ranked from most to least effective from top to bottom. The Table has the same columns as Table 40, with the addition of the NPI present for limiting the number of people allowed in the meeting rooms. For the NPI columns, a ‘P’ stands for that NPI being present, and ‘NP’ stands for the NPI not being present in that policy. The color coding is also the same as in Table 40, with the addition of the orange color in the percentual decrease column meaning a significant increase in infections.

Similar to set-up 4, the most effective policy is the policy in which all NPIs are present, similar to the results for set-up 4. In contrast with set-up 4, however, it can be seen that the variance in set-up 14 reaches higher values.

Table 43: Number of infections per NPI policy in set-up 14, with darker green meaning a higher decrease in infections and orange meaning a significant increase in infections

Wear mask while walking	Increased Cleaning rate	Increased ventilation rate	Number of people allowed at the office at 50 %	Occupancy rate meeting rooms of 50 %	Average number of infected/ R-value	Relative standard deviation (%)	Absolute decrease in infections	Percentual decrease in infections (%)
P	P	P	P	P	0.084	34.52	2.027	96.0
P	NP	P	P	P	0.090	35.6	2.021	95.7
P	P	P	NP	P	0.112	35.7	2.000	94.7
P	P	NP	P	P	0.112	40.2	1.999	94.7
P	NP	NP	P	P	0.119	30.3	1.993	94.4
P	NP	P	NP	P	0.124	27.4	1.987	94.1
P	P	NP	NP	P	0.137	35.0	1.974	93.5
P	NP	NP	NP	P	0.150	34.7	1.962	92.9
P	P	P	P	NP	0.159	13.8	1.952	92.5
P	NP	P	P	NP	0.162	14.8	1.950	92.3
P	P	P	NP	NP	0.165	19.4	1.946	92.2
P	NP	P	NP	NP	0.169	20.1	1.942	92.0
P	P	NP	NP	NP	0.224	18.3	1.888	89.4
P	P	NP	P	NP	0.228	18.0	1.882	89.2
P	NP	NP	NP	NP	0.228	12.3	1.884	89.2
P	NP	NP	P	NP	0.230	12.2	1.880	89.1
NP	P	P	P	P	0.738	27.9	1.374	65.1
NP	NP	P	P	P	0.746	29.1	1.366	64.7
NP	P	P	NP	P	0.752	41.5	1.359	64.4
NP	NP	P	NP	P	0.765	40.7	1.347	63.8
NP	P	NP	NP	P	0.948	40.3	1.164	55.1
NP	NP	NP	NP	P	0.960	39.6	1.152	54.6
NP	P	NP	P	P	0.976	25.8	1.138	53.9
NP	NP	NP	P	P	0.998	24.4	1.113	52.7
NP	P	P	NP	NP	1.630	27.6	0.481	22.8
NP	NP	P	NP	NP	1.635	27.3	0.476	22.6
NP	P	P	P	NP	1.999	14.7	0.113	5.3
NP	NP	P	P	NP	2.018	14.5	0.094	4.4
NP	P	NP	NP	NP	2.064	27.0	0.048	2.3
NP	NP	NP	NP	NP	2.112	27.9	0	0
NP	NP	NP	P	NP	2.473	12.8	-0.362	-17.1
NP	P	NP	P	NP	2.537	12.9	-0.426	-20.2

A difference from set-up 4 is that the policy with the highest amount of infections is not the policy in which no NPIs are present. In set-up 14, two policies have an even higher number of infections than the policy in which no NPIs are present. Both policies have the NPI limiting the number of people present in the office present, while one of them also has the additional cleaning NPI present. In the analyses in set-up 4, it was found that the NPI of cleaning does not provide a significant effect on the number of infections. Therefore, it is expected that the number of infections is mostly related to the NPI of limiting the number of people allowed at the office. An explanation could be that the NPI of limiting the employees present at the office gives inconsistent results, as it on the one hand limits the number of people present to potentially get infected, but on the other hand potentially increases the infection risk of the people present, as these people now attend more meetings. The reason this inconsistency is more apparent in set-up 14, is that this set-up has larger and more frequency meetings than set-up 4, enlarging the effect of the increased number of meetings. The inconsistency in the results could also possibly be explained by the influence of the other NPIs present, as based on comparison of the number of infections per scenario, the effectiveness of the limiting the number of people allowed at the office seems to increase when more other NPIs are present. This could be explained by the fact that if more other NPIs are present, the people present in the office have less risk of infection in the increase number of meetings, while the effect of less people being present in the office to get infected is still present.

The effect of the NPI of limiting the number of people is thus inconsistent for set-up 14, and its effect is also influenced by the other NPIs present. Therefore, the addition of the other NPIs in this set-up does not only have the potential to lower the number of infections due to their own direct impact, but also due to the interaction effect with the limited office capacity NPI. For example, adding only the NPI of mask wearing to a policy in which no NPIs are present reduces the infections with 89.21 %. However, adding the mask-wearing NPI to a policy in which the office capacity NPI is already present brings out the effectiveness of this NPI. This effectiveness is brought out as now the people present have less risk of getting infected, but less people are present to get infected. This interaction effect leads to the mask-wearing NPI in this scenario, in which the office capacity NPI is already present, reducing the number of infections by more than 100 %.

To learn more about the effect of each of the NPIs in this set-up, next, the impact of each NPI separately is assessed based on the results of the statistical tests given in Table 44.

Table 44: Overview of the effect per NPI in set-up 14

NPI	Coefficient	Significant coefficient	Highest Decrease	Lowest decrease	Significance decreases
Mask wearing	-1.3037 (61.672 %)	Yes	109.38 %	30.93 %	Always
Cleaning	-0.0071 (0.03 %)	No (0.954)	3.8 %	3.14 %	Never
Ventilation	-0.1968 (9.32 %)	No (0.114)	25.51 %	1.31 %	Sometimes
Limited capacity office	0.0934 (4.42 %)	No (0.444)	1.47 %	17.13 %	Sometimes
Max capacity meeting room	-0.6389 (30.25 %)	Yes	74.06 %	0.14 %	Always

Based on the results, it seems that wearing a mask is again the most effective NPI. The consistent effectiveness of mask-wearing could be expected, due to its effectiveness not being influenced by other NPIs or contextual factors. However, the high effectiveness is deemed somewhat surprising, as the masks are only worn while walking and thus not during the meetings, which also plays a big role in the virus spread. However, it does effectively protect employees each time they get coffee, use the restroom or walk through the corridors, which happens regularly during the day. The NPI of increasing the ventilation seems less effective in set-up 14 compared to set-up 4, as its effect is not always significant in this set-up. In contrast, in set-up 4 its effect always was significant. However, it still provides a significant decrease in policies in which no or a few NPIs are already present and could therefore still be useful. Increasing the cleaning frequency does not provide any significant impact, which is in line with the results in set-up 4. For the NPI of the office capacity, the insignificant coefficient of 0.09, which is very close to zero, reaffirms the expectation that this NPI has an inconsistent effect on the number of infections.

Finally, the NPI of having the capacity of the meeting rooms reduced to 50 % is tested for the first time in this set-up. The relatively high effectiveness of limiting the number of people in meetings is expected to be due to the large group size of the meeting, 14 people, in this set-up. Reducing this to 7 people per meeting for the 8 meetings per day causes the number of people employees come into contact with through meetings to be much lower. Lowering the number of people in the meeting rooms to 7 in effect makes these policies an implementation of set-up 11, which, even without NPIs present, has a significantly lower average number of infections compared to set-up 14, 0.64 as opposed to 2.01 infected. This NPI can thus be expected to significantly help lower the number of infections, when added to any policy. This can also be seen in the results of Table 44, which shows this NPI has a significant impact on lowering the number of infections. However, the decrease of the infections caused by this NPI seems to be very variable, influenced by other NPIs present.

Compared to the results of set-up 4, the differences in the number of infections for the NPIs seem to be larger in set-up 14. This could be due to various reasons. First of all, a reason could be the larger role of interaction effects, as was noted for the NPI of limiting the number of people in the office. Other reasons could be the addition of the NPI of the maximum amount of people in meeting rooms in set-up 14 or the different size and frequency of the meetings of both set-ups, which also play a role in the number of infections.

Finally, the impact of all NPIs on the number of infections in set-up 14 is assessed based on the Adjusted R-Squared value, shown in Table 45. The Adjusted R-Squared value shows that roughly 82 % of the observed variance in the number of infections can be explained by the presence of NPIs in this set-up. This value is lower compared to the 94 % found for set-up 4. A possible explanation for this could be the inconsistent effect the NPI of limiting the number of people allowed at the office has. However, despite the lower value compared to set-up 4, it can still be assumed the NPIs present can explain the number of infections quite well in this set-up.

Table 45: Adjusted R-squared value for NPIs in set-up 14

Adjusted R-Squared	Significance
0.823	5.47^{-10}

8.4 Conclusions

Having gathered data on the number of infections given various NPI policies and analyzing this data statistically, some conclusions can be drawn on the NPIs' effectiveness. However, before this is done, the certainty of these results is shortly reviewed. It should be noted that for many of the policies, there exists quite some variation in the number of infections. Various instances were found where two values for the number of infected within one policy differed from one another with a relative standard deviation of 100 %, meaning one is twice as big as the other. The only difference between the runs is which employee is the infected individual, with each employee having a different schedule. This points to the number of infections and the effectiveness of the NPIs also being dependent on which employee actually is the infectious one. For example, this could mean that within the same setting and with the same NPI policies present, one sick employee could walk in and infect no one, but another sick employee could walk in and infect at least one person. This variance leads to more uncertainty in the results (Zidek & Van Eeden, 2003). This could also be the reason that some statistical tests did not provide definitive results on the significance of the difference between the number of infections per policy or on the coefficients for variables or a higher explanatory capacity of the NPIs on the number of infections (Bangdiwala, 2016). Therefore, the conclusions on the effectiveness of the NPIs have to be given with some uncertainty.

8.4.1 Measuring the effectiveness using the Reproduction value

Having given some information about the certainty of the results, the next step is how to interpret them. This could be done purely based on the average number of infections and the t-tests and OLS Linear Regression analyses performed. However, as the results relate to an office situation and the social context also plays an important role in which infection rate is acceptable, other considerations are also important. The clearest guidelines for COVID-19 prevention come from the government. The Dutch government does not give hard guidelines when the number of infections or spread of the COVID-19 virus is too high/large, as it also depends on how sick people get from the virus (also depending on the vaccination rate), how many people need to go to the hospital (NOS, 2022) and how much capacity the hospital has to harbor and care for these people (NOS, 2021). However, a general government guideline is to try to keep the Reproduction value (R-value) below 1, meaning the virus is not spreading any further (Blauw, 2020).

The guideline of keeping the R-value below 1 seems to also be a good guideline for an office, if they want to make sure their COVID-19 prevention plan is in line with government regulations. Based on this, it is assumed in this thesis that the maximum acceptable R-value is just below 1. This maximum value would then apply to the average number of infections a policy scenario/ NPI brings about. However, as mentioned in section 8.4, the R-value can heavily vary for different employees being the infectious individual, within the NPI policy implemented. Therefore, if an office manager wants to be even safer, the highest number of infections for each NPI policy, the 'worst-case scenario', should be considered. Both the average case and worst-case can be used to compare the R-value and are used to compare NPI policies in 8.4.3.

8.4.2 Effectiveness of the individual NPIs

Having set a straightforward way to interpret the results of the analysis, both the separate effectiveness of the separate NPIs and the effectiveness of the different combinations (policies) of NPIs can be reviewed now. First, the individual effectiveness per NPI is reviewed based on the policy analysis. The ranking of the NPIs from most to least effective can be seen in Table 46. This is only done based on the average number of infections and not on the worst-case scenario, to base the results on more data and to make full use of the policy analysis performed.

Table 46: Overview of the effect of the NPIs using the results of both set-ups

NPI	Coefficient set-up 4	Coefficient set-up 14	Significant influence	Overall maximum decrease	Overall minimum decrease
Mask wearing	75.70%	61.672 %	Yes	109.4 %	30.9 %
Max capacity meeting room	No data	30.25 %	Yes	74.1 %	0.1 %
Ventilation	11.56 %	9.32 %	Sometimes	25.5 %	1.3 %
Cleaning	0.53%	0.03 %	Never	3.8 %	3.1 %
Limited capacity office	- 3.77 %	4.42 %	Sometimes	7.0%	17.1 %

According to the results, the most effective NPI is the obligation to wear a mask while walking. This NPI has the largest impact in terms of percentual decrease, which is significant in combination with any other combination of NPIs. The second best NPI seems to be limiting the number of people in the meeting rooms, as this shows a significant decrease in the number of infections in all policies. However, the impact of this NPI is less significant than the mask-wearing and is only tested in one set-up, because it is only applicable to set-ups containing large meetings. Ventilation is rated third, because the NPI does both have a lower impact than the first two NPI, in addition to not always providing a significant impact. This was seen in the case of set-up 14, in which more and larger meeting room meetings are held. Its impact also strongly declines when more other NPIs are present. Therefore, usefulness in preventing infections of this NPI is not guaranteed in settings with many and large meetings, or where other effective NPIs are already present. Increasing the cleaning frequency is ranked fourth, as it does not have any impact on both set-ups. This is in line with the finding of Duives et al. (2022) on the fomites not playing a large role in virus transmission. It is thus advised office managers not invest too much money in increasing cleaning.

Reducing the number of people allowed in the offices is ranked as less effective because despite having the potential to decrease the number of infections, its effect is very inconsistent and dependent on other factors. Because it is assumed that office managers value the trustworthiness of the effectiveness of NPIs very much, this NPI is ranked least effective.

8.4.3. Effectiveness of the combined NPIs

Next, the effectiveness of the different combinations of NPIs is assessed using the different policies formulated in the policy analysis in 8.2 & 8.3. This is first done using the average number of infections per policy and then done using the worst-case scenario. When looking at the results of the policy analysis based on the average number of infections, for both set-up 4 and 14, the policies with the lowest R-value are the policies in which all NPIs are present. Therefore, if an office manager has unlimited resources, implementing all NPIs would ensure the lowest risk. However, the cleaning NPI never provides any significant effect based on the data. Therefore, leaving this NPI out is expected to provide the same effect. However, if the office manager has more limited resources, it might be more beneficial to see which NPIs are needed at a minimum to bring the R-value below 1. For the two set-ups, different policies apply, as the effect of the NPI regarding the maximum meeting room could only be analyzed in set-14. For set-up 14, the average number of infections/ R-value is below 1.0 if either at least the mask-wearing NPI is present or if this is not the case, the restriction of the number of people in the meeting rooms NPI is present. For set-up 4, the R-value of the policies is under 1 if at least either the mask-wearing NPI or the increased ventilation NPI is present.

However, in the case that an office manager wants to take even less risk of having a potential outbreak of COVID-19 in the office, he could also base his choice of NPIs on the maximum number of infections found per scenario. The maximum number of infections for all policies for both set-up 4 and set-up 14 can be seen in Table 47. Here, the values found for set-up 4 are attributed to the policies where the occupancy rate of the meeting rooms is 50 % because although this NPI is not explicitly implemented there, it is complied with due to the standard sizes of the meeting room meetings in this set-up.

Table 47: Maximum number of infections for policies in both set-ups (P = present, NP = not present)

Wear mask while walking	Increased Cleaning rate	Increased ventilation rate	Number of people allowed at the office at 50 %	Occupancy rate meeting rooms of 50 %	Maximum number of infections in set-up 14	Maximum number of infections in set-up 4
P	P	P	P	P	0.155	0.160
P	P	P	NP	P	0.184	0.142
P	P	NP	P	P	0.195	0.190
P	NP	NP	P	P	0.198	0.206
P	NP	P	NP	P	0.197	0.146
P	P	NP	NP	P	0.232	0.186
P	NP	NP	NP	P	0.244	0.187
P	P	P	P	NP	0.205	
P	NP	P	P	NP	0.210	
P	NP	P	P	NP	0.159	0.174
P	NP	P	NP	NP	0.247	
P	NP	P	NP	NP	0.253	
P	P	NP	NP	NP	0.304	
P	P	NP	P	NP	0.287	
P	NP	NP	NP	NP	0.313	
P	NP	NP	P	NP	0.291	
NP	P	P	P	P	1.196	1.061
NP	NP	P	P	P	1.206	1.060
NP	P	P	NP	P	1.424	1.519
NP	NP	P	NP	P	1.446	1.528
NP	P	NP	NP	P	1.814	1.869
NP	NP	NP	NP	P	1.823	1.879
NP	P	NP	P	P	1.451	1.376
NP	NP	NP	P	P	1.463	1.381
NP	P	P	NP	NP	2.281	
NP	NP	P	NP	NP	2.281	
NP	P	P	P	NP	2.639	
NP	NP	P	P	NP	2.643	
NP	P	NP	NP	NP	2.986	
NP	NP	NP	NP	NP	2.983	
NP	NP	NP	P	NP	3.237	
NP	P	NP	P	NP	3.225	

Using the maximum number of infections, an office manager would then again want the R-value, which is equal to the number of infections, to be below 1. Based on this guideline, for both set-up 4 and set-up 14 this entails that at least the obligation of wearing a mask while walking needs to be present to have an R-value below 1. Compared to using the average number of infections, it thus does not suffice anymore to combine either the increased ventilation in set-up 4 or the maximum size of meeting rooms in set-up 14 with the mask-wearing.

Finally, it should be noted that compliance with the NPIs is incorporated into their effectiveness. Therefore, it is not expected that the effectiveness of the NPIs in real life significantly differs from these results due to people possibly not adhering to them. For example, the cleaning NPI is not subjected to compliance of employees while the masking wearing is. However, as the compliance is already incorporated, it is not expected that the increased cleaning NPI's effectiveness comes closer to the effectiveness of mask-wearing on this account.

9 Discussion

This chapter covers different types of reflection on the performed research. Section 9.1 starts with the implications of the results found in chapter 8 in relation to current literature. Next, section 9.2 discusses some of the choices and assumptions made in this research and recommendations are made based on this. Then, in section 9.3, the insights for policy-makers are given. In 9.4 the scientific and societal contributions of this thesis are discussed. Finally, in 9.5 the thesis' suitability with the CoSEM master program is described.

9.1 Literature implications

In this section, the results of this research are discussed in the perspective of the findings of the literature review. The influence of the NPIs and the parameters related to the meeting part of activity scheduling found in this research are compared with the influences found in the literature. In some cases, this confirms the results found, while in other instances different results were found in literature. When the results of this thesis research and literature differ, an explanation for this is sought, which could provide further insight into the effectiveness of this parameter or NPI.

In this research, no significant impact of increasing the amount of cleaning of surfaces was found. There is little literature on the separate impact of increased cleaning. However, the case study of Duives et al. (2022), which also used the PeDViS model, found that fomites do not present a large contribution to virus transmission. In a case study using PeDViS in a festival setting, similar results were found. Therefore, for the PeDViS model, fomites do not seem to play a large role in virus transmission under both restaurant, outdoor festival and office settings. However, due to this still only referring to the PeDViS model, it would be better to put this finding into the perspective of additional literature. Empirical findings and correspondence on the role of fomites also found that the role of fomites in virus transmission is very limited (Zhang et al., 2021; Colaneri et al., 2020; Dowell et al., 2004). By extension, it can be assumed that increased cleaning, which is the removing of fomites from surfaces, does not have much impact on decreasing the number of infections. This is in line with the findings of this thesis.

Next, increasing the ventilation rate was shown to often bring down the number of infections, although this effect is limited when other effective NPIs are present or the number of infections is relatively high. The reduction in the average number of infections found in this reach is between 25.51% and 1.31%. This range of values lies below the reduction found by Duives et al. (2022). Duives et al. found that increasing the ventilation by an amount roughly similar to the increased ventilation rate of this research (without including compliance) reduced the number of infections by 41 %. In further literature, there seems to be no consensus on the effectiveness of increased ventilation. Li & Yin (2021) also found no significant impact in making the ventilation rate twice as high, while Salmenjoki et al. (2021) 'point to the crucial role of ventilation in order to reduce aerosols'. It is therefore assumed that increased ventilation can play an important role, but its role is heavily influenced by various external circumstances.

The wearing of masks while walking was found to be very effective in this research, under all different policies tested for the office. The reduction it caused in the number of infections was found to be at least 30.93 %, and when added as the only NPI present to be up to 89.21 %. The lower value of 30.39 % is in line with the research of Alvarez Castro & Ford (2021), which found that 'facemasks could reduce infection peak by 30% if worn by everybody'. It was also found that by solely obligating everyone to wear a mask while walking, sufficient safety could be guaranteed in the office. This is partly in line with the work of D'Orazio et al. (2021), which states virus spread could be limited by everyone wearing FP3 masks. They also found that a combination of wearing surgical masks with limiting the number of people in the building guarantees sufficient safety. This finding is in line with the finding of this research that the combination of these two NPIs often gives an R-value below 1, thus providing sufficient safety according to the guidelines used.

An interesting comparison can be found comparing the results to the results of the impact of facemask use of Duives et al. (2022). Although the findings of Duives et al. are also based on the PeDViS model, as the configuration for the mask filter efficiency and wearing of the mask being the same, they have found the impact of facemasks to be 'really minor'. This is in contrast with the finding of this adaption of the PeDViS model, which finds a great impact for wearing masks. A possible explanation for this difference could be found in the different runtime and layout of the model. First of all, the restaurant version of PeDViS only simulated an indoor space for a couple of hours, while this office adaption simulates an entire workday. This leads to a higher build-up of fomites and droplets in the building, especially in the corridors, which employees can exposed to when walking through the office. In addition, in the layout no doors are modelled, due to this not being possible in PeDViS. This leads to the effect of all doors being open all the time. Therefore virus particles can more easily spread through the entire office. The combined effect over higher buildup and more virus spread throughout the building could then possibly explain this difference. However, the increased effectiveness of mask in the office adaption could also be (partly) due to other reasons, such as the difference in lay-out or in activities compared to the restaurant version. Further research is needed to learn more about the cause of the difference, on which more information is given in section 9.2.

Reducing the number of people allowed at the office was found to have a varying effect on the number of infections, possibly due to the duality of having fewer people present but these people having more meetings and thus coming into contact with more people. In both, Harweg, Bachmann & Weichert (2020) and Salmenjoki et al. (2021) limiting the number of people was found to reduce virus transmission. However, in both cases, the cause for this reduction was the resulting lower density of people. The reduction of the number of people in the office could cause a lower density of people in the coffee area or restroom. However, this is not the case for meetings, as these stay the same size. The amount of meetings the people present in the office have even rises, which can in turn lead to a higher infection risk. It could therefore be possible that the reduction of the number of people allowed in an office is only effective if this leads to a lower density of people. The principle of ensuring lower density when implementing an NPI can be extended to the NPI of reducing the number of people allowed in a meeting room. For example, when only 50 % of the meeting room's capacity is allowed in the room, this could be combined with the obligation to leave one chair unoccupied at each side of a person. This would lead to a lower density of people in the room. In comparison, if all people would all sit right next to each other when at 50 % capacity, this would potentially lower the number of people with an increased infection risk, but keep the density of people at roughly the same level. This theory also seems in line with the work of Li & Yin (2021), which found that the reduction of the number of people is only effective in combination with other NPIs.

In addition, the size and frequency of meetings held in meeting rooms are found to play a large role in virus transmission in the office. This is in line with the work of Ford (2021), which found that the most dangerous places are those where many people interact for a long time, which in the office context refers to the meeting rooms. In Duives et al. (2022) it is also stated that locations like the bathroom 'provide a low risk of infection, due to the short time people spend there.' This further confirms the idea that meetings are one of the most important sources of infections.

Finally, it was generally found that the highest reduction of the average number of infections was attained when all NPIs were present, in comparison to only one or part of the NPIs being present. This is in line with Li & Yin (2021), who found that the combination of NPI is superior to any single intervention in reducing the number of infections.

9.2 Limitations & recommendations for future work

Using the modelling approach, various choices and assumptions have been made through this thesis that can lead to limitations of the usability of the model and other parts of the research. Therefore, in this section, the assumptions and limitations of this research are discussed and recommendations are done for future work. These limitations are categorized to the different parts of the thesis research they apply to.

9.2.1 Modelling choices

There are various limitations of the usability of the model and possible future extensions related to the modelling choices made. First, the choice was made to design personal offices allotted to one person. This was done to provide an initial setup of an office, that was relatively easy to model and schedule the activities for. This way, for example, there would be no additional scheduling needed for small meetings being held in an office only when no other employee is working there, so as not to disturb them. In addition, already having these personal offices, also adding shared work spaces was though to add less new insights than adding other types of functional office areas to the layout. Due to only having the time to include a limited number of functional areas, shared work spaces are not included in the office lay-out and behavior. However, despite personal offices being observed to exist in offices, it was more commonly observed that office spaces were shared, often by 2 or possibly 3 people. Modelling the office infrastructure would thus possibly be more representative, in addition to providing much insight into the virus transmission within these shared offices, which could be a large part of virus transmission. This could unfortunately not be added anymore due to limited time and resources. Therefore, changing the infrastructure to contain shared offices would be a good expansion for future work.

Another limitation is that the model is scoped to not include the activity of having lunch. This choice was made because the restaurant version of the PeDViS model already simulates transmission in a context similar to having lunch. However, having this type of activity incorporated in the office version, as an activity most employees have in the middle of the day, would be of added value. As this is an activity in which a lot of human contact happens, this activity does have a high potential for causing infections and would be of added value to include.

In addition, a limitation of the model regarding the first 30 minutes modelled should be noted. The model starts the simulations at 8:30, from which time till 9:00 people come into the offices. The first possible meeting starts at 9:00. People are instructed to always head to their office as they first come in. However, due to how the code underlying the model is set up when people have their office set as their next destination, the time they spend there will either be till the first meeting of the day at 9:00, if they happen to be one of the people attending that meeting or for a period of either 30, 60 or 120 minutes. People will thus not just hop in their office quickly and go to another location fairly. In addition, during the data collection there was no data gathered on which location people headed to first when entering the office. Therefore, different probabilities of going to the various locations upon entering the office couldn't be modeled. For these reasons, the first 30 minutes of the simulations mainly consist of people sitting in their personal offices, and not walking around much. Due to more activity happening in real life in this time period, these first 30 minutes are expected to have a lower amount of virus transmission than would be present in real life. A possible extension for future work could be to allow employees to choose from all activities upon first entering the offices or letting them work in their personal office for shorter amounts of time at the start of the day. Previous to this extension of the modelled, data would need to be gathered on which location employees go to when they first come into the office and their behavior for their first 30 minutes in the office.

Next, a possible limitation could results from the activity scheduling being scoped to be per minute. This choice was made because this causes a lot less computing power to be needed, making it easier to run the model. However, choice caused certain processes/activities to not get assigned the duration they were observed to have in real life or that would be most fitting based on calculations. An example is the activity of walking from one location to another, for which a duration in terms of seconds instead of per minute would be more fitting. If the activity scheduling is scoped to assign durations in seconds, the walking time could also be a calculated value based on the distance between two places and a person's individual standard walking speed. However, this would altogether be a quite large extension, which might not be a very plausible direction for future work unless extensive computational power is available. An alternative way to get the walking time better represented would be by having the walking speed not drawn from a uniform distribution anymore but set to a constant value in the middle of that distribution. This way, the overall walking time of employees only depends on the distances they have to walk and are expected to lie closer together, as there are no employees that walk very fast/slow that deviate a lot in their average walking time. In consequence, the configuration that would then be found to incorporated the average walking time is expected to lie closer to the average walking time many employees have. This way, employees would have smaller deviations from their work schedules. However, the downside of having a universal walking speed for all employees would be that less stochasticity would be included in the model and the walking behavior of individuals is less accurately represented.

Third, it should be considered that the size of the office meetings is set to a constant value of 3 people, which is equal to the number of seats available at the meeting tables in offices. However, first of all, the data gathering did not provide much information about these office meetings, as these were hard to observe behind the closed doors of the offices. More information on their sizes especially would be beneficial. This could for example be obtained by letting TU Delft employees fill in per day how many people they have in their office. In addition, the constant value of 3 might not have been the most representative, as people could also have small meetings consisting of two people. Looking back, having some meetings of 2 people and some meetings of 3 people would have probably been a better value based on the data available. However, due to limited time and the added complexity of incorporating this, this was not incorporated in the model anymore, but it would be a good addition for newer versions of this model.

Next, an point for improvement of the QVEmod part of the model could be to do sensitivity testing of fine-tuning for the touching frequency values implemented. Some of the touching frequencies are based on the PeDViS version of Duives et al. (2022) and can be expected with more certainty to be fitting. However, other frequencies are based on literature of calculations based on assumptions. A possible extension would be to first of all gather more data on the actual touching frequencies at the different locations at the office, for example through observation. Second, it could be tested if the touching frequencies has a strong impact on the infection risk, and the number of infection resulting from having these values implemented are in accordance with expectations. However, as the role of fomites in virus is expected to be limited, it is expected the touching frequencies do not play a big role in the number of infections. Therefore, it should be kept in mind that the touching frequencies implemented might be different from real life situations, but exploring their values are not recommended as a priority for future research.

The final limitation within this category related to the NPI of restricting the number of people allowed in the meeting rooms. When this NPI is present, the people that are attending the meeting but do not fit into the meeting room anymore are joining these meetings online in their office. However, these people are not bound to their office for the duration of the meeting. This could be implemented in future versions, but is not included in this thesis research due to limited time to make these changes. People not being bound to their offices results in these employees occasionally getting coffee or going to the restroom during these meetings. However, as these are both short activities it is expected that this does not make much of a difference in terms of their activity schedule. It lowers the coffee and restroom crowdedness peak, as it reduces the number of people going to these activities straight from the meeting rooms. This lower crowdedness at these locations is in line with that normally people joining online would not be in the same building, and would thus also not get coffee or use the same restroom with the people in the meeting. This behavior and them not sitting in the same room with the other people in the meeting larges simulates the same decrease in virus transmission as having these people bound to their office would have.

9.2.2 Model Limitations

Based on the model verification performed, various flaws are found in the simulations of the activity scheduling part of the model. These limitations are described in this section. First of all, while monitoring the employee's walking behavior, it is found that at times one employee stands on top of another employee in one of three locations: on the toilet, the sink or the coffee area. The model schedules people's behavior in a way that they check if the spot they want to go to is already taken at the moment. However, due to people deviating up to 8 minutes from their schedule, it is not yet possible to fully prevent people from being at the same place at the same time. This problem mostly occurs in meeting set-ups where the crowdedness at the coffee area and restroom is happening more in peaks, due to many and large meetings taking place. In this case, it happens up to 15 times per day in the coffee area and 4 times a day in the restroom. This problem should not have too grave an impact on the results, as them standing on top of each other has the same effect on the virus transmission as two people standing close to each other. For the coffee area and the sinks, this is relatively similar to how it would be in real life. For the bathroom stalls, the employees standing on top of each other is not similar to real life, as the restroom stalls are supposed to be in a closed-off environment. However, the impact should still be limited due to both the toilet activity only taking one minute, and the overlap usually not even happening for the full minute, in addition to it only happening up to 4 times a day in the busiest set-up.

Next, it should be noted that the modelled coffee and restroom area might sometimes busier than they would be in real life. This crowdedness is caused by many meetings still ending at similar times, despite the measures of two office meeting shifts and the meeting room meetings having different durations that are put in place to vary the meeting times. This also explains some of the overlaps of people in the coffee area and restroom explained in the previous point. However, the randomness of the amount of deviation each employee has from their schedule due to walking time does allow for the peaks in crowdedness to be more spread out. In addition the highest number of people in restrooms and coffee areas observed in the model[2] is quite similar to what was observed in real life. Also, it could be argued that at least the similar starting times of meetings in the model mirror real life well, as meetings often start at the full hour or half hour, causing more meetings to start as the same time.

Another limitation is a side effect of one of the measures to prevent crowdedness in the coffee area, which entails that an employee who wants to get coffee picks another activity if 6 other people are already scheduled to get coffee at that time. A flaw of the model is that due to the schedules being constructed one at a time in the order in which the employees enter the office, the employees that enter the office later have a higher probability of having to skip getting coffee due to crowdedness. Based on the verification checks in set-ups with high crowdedness at the coffee area, it was noted that employees skipping getting coffee starts from around employee 20 (the 20th employee that enters the office). However, it was also discovered that after employee 20, the number of times an employee had to skip getting coffee did not increase linearly. However, the probability of an employee having to skip coffee at least once did seem to increase the later the employee entered the building. In general, it was noted that employees had to skip getting coffee mostly between 0 and 3 times a day for the busiest set-up, with a sporadic outlier of 5 times a day. In set-ups with fewer meetings and the peaks of crowdedness at the coffee area thus being flatter, this happened less often for all employees. In general, the problem is that the first half of the employees that enter the office building in the morning never have to skip getting coffee, and thus being in a crowded coffee area more often, which can influence their probability of getting infected.

In addition, the schedules being filled in in the order in which employees enter the office also influences the exact location employees go to in the coffee area and rest room. This mechanism causes the choice location within the areas not to be fully random. The people who enter the office earlier in the

[2] A maximum of 7 people at the coffee area and 4 at once at the restrooms, but happening very sporadically

morning are the ones that have their schedule filled in earlier, causing them to be more likely to have the spot closest to the coffee machine or the first restroom, while the people start their workday a few minutes later are more likely to have another spot the entire day. This causes the place assignment not to be fully random but influenced by the time in which they got in the office, which would not be the case in real life.

A fourth limitation, also related to the overlap of employees at certain spots, is that if an employee sees that all toilets are occupied, he is not programmed to wait until one of the stalls is free and then go into that stall. An attempt was made to include this. However, because the activity scheduling does not translate very precisely into Nomad timewise due to employees deviating up to a few minutes from their schedule, this only led to people unnecessarily waiting at the hall of the restroom while one or more restrooms are available and overlap happening in the queue. In addition, it was observed that all the toilets being occupied happens only very sporadically. The overlap of people in the restroom also did mostly happen while not even all toilets were occupied. Therefore, the choice was made to not include the waiting for a free restroom mechanism, as this led to overall more logical behavior.

The final limitation is that due to there being some deviation from the activity schedules, it can happen that two people are already waiting at the office meeting table of another office, while the third person, to whom the office belongs, is still sitting at their desk for a few minutes. In real life, this seems unlikely to happen, as it is expected a person would stop working or start to wrap up their work once the first person is there for the meeting. This thus does not represent that behavior very accurately. However, it is not expected to influence the virus transmission very dramatically. This, because the person is still in the same room but roughly 1 meter further away than when sitting at the table and these minutes are only a small percentage of the entire duration of the 45-minute meeting.

9.2.3 Sensitivity and policy analyses

For the sensitivity and policy analysis, similar approaches are used to retrieve and analyze their results. In this section, some of the flaws and possible improvements relating to these aspects are mentioned for either the sensitivity analysis, policy analysis or both.

First of all, for both the sensitivity analyses and the policy analysis, the model is run the number of times needed per scenario/set-up to comply with the formula of Truong et al. (2015). However, these runs all had the same Nomad run and random seed. This entails that the employees had the same schedule each time, and the main difference between the runs is which employee is the infectious one. This choice was made as it cuts the runtime required down significantly given the limited time of this research. In addition, it cuts down the computing power and storage needed. In addition, due to the QVEmod part being different each time, and therefore the infected person being another person each time, there is still a fair amount of complexity in the results. However, to further support the reliability of the results, additional runs with variation in the Nomad runs and/or the random seed would be recommended. This way, it could be better understood how a difference in work schedule or randomness influence the virus spread in the office.

Next, it is important to note that in addition to the necessary complexity the varying of which employee is the infected one brings, it also increases the variance found in the results. The number of infections resulting from one employee being infected versus another employee, within the same policy can differ. This variance is therefore important for policy-makers and office managers to understand when choosing NPIs. It would be very useful to get more information on the exact reasons behind these differences in infections for different infected employees. Future work on this aspect is therefore recommended. In future work, a starting point would be to look into the schedules of the employees, and how the differences in schedules influence the number of infections. For example: Are there common activities that employees causing a high infection risk do more often than other employees? Or is there a specific location in the office they tend to come more often? A possibility would also be to compare these schedules to a map of the office that shows the distribution of virus particles in the office. Another possible direction for future work could be to test the NPIs with more than one employee being the infected one, as in this research it is only ever tested with one infected employee.

In addition, the higher variability caused by the different employees being the infected one causes less certainty in the effectiveness of the NPI policies based on the statistical tests. These statistical tests therefore provided less definitive results. Examples of these less definitive results are no proof the number of infections differ significantly between policies, insignificant values for the coefficient per NPI, the adjusted R-value being under the guideline of 85 % for set-up 14 (Bangdiwala, 2016) and less precise results for the OLS Linear regression. The final example of less precise results of the OLS linear regression is due to the high variance also leading to heterogeneity in the data. The presence of heterogeneity in data used for OLS Linear regression leads to less precise results of the analysis. In addition, there were also instances where there seems to be a clear difference between the number of infections, for example, a decrease of 10 %, no significant difference is found. A lower variability could have provided clearer results. For this research, this problem of high variability could have been mitigated by increasing the number of runs. The number of runs could for example be the number of runs required by Truong to have a percentage error of 1%, instead of the used percentage error of 5 %. In conclusion, either lower variance of more runs could lead to more insight into the NPIs that now did not always provide significant: increased ventilation, cleaning and restricting the number of people in the office. Performing the policy analysis again with either more runs or with an approach that causes there to be less variance would therefore be another recommendation for future work.

Next, a limitation focusing on the approach of the sensitivity analysis is that it currently only explores the influence of the amount and the size of the meetings, as meetings play a large role in virus transmission because of their larger group size and frequency. While looking into the impact of the meeting part of activity scheduling provides useful results, other parts of activity scheduling are not explored in this thesis research. Both the durations of the meetings and the probabilities used in the transition matrix have been solely based on the data gathering process of this thesis. Sensitivity testing or further verification of these variables would therefore provide additional value to the results of this thesis, but could not be incorporated in the limited time of this thesis. For future work, this would therefore also be a good starting point.

Finally, for the policy analysis, only two meeting set-ups are used, based on their average number of infections, to test the impact of the NPIs on: the set-up of which the number of infections is closest to the overall average number of infections based on all set-ups and the highest risk set-up in terms of infections. However, in case of unlimited time and resources, it would have been best to test the impact of the NPIs on all of the set-ups created. This would provide more robust insight into the impact of the NPIs in a broader context. The priority in this would be to use a set-up in which the initial number of infections is very low, to learn more about which NPIs could have unexpected side effects, that could possibly lead to an increased infection risk in such situations. Another interesting set-up to use next would be a set-up in which there is a relatively high number of infections, but in which the size of the meetings is not very large. This would be interesting as the size of meetings was shown to be the most important indicator of the number of infections. Leaving the size of the meetings out of the equation would provide more focus on other aspects with regards to effects of the other NPIs.

9.2.4 Results

Finally, some limitations related to the results of the thesis research should be noted and recommendations for future work are discussed. In section 9.1, it was found that the effectiveness of mask-wearing found it this model is significantly higher than the effectiveness often found in literature. Different reasons could be the cause of this, such as the high build-up of virus particles in the office and the long duration of an office day that is simulated in the model. Further research is needed to learn more about the cause of the difference in effectiveness. As the PeDViS model also produces information on the amount of virus at each location over time, a video could be made about the amount of fomites per location over time. This way it can be seen how the virus particles build up, spread through the building and possibly infect employees.

Second, in the model, the impact of the NPIs can only be tested for premade set-ups for the number and size of meetings, given various specification such as the size, layout and general behavior in the office. The layout of the office or the duration of activities cannot yet me easily adjusted in the model. To provide office managers with advice on NPI use suitable to their offices, the next step for a model could be to not provide such set-ups, but to allow office managers to input data specific to the activities in and lay-out their office, to test the impact of the various NPIs on and allow comparison, similarly to what the SamenSlimOpen (SSO) webtool now offers.

However, it could be expected that this model and the finding based on this specific office context, can already prove useful to a broader set of office spaces. The size of meeting room meetings and the frequency of these meetings and smaller meetings held in offices are simulated over a wide range of values, allowing the model to simulate the virus transmission for different types of behavior in offices. As high maxima and low minima were taken, it is not expected that many offices have values outside these ranges. For example, it seems unlikely offices often have meetings of more than 14 people. In addition, the crowdedness at other locations in the offices such as the coffee area and restroom area is not expected to differ much from much smaller/larges offices, as the number of coffee machines and restrooms is usually based on the number of people working in an office. Therefore, the activity scheduling of the model provides a good starting point for offices of different sizes, provided with a different office layouts. However, due to certain choices made for the conceptual model, such as having no shared workspaces and not including certain activities such as having lunch, the number of infections found in the model could diverge from the number found in certain offices.

Next, it should be noted that this model currently only includes 5 NPIs, while there are many more NPIs available for office managers. It would thus be in the interest of office managers and office managers if there was a larger assortment of NPIs to implement in a model. This could be an additional next extension of the model, combined with the input of data specific to offices as mentioned earlier.

A final limitation is that the impact of the NPI of restricting the percentage of people allowed at the office to 50 % has a strongly varying effect on the number of infected. It is expected that the inconsistent effectiveness of this NPI is caused by the fact when this NPI is present, the people that remain at the office have even more meetings per day, as they plan all their meetings to be on the days they are at the office. However, this is a strong assumption that seems to heavily influence the impact of this NPI. Therefore, it would be better in future research to first of all learn if this assumption can be supported by real-life data, based on observations or desk research. Second, additional simulations with the number of meetings per person staying the same if fewer people are allowed in the office would be beneficial and provide a broader insight into this NPI.

9.3 Insights for policy-makers

In this section, insight for policy-makers based on this research are briefly presented.

- The right NPI(s) can have a significant impact in reducing virus transmission in an office. Some NPIs in this research, such as wearing a mask or reducing the size of meetings, consistently showed to have a significant impact on reducing virus transmission. However, other NPIs showed no or inconsistent effectiveness in decreasing the number of infections, such as increasing the cleaning frequency or limiting the number of people allowed at the office respectively. Therefore, using the right NPI to reduce virus spread is important.

- Although some NPIs showed better performance than others, a policy that obligates the implementation of one single NPI without considering the context or providing some specifications is not recommended. The context relates to two aspects. First of all, to the other NPIs already present. Second, the context related to the size of the office and the behavior in the office relating to the frequency and group size of meetings. Some examples of what could happen when obligating the use of an NPI without further specification or considering the context are given below.

  – The obligation to have a certain ventilation rate in offices cannot guarantee a certain level of safety. This is because some NPI like ventilation showed to be less effective if more other NPIs are already present. This would therefore not give uniform results among offices.

  – Setting a limit on the maximum size of meetings is more effective in big offices with large meetings or in offices where more meetings are held. Reducing the number of people allowed to 50 % of the capacity of the meeting rooms would therefore have a different effect in different offices.

  – An NPI for which specification of the implementation is especially needed is limiting the number of people allowed at the office. This NPI is shown to have a varying effect on the number of infections based on the context in which it was implemented, such as the NPIs present and the number and size of meetings. It is recommended that when COVID-19 restriction, such as 50 % of employees having to work at home is enforced, some additional specifications are given. Such specifications could be a maximum group size per room in the office or a maximum meeting frequency or size.

- The effectiveness of one or more NPIs was not only shown to depend on the behavior in an office and the other NPIs present, but also on which of the employees is infectious and what their schedule for that day is. As this is something that cannot be predicted or controlled, this uncertainty on the number of infections given an NPI policy should be kept in mind. No guarantees can be given about the effectiveness of any set of NPIs.

- The NPI of increasing the cleaning frequency was shown to not have much effect. As similar results were found in research regarding the role of fomites, it could be argued that the role of fomites in virus transmission is limited. The recommendation is therefore for policy-makers to not focus on measures to prevent virus transmission via virus particles on surfaces and hands, such as cleaning or handwashing policies. Focusing on measures that prevent virus transmission through the routes of aerosols or droplets is advised.

- Further government funded research into the spread of COVID-19 in offices and the effectiveness of prevention measures is advised, as there are still many unknown factors. Two important subjects for further research the government could facilitate are mentioned here. First, further research into the role of other office activities in virus transmission, such as working in shared work spaces, during lunch or when getting coffee is recommended. Second, other NPIs than the ones in this research are also available (mask while sitting, walking directions, group gathering coffee/toilet etc.). Funding for models allowing an even broader range of NPIs and easy comparison between them is recommended.

9.4 Contribution of this research

This thesis aims to make different scientific contributions to fill in various knowledge gaps, as stated in section 2.3.4. The knowledge gaps found in the field of the estimation of indoor virus transmission for offices are:

- *Most models in current research only allow for the implementation of one or two NPIs, if any at all.*
- *In addition, the results of these models still have relatively high uncertainty for two reasons.*
    - *The limited extent of movement modelling they incorporate.*
    - *The models are not explicitly designed for measuring infection risk in real-life office scenarios, or even in office spaces in general.*

Various scientific contributions have been made through this thesis. First of all, a virus transmission model focusing on an office context has been developed. In current literature, not many virus transmission models focusing on offices were found, especially not including extensive movement modelling. For this thesis, observations are held in offices, in addition to desk research to gain insight into office behavior, and this behavior has been included in the model in the form of different activities employees can perform. In addition, the locations, durations and order in which these activities are performed are all based on the data gathering performed in this thesis. The impact of some aspects of the activities relating to the office on the number of infections has been explored through sensitivity testing. This also contributes to the knowledge of the role of meetings, particularly their size and frequency, on the infection risk in an office. This contributes to the knowledge gaps of the high uncertainty of using the existing transmission model for virus transmission estimations, as this model is suited for an office context, while also containing movement modelling.

Another scientific contribution the model provides is that within the model, the effects of more than one NPI are included and analyzed. In addition, it is also possible to include all possible combinations of these NPs to learn how they perform in combination with one another and to provide office managers with more options to keep their offices safe. In the current literature on indoor virus transmission, only the effects of one or two NPIs were included, if any at all, while this model includes up to 5 NPIs. In addition, the thesis also provides a potential strategy that can be used to compare the effectiveness of the NPIs, both separately and in combination with one another. This strategy consists of various statistical tests and considerations regarding the acceptable numerical values required for an NPI to provide significant or acceptable prevention of virus transmission risk. Finally, the filling in of these research gaps allows the final scientific contribution the regarding conclusion about the effectiveness of the NPIs for certain office contexts. Various numerical results for the effectiveness of the 5 NPIs included in this thesis have been given, both in the context of other NPIs present and in two different contexts regarding the frequency and size of meetings in the office.

In addition to scientific contributions, this thesis aims to provide societal contributions. This is done by providing insight into the effectiveness of implementing COVID-19 containment policies in offices. Application of these insights into implemented policies can in turn contribute to reducing the number of afflicted by the COVID-19 virus, through either sickness or even death, through infection at the office. Second, providing offices with a way to choose the right NPIs to safely stay open, allows offices to keep operating which is both economically advantageous and psychologically advantageous for their employees who experience physical drawbacks from having to work at home all day.

9.5 Suitability with CoSEM program

This research thesis is conducted in the Complex System Engineering and Management (CoSEM) master program for the following reasons. First, it addresses both the technical perspective of the assessment models and the social perspective in considering people's behavior to configure these simulations, making it a study of a socio-technical system. The technical component in this thesis can be found in the modelling through the use of coding with Python. Several technical issues are addressed, both in implementing certain aspects into the model and technical issues still present in the existing model and their impact.

In addition, this thesis research represents the values of the CoSEM program of working towards a clear design, the design in this case being a comparative model for COVID-19 containment strategies of offices and the design of various COVID-19 prevention policies. The modelling approach used to design the model contains various Systems Engineering processes, such as conceptualization and implementation. In the overall approach used in this thesis, various methods learned during the CoSEM master program are used, such as agent-based modelling, policy design and analysis and multi-criteria analysis (for selecting the NPIs).

Finally, the subject of this thesis covers both values from the public and private domains. The private domain is represented in the focus on offices and helping office managers keep their offices open. The model is already partly tailored toward use by office managers, as they can adjust some of the settings to make the model better represent their office. The public domain is represented by also focusing on how this model and the knowledge on the effectiveness of COVID-19 prevention measures retrieved from it can help policy-makers. Policy-makers also strive for offices to be safely open due to the positive effects on the economy and people's wellbeing. Therefore, they are interested in knowing which advice or guidelines they should give offices regarding their prevention measures. The government's guideline of focusing on the R-value is therefore also used to let the results of this thesis better translate into acceptable policies that can be used by policy-makers.

10 Conclusion

It is important that a safe way to keep offices open amid the pandemic is found. Non-Pharmaceutical Interventions (NPIs), such as wearing masks and social distancing, are an important way to do this. However, to pick the right NPIs for one's office, knowledge is needed of the effects of the different NPIs. However, as of now, there are not many models that can provide a detailed estimation of COVID-19 spread and quantify the impact of NPIs in office spaces. Based on the state-of-the-art of models for this purpose, various knowledge gaps were found:

- Most models in current research only allow for the implementation of one or two NPIs, if any at all.
- In addition, the results of these models still have relatively high uncertainty for two reasons.
    - The limited extent of movement modelling they incorporate.
    - The models are not explicitly designed for measuring infection risk in real-life office scenarios or even in office spaces in general.

Therefore, the purpose of thesis research is to help offices safely stay open by providing a model that allows for the comparison of the effects of various Non-Pharmaceutical Interventions (NPIs) to contain COVID-19 in real-life office settings, which fills these knowledge gaps. Four sub-questions are answered to eventually answer the main research question. The first sub-question states: *What is the typical setup of office spaces and the behavior of the people working there?* A conceptual model for virus transmission and NPI comparison is made, based on data gathered on the typical setup of office spaces and the behavior of its employees. Resulting output of this conceptual model includes the design of the office fitting for 40 employees that can get coffee, go to the restroom, work in their office and attend meetings. In addition, the NPIs to be tested and thus included in the model were chosen. The chosen NPIs are: (1) increasing the cleaning frequency, (2) increasing the ventilation rate, (3) obligate all people to wear a mask while walking, (4) limit the number of people allowed at the office to 50% and (5) limit the number of people allowed at the meeting rooms to 50% of its capacity.

The next step in this research was the implementation of this conceptual model into a computerized model. This answers sub-question 2: *How to quantitatively model the behavior of people working in an office and the resulting virus spread accordingly?* The result was a working model in Python, that can simulate employee behavior and the virus transmission resulting from this behavior. In addition, the NPIs chosen in the previous phase can be implemented in the model.

Before this model was used to test the impact of the NPI, model experimentation was performed. Here, the model parameters, related to the part of activity scheduling focusing on the meetings held in the office, are explored in terms of their effect on virus transmission. With sensitivity analyses, the third sub-question was then answered: *What is the influence of the parameter values, related to the part of activity scheduling focusing on meetings, on the virus spread?* The impact of (1) the size of meetings held in meeting rooms and the frequency of both (2) meetings held in offices and (3) meeting rooms on the number of infections were explored. Although the variance in the number of infections could not fully be described by the three parameters, they were shown to influence the number of infections. Of the three meeting variables, the number of infections is found to be most sensitive to the number of meeting room meetings held on a day. However, based on the values deemed representative for these variables, the group size of the meeting room meeting is found to have the biggest overall impact. This is due to it having a broader value range (up to 14 people in a meeting versus up to 8 meetings on a day). The frequency of office meetings was found to also impact the infection numbers, but it has a lower impact than the other two variables. This could potentially be explained by the smaller group size of 3 people in the office meetings. This smaller size could have two effects. First, it leads to fewer direct infections during the meeting itself due to the fewer other people a person comes into contact with during the meeting. Second, it leads to less build-up of virus in the office meeting areas, as there is a smaller probability the infected person had a meeting in the same area earlier. Besides the effect of the smaller group size, office meetings also have a shorter duration than meetings in meeting rooms, which also makes them less risk prone.

The final sub-question states: *What is the effect of the NPIs on the office space scenarios being simulated?* Using policy analysis, it is attempted to learn more about the effectiveness of the different NPIs policies in the office situation simulated in the model. These policy analyses are performed in two different contexts relating to meetings. The first context is one with a lot of meetings, which causes a high number of infections. The second context has a number of meetings that is observed to be more average, which causes a number of infections that is the closest to the average of all number of infections found in different contexts. Based on these analyses in the two contexts, some conclusions could be drawn on the effect the various NPIs had both on their own and in combination with the other NPIs present in the model. Among other things, it was found that mask-wearing is the most effective NPI, causing a substantial decrease in infection numbers in all tested situations. The NPIs of increased ventilation and a maximum capacity for meeting rooms often showed to provide a decrease in the number of infections, but not in all situations. For increasing the cleaning frequency, no significant decrease in infections was found. Finally, for limiting the number of people allowed in the office, an inconsistent effect on the number of infections was found. This is expected to be due to the combined effect of less people being present in the office that can be infected and the increased number of meetings the people present in the office attend, which causes them to be in contact with more people.

Next, it was found that if the goal is to keep the Reproduction value (R-value) of COVID-19 in the office below 1, at least one of the following NPI policies needs to be present

- All employees wear masks while walking

or

- The maximum capacity of the meeting room is set to 50 %. This has to be combined with the NPI of increasing the ventilation to the government-recommended levels in contexts that give a reason for a higher initial infection risk (such as many meetings held on a day)

When the goal is to the R-value of worst-case scenario below 1, the mask-wearing NPI always needs to be present in any combination of NPIs.

The answers to these sub-questions have led to an answer to the main research question of this thesis:

What is the effect of Non-Pharmaceutical Interventions on COVID-19 spread in office spaces?

This research has tested the effect of a selection of five NPIs: increased cleaning, mask-wearing while walking, increased ventilation and limiting the number of people allowed in the office and the meeting rooms to 50%. The effect of these NPIs was tested for specific office settings, chosen for this model, such as the lay-out, size and behavior. The exact outcomes for the effects of the NPIs are therefore based on these specific office settings, with various limitations existing in the research, as discussed in section 9.2.

Besides the applicability of the model, there was also some uncertainty found in the results due to variability of the number of infections within an implemented NPI policy. This variability is due to it playing a big role in the virus transmission which employee is infectious and what their schedule is. Therefore, there is some uncertainty in stating the effectiveness of the policies. In conclusion, there is both uncertainty in the results and limited generalization. Therefore, it should be kept in mind that the results of this research on the effectiveness of NPIs and factors influencing the number of infections cannot be guaranteed to translate to or be present in every other office situation.

However, despite these potential differences between offices implementing virus spread and uncertainty on the effectiveness of NPIs, some general observations on the use of NPIs in the offices can be given. First of all, obligating employees to wear a mask while walking seems very effective, even more effective than shown in other research. A possible explanation for this could be that the model simulates virus transmission for a longer period of time than most other models, simulating an entire work day each run. This leads to a higher amount of virus build-up, which can be spread through the entire office due to there being no doors modelled. These factors could lead to the increased effectiveness of wearing masks while walking, as employees then have to walk through different areas with a high build-up of virus particles at the end of the day.

Second, it was found that the size of meetings in meeting rooms plays a large role in the number of infections. This finding is in line with the result that the NPI of restricting the size of these meetings often proves to be effective in reducing infection numbers. Based on this, it seems like restricting the size of meetings, especially preventing very large meetings, would be a good way for offices to lower virus transmission.

Next, increasing the ventilation was shown to often bring down the number of infections, although this effect is limited when other effective NPIs are present or when the initial situation of the office does not have a particularly high infection risk. In both cases, this means there are less virus particles present in the air. Therefore the increased ventilation can bring about a lower absolute decrease in virus aerosols and thus in the amount of virus people are exposed to.

Cleaning was shown to have limited effect in reducing virus transmission in these offices. This finding is expected to also apply to other offices because based on literature, virus transmission via virus particles on surfaces and hands seems to play a small role in virus transmission. This in extend means cleaning these surfaces often cannot be expected to have much effect.

Finally, it was found that limiting the number of people allowed to come to the offices can have inconsistent effectiveness. This inconsistent effect is expected to be due to the increased number of meetings the employees present in the offices have when this NPI is present. The modelled increased number of meetings per employee is based on the assumption that when people can only come to the office on certain days, they plan more of their meetings on the days they are at the office. In other research, limiting the number of people in a building did reduce the number of infections significantly. This was due to the lowering of the density of people in these cases, which was not the case in the office situation of this thesis. Therefore, it is advised that when this NPI is implemented, employees get encouraged to do their meetings online when at home and to set a limit for the size and frequency of meetings at the office. This should ensure a lower density of people in the office instead of a potentially higher one. Also, due to the interaction effect found of the NPI with other NPIs, it is advised that this NPI is implemented in combination with other NPIs, to reduce the virus spread during meetings.

It should be noted that besides these NPIs, there are more NPIs that could be implemented in offices, that are not included in this research and the effects of which can therefore not be described. For the NPIs included in this thesis, various recommendations for future research are made to further explore their effectiveness. Ideally, a webtool is made that is similar to the SamenSlimOpen webtool (Samen Slim Open Project, 2022). This web tool should then be tailored towards office spaces and allow users to enter the specifics of their office and compare the effectiveness of various NPIs.

Finally, this thesis helps to fill in the knowledge gaps and provides scientific contributions in the field of indoor virus transmission in offices in various ways. First of all, this thesis provided data on office layouts and employees' behavior through a data gathering process. Second, a method was provided to create a computerized model to simulate these aspects and their impact on virus transmission. In addition, some insight into important factors and recommendations for further steps are given. Finally, the research provides an example of a method to learn more about the influence of both different behavioral aspects and NPIs on virus transmission. Therefore, the results of this thesis hopefully contribute to offices being able to safely stay open by increasing the knowledge of the effects of various NPIs and allowing office managers tools to choose fitting NPIs to implement.

Reference list

Adam, D. C., Wu, P., Wong, J. Y., Lau, E. H. Y., Tsang, T. K., Cauchemez, S., Leung, G. M., & Cowling, B. J. (2020). Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. Nature Medicine, 26(11), 1714–1719. https://doi.org/10.1038/s41591-020-1092-0

Alvarez Castro, D. & Ford, A. (2021). 3D agent-based model of pedestrian movements for simulating COVID- 19 transmission in university students. *ISPRS International Journal of Geo-Information*, *10*(8), 509-509. https://doi.org/10.3390/ijgi10080509

Antonini,G., Bierlaire, M. & Weber, M. (2004). Simulation of pedestrian behaviour using a discrete choice model calibrated on actual motion data. *4th STRC Swiss Transport Research Conference*, *7*(2004), 249–258.

Asadi, S., Cappa, C. D., Barreda, S., Wexler, A. S., Bouvier, N. M., & Ristenpart, W. D. (2020). Efficacy of masks and face coverings in controlling outward aerosol particle emission from expiratory activities. Scientific Reports, 10(1), 15665.

Bandini, S., Gorrini, A. & Vizzari, G. (2014). Towards an integrated approach to crowd analysis and crowd synthesis: a case study and first results. *Pattern Recognition Letters*, *44*(2014), 16–29. https://doi.org/10.1016/j.patrec.2013.10.003

Bangdiwala, S. I. (2016). Understanding significance and p-values. *Nepal Journal of Epidemiology, 6*(1), 522–524. https://doi.org/10.3126/nje.v6i1.14732

Banks, J. (Ed.). (1998). *Handbook of simulation: principles, methodology, advances, applications, and practice*. John Wiley & Sons.

Bellomo, N. & Dogbe, C. (2011). On the modeling of traffic and crowds: A survey of models, speculations, and perspectives. *SIAM review*,*53*(3), 409–463. https://doi.org/10.1137/090746677

Bellomo, N., Piccoli, B. & Tosin, A. (2012). Modeling crowd dynamics from a complex system viewpoint. *Mathematical models and methods in applied sciences, 22*(2).

Bernardi, N. F., Snow, S., Peretz, I., Orozco Perez, H. D., Sabet-Kassouf, N., & Lehmann, A. (2017). Cardiorespiratory optimization during improvised singing and toning. *Scientific Reports, 7*(1), 8113.

Bland, W.L. (1999). Toward integrated assessment in agriculture. *Agricultural Systems 60*(3), 157-167. https://doi.org/10.1016/S0308-521X(99)00025-6

Blauw, S. (2020, April 29). Dit is het belangrijkste getal van deze epidemie – hoe wordt het berekend? De Correspondent. Retrieved August 11, 2022, from https://decorrespondent.nl/11149/dit-is-het-belangrijkste-getal-van-deze-epidemie-hoe-wordt-het-berekend/162248208386-a0a5fdf8

Brugnach, M., Neilson, R., & Bolte, J. (2005). A Sensitivity Analysis Method to Study the Behavior of Complex Process-based Models. *Ecological Modelling*, *187*(2021), 99–120. https://doi.org/10.1016/j.ecolmodel.2005.01.044

Burgio, K. L., Engel, B. T., & Locher, J. L. (1991). Normative patterns of diurnal urination across 6 age decades. *Journal of Urology*, *145*(4), 728–731. https://doi.org/10.1016/S0022-5347(17)38436-7

Campanella, M. C. (2016). *Microscopic modelling of walking behaviour*. Ph.D. thesis, Delft University Press.

Carnevale, C., Finzi, G., Pisoni, E., Volta, M., Guariso, G., Gianfreda, R., Maffeis, G., Thunis, P., White, L., Triacchini, G. (2012). An integrated assessment tool to define effective air quality policies at regional scale. *Environmental Modelling and Software 38*(2021), 306-315.

Cartenì, A., Di Francesco, L., & Martino, M. (2021). The role of transport accessibility within the spread of the Coronavirus pandemic in Italy. *Safety science*, *133*(2021), 104999-104999. https://doi.org/10.1016/j.ssci.2020.104999

CDC. (2020, November 17). Social Distancing - Keep a safe distance to slow down the spread. Centrum for Disease Control. Retrieved February 15, 2022, from https://stacks.cdc.gov/view/cdc/90522.

CDC. (2021, January 5). *Disinfecting your facility*. Centers for Disease Control and Prevention. Retrieved February 15, 2022, from https://www.cdc.gov/coronavirus/2019-ncov/community/disinfecting-building-facility.html

Central Bureau for Statistics. (2022, March 29). Wat zijn de economische gevolgen van corona? Retrieved April 30, 2022, from https://www.cbs.nl/nl-nl/dossier/cbs-cijfers-coronacrisis/wat-zijn-de-economische-gevolgen-van-corona-#:%7E:text=In%20het%20tweede%20kwartaal%20van,product%20(bbp)%20weer%20terug.

Chen, F., Gao, Y., Wang, Z., & Liu, Y. (2020). Analysis on Alighting and Boarding Movement Laws in Subway Using Modified Social Force Model. *Collective Dynamics, 5*, 307–315. https://doi.org/10.17815/CD.2020.64

Chin, A. W. H., Chu, J. T. S., Perera, M. R. A., Hui, K. P. Y., Yen, H.-L., Chan, M. C. W., Peiris, M., & Poon, L. L. M. (2020). Stability of SARS-CoV-2 in different environmental conditions. *The Lancet. Microbe, 1*(4), e145-e145. https://doi.org/10.1016/S2666-5247(20)30093-8

CIRES. (2020, June 25). COVID-19 Airborne Transmission Tool Available. CIRES. Retrieved on June 7, 2022, from https://cires.colorado.edu/news/covid-19-%20airborne-transmission-tool-available

Colaneri, M., Seminari, E., Novati, S., Asperges, E., Biscarini, S., Piralla, A., ... & Vecchia, M. (2020). Severe acute respiratory syndrome coronavirus 2 RNA contamination of inanimate surfaces and virus viability in a health care emergency unit. *Clinical Microbiology and Infection, 26*(8), 1094-1-95.

Coleman, K. K., Tay, D. J. W., Tan, K. S., Ong, S. W. X., Than, T. S., Koh, M. H., Chin, Y.Q., Nasir, H., Mak, T.M., Chu, J. J. H., Milton, D. K., Chow, V. T. K., Tambyah, P. A., Chen, M., & Tham, K. W. (2022). Viral load of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in respiratory aerosols emitted by patients with coronavirus disease 2019 (COVID-19) while breathing, talking, and singing. *Clinical Infectious Diseases, 74*(10), 1722-1728. https://doi.org/10.1093/cid/ciab691

Cuevas, E. (2020). An agent-based model to evaluate the COVID-19 transmission risks in facilities. *Computers in biology and medicine, 121*, 103827.

Cui, Z., Cai, M., Xiao, Y., Zhu, Z., & Yang, M. (2021). Forecasting the Transmission Trends of Respiratory Infectious Diseases with an Exposure-Risk-Based Model at the Microscopic Level. *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS.* https://doi.org/10.36227/techrxiv.17046236.v1.

Dai, E., Ma, L., Yang, W., Wang, Y., Yin, L., & Tong, M. (2020). Agent-based model of land system: Theory, application and modelling framework. *Journal of Geographical Sciences*, *30*(10), 1555–1570. https://doi.org/10.1007/s11442-020-1799-3

Dai, H., & Zhao, B. (2020). Association of infected probability of COVID-19 with ventilation rates in confined spaces. *Building Simulation 13*(6), 1321-1327. https://doi.org/10.1007/s12273-020-0703-5

Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P., & Meester, L. E. (2005). *Modern Introduction to Probability and Statistics : Understanding Why and How*. Springer.

Delval, T., Sauvage, C., Jullien, Q., Viano, R., Diallo, T., Collignan, B., & Picinbono, G. (2021). *A BIM-Based Approach to Assess COVID-19 Risk Management Regarding Indoor Air Ventilation and Pedestrian Dynamics,15*(4), 187-196. https://www.researchgate.net/publication/351632327_A_BIM-Based_Approach_to_Assess_COVID-19_Risk_Management_Regarding_Indoor_Air_Ventilation_and_Pedestrian_Dynamics/link/60a27c78a6fdcc21dfdd2d35/download

Deng, W., Bao, L., Gao, H., Xiang, Z., Qu, Y., Song, Z., Gong, S., Liu, J., Liu, J., Yu, P., Qi, F., Xu, Y., Li, F., Xiao, C., Lv, Q., Xue, J., Wei, Q., Liu, M., Wang, G., … Qin, C. (2020). Ocular conjunctival inoculation of SARS-CoV-2 can cause mild COVID-19 in rhesus macaques. *Nature Communications, 11*(1), 4400.

Di Francesco, M., & Rosini, M. D. (2015). Rigorous derivation of nonlinear scalar conservation laws from follow-the-leader type models via many particle limit. *Archive for rational mechanics and analysis*, *217*(3), 831-871. https://doi.org/10.48550/arXiv.1404.7062

D'Orazio, M., Bernardini, G. & Quagliarini, E. (2021). A probabilistic model to evaluate the effectiveness of main solutions to COVID-19 spreading in university buildings according to proximity and time-based consolidated criteria. *Building simulation. 14*(6), 1795-1809. https://doi.org/10.1007/s12273-021-0770-2

Dowell, S. F., Simmerman, J. M., Erdman, D. D., Wu, J. S. J., Chaovavanich, A., Javadi, M., ... & Ho, M. S. (2004). Severe acute respiratory syndrome coronavirus on hospital surfaces. *Clinical infectious diseases, 39*(5), 652-657.

Drewnick, F., Pikmann, J., Fachinger, F., Moormann, L., Sprang, F., & Borrmann, S. (2021). Aerosol filtration efficiency of household materials for homemade face masks: Influence of material properties, particle size, particle electrical charge, face velocity, and leaks. *Aerosol Science and Technology, 55*(1), 63–79. https://doi.org/10.1080/02786826.2020.1817846

Duives, D., Chang, Y., Sparnaaij, M., Wouda, B., Boschma, D., Liu, Y., ., Yuan, Y., Daamen, W., de Jong, M., Teberg, C., Schachtschneider, K., Sikkema, R., van Veen, L. & ten Bosch, Q. (2022). The multi- dimensional challenges of controlling SARS-CoV-2 transmission in indoor spaces: Insights from the linkage of a microscopic pedestrian simulation and virus transmission models. Preprint MedRxiv. https://doi.org/10.1101/2021.04.12.21255349

Dutch Central Government. (2021a, June 7). *Compliance and support | Coronavirus Dashboard | Government.nl*. Corona Dashboard. Retrieved June 3, 2022, from https://coronadashboard.government.nl/veiligheidsregio/VR19/gedrag

Dutch Central Government. (2021b, February 2). *Routekaart coronamaatregelen bestuurdersversie* [Press release]. https://open.overheid.nl/repository/ronl-d7d0a6f9-78a9-4695-a3a3-89905155f439/1/pdf/Routekaart%20coronamaatregelen%20bestuurdersversie.pdf

Dutch central government. (2022a, February 7). *Sterfte | Coronadashboard | Rijksoverheid.nl*. Coronadashboard. Retrieved from: https://coronadashboard.rijksoverheid.nl/landelijk/sterfte

Dutch central government. (2022b, February 7). *Positief geteste mensen | Coronadashboard | Rijksoverheid.nl* [Press release]. Retrieved from: https://coronadashboard.rijksoverheid.nl/landelijk/positief-geteste-mensen

Dutch central government. (2022c, February 7). *Vaccinaties | Coronadashboard | Rijksoverheid.nl* [Press release]. Retrieved from:https://coronadashboard.rijksoverheid.nl/landelijk/vaccinaties

Dutch Central Government. (2022d, June 27). *Compliance and support | Coronavirus Dashboard | Government.nl*. Corona Dashboard. Retrieved July 3, 2022, from https://coronadashboard.government.nl/veiligheidsregio/VR19/gedrag

Dutch Central Government. (2022e, June 27). *Compliance and support | Coronavirus Dashboard | Government.nl*. Corona Dashboard. Retrieved July 3, 2022, from https://coronadashboard.government.nl/veiligheidsregio/VR19/gedrag

Dutch Central Government. (2022f). Reproduction number | Coronavirus Dashboard | Government.nl. Corona Dashboard. Retrieved August 1, 2022, from https://coronadashboard.government.nl/landelijk/reproductiegetal

Dutch National Institute for Health and Environment. (2020, April 3). *Generiek kader coronamaatregelen* [Press release]. Retrieved on March 7, 2022, from https://www.rivm.nl/coronavirus-covid-19/kaders-coronamaatregelen/generiek-kader

Dutch National Institute for Health and Environment. (2022, July 28). *Hoe rekenmodellen bijdragen aan de bestrijding van COVID-19*. RIVM. Retrieved August 20, 2022, from https://www.rivm.nl/coronavirus-covid-19/onderzoekrekenmodellen

Dzau, V. J., Kirch, D., & Nasca, T. (2020). Preventing a parallel pandemic—a national strategy to protect clinicians' well-being. *New England Journal of Medicine, 383*(6), 513-515.

EenVandaag. (2020, May). *Waarom je binnen meer risico loopt op een coronabesmetting dan buiten* [Press release]. Retrieved March 15, 2022, from https://eenvandaag.avrotros.nl/item/waarom-je-binnen-meer-risico-loopt-op-een-coronabesmetting-dan-buiten/

European Commission. (2020). *SME definition*. Official website of the European Commission. Retrieved March 24, 2022, from https://ec.europa.eu/growth/smes/sme-definition_en

Fang, Z., Huang, Z., Li, X., Zhang, J., Lv, W., Zhuang, L., Xu, X. & Huang, N. (2020). *How many infections of covid-19 there will be in the" diamond princess"-predicted by a virus transmission model based on the simulation of crowd flow*. In PrePrint. https://arxiv.org/abs/2002.10616

Fouda, B., Tram, H. P. B., Makram, O. M., Abdalla, A. S., Singh, T., Hung, I.-C., Raut, A., Hemmeda, L., Alahmar, M., ElHawary, A. S., Awad, D. & Huy, N. T. (2021). Identifying SARS-CoV2 transmission cluster category: An analysis of country government database. *Journal of Infection and Public Health*. *14*(4), 261-467. https://doi.org/10.1016/j.jiph.2021.01.006

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., . . . DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1–2), 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Hallett, S., Toro, F., & Ashurst, J. V. (2020). Physiology, Tidal Volume. In *StatPearls*. Treasure Island (FL): StatPearlsPublishing.

Harweg, T., Bachmann, D. & Weichert, F. (2021). Agent-based simulation of pedestrian dynamics for exposure time estimation in epidemic risk assessment. *Journal of Public Health*, (2020) 1-8. https://doi.org/10.1007/s10389-021-01489-y

Helbing, D., & Molnár, P. (1995). Social Force Model for Pedestrian Dynamics. *Physical Review, 51*(5), 4282–4286. https://doi.org/10.1103/PhysRevE.51.4282

Hinds, W. C. (1999). Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles. John Wiley & Sons.

Hofman, E. (2020) Long Covid. *TvPO, 17*(2022), 20–22. https://doi.org/10.1007/s12503-022-0927-

Hoogendoorn, S. P. & Bovy, P. H. (2004). Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological, 38*(2), 169–190.

Huang, Q., Mondal, A., Jiang, X., Horn, M. A., Fan, F., Fu, P., Wang, X., Zhao, H., Ndeffo-Mbah, M. & Gurarie, D. (2021). SARS-CoV-2 transmission and control in a hospital setting: an individual-based modelling study. *Royal Society Open Science*, *8*(3). https://doi.org/10.1098/rsos.201895

Investopedia. (2021, September 12). What Is R-Squared? Reviewed August 20, 2022, from https://www.investopedia.com/terms/r/r-squared.asp#toc-what-r-squared-can-tell-you

Islam, M. T., Jain, S., Chen, Y., Chowdhury, B. D. B., & Son, Y. J. (2021). An Agent-Based Simulation Model to Evaluate Contacts, Layout, and Policies in Entrance, Exit, and Seating in Indoor Activities Under a Pandemic Situation. *IEEE Transactions on Automation Science and Engineering*, (2021) 1-17. https://doi.org/10.1109/TASE.2021.3118008.

Julian, T. R., Leckie, J. O., & Boehm, A. B. (2010). Virus transfer between fingerpads and fomites. Journal of Applied Microbiology, 109(6), 1868–1874.

Klatsky, A. L., Armstrong, M. A., & Friedman, G. D. (1993). Coffee, tea, and mortality. A*nnals of Epidemiology*, *3*(4), 375–381. https://doi.org/10.1016/1047-2797(93)90064-B

Kraay, A. N., Hayashi, M. A., Hernandez-Ceron, N., Spicknall, I. H., Eisenberg, M. C., Meza, R., & Eisenberg, J. N. (2018). Fomite-mediated transmission as a sufficient pathway: a comparative analysis across three viral pathogens. *BMC infectious diseases, 18*(1), 1-13. https://doi.org/10.1186/s12879-018-3425-x

Kraay, A. N., Hayashi, M. A., Berendes, D. M., Sobolik, J. S., Leon, J. S., & Lopman, B. A. (2021). Risk for Fomite-Mediated Transmission of SARS-CoV-2 in Child Daycares, Schools, Nursing Homes, and Offices. *Emerging Infectious Diseases, 27*(4), 1229-1231. https://doi.org/10.3201/eid2704.203631.

Lansink, J. (2019, March 7). *Zoveel vierkante meter kantoorruimte heb je nodig per persoon*. SKEPP. Retrieved March 24, 2022, from https://skepp.nl/nl/blog/kantoorruimtetips/zoveel-vierkante-meter-kantoorruimte-heb-je-nodig-per-persoon

Lei, H., Xiao, S., Cowling, B. J., & Li, Y. (2020). Hand hygiene and surface cleaning should be paired for prevention of fomite transmission. Indoor Air, 30(1), 49–59. https://doi.org/10.1111/ina.12606

Li, C. Y., & Yin, J. (2021). A pedestrian-based model for simulating COVID-19 transmission on college campus. *Transport metrica A: Transport Science*, (2020), 1-25. https://doi.org/10.1080/23249935.2021.2005182

Li, S., Xu, Y., Cai, J., Hu, D., & He, Q. (2021). Integrated environment-occupant-pathogen information modeling to assess and communicate room-level outbreak risks of infectious diseases. *Building and Environment, 187*(2021), 107394. https://doi.org/10.1016/j.buildenv.2020.107394.

Ma, J., Qi, X., Chen, H., Li, X., Zhang, Z., Wang, H., Sun, L., Zhang, L., Guo, J., Morawska, L., Grinshpun, S. A., Biswas, P., Flagan, R. C., & Yao, M. (2020). Coronavirus Disease 2019 Patients in Earlier Stages Exhaled Millions of Severe Acute Respiratory Syndrome Coronavirus 2 Per Hour. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America. 72*(10), e652-e654 https://doi.org/10.1101/2020.05.31.20115154

Malleson, N, Heppenstall, A., & Crooks, A. (2018). Place-Based Simulation Modeling: Agent-Based Modeling and Virtual Environments. In H. N., Pontell, (ed.) *Criminology & Criminal Justice: Oxford Research Encyclopedias*. Oxford, UK: Oxford University Press.

Ministry of General Affairs. (2021, November 23). *1,5 meter afstand houden is weer verplicht*. Nieuwsbericht | Rijksoverheid.nl. Retrieved April 28, 2022, from https://www.rijksoverheid.nl/actueel/nieuws/2021/11/23/15-meter-afstand-houden-is-weer-verplicht#:%7E:text=Blijf%20de%20basisregels%20volgen,voor%20frisse%20lucht%20in%20binnenruimtes

Ministry of General Affairs. (2022a, January 28). *Thuiswerken in coronatijd* [Press release]. Retrieved om February 10, 2022, from https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/ondernemen-en-werken-in-coronatijd/thuiswerken-in-coronatijd

Ministry of General Affairs. (2022b, July 7). *Alles kan weer open in 3 stappen*. Nieuwsbericht | Rijksoverheid.nl. Retrieved June 3, 2022, from https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/nieuws/2022/02/15/alles-kan-weer-open-in-3-stappen

Ministry of General Affairs. (2022c, April 15). *Ondernemen en werken in coronatijd*. Coronavirus COVID-19 | Rijksoverheid.nl. Retrieved April 28, 2022, from https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/ondernemen-en-werken-in-coronatijd

Ministry of General Affairs. (2022d, February 22). *Overzicht coronamaatregelen*. Publicatie | Rijksoverheid.nl. Retrieved April 28, 2022, from https://www.rijksoverheid.nl/documenten/publicaties/2022/02/15/overzicht-coronamaatregelen

Ministry of Social Affairs and Employment. (2018, August 3). *Luchtverversing*. Arboportaal. Retrieved June 26, 2022, from https://www.arboportaal.nl/onderwerpen/luchtverversing

Moore, A. D., Holzworth, D. P., Herrmann, N. I., Huth, N. I., & Robertson, M. J. (2007). The Common Modelling Protocol: A hierarchical framework for simulation of agricultural and environmental systems. *Agricultural Systems*, *95*(1-3), 37-48.

Morawska, L., G. R. Johnson, Z. D. Ristovski, M. Hargreaves, K. Mengersen, S. Corbett, C. Y. H. Chao, Y. Li, and D. Katoshevski. (2009). Size Distribution and Sites of Origin of Droplets Expelled from the Human Respiratory Tract during Expiratory Activities. *Journal of Aerosol Science 40* (3), 256–269. https://doi.org/10.1016/j.jaerosci.2008.11.002

Motmans, R. (2017, October 16). NEN1824 – Kantoorruimte. Ergonomie site. Retrieved April 5, 2022, from https://www.ergonomiesite.be/nen1824-kantoorruimte/

MV Kantoor. (2018, May 25). *MV Kantoor*. MV Kantoortechniek B.V. Retrieved April 25, 2022, from https://www.mvkantoor.nl/nieuws/aantal-personen-en-afmeting-van-de-tafel/

Nasir, M., Nahavandi, S. & Creighton, D.(2012, June 10-15). *Fuzzy simulation of pedestrian walking path considering local environmental stimuli*. [Paper presentation] 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Brisbane, Australia. https://ieeexplore.ieee.org/document/6250764

National Institute for Public Health and the Environment,. (2019, December 31). *Algemene hygiënerichtlijn*. Retrieved July 3, 2022, from https://www.rivm.nl/hygienerichtlijnen/algemeen

National Institute for Public Health and the Environment. (2022, February 18). Variants of the coronavirus SARS-CoV-2 [Press release]. https://www.rivm.nl/en/coronavirus-covid-19/virus/variants

Nicas, M. (1996). An analytical framework for relating dose, risk, and incidence: an application to occupational tuberculosis infection. Risk Analysis: An Official Publication of the Society for Risk Analysis, 16(4), 527–538.

Nicas, M., & Sun, G. (2006). An Integrated Model of Infection Risk in a Health-Care Environment. *Risk Analysis: An Official Publication of the Society for Risk Analysis 26* (4), 1085–96.

NOS. (2021, December 11). "*Opschalen capaciteit is weg uit crisis", en drie andere misverstanden over de IC.* NOS.nl. Retrieved August 21, 2022, from https://nos.nl/artikel/2409039-opschalen-capaciteit-is-weg-uit-crisis-en-drie-andere-misverstanden-over-de-ic

NOS & I & O Research. (2020, November 19). *Bijna helft jongeren zegt zich slechter te voelen door Coronacrisis.* [Press release]. Retrieved February 20, 2022, from: https://nos.nl/artikel/2357164-bijna-helft-jongeren-zegt-zich-slechter-te-voelen-door-coronacrisis.html

NOS. (2022, January 18). *RIVM: nog altijd stijging besmettingen, steeds minder ziekenhuisopnames*. [Press release]. Retrieved August 21, 2022, from https://nos.nl/artikel/2413573-rivm-nog-altijd-stijging-besmettingen-steeds-minder-ziekenhuisopnames

Padilla, J. J., Shuttleworth, D., & O'Brien, K. (2019). Agent-Based Model Characterization Using Natural Language Processing. *2019 Winter Simulation Conference (WSC)*. https://doi.org/10.1109/wsc40007.2019.9004895

Popa, A., Genger, J.-W., Nicholson, M. D., Penz, T., Schmid, D., Aberle, S. W., Agerer, B., Lercher, A., Endler, L., Colaço, H., Smyth, M., Schuster, M., Grau, M. L., Martínez-Jiménez, F., Pich, O., Borena, W., Pawelka, E., Keszei, Z., Senekowitsch, M., … Bergthaler, A. (2020). Genomic epidemiology of superspreading events in Austria reveals mutational dynamics and transmission properties of SARS-CoV-2. *Science Translational Medicine, 12*(573). https://doi.org/10.1126/scitranslmed.abe2555

Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N., Jit, M., Klepac, P., Flasche, S., Clifford, S., Pearson, C. A. B., Munday, J. D., Abbott, S., Gibbs, H., Rosello, A., Quilty, B. J., Jombart, T., Sun, F., Diamond, C., … Hellewell, J. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet Public Health, 5*(5), e261–e270. https://doi.org/10.1016/s2468-2667(20)30073-6

Raza, A., Ali, Q. & Hussain, T. (2021). Role of knowledge, behavior, norms, and e-guidelines in controlling the spread of COVID-19: evidence from Pakistan. *Environmental Science and Pollution Research*, *28*(30), 40329-40345. https://doi.org/10.1007/s11356-020-10931-9

Romero, V., Stone, W. D. & Ford, J. D. (2020). COVID-19 indoor exposure levels: An analysis of foot traffic scenarios within an academic building. *Transportation Research Interdisciplinary Perspectives, 7*(2020). https://doi.org/10.1016/j.trip.2020.100185

Ronchi, E. & Lovreglio, R. (2020). EXPOSED: An occupant exposure model for confined spaces to retrofit crowd models during a pandemic. *Safety Science, 130* (2020), 104834-104834. https://doi.org/10.1016/j.ssci.2020.104834

Sajjadi, S., Hashemi, A. & Ghanbarnejad, F. (2021). Social distancing in pedestrian dynamics and its effect on disease spreading. *Physical Review E, 104*(1), 10-10. https://doi.org/10.1103/PhysRevE.104.014313

Salmenjoki, H., Korhonen, M., Puisto, A., Vuorinen, V. & Alava, M. J. (2021). Modelling aerosol-based exposure to SARS-CoV-2 by an agent based Monte Carlo method: Risk estimates in a shop and bar. *Plos one, 16*(11), e0260237. https://doi.org/10.1371/journal.pone.0260237.

Samen Slim Open project. (2022, April 12). *De tool*. SamenSlimOpen. Retrieved July 21, 2022, from https://www.samenslimopen.nl/de-tool/

Te Grotenhuis, M., & Van der Weegen, T. (2013). *Statistiek als hulpmiddel* (3rd edition). Koninklijke van Gorcum.

Tomar, A., & Gupta, N. (2020). Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Science of The Total Environment*, *728*, 138762. https://doi.org/10.1016/j.scitotenv.2020.138762

Truong, L. T., Sarvi, M., Currie, G., and Garoni, T. M. (2015). How many simulation runs are required to achieve statistically confident results: a case study of simulation-based surrogate safety measures. In 2015 IEEE 18th International Conference on Intelligent Transportation Systems, 274–278. IEEE.

TU Delft. (2020, October 1). *Corona Virus Information TU Delft*. Corona Virus Information TU Delft. Retrieved March 28, 2022, from https://www.tudelft.nl/2022/tu-delft/coronavirus

TU Delft. (2021, June 1). *Nomad Wiki*. GitLab. Retrieved May 11, 2022, from https://gitlab.tudelft.nl/users/sign_in#restaurant-scheduler

TU Delft. (2022, February 1). *Corona maatregelen TU Delft*. Retrieved April 28, 2022, from https://www.tudelft.nl/ewi/actueel/informatie-coronavirus/looproutes

Tuite, A. R., Fisman, D. N., & Greer, A. L. (2020). Mathematical modelling of COVID-19 transmission and mitigation strategies in the population of Ontario, Canada. *Canadian Medical Association Journal, 192*(19), 497–505. https://doi.org/10.1503/cmaj.200476

van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., Tamin, A., Harcourt, J. L., Thornburg, N. J., Gerber, S. I., Lloyd-Smith, J. O., de Wit, E., & Munster, V. J. (2020). Aerosol and surface stability of SARS-CoV- 2 as compared with SARS-CoV-19. *The New England Journal of Medicine, 382*(16), 1564–1567. https://doi.org/10.1056/NEJMc2004973

Verbeek, M. (2008). *A guide to modern econometrics.* John Wiley & Sons.

Voinov, A., Bousquet, F., (2010). Modelling with stakeholders. *Environmental Modelling and Software, 25*(11), 1268-1281. https://doi.org/10.1016/j.envsoft.2010.03.007

Vuorinen, V., Aarnio, M., Alava, M., Alopaeus, V., Atanasova, N., Auvinen, M., Balasubramanian, N., Bordbar, H., Erästö, P., Grande, R., Hayward, N., Hellsten, A., Hostikka, S., Hokkanen, J., Kaario, O., Karvinen, A., Kivistö, I., Korhonen, M., Kosonen, R., Österberg, M. (2020). Modelling aerosol transport and virus exposure with numerical simulations in relation to SARSCoV-2 transmission by inhalation indoors. *Safety Science, 130*, 104866.

Wang, X. (2021). Quantifying transmission risks of SARS-CoV-2 in pedestrian interactions at large events. Ph.D. thesis, Delft University Press.

Watkins, J. (2020). Preventing a covid-19 pandemic*. BMJ: British Medical Journal*, *368*(2020). doi: https://doi.org/10.1136/bmj.m810

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *Nature*, *393*(6684), 440-442.

WHO. (2021, January 25). *WHO Coronavirus Disease (COVID-19) Dashboard*. World Health Organisation. Retrieved February 18, 2022 from https://covid19.who.int/

Wilensky, U., & Rand, W. (2015). *An introduction to agent-based modeling*. The MIT Press.

Winther, B., McCue, K., Ashe, K., Rubino, J. R., & Hendley, J. O. (2007). Environmental contamination with rhinovirus and transfer to fingers of healthy individuals by daily life activity. *Journal of Medical Virology, 79*(10), 1606–1610.

Xiao, Shenglan, Yuguo Li, Minki Sung, Jianjian Wei & Zifeng Yang. (2018). A Study of the Probable Transmission Routes of MERS-CoV during the First Hospital Outbreak in the Republic of Korea. *Indoor Air 28 (1)*, 51–63.

Xiao, Y., Yang, M., Zhu, Z., Yang, H., Zhang, L. & Ghader, S. (2021). Modeling indoor-level non-pharmaceutical interventions during the COVID-19 pandemic: a pedestrian dynamics-based microscopic simulation approach. *Transport policy, 109* (2021), 12-23. https://doi.org/10.1016/j.tranpol.2021.05.004d

Xu, Q. & Chraibi, M. (2020). On the effectiveness of the measures in supermarkets for reducing contact among customers during COVID-19 period. *Sustainability*, *12*(22). https://doi.org/10.3390/su12229385

Zhang, N., & Li, Y. (2018). Transmission of Influenza A in a Student Office Based on Realistic Person-to-Person Contact and Surface Touch Behaviour. *International Journal of Environmental Research and Public Health*, *15*(8), 1699. https://doi.org/10.3390/ijerph15081699

Zhang, N., Chen, X., Jia, W., Jin, T., Xiao, S., Chen, W., ... & Kang, M. (2021). Evidence for lack of transmission by close contact and surface touch in a restaurant outbreak of COVID-19. *Journal of Infection, 83*(2), 207-216.

Zhao, Z., Li, X., Liu, F., Zhu, G., Ma, C., & Wang, L. (2020). Prediction of the COVID-19 spread in African countries and implications for prevention and control: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. *Science of the Total Environment, 729*. https://doi.org/10.1016/j.scitotenv.2020.138959

Zidek, J. V., & Van Eeden, C. (2003). Uncertainty, entropy, variance and the effect of partial information. *Lecture Notes-Monograph Series*, 155-167.

Zipf, G.K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley Press.

Zuo, Y. Y., Uspal, W. E., & Wei, T. (2020). Airborne Transmission of COVID-19: Aerosol Dispersion, Lung Deposition, and Virus- Receptor Interactions. ACS Nano. https://doi.org/10.1021/acsnano.0c08484

Appendix A data collection

This appendix explains the data gathering performed to gain the data used for the conceptualization. This data is used for the conceptualization and model implementation phases. In addition to the desk research performed and the personal experience used to set up the conceptual model, observation was performed and employee work scheduler were analyzed to gather data on office settings and behavior to adapt the model to an office setting.

Observations

Three observation locations were used. The first location, the third floor of Civil engineering was chosen to gather the information on the movement behavior with regard to the meeting room functional areas. The location lends itself very well to this objective, as it had a view on 4 meeting rooms and the nearby coffee area and toilets, also allowing observations on the movements between these areas. The second observation spot was the third floor at the TPM faculty, which contains the employee's personal offices, a coffee area and two bathrooms. Third spot was a spot at TPM, from which an entire meeting room meeting could be observed.

During the observations at the TU Delft, data on the following aspects was gathered: the size of meetings held in meeting rooms and in offices, the duration of meetings, the duration of getting coffee and using the restroom, the number of people present in a functional area at the same time, the frequency and the order in which employees do the different activities in an office. An overview of the data can be found in Table 48

Table 48: Number of people observed per functional area

Functional Area	Average value	Maximum Value	Minimum Value
Meeting room meetings	6.16	14	3
Coffee	0.22461	7	0
Restroom	0.306818	2	0

Also, for each activity space, the duration for which pedestrians visited them was recorded during the observations. The average time spend in the functional areas, based on this data, can be found in Table 49.

Table 49: Time spend in each of the functional areas (in minutes)

Activity space	Meeting room	Personal Office	Meeting area personal office	Coffee	Toilet
Mean	77.75	X	48.18	1.71	2.03

Order of activities

From the observation spot at TPM, the order in which people moved between the restrooms and coffee areas, and to some extent (for the individuals for which the personal office was close to the observations spot) between their personal offices and the coffee area and restrooms. At the faculty of Civil Engineering, the movement between the coffee area, nearby restrooms and the meeting rooms could be seen. Unfortunately, no good spot could be found to study the movements between personal offices and meeting rooms. All movements were written down and counted. It was observed that there is roughly an equal probability employees leave their office to get coffee as to use the restroom, with the probability they are getting to get coffee being slightly higher.

From the observations at Civil Engineering, it was concluded that most people coming out of the meeting rooms at the end of a meeting left the area straight away, not visiting the nearby coffee area and only very few going straight to the restrooms. However, based on the observation spot it could not be observed if they perhaps visited a restroom or coffee area further away and closer to their personal office. Visiting the coffee area or bathrooms coming out of the meeting rooms did happen frequently if a meeting had a small break, but breaks were observed only twice in the ten full meetings observed, being 20 percent of the time.

It was observed people mostly got coffee either before going to a meeting, but also often returned to their own office after getting some coffee. Going to the restroom after getting coffee was not observed often, which is also logical as this would be inconvenient to have to put down the coffee in the restroom.

At TPM, is was observed that people were twice as likely to return to their office as they were to get coffee after using the restroom. There were barely any instances observed that people walked from the restroom straight to a meeting in a personal office. The movement from the restroom to the meeting rooms was observed at Civil Engineering and can therefore not be directly compared to the other frequencies, but it was observed going straight from the restroom to the meeting room did not occur that often. However, as there was not one spot to observe people going from the restroom to either their personal office, meeting room, meeting in a personal office or coffee area, some degree of estimating is needed.

Not much data on the movements visiting each of the functional areas after coming out of a meeting in a person office was available, as only a few of these types of meeting were observed in real-life. Based on these few meetings, it was observed people usually left the office together and either got some coffee together or walked to their own office. In the few instances observed, it was not noted people went to the restroom or to another meeting straight after.

Work schedules

The work schedules for a period of two employees were analyzed, to learn additional information on the use of functional office spaces. Based on these schedules, data was gathered on the duration, size and frequency of both meetings in office and in meeting rooms. An overview of this data can be seen in Table 50. Among other things, the number of meetings employees have in meeting rooms on a daily basis was researched. Currently, in the aftermath of the COVID-19 restrictions, many meetings are still held online. In this research, it is assumed all meeting are held in real-life, to better represent a “normal” working situation. Based on a work schedules obtained, in which the number of people per meeting could also be seen, meetings consisting of less than 4 people meeting rooms. This led to the estimation of employees having 0.875 meetings in an meeting room per day, and 2.733 meetings in a personal office per day.

Table 50: Data found in the work schedules

	Average Value	Maximum value	Minimum value
Duration meeting room meeting	65.36	180	30
Duration office meeting	48.19	180	30
Size office meeting	2.27	3	2
Size meeting room meeting	8.312	22	4
Number of office meeting a person has per day	2.733	6	0
Number of meeting room meetings a person has per day	0.875	4	0
Frequency Personal office working space	4.83	x	x

Desk research

Frequency of visits to functional areas

To model the activity scheduling of office employees, data is also needed on the frequency of how often people visits each functional area, defined per day. As the observations were performed from one spot per observation space, no data was gathered on the behavior of one individual for one entire day. Therefore, the frequency of functional areas visited per workday is based on desk research. The desk research consisted of both a literature study and the received work schedules of the two office workers. The values for the frequency can be found in Table 51. Below a more detailed explanation of how every values was found is given.

First, the number of times the coffee area is visited is determined. According to Burgio, Bernard et al. (1991), 63 % of people drink coffee, drinking on average 3.1 cups per day. In addition, they found around 38 % of people drink tea, drinking on average 2 cups a day. This overlaps as people can drink both coffee and tea on a day. This translates to people in general drinking (3.1 x 0.63 + 2 x 0.38 =) 2,713 cups of coffee and/or tea on a day, the duration of a day in the research being from 6 am till 12 am (18 hours). To translate this to an 8 hour workday the value of 2,713 could simply be multiplied by 8/18, resulting in the value of 1.21 cups of coffee or tea during a workday. However, this value seems to be quite low based on personal experience. This could be due not every hour in the whole 18 hour span having an equal probability of people getting coffee, and the working hours (9am-5pm) have a relatively high probability of people consuming their coffee or tea then. Therefore it is assumed in this paper that people who drink coffee or tea have one of their coffee or tea drinks at home, during breakfast of after dinner, and the others at work. This would give an average of 2,713 – (0.63 x 1) = 2.08 cups of coffee or tea at work per employee, also meaning 2.08 visits to the coffee area per day. This value seems to be in line with the research of (Klatsky et al., 1993), which shows that the biggest group of people had 1-3 cups of coffee a day (so also outside of work hours).

Next, the frequency of bathroom visits during the workday was estimated. Burgio, Bernard et al. (1991) found that people go to the bathroom on average 5,1 times a day. Taking into account sleeping hours (for which a value of 7 was taken) and nocturnal bathroom visits, an average on 4.52 bathroom visits during the waking hours could be derived. The resulting average number of bathroom visits during the work they were then calculated as: 4.52 x (8/(24-7)) = 2.13.

Table 51: Frequency of visits to each functional area per workday

Activity space	Number of visits (per workday)
Toilet	2.3
Coffee	2.08

Overall values found based on all data gathering

Finally, the data found through both observation, analyzing work schedules and desk research can be combined to determine the final values to be used in this thesis. For some variables, a value is found through two different methods of data gathering. Therefore, for these values an overall value established. These values and the overall values established for them can be seen in Table 52. These values are calculated by taking the average of all of the data found. The total average could thus lay closer to the average found in one of the methods, if that average was based on more data samples.

Table 52: Values based on data from multiple methods

	Average	Maximum	Minimum
Duration meeting room meeting	68	180	20
Duration office meeting	48	180	3
Size meeting room meetings	7.14	22	3

Finally, an overview of all other data combined can be seen in Table 53.

Table 53: Overview of all other data found in data gathering process

	Average value	Maximum Value	Minimum Value
Duration getting coffee	1.71		
Duration restroom	2.03		
Number of people present restroom	0.306818	2	0
Number of people present coffee	0.22461	7	0
Number of office meeting a person has per day	2.733	6	0
Number of meeting room meetings a person has per day	0.875	4	0
Size office meeting	2.27	3	2
Frequency personal office	4.83	x	X

Appendix B layout/sizes

In this appendix, the determination of the sizes and layout of the different functional areas is described. To determine the layout, desk research was used, supplemented by the data found in the data gathering phase.

Personal offices

First of all, this was done for the personal offices. As the offices are individual, the same number of personal offices as maximum number of employees are needed, which is 40. The Dutch norm NEN 1824 regarding office spaces, states the square meter needed in an office given certain facilities being present. The minimum periphery of an office should be 4m^2, with an additional 1 m^2 added each time one of the following items is present: working surface with flatscreen/ read- and writing surface/a closet/a door. In addition, if the office has space for small meetings (less than 6 people), 1.5m^2 should be added per person (Motmans, 2017).

The choice was made to design offices that included a surface with a flatscreen and a read-and writing surface (a desk) a door and a space to meet with a maximum of three people. Therefore, according to the NEN 1824 the office should have a minimum surface area of 11.5 m^2. For simplicity purposes in designed the overall office, all offices in the layout are exactly 11.8 m^2, with 0.2 of additional width of the wall separating it from the next office. Based on the information provided by an employee alongside their work schedule, the knowledge was added that in general, meetings with less than four people were held in the meeting area of personal offices instead of in meeting rooms. As the meeting area in a personal offices has a different main function than the rest of the personal office, it is regarded as a separate functional area from here on out. The personal offices are thus all split in two functional areas: personal office, the place where employees work on their own, and the meeting area.

The layout of the personal office with meeting area can be seen in Figure 17. The sizes of the furniture are based on personal experience and recommendations for the sizes of certain office furniture (MV Kantoor, 2018) and are also represented to scale in Figure 17. In one half of the office, the “working area” of the employee can be found, with a desk and chair, where their screen and writing/reading area is, also having the required amount of space needed. In the other half, the meeting area can be found, which consist of a meeting table and three chair, also covering the minimum surface area required by law. As the offices are personal, it can be assumed it is always possible to hold the meeting in one own’s meeting area, and the occupancy rate thus doesn’t go over 100 %.

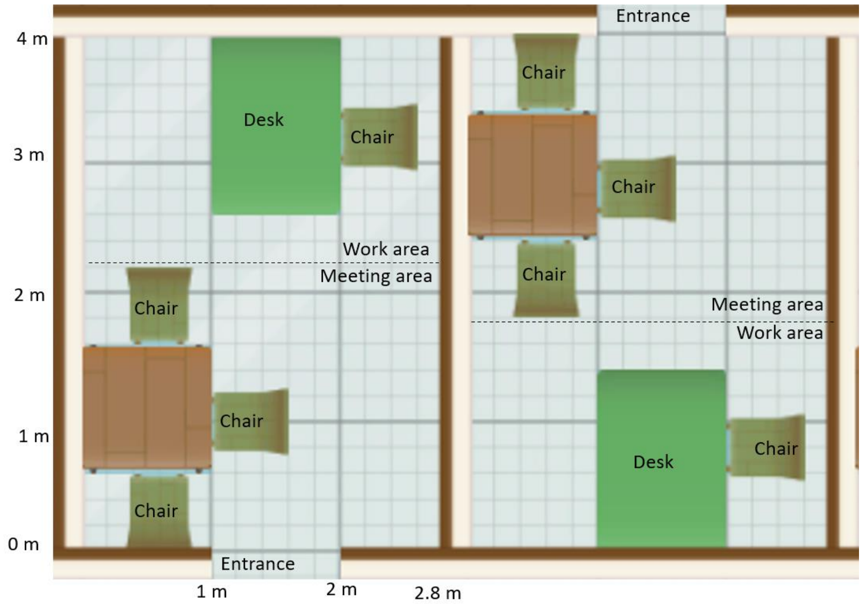

Figure 17: Layout of personal office with meeting area

The meeting rooms

For the meeting rooms, the NEN 1824 states that 2 m^2 of meeting room per office employee should be available at an office (Lansink, 2019). With a maximum of 40 office employees this would lead to a required minimum of 80 m^2 of meeting room at the office. Based on personal experience, many meeting rooms have a size of roughly 7 x 7 m, being around 50 m^2. Based on the analysis of employee schedules during the data collection phase, it was noted that the maximum number of people invited to a meeting was around 20. However, additional information given by the employees stated that often in the cases of big meetings, not all of the invited actually attended the meeting, and attendance was more around 10 people in these cases. In the observations of the meeting rooms, a maximum of 14 people in attendance was seen. Therefore, a maximum of 14 people needing to fit inside a meeting is taken as the guideline. According the NEN 1824 law, this would require a meeting room of at least 28 m^2. Combining all this information, it was envisioned that two meeting rooms could be placed in the layout, both being 30 m^2 (5 m x 6 m), so they are both able to host the maximum amount of employees to be expected.

To double check this number, it is calculated how much meeting room capacity is needed given the expected demand for the meeting rooms in an office of 40 employees. Based on agendas received by office employees, employees have a meeting in meeting rooms on average 0.88 times a day. In an office of at maximum 40 present at once this entails there are 0,88 * 40 = 35 instances in which someone is in a meeting during the day at a maximum. The average size of meetings in meeting rooms were observed to be 7.14. Therefore, dividing the number of times all of the employees are in a meeting during the day by the sizes of the meeting (35/7.14), means that there should be 4.9 meeting per day on average. The average time people spend in meeting rooms was observed to be 68 minutes. If the office only had one meeting room, this would lead the average time the meeting room is occupied to be 4.9 * 68 = 5.5 hours. This is lower than the number of hours the office is open, which is 8, so it could in theory be sufficient. However, an occupancy rate of 0.68 seems relatively high and it could be very likely that two meetings need to happen at the same time. In addition, the number and duration of meetings is not equal each day of the week, so the occupancy rate could be even higher on busy days. Therefore, having the earlier defined two meeting rooms still seems like a good number. The meeting rooms are identical, though mirrored, and their layout can be seen in Figure 18. The sizes of the furniture are also represented in the figure and are based on real life experience.

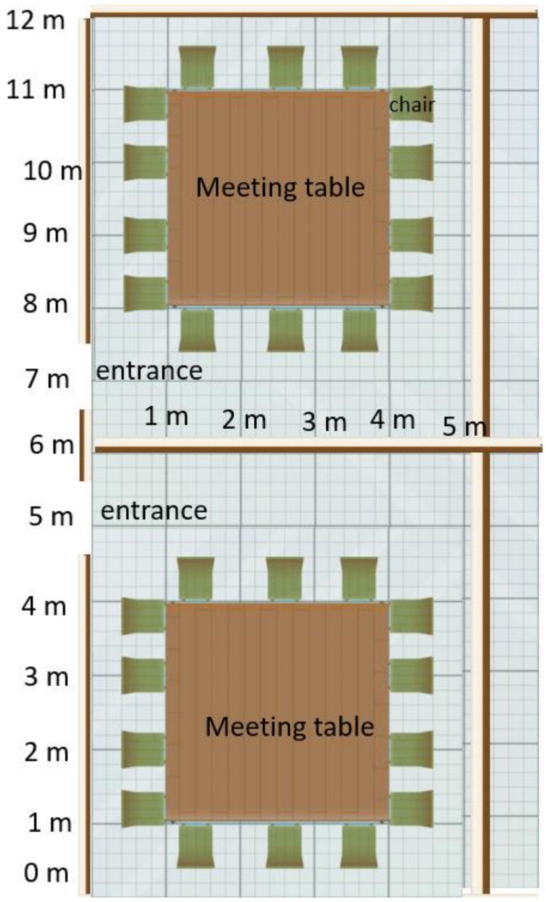

Figure 18: Meeting rooms layout

The restrooms

The Dutch law regarding office spaces (*de Arbo wet*), recommends there being at least one toilet for each 15 people working in the office, per gender. These would lead to 3 bathrooms being needed. As often offices have an even number of toilets, due to having gender separated restroom, the choice is made that the bathroom area consists of 4 bathroom stalls. The restroom area consequently has toilet stalls of 0.8 m x 0.8, and a hand washing area of 0.4 m x 0.4 m, the sizes based on the existing standardized sizes in the SSO webtool and personal experience. The layout of the restrooms at the office can be seen in Figure 19. Four bathroom stalls can be seen on the right and on the left four sinks are located. The sizes of the furniture are also represented in Figure 19 and based on real life experience.

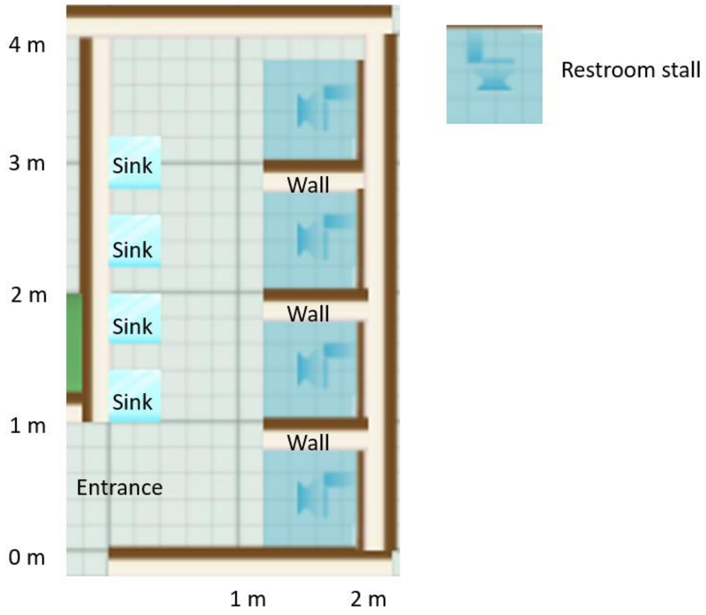

Figure 19: Layout of restrooms at the office.

The coffee area

Finally, the number of coffee areas and their size is determined. Based on the study of Burgio, Bernard et al. (1991) was derived that office workers on average get a drink (coffee or tea) 1,7 times during a workday. For 40 employees and a workday of 8 hours, this would mean there are 40 * 1,7 / 8 = 8,5 people/hour getting a drink at the coffee area. Based on observations, it was determined it takes on average 1,71 minutes to get a coffee (including the time people chat at the coffee area). This means that the coffee area is occupied 8,5 * 1,71 = 15,04 minutes/ hour, the coffee area then being occupied around 25 % of the time. This seems like an acceptable occupancy rate, especially since many people like to have a little talk with colleagues at the coffee area. It is thus decided only one coffee area is needed. The size of the coffee area is 2 m x 3 m, based on personal experience and the observed sizes of TU Delft's coffee areas for employees, used for comparison. The layout of the coffee area can be seen in Figure 20. The entrance of the area is located at the top left in the figure. At the bottom, the coffee and tea machine can be found. The sizes of the objects inside are also represented in Figure 20.

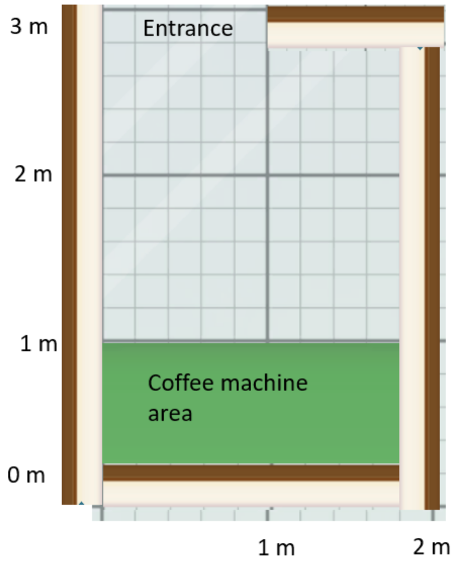

Figure 20: Layout of coffee area at the office.

Appendix C Mathematical framework

Mathematical framework underly the Nomad, QVEmod and risk identification parts of the PeDViS model. The parameterization of the models and the equations representing these models are given in this appendix.

Nomad

First of all, the walking behavior in the Nomad model is simulated, based on a couple of formulas, that are explained below

$$\frac{d}{dt}\vec{r}_p(t) = \vec{v}_p(t) \tag{1}$$

First of all, the velocity of a pedestrian, $\vec{v}_p(t)$, is the derivative of the position of vector of a pedestrian at time t, $\vec{r}_p(t)$.

$$\frac{d}{dt}\vec{v}_p(t) = \vec{a}(t) \tag{2}$$

Second, the acceleration of a pedestrian, $\vec{a}(t)$, is the derivative of their speed $\vec{v}_p(t)$.

$$\vec{a}(t) = \vec{a}_c(t) + \vec{a}_p(t) + \vec{\varepsilon}(t) \tag{3}$$

Third, the acceleration of a pedestrian is based on three components:

-their controlled acceleration, $\vec{a}_c(t)$

-their partly uncontrolled acceleration $\vec{a}_p(t)$

-and the noise term, $\vec{\varepsilon}(t)$, that simulates the natural fluctuations of pedestrian movements

$$\vec{a}_c(t) = \vec{a}_s(t) + \vec{a}_O(t) + \vec{a}_{pq}(t) \tag{4}$$

The controlled acceleration component $\vec{a}_c(t)$ is in turn the result of the pedestrian's desire for their preferred velocity $v_0(t)$ and physical interaction with surrounding objects and other pedestrians. These are represented by the following parameters, in that order:

-the path straying component $\vec{a}_s(t)$

-the obstacle interaction component $\vec{a}_O(t)$

-and the pedestrian interaction component $\vec{a}_{pq}(t)$

$$\vec{a}_s(t) = \frac{(v_0(t) \cdot \vec{e}_g) - \vec{v}(t)}{\tau} \tag{5}$$

The first of these components, the path straying component $\vec{a}_s(t)$, focuses on the increasing costs pedestrians find when straying from their optimal speed and direction (towards their destination). It is defined as the difference between their optimal speed velocity $(v_0(t) \cdot \vec{e}_g)$ and their current speed $\vec{v}(t)$. How fast the pedestrians can alter their speed and direction is represented by the constant acceleration time τ.

$$\vec{a}_{pq}(t) = -\vec{e}_{pq} \cdot A_0 \cdot e^{\frac{-d_{pq}}{d_i}} \quad (6)$$

The pedestrian interaction component simulates the costs of interaction between two individuals. The closer two individuals are, the larger the collision avoidance forces in opposite direction of the interaction. This is represented by formula 6, in which $\vec{e}_{pq}$ is the unit vector pointing in the direction of the other individual, A_0 represents the interaction strength, d_i the interaction distance and d_{pq} the anticipated distance.

$$\vec{a}_O(t) = -\vec{e}_O \cdot A_O \sum_{o \epsilon O} \begin{cases} 1 & \text{for } 0 < d_{pO} < d_0 \\ 1 - (d_{pO} - d_0) & \text{for } d_0 < d_{pO} < 2d_0 \\ 0 & \text{for } d_{p0} > 2d_0 \end{cases} \quad (7)$$

The strength of the interaction with an obstacle depends on the anticipated distance to the nearest object d_{pO}, the general interaction strength of objects A_O and the direction of the object $\vec{e}_O$. A step-based approach determines the influence of the distance. Here, close objects, within the distance threshold d_0 have maximum influence, while objects as far away as twice the threshold have no influence. Objects placed between d_0 and $2d_0$ have a gradual linear decline of obstacle avoidance force the further away they are. Using this formula, pedestrians are only reactive to obstacles that are very close.

QVEmod

The seven processes relating to QVEmod described in step 4, can be explained in detail by demonstrating the one or more formulas underlying them. As all of the seven processes that are detailed below are continuous, they are simulated in small time steps. Thus, the equations require small time steps (1 minute used in this application)

Process 1: Infectious individuals emit virus

Infectious individuals emit virus throughout the simulation. One portion of the virus is excreted in the infectious individual's hands and the rest is emitted into the air. Virus can be emitted through three respiratory activities: breathing, speaking or singing. These activities differ in the amount of virus they emit and the partition of aerosols and droplets.

The virus emission rate that an infectious agent *i* sheds into the air distributed over aerosols ($v^i_{emission\text{-}aerosols}$) and droplets ($v^i_{emission\text{-}droplets}$) are given by the following formulas:

$$r^i{}_{emission-aerosols} = \omega(1-\eta)\delta\sigma p_{aerosols}(1 - FE_{aerosols})\Delta t \quad (8)$$

$$r^i{}_{emission-droplets} = \omega(1-\eta)\delta\sigma p_{droplets}(1 - FE_{droplets})\Delta t \quad (9)$$

An individual's "standard" emission rate while spending half of the time breathing and half of the time talking is given by the parameter *ω*. The parameter *η* represents that proportion of virus excreted to the individual's hands, with *(1 – η)* therefore representing the proportion emitted into the air. The emission rates of the two respiratory activities are scaled relative to *ω* and can be set through the *δ.* The respiratory activities also differ in terms of the proportion of virus emitted in the form of aerosols (instead of droplets), which is represented through p_i. How infectious the individual is, is set by the individual scaler *σ*. Finally, the effect of masks in diminishing the amount of virus emitted is given by the mask filter efficiency *FEi.* The calculation only applies to cell (x, y) in which the individual is at the time t.

Process 2: Viral-laden droplets fall onto surfaces

Droplets can contaminate surfaces through the process of sedimentation, with the then contaminated surfaces called fomites. The formula below gives the rate of viruses-laden droplets transferring into fomites for cell (x, y) at time t ($v_{sedimentation}$). In this formula, $\mu_{droplets}$ represent the unit sedimentation rate of the droplets.

$$r_{sedimentation}(x,y,t) = V_{droplets}(x,y,t)\mu_{droplets}\Delta t \quad \textbf{(10)}$$

Process 3: Virus decay in the air and on surfaces

The COVID-19 virus is assumed to undergo an exponential decay. As the decay rate differs in aerosols and on different surface materials, two formulas are used to represent the virus decay.

Viruses in aerosols lose infectivity at a constant rate, with the air change rate (ACH) having an increasing impact on their decay. Virus-laden droplets are assumed to fall on surfaces rapidly (equation above). Therefore, the decay of droplets is assumed negligible. The virus decay on fomites happens at a constant rate, which depends on the material of the fomite. The formulas below show the aerosol decay ($v_{decay\text{-}aerosols}$) and fomite decay ($v_{decay\text{-}fomites}$) for cell (x, y) at time t. Here, $\mu_{aerosols}$ and $\mu_{fomites}$ represent the respective decay rates of aerosols and fomites.

$$r_{decay-aerosols}(x,y,t) = V_{aerosols}(x,y,t)(1 - e^{-\mu_{aerosols}\Delta t - ACH\Delta t}) \quad \textbf{(11)}$$

$$r_{decay-fomites}(x,y,t) = V_{fomites}(x,y,t)(1 - e^{-\mu_{fomites}\Delta t}) \quad \textbf{(12)}$$

Process 4: Virus-laden aerosols and droplets diffuse in the air

The diffusion of virions in the air, in the x and y direction at time t, is represented for aerosols ($v_{diffusion\text{-}a}$) and droplets ($v_{diffusion\text{-}d}$) in formulas 13 and 14 respectively.

$$r_{diffusion-a}(x,y,t) = D\frac{(V_a(x-\Delta x,y,t) + V_a(x+\Delta x,y,t) + V_a(x,y-\Delta y,t) + V_a(x,y+\Delta y,t) - 4V_a(x,y,t))\Delta t}{\Delta x \Delta y} \quad \textbf{(13)}$$

$$r_{diffusion-d}(x,y,t) = D\frac{(V_d(x-\Delta x,y,t) + V_d(x+\Delta x,y,t) + V_d(x,y-\Delta y,t) + V_d(x,y+\Delta y,t) - 4V_d(x,y,t))\Delta t}{\Delta x \Delta y} \quad \textbf{(14)}$$

Here, *D* is the diffusion coefficient per cell size per time step. The assumption is made the particles are well mixed in the volume of the grid cell (the first half of the upper part of the formula), after which aerosols diffuse. Δx and Δy represent the length units of the cells.

Process 5: Susceptible individuals inhale air with viral-laden droplets and aerosols

Susceptible individuals can get exposure to the virus by inhaling part of the airborne viruses, consisting of virus-loaded aerosols (*$v_{aerosol(x,y,t)}$)* and virus-laden droplets (*$v_{droplets(x,y,t)}$)* from the cell they are in (x, y) at time t.

A susceptible individual's accumulated inhaled amount of virus in the forms of aerosols and droplets is represented by $E_{aerosols}(t)$ and $E_{droplets}$,. This inhaled amount is given by the summation of the inhaled amount of viruses ($v^s_{inhalation\text{-}aerosols}$ and $v^s_{inhalation\text{-}droplets}$) .

The inhalation of virus in both the form of aerosols and droplets is the ratio of the human tidal volume per time step, represent by inhalation rate *ρ,* over the cell volume (*L*).In addition, the effect of a person wearing a mask is given by FE_i, representing the mask filter efficiency.

$$r^s{}_{inhalation-aerosols}(t) = V_{aerosols}(x, y, t)\frac{\rho}{L}(1 - FE_{aerosols})\Delta t \quad \textbf{(15)}$$

$$r^s{}_{inhalation-droplets}(t) = V_{droplets}(x, y, t)\frac{\rho}{L}(1 - FE_{droplets})\Delta t \quad \textbf{(16)}$$

$$E^s{}_{aerosols}(t) = \sum_0^t r^s{}_{inhalation-aerosols}(t) \quad \textbf{(17)}$$

$$E^s{}_{droplets}(t) = \sum_0^t r^s{}_{inhalation-droplets}(t) \quad \textbf{(18)}$$

Process 6: Infectious individuals contaminate surfaces

Infectious individuals can contaminate surfaces by touching them, transferring the virus on their hands to the surface (Winther et al. 2007).

In this thesis, surfaces are assumed to be touched by proximate individuals at a constant rate, if the surface is within the reachable distance of 0.5 meter of the infectious individual *i*. For a grid cell (x, y) that contains a surface element, the surface contamination rate at time t ($v^i_{contamination}(x,y,t)$ is given by the transfer efficiency *ϑ*, the ratio of finger pads surface relative to the contaminated area *π* , the touching frequency ϒ and the amount of virus on the infectious agent's hands (V^i_{hand}) . V_{hand} is set to zero at timestep 0 and is based on the emission rate (*ω*) times the proportion of the virus that gets excreted into the individual's hands (*η*) instead of into the air

The decreasing rate (due to decay or transfer) of V^i_{hand} is assumed to be constant throughout the event, resulting in V^i_{hand} also being constant throughout the event.

$$r^i{}_{contamination}(x, y, t) = V^i{}_{hand}(t)\gamma\theta\pi\Delta t \quad \textbf{(19)}$$

$$V^i{}_{hand}(t) = V^i{}_{hand}(0) = \omega\eta \quad \textbf{(20)}$$

Process 7: Susceptible individuals touch virus on surfaces

Virus exposure through fomites happens when susceptible individuals pick up virus on their hands by touching fomites and then sending it to their facial membranes. First, a certain amount of virus transfers from the fomites to a susceptible individual's hands ($v^s_{pick\text{-}up}$), when they touch contaminated surfaces. This amount is calculated based on the amount of virus on the surface ($V_{fomites}$) the transfer efficiency ϑ, the ratio of finger pads surface relative to the contaminated area π, and the touching frequency Υ. The virus on this individual's hands (V^s_{hand}) is accumulated over time. Finally, the individual's exposure through fomites is calculated as the amount of virus on his hands times the proportion of virus that is assumed to transfer from hands to facial membranes (ε).

$$r^s{}_{pick-up}(t) = V_{fomites}(x, y, t)\gamma\theta\pi\Delta t \quad \textbf{(21)}$$

$$V^s{}_{hand}(t + \Delta t) = V^s{}_{hand}(t) + r^s{}_{pick-up}(t) \quad \textbf{(22)}$$

$$E^s{}_{fomites}(t) = \sum_{0}^{t} V^s{}_{hand}(t)\varepsilon\Delta t \quad \textbf{(23)}$$

Risk identification

In the QVEmod part, an individual's exposure is calculated over three transmission routes. A susceptible individual's exposure E^s is a value that is scaled as the emission rate ω, which influences the exposure, is also scaled to the value one. Therefore, the exposure is multiplied by the emission rate of an average infectious individual (φ).

The relationship between this exposure and the risk of getting infected is different between the various transmission routes (Deng et al., 2020). This relationship is modeled by the following formula, using the exponential dose-response relationship (Nicas, 1996).

$$P^s = 1 - e^{-\left(\frac{\phi E^s{}_{aerosols}}{k_{aerosols}} + \frac{\phi E^s{}_{droplets}}{k_{droplets}} + \frac{\phi E^s{}_{fomites}}{k_{fomites}}\right)} \quad \textbf{(24)}$$

P^s represent the probability of a susceptible individual getting infected, $E_{aerosols}$, $E_{droplets}$, and $E_{fomites}$ are an individual's exposure over the different routes, and $k_{aerosols}$, $k_{droplets}$ and $k_{fomites}$ is the exposure per route at which individuals have a 63% probability of getting infected. K_{route} in turn depends on two parameters, the relationship between them shown in formula 25.

$$k_{route} = \frac{D_{inf}}{\phi * c_{route}} \quad \textbf{(25)}$$

Here, the exposure per route k_{route} depends on the infectious dose D_{inf}, and the proportional of viral particles a person is exposed to that reach the respiratory tract cells c_{route} also influences the exposure.

These formulas can thus be used to estimate an individual's exposure and the risk of someone getting infected. To determine if someone is infected, a random number is drawn from a uniform distribution [0,1] which is then compared this number to their infection risk P^s. An infection is assumed if the individual's infection probability is larger than the randomly generated number. The total number of infections is then the summation of the number of infections realized. Due to the stochastic nature of these calculations, these calculations have to be repeated often. The resulting mean number of infections can then be seen as the reproduction number of the event *R:* the number of infections resulting from one event with an infectious individual present.

Appendix D Nomad input

This appendix provide some additional details regarding the file used to provide the office layout.

The infrastructure is inputted into the model as an XML file. After initially creating this XML file using the SSO webtool, some final adjustment were made to it by hand. These adjustments included:

-Deleting the kitchen door, register and coatrack from the file. In the SSO webtool is was obligated to include these items. However, as these are not items present in the office, these later had to be removed by hand.

-Destinations can have the attribute of being 'virtual' in Nomad, which entails that for routing purposes this object is seen as an obstacle when it is not someone's destination at that time. As it could not be defined in the webtool whether or not an object is virtual, this later had to be changed by hand for some destinations.

For obstacles, the same was done for the 'see through' attribute, which entails that if true the obstacle does not obstruct the sightline of a pedestrian. In addition, the names of several locations in the XML file were altered to allow for an easier understanding of which exact destination they refer to and to make the activity scheduling easier. For this purpose, the offices, restrooms, meeting rooms and meeting room chairs were numbered. The numbering of the meeting room chairs can be seen in Figure 21. The chairs behind the desks in the personal offices were attributed the number corresponding to the offices, with the offices being numbered from left to right, starting with the highest row.

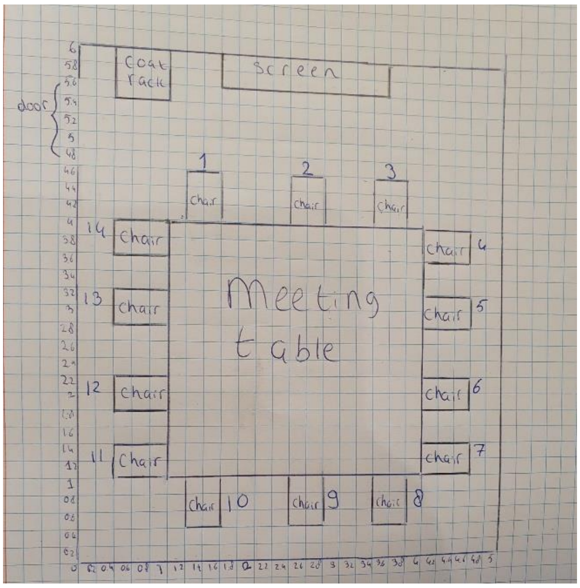

Figure 21: Numbering of meeting room chairs

Beside the infrastructure xml file, the Nomad part of the model also requires a file containing the pedestrian parameters to be used. The pedestrians parameters, a description of what they entail and the value(s) used can be found in Table 54.

Table 54: Important pedestrian parameters in Nomad and their values

Parameter	Description	Average value
V_0	Free speed, speed at which pedestrian would walk unhindered	Normal Gaussian distribution (mean 0.9 and stand. dev 0.2)
τ	Acceleration time	0.5 s
A_0	Interaction strength, intensity in which pedestrians are avoiding each other	2 m/s^2
d_i	Interaction distance, controls how sensitive avoidance accelerations are to the distance between pedestrians	0.4 m
d_0	Shy away distance	0.1 m

Appendix E source code and files

In this appendix, the GitHub link to the source code and files are given.

The GitHub link is the following:

https://github.com/EmmaDeijkers/Code-activity-scheduling/blob/main/Code%20Emma%20Deijkers%20for%20activity%20scheduling%20office%20PeDViS.py

First of all, the activity scheduling code made for this thesis can be found. This code is formatted in Python.

In addition, the following input files can be found in the link

-the infrastructure.xml (same file always used)

-the model_without_agents.json (used with different ventilation values)

-pedestrian_parameter_sets.xml (same file always used_

-simulation_scenario file (different files used for different set-ups or NPIs present, this file represent set-up 14, with no NPIs present)

-the scenario.json file (used with different values of mask wearing present, cleaning interval and ventilation)

Appendix F Activity Scheduling

In this appendix, some coding additional details for the activity scheduling in Nomad are given, for two aspects: the demand patterns and the walking time.

To incorporate these schedules into the Nomad model, a discrete demand pattern is used to model each individual separately, so each pedestrian can get their personal activity schedule assigned to them. The demand patterns are thus defined per office employee, and, therefore, consist of only one pedestrian each time. All of these demand patterns have the same source, the main entrance of the office, which is the sole source in the model. Also, each demand pattern has the same parameter set, the one specified in 6.1.2.

Incorporating this walking time of one minute between the activities is done in various ways, differing per activity. For both types of meetings, the duration of the office activities defined in Nomad is reduced by one minute, to allow them to be at their next activity after the meeting at the right time. For the activities of getting coffee and using the restroom, the reverse was done, making them one minute longer in the activity schedule than they are and how they are defined in Nomad, as this allowed for easier scheduling. For the activity of working in one's office for a certain number of minutes, using the one minute of working in one's office multiple times, after the final minute an empty spot is left in the schedule, to account for the time to walk to the next activity.

Appendix G Transition matrix

In this appendix, it is described how the values inside the transition matrix were determined.

Several transition matrix has been developed, that determine the order in which employees visit the different functional areas. The values in the transition matrix are based on the observations at the TU Delft, some data retrieved during the desk research part of the data collection and some assumptions. From the observation spot at TPM, the order in which people moved between the restrooms and coffee areas, and to some extent (for the individuals for which the personal office was close to the observations spot) between their personal offices and the coffee area and restrooms could be observed. At the faculty of Civil Engineering, the movement between the coffee area, nearby restrooms and the meeting rooms could be seen**.** All movements were written down and counted to compare the frequencies of movement between each pair of areas. This information, combined with general patterns seen in behavior during the observations and personal experience, were translated into the values found in Table 55. As the meetings employees have are not assigned to their work schedule based on transition matrices, but based on preplanned meeting, going towards meetings is not included in the matrices. Therefore, there are not column for going towards office meetings and meeting room meetings.

The cells representing movement between the same functional areas are excluded, as these are not actual likely relocations. It is highly unlikely that people go straight back to the restroom after just coming out of one for example. Another example: even if people have back to back meetings in meeting rooms, they often get some coffee or go to the restroom in between those meetings. By extension, it was assumed people do not go from any time of meeting (in a personal office or meeting room area) straight into another meeting.

Table 55: Full transition matrix

	Work in office	Coffee	Toilet
Work in office	0	0.5	0.5
Meeting room meeting	0.7	0.15	0.15
Coffee	0.75	0	0.25
Restroom	0.67	0.33	0
Office meeting	0.5	0.45	0.05
Entrance	1	0	0

First, the probability employees move towards each of the activity spaces, having their personal office as origin, was determined. It was observed that there is roughly an equal probability employees leave their office to get coffee as to use the restroom, with the probability they are getting to get coffee being slightly higher. The frequency of visits to each of the functional areas in Table 56 were also used.

Table 56: Frequency of visits to each functional area per workday

Activity space	Number of visits (per workday)
Toilet	2.3
Coffee	2.08

These values were converged to the probabilities of visiting the restroom and coffee area starting from the personal office, setting the probabilities of going to the coffee area or bathroom equal as they had been shown to be similar.

Next, the probabilities of going to each location, having a meeting room as origin, was determined. From the observations at Civil Engineering, it was concluded that most people coming out of the meeting rooms at the end of a meeting left the area straight away, not visiting the nearby coffee area and only very few going straight to the restrooms. However, based on the observation spot it could not be observed if they perhaps visited a restroom or coffee area further away and closer to their personal office. Visiting the coffee area or bathrooms coming out of the meeting rooms did happen frequently if a meeting had a small break, but breaks were observed only twice in the ten full meetings observed, being 20 percent of the time. Therefore, the probabilities of going to the bathroom or coffee area is roughly split into a probability of 10 % each, making the assumptions that all people either use the restroom of get coffee during a break, with an equal probability of choosing each activity. For each of the two locations, an additional 5 percent probability is added for the times it could to be observed they went to a restroom or coffee machine further away. This results in a 15 % probability of going to the restroom or coffee area after a meeting. For the resulting 70% of the time, it is assumed people return to their office after a meeting most of the time, as people often have a little spare time between meetings.

Third, the probabilities of having each area as a destination after getting coffee is estimated. It was observed people mostly got coffee either before going to a meeting, but also often returned to their own office after getting some coffee. Going to the restroom after getting coffee was not observed often, which is also logical as this would be inconvenient to have to put down the coffee in the restroom. This probability was therefore set at a low value of 10 percent. Of the resulting 90 percent, it was then roughly estimated that people are twice as likely to get coffee before heading into a meeting as they are before returning to their office, the split then being 60 and 30 percent. Finally, taking the heading to meeting rooms out of the equation as this is not included in the matrices, the probability of going to the restroom or one's after are reschedule to 25 % and 75 % respectively.

Next, the probabilities of going to each functional area after visiting the restroom were determined. This was based on a combination of the observations at Civil Engineering, from which visiting the coffee area or meeting room after using the restroom could be observed, and the observations at TPM, from which visiting the coffee area, returning to their office or going to a meeting in a personal office could be observed. At TPM, is was observed that people were twice as likely to return to their office as they were to get coffee after using the restroom. Therefore, the probabilities of going to one's office and getting coffee were set to 67 % and 33 % respectively.

Finally, the probabilities of visiting each of the functional areas after coming out of a meeting in a personal office were determined. Not much data on these movements was available, as only a few of these types of meeting were observed in real-life. Based on these few meetings, it was observed people usually left the office together and either got some coffee together or walked to their own office. In the few instances observed, it was not noted people went to the restroom straight after. Therefore, this probability is set to the very low value of 5 percent. As going to get coffee or going back to their own office was observed roughly at the same frequency, the resulting 90 percent is roughly split to 50 % and 45 % of going to either one's office or to get coffee.

Having the full transition matrix ready as stated in Table 55, adaptions of the transition matrix can be made for when an employee can choose from only two locations to go to, such as in Table 57, Table 58 and Table 59. The probabilities of these tables are derived from the probabilities in Table 55 .

Table 57: Transition matrix choice Coffee Toilet

	Coffee	Toilet
Work in office	0.5	0.5
Meeting room meeting	0.5	0.5
Coffee	0	1
Restroom	1	0
Office meeting	0.9	0.1

Table 59: Transition matrix choice Coffee Work

	Work in office	Coffee
Work in office	0	1
Meeting room meeting	0.82	0.18
Coffee	1	0
Restroom	0.67	0.33
Office meeting	0.5025	0.4525

Table 58: Transition matrix choice Work Toilet

	Work in office	Toilet
Work in office	0	1
Meeting room meeting	0.82	0.18
Coffee	0.75	0.25
Restroom	1	0
Office meeting	0.725	0.275

Appendix H QVEmod input

This appendix shows the values for some of the most important parameters for the QVEmod part of the PeDViS model. In Table 60, the various parameters of the Qvemod part are stated, along with the values they take and the sources on which these values are based. All of these values are the values used in the work of Duives et al. (2022) and are also used for the adapted model of this thesis.

Table 60: QVEmod parameters, their values and sources used

Parameter	Value	Source
Emission rate (ω)	Scaled to 1 unit per hour (Typical infectious individual, half of the time breathing and half of the time talking)	
Emission rate by average infectious individual (ϕ)	10^6 RNA copies per hour (used for informing dose-response relationships)	Ma et al. (2020)
Activity infectiousness scaler (δ)	1.86 singing 0.14 breathing 2.4 talking (relative to baseline)	Coleman et al., (2022)
Individual infectiousness scaler (σ)	1 (A typical infectious individual) 0 (Susceptible individual)	
Proportion of viruses emitted in the form of aerosols ($P_{aerosols}$)	0.978 (Breathing) 0.0652 (Singing) 0.171 (Talking)	Chen et al. (2021)
Proportion of pathogen exerced to hands (η)		Shuai Li et al. (2021) Kraay et al. (2018)
Transfer efficiency (θ)	0.25 per touch	Julian et al. (2010) Liu et al. (2013)
Ratio of finger pads size to the cells size (π)	0.0196	Calculated by Duives et al (2022) based on (Beamer et al. 2015; U.S. Environmental Protection Agency 2011; Wilson et al. 2021; AuYeung et al., 2008)
Fractional transfer rate from hands to facial membranes (ε)	0.01 per hour	Calculated by Duives et al 2022) based on: (Harvey et al. 2021; Pitol and Julian 2021; AuYeung et al.,2008; U.S. Environmental Protection Agency, 2011; Mark Nicas and Best, 2008; Kwok, Gralton, and McLaws, 2015 ; Pitol et al., 2017; Rusin et al., 2002;. Xiao et al. 2018; Wolff et al. 2005; Kraay et al., 2021; Weber and Stilianakis, 2008; Mark Nicas and Jones 2009)
Unit decay rate of viruses in aerosols ($\mu_{aerosols}$)	1.51 per hour	Van Doremalen et al. (2020)
Unit sedimentation (deposition) rate of droplets ($\mu_{droplots}$)	37.93 per hour (for baseline scenario of 50 % talking and 50 % breathing)	Value set by Duives et al 2022) based on the work of Vuorinen et al. (2020) Morawska et al. (2009)
Diffusion coefficient (D)	0.0016 m^2/s	Kudrayashova et al. (2021). Also a sensitivity analysis performed on the value by Duives et al (2022)

Decay rate of viruses on surfaces ($\mu_{surface}$)	Wood: 0.969 per hour, Plastic: 0.279 per hour,	Chin et al. (2020), Liu et al. (2021) and Van Doremalen et al. (2020).
Inhalation rate (ρ)	288 L per hour (breathing, talking) 432 L per hour (singing)	Calculated by Duives et al (2022) based on Hallett & Ashurst (2020) and Bernardi et al. (2017)
Volume of a cell (L)	125 L	
Infectious dose (D_{inf})	1000 RNA copies	Popa et al. (2020)
The proportion of virions reaching respiratory cells ($c_{aerosols}$, $c_{droplets}$, $c_{fomites}$)	10% (aerosols) 10% (droplets) 10% (fomites)	Zuo et al. (2020), Hinds (1999), Kraay et al. (2021), Adam et al. (2020). Comparison of different values have been performed by Duives et al (2022)

For a few variables, some additional explanation is given here. The emission rate represents the viral load an infector emits in an hour. The emission of a typical infectious individual per hour, of which he spends half of the time talking and half of the time breathing, is scaled to one. The emission of an individual having another distribution of respiratory activities, can be represented by scaling the emission rate up or down using the activity scaler. The values of the activity scaler have therefore specifically been chosen relative to the emission when both speaking and breathing half of the time. The cell volume L is not based on literature, but is calculated as 0.5 m x 0.5 m x 0.5 m = 125 L, as the cells have unit sizes of 0.5 m (Duives et al., 2021). The ratio of finger pads size over reachable surface is based on a calculation, between the average hand area (U.S. Environmental Protection Agency 2011; Beamer et al. 2015; A. M. Wilson et al. 2021), fraction of a hand used for transfer (AuYeung et al., 2008) and the reachable surface area. The researchable surface area is set to 0.5*0.5 = 0.25 meter squares based on the reachable distance assumption of 0.5 meters. The fractional virus transfer rate from hands to facial membranes has been defined as the aggregate value of four dimensions : (i) touching surface ratio between the hands and facial membranes, (ii) face touching frequency, (iii) virus transfer efficiency from hands to face and (iv) inactivation rate of virus on the skin. The values for these dimensions can be found in the literature stated in the table.

Appendix I NPIs

In this appendix, more details on the process of selecting the NPIs to be incorporated in the model are given.

As the main objective of this thesis is to study the effect of NPIs on COVID-19 spread, the effect of NPIs on behavior is an important aspect to consider. More than two years after the discovery of the COVID-19 virus, many NPIs to contain the virus have been found. To test such measures in an office setting, first, a selection should be made of the NPIs that are most common and/or useful in an office scenario. To initiate forming this selection, a list of NPIS was made, based on personal experience, existing NPIs in the PeDViS restaurant scenario (Duives et al., 2022). In addition, desk research is done into the NPIs that are or were effective at the TU Delft and NPIs that were/ are advised by the Dutch central government. This full list can be seen below.

- Have hand sanitizer available at entrance (Personal experience)
- Advice to wash hands often (Ministry of General Affairs, 2022c)
- Allow maximum number of people in office / area (Ministry of General Affairs, 2022d)
- Increase cleaning frequency (Duives et al., 2022)
- Increase ventilation rate (Duives et al., 2022)
- Obligation to wear mask while walking (TU Delft, 2020)
- Obligation to wear mask at all times (TU Delft, 2020)
- 1.5 meter distance (Ministry of General Affairs, 2021)
- Ask people to cough and sneeze in elbow (Ministry of General Affairs, 2022c)
- Walking routes (TU Delft, 2022)
- 50 % people working at home (Ministry of General Affairs, 2022b)
- Maximum groups size/no group gatherings (Dutch Central Government, 2021b)

These NPIs were than reviewed using a multi-criteria matrix, which can be seen in Table 61. Three criteria were used to review the NPIs. First of all the expected level of impact, that entails whether it is expect this NPIs can have a significant attribution to decreasing the spread of Covid. The second criteria is if the NPI already exists in the existing version of PeDViS, and therefore is easy to also incorporate in the experiments for this thesis, however it not yet existing it not a negative trait for an NPI per se. The third criteria is for the NPIs that are not already incorporated in the existing version of PeDViS, and reviews the level of difficulty adding the NPI to the model. The level of impact of each NPI were based on personal estimations, assumptions and sources. Whether something was already present in the model and how difficult it would be to at the NPI if not, was based on studying the received original restaurant version of the PeDViS model and personal communication with the developers of the model or people who have worked on the model and the papers they have written about it.

Table 61: Multi-criteria assessment of the NPIs

NPI	Expected level of impact	Already in model?	If not in model, level difficulty to incorporate
Hand sanitizer at entrance	Impact the transmission through fomites, but fomites is not the most important route so medium impact. Also would only lower the amount fomites at one point when they enter	No	Make the level of fomites on persons hand lower on entering for some, who use it

Advice to wash hand often	Fomites are not expected to have the highest impact on infection rate, as their contribution to transmission was small in the restaurant version (Duives et al., 2022). As this NPI focuses on fomites, the effect is expected to be relatively small. In addition, the impact is very similar to the previous NPI of having hand sanitizer at the entrance.	No	Add more handwashing to the personal schedules, and add the activity of handwashing to the model
Ask people to cough/sneeze in elbow	Reduces the probability of people having fomites on their hands, and then spreading these fomites by touching objects. Some impact, but limited due to fomites not being expect to be the most important source of infection.	No	Not feasible, coughing was disabled in the original version because the direction of emission was too hard to handle
Maximum number of people	Less people in the office would in general mean possible infections. Would also cause more space available per person and thus more distance and less infection risk	Yes, through demand	x
Frequent cleaning	Impact the transmission through fomites, but fomites is not expected to be the most important route so medium impact	Yes	x
Ventilation rate	"Increasing ventilation affects the amount of viral-laden aerosols" (Duives et al., 2022), which is an important transmission route	Yes	x
Face mask required while walking	Various types of masks can be used. Duives et al. (2022) found that for mask with 40 % filter efficiency worn while walking, virus exposure can be significantly reduced (20 %). Walking situations in which it would make a difference are crossing on another or talking while being in the same place	Yes	x
Face mask required at all times	Would extent the NPI mask required while walking to also wearing them in meetings, which would then also reduce the exposure there with 20 %	No	Can be realized by adjusting some lines of code in the model
1.5 meter distances	Has impact on virus transmission through aerosols and droplets, which are important transmission ways. Risk of infection very low if 1.5 meter distance	Yes, but not used in original version (Duives et al., 2022) and gave problems with pedestrians getting stuck in festival	Fixing the problems if pedestrians get stuck in corners due to the 1.5 distances is expected to bring quite some additional work

		adaption (Wang, 2021)	
Walking routes	Would reduce the emission through people being within 1.5 of one another, but transmission through walking through aerosol left behind by person walking in front of you still present	No	Possible to make walking routes in Nomad through waypoints and additional modelling, but would be quite some additional work
Maximum capacity of meeting room meetings	Expected to have a large impact, as large meeting are the biggest gathering of people in offices. Reducing the size could therefore make a big difference.	Existing model does not have meetings as it is not made for offices, therefore this NPI is also not present	The schedules of people would then be adjusted to simply not join the meeting, which results in them sitting in their office most of the time and maybe getting coffee or using the restroom

Chosen NPIs

Finally 5 NPIs were chosen that were expected to have considerable impact in reducing the number of infections and allowing for relatively easy incorporating into the model. These NPIs can be seen in Table 62. The NPIs that were chosen to be incorporated can be seen in Table 62. On the right column, in can be seen in which functional area(s) these NPIs have an effect.

Table 62: Overview of chosen NPIs and the functional area they affect

NPI	Functional area
Maximum number of people	Applies to whole office, as it effects the number of people coming in and people entire all functional areas during the workday
Frequent cleaning	Applies to whole office, effect on all functional areas except personal office, as only one employees touches that area
Ventilation rate	Applies to all office, and expected to have effect on all functional areas
Face mask required while walking	Applies when walking from one functional area to the next, and being in the coffee area and restrooms.
Maximum capacity of meeting room meetings	Only applies to and has effect on transmission during meeting room meetings, which in the functional areas of the two meeting rooms.

Appendix J Sensitivity analysis

This appendix provides some additional details on the sensitivity analysis, for both the model runs and the statistical tests performed on the results of these model runs.

Sensitivity runs

The simulator.py was run for each policy as simulator.py local 1 'nr' 1. The first numerical parameter of 1 entails that one instance of the Nomad model, due to it being stochastic, is used. This value was chosen both because the simulator did not yet work for other values, and that multiple Nomad runs takes more time than multiple QVEmod runs, while time was limited. The second numerical parameter, entails the number of QVEmod runs. The PeDViS model was thus run for 1 instance of Nomad, followed by various runs with QVEmod, for which each run had a different person being the infected one. The simulator works in a way it randomly picks the infected person each time. It was afterwards checked which agents/employees this were, also to check if the same agent/employee was not picked twice. The third and last numerical parameter of 1 determines the randomness, and is the random seed. This was set to 1, instead of letting it be randomly drawn each run, to ensure the only difference to be the infected agent/employee for each scenario. The same set-up was used for the policy analysis.

Requirements statistical tests

To perform a statistical tests, the data needs to comply with the requirements set for that specific test. In this section, the compliance of the data with the requirements for both the independent t-tests and the Ordinary Least Squares (OLS) Linear Regression are described.

To perform the t-test, two requirements need to be fulfilled. First, the variable to be compared needs to be of at least an interval scale. As the number of infections is on a ratio scale, this requirement is fulfilled. The second requirement is that the average values compared with one another can be stated with some certainty. Often the rule of normality is used for this (Te Grotenhuis & Van der Weegen, 2013), but as the certainty of the average value for the number of infections is calculated by the number of required runs, this requirement is already fulfilled.

Finally, it needs to be known whether the variances of the samples can be assumed to be equal, to know whether the t-test needs to be performed assuming equal or unequal variance (Dekking et al., 2005). Whether the variances of two set-ups can be assumed equal can be tested using Levene's test (Te Grotenhuis & Van der Weegen, 2013). This test is therefore done for each combination of set-ups, to consequently apply the fitting t-tests.

Next, the compliance with the assumptions for the OLS Linear Regression analysis need to be reviewed. These are:

1. The regression model is linear in the coefficients and the error term
2. Homoscedasticity of errors /The error term has a constant variance (no heteroscedasticity)
3. Observations of the error term are uncorrelated with each other
4. The error term has a population mean of zero Independence of observations
5. All independent variables are uncorrelated with the error term
6. No independent variable is a perfect linear function of other explanatory variables (Frost, 2022)

The first assumption of the linearity of the coefficients and error term is tested using a line diagram, seen in Figure 22. In the diagram the combined values of the independent variable is shown against the dependent values. It can be seen there that the shape of the relationship is inherently linear. Hence, the OLS regression analysis is fitting to perform.

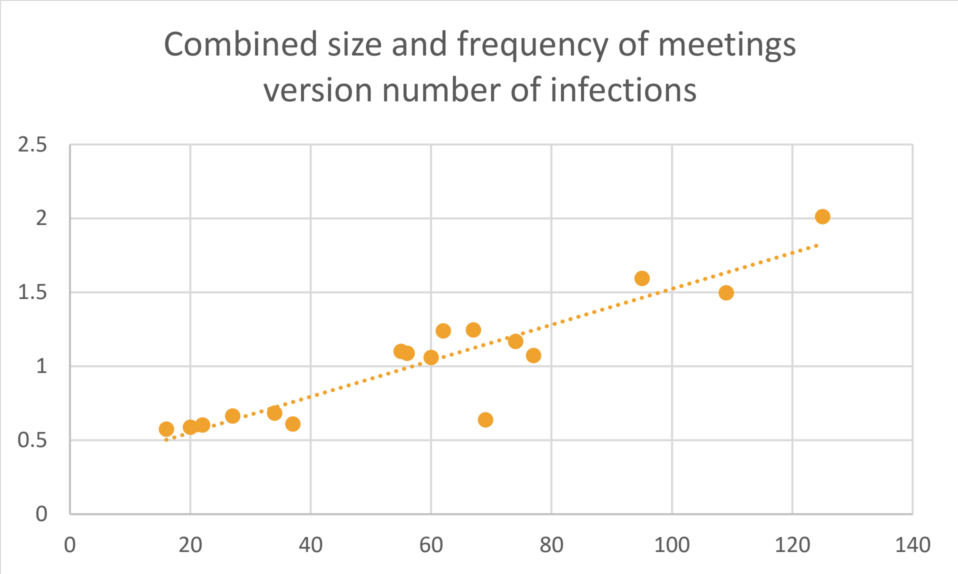

Figure 22: Plot of influence meeting variables on infections

According to the second assumption, for a model to be unbiased, the average value of the error term must be equal to zero. However, when a constant is included in the regression model, this assumption should be fulfilled (Frost, 2022). As the analysis that are performed with a constant included, this assumption should be complied with. Another way this is often tested is checking if the data seems to be equally distributed amongst the regression line (Te Grotenhuis & Van der Weegen, 2013). In Figure 22 this seems to be the case.

The third assumptions states that the observations of the error term are uncorrelated with each other. Such a correlation can be caused by simultaneity between the independent and dependent variables or a measurement error and can bias the coefficient estimate (Frost, 2022). As no error term was given, the correlation could not be measured. However, as measurement are based on an existing calibrated model measurement error are not assumed to be present. In addition, simultaneity between the dependent and independent variables is also not present, as the infections are calculated after all the meetings.

The fourth assumption states the observations must be independent from one another and the result of one observation should not affect the result of another observation. For this research, the observations are assumed to be independent, as the groups of employees representing each set-up are not the same, as in each set-up people have a different activity schedule that defines them and the set-ups do not influence one another but are created independently.

The fifth assumption states that there is no heteroscedasticity. This entails that the variance does not change for each or for some observations (Frost, 2022). As for each observation a variance could be found, this variances can be visually compared. Here, it could be seen that the largest variance was twice the size of other variances. Therefore, there could be some heteroscedasticity, which reduces the precision of the estimates in the OLS linear regression.

The final assumption is that no independent variable is a perfect linear function of other explanatory variables. This can be assumed to be correct as the three independent variables are completely independent from one another, the size of the meeting rooms does not in any way influence the number of meetings held in this model, or the number of office meetings held etc.

Statistical tests

In this appendix the statistical results of the Levene's tests and independent t-tests performed to compare the average number of infections of each combination of set-ups is given. The output of the Levene's tests, was used to perform the correct type of t-test. As a very large number of statistical tests are performed to compare all-setups, the results of these tests are summarized in the form of a matrix in Table 63. For the Levene's test, each cell in the matrix represents whether the set-up of the column and the set-up of the row are data samples for both set-ups are homogenous. In the states 'non', the combination is non homogenous, if the cell is empty the combination is homogenous.

For the matrix of the independent t-test, Table 64, each cell represent whether the average number of infections of set-up of the column and the set-up of the row significantly differ from one another. When the cells says 'diff' the set-ups significantly differ from one another (the P value was lower than 0.05). In case the p value is lower, the difference cannot be proven and the p value of the t-test is given in the cell.

Table 63: Visualized results Levene's test for sensitivity analysis

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	x		non	non	non	non		Non		non	non		non	non				non
2		x																
3			x				non	non	Non			non				non	non	Non
4	non			x												non		
5	non				x		non	non	non	non	non	non				non	non	non
6	non					x	non		non			non				non	non	non
7			non		non	non	x						non	non				
8	non		non		non			x					non	Non				
9			non		non	Non			x				non	non				
10	non				non					x		non	non			non		
11	non				non						x		non	non				
12			non		non	non				non		x	non	non				
13	non						non	non	non	non	non	non	x			non	non	non
14	non						non	non	non		non	non		x		non	non	non
15															x			
16			non	non	non	non				non			non	Non		x		
17			non		non	non							non	Non			x	
18	non		non		non	non							non	non				x

Table 64: Results independent T-tests for sensitivity analysis

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	x	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif	dif
2	dif	x	0.06	0.39	dif	0.61	0.50	0.41	0.44	dif	dif	dif	0.72	dif	dif	dif	dif	dif
3	Dif		x	dif	0.65	0.19	0.11	dif	dif	dif	dif	dif	0.09	dif	dif	dif	dif	dif
4	Dif		Dif	x	Dif	0.23	0.10	0.92	0.79	dif	dif	dif	0.84	dif	dif	dif	dif	dif
5	Dif	dif		dif	X	0.10	0.06	dif	dif	dif	dif	dif	dif	0.08	dif	dif	dif	dif
6	Dif					x	0.96	0.24	0.26	dif	dif	dif	0.50	dif	dif	dif	dif	dif
7	Dif						x	0.09	0.06	dif	dif	dif	0.45	dif	dif	dif	dif	dif
8	Dif		Dif		Dif			x	0.87	dif	dif	dif	0.88	dif	dif	dif	dif	dif
9	Dif		dif		Dif				x	dif	dif	dif	0.95	dif	dif	dif	dif	dif
10	Dif	dif	dif	dif	Dif	dif	dif	dif	Dif	x	0.66	0.26	dif	dif	0.87	0.35	0.39	0.26
11	Dif	dif	dif	dif	Dif	dif	dif	dif	dif		x	0.51	dif	dif	0.84	0.65	0.72	0.49
12	Dif	dif	dif	dif	Dif	dif	dif	dif	dif			x	dif	dif	0.46	0.72	0.66	0.86
13	dif				Dif					dif	dif	dif	x	dif	dif	dif	dif	dif
14	Dif	dif	dif	dif		dif	dif	dif	dif	dif	dif	dif	Dif	x	dif	dif	dif	dif
15	Dif	dif	dif	dif	dif	dif	dif	dif	Dif				dif	Dif	x	0.56	0.52	0.44
16	Dif	dif	dif	dif	dif	dif	dif	dif	Dif				dif	Dif		x	0.90	0.66
17	Dif	dif	dif	dif	dif	dif	dif	dif	dif				dif	Dif			x	0.61
18	dif	dif	dif	dif	dif	dif	dif	dif	Dif				dif	dif				x

Results OLS Linear Regression analysis

In Table 65 and Table 66, the output of the OLS Linear Regression analysis can be found.

Table 65: OLS Regression Results

Dep. Variable	infections
Model	OLS
Method	Least Squares
No. Observations	18
Df Residuals	14
Df Model	3
Covariance Type	nonrobust
R-squared	0.702
Adj. R-squared	0.639
F-statistic	11.01
Prob (F-statistic)	0.000559
Log-Likelihood	0.78351
AIC	6.433
BIC	9.994

Table 66: Results on coefficient OLS Linear Regression analysis

Parameter	**Coefficient**	**std err**	**T statistic**	**P value**
constant	0.1492	0.160	0.935	0.366
amount_people	0.0464	0.013	3.479	0.004
amount_mr_meetings	0.0488	0.021	2.292	0.038
nr_office_meetings	0.0110	0.003	3.674	0.003

Appendix K model assessment

In this appendix the number of runs per scenario, the compliance with the assumptions for the different statistical tests is described and the results of the statistical tests are described.

Number of runs per scenario

In this section the number of runs per policy is given, for both set-up 4 and set-up 14. In Table 67, the number of runs per policy is set-up 4 is given. Which set of NPIs a policy refers to can be found in Table 69.

Table 67: Number of runs per policy set-up 4

Policy	Number of runs
1	20
2	20
3	20
4	20
5	20
6	20
7	20
8	20
9	20
10	20
11	20
12	20
13	20
14	20
15	20
16	20

In Table 68, the number of runs per policy is set-up 4 is given. Which set of NPIs a policy refers to can be found in Table 74.

Table 68: Number of runs per policy set-up 14

Scenario	Number of runs
1	20
2	20
3	20
4	20
5	20
6	20
7	20
8	15
9	18
10	18
11	18
12	18
13	16
14	15
15	15

16	20
17	15
18	15
19	15
20	15
21	15
22	15
23	15
24	15
25	17
26	17
27	16
28	16
29	15
30	16
31	15
32	15

Requirements statistical tests

To perform a statistical tests, the data needs to comply with the requirements set for that specific test. In this section, the compliance of the data with the requirements for both the independent t-tests and the Ordinary Least Squares (OLS) Linear Regression are described.

To perform the t-test, two requirements need to be fulfilled. First, the variable to be compared needs to be of at least an interval scale. As the number of infections is on a ratio scale, this requirement is fulfilled. The second requirement is that the average values compared with one another can be stated with some certainty. Often the rule of normality is used for this (Te Grotenhuis & Van der Weegen, 2013), but as the certainty of the average value for the number of infections is calculated by the number of required runs, this requirement is already fulfilled.

Finally, it needs to be known whether the variances of the samples can be assumed to be equal, to know whether the t-test needs to be performed assuming equal or unequal variance (Dekking et al., 2005). Whether the variances of two set-ups can be assumed equal can be tested using Levene's test (Te Grotenhuis & Van der Weegen, 2013). This test is therefore done for each combination of set-ups, to consequently apply the fitting t-tests.

Next, the compliance with the assumptions for the OLS Linear Regression needs to be reviewed. There are six mandatory assumptions that are reviewed for the analyses for both set-up 4 and set-up 14 here:

1. The regression model is linear in the coefficients and the error term
2. Homoscedasticity of errors /The error term has a constant variance (no heteroscedasticity)
3. Observations of the error term are uncorrelated with each other
4. The error term has a population mean of zero Independence of observations
5. All independent variables are uncorrelated with the error term
6. No independent variable is a perfect linear function of other explanatory variables (Frost, 2022)

The first assumption addresses that the relationship between the dependent en independent variables are linear, in additional to the error term. For each of the NPIs, the value 'on' or 'off' only exists, so only a linear relationship between each of them and the number of infections would be possible. For the combination of the NPIs, the diagram in Figure 23 and Figure 24 are used to see the shape of the relationship between the number of NPIs present and the number of infections. It can be seen there that the shape of the relationships for set-up 4 are somewhat linear, and do not clearly any other shape very strongly. For set-up 14, the shape in even less clear. An OLS linear regression can be done to at least see if a linear relationship can be found for some of the NPIs.

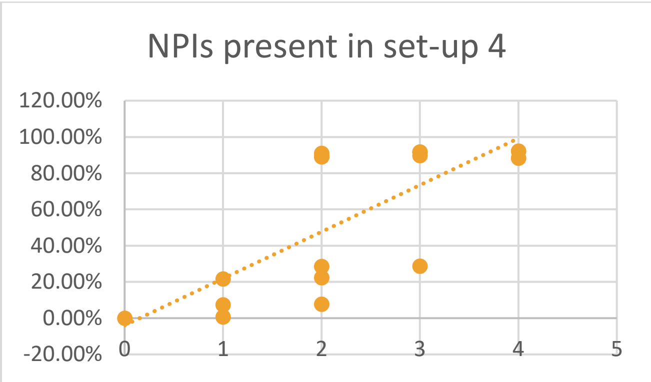

Figure 23: Plot on influence NPIs on infection set-up 4

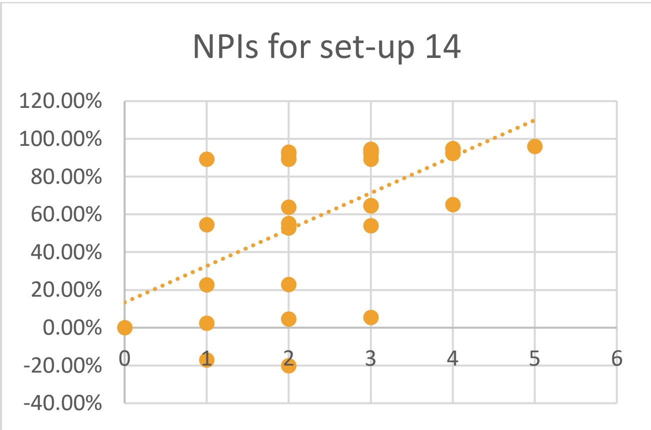

Figure 24: Plot on influence NPIs on infection set-up 14

According to the second assumption, for a model to be unbiased, the average value of the error term must be equal to zero. However, when a constant is included in the regression model, this assumption should be fulfilled (Frost, 2022). As the analysis are performed with a constant included, this assumption should be complied with.

The third assumptions states that the observations of the error term are uncorrelated with each other. Such a correlation can be caused by simultaneity between the independent and dependent variables or a measurement error and can bias the coefficient estimate (Frost, 2022). As no error term was given, the correlation could not be measured. However, as measurement are based on an existing calibrated model measurement error are not assumed to be present. In addition, simultaneity between the dependent and independent variables is also not present, as the NPIs are installed before any of the infections take place.

The fourth assumption states the observations must be independent from one another and the result of one observation should not affect the result of another observation. For this research, the observations are assumed to be independent, as the groups of employees representing each set-up are not the same, as in each set-up people have a different activity schedule that defines them and the set-ups do not influence one another but are created independently. In addition, as it is not a time series model, serial correlation is less likely to occur.

The fifth assumption states that there is no heteroscedasticity. This entails that the variance does not change for each or for some observations (Frost, 2022). As for each observation a variance could be found, this variances can be visually compared. Here, it could be seen that the largest variance was twice the size of other variances, however all variances are relatively small. Therefore, there could be some heteroscedasticity, which reduces the precision of the estimates in the OLS linear regression.

The final assumption is that no independent variable is a perfect linear function of other explanatory variables. This can be assumed to be correct if the five independent variables, the five NPIs, are completely independent from one another. This is in essence the case, for example higher ventilation does not change the effectiveness of a mask. However, due to one NPI possibly lowering the amount of virus present, this can reduce the percentual effectiveness of another NPI, as there is then less virus present to lower/prevent contact with. This should be remember in interpreting the results.

Results statistical tests

The statistical results of the Levene's tests and independent t-tests performed to compare each combination of policies is given for both set-up 4 and set-up 14. The results of these tests are summarized in the form of matrices. For the Levene's test, each cell in the matrix represents whether the set-up of the column and the set-up of the row are data samples for both set-ups are homogenous. In the (red) cells with 'non', the combination is non homogenous, if the (green) cell is empty the combination is homogenous.

For the matrix of the independent t-test, each cell represent whether the average number of infections of set-up of the column and the set-up of the row significantly differ from one another. When the cells says 'diff' the set-ups significantly differ from one another (the P value was lower than 0.05). In case the p value is lower, the difference cannot be proven and the p value of the t-test is given in the cell

For set-up 4

For set-up 4, what NPIs present the policy numbers refer to can be seen in Table 69. These policies numbers are used in Table 70 and Table 71 to show the results of the Levene's tests and t-tests.

Table 69: Policies used for set-up 4

Policy nr	Wearing a mask	Cleaning rate	Ventilation rate	Number of people allowed
1	No mask	High	High	Maximum
2	No mask	Normal	High	Maximum
3	No mask	High	Low	Maximum
4	No mask	Normal	Low	Maximum
5	No mask	High	High	50 %
6	No mask	Normal	High	50 %
7	No mask	High	Low	50 %
8	No mask	Normal	Low	50 %
9	Mask while walking	High	High	Maximum
10	Mask while walking	Normal	High	Maximum
11	Mask while walking	High	Low	Maximum
12	Mask while walking	Normal	Low	Maximum
13	Mask while walking	High	High	50 %
14	Mask while walking	Normal	High	50 %
15	Mask while walking	High	Low	50 %
16	Mask while walking	Normal	Low	50 %

For the Levene's test, each cell in the matrix represents whether the set-up of the column and the set-up of the row are data samples for both set-ups are homogenous. In the states 'non', the combination is non homogenous, if the cell is empty the combination is homogenous.

For the matrix of the independent t-test, each cell represent whether the average number of infections of set-up of the column and the set-up of the row significantly differ from one another. When the (red) cell is empty, the set-ups significantly differ from one another (the P value was lower than 0.05). In case the p value is higher, the difference cannot be proven and the p value of the t-test is given in the (green) cell.

Table 70: Results Levene's tests for the policy analysis using set-up 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	X								non	non	non	non	non	non	non	non
2		X							non	non	non	non	non	non	non	non
3			X						non	non	non	non	non	non	non	non
4				X					non	non	non	non	non	non	non	non
5					X				non	non	non	non	non	non	non	non
6						X			non	non	non	non	non	non	non	non
7							X		non	non	non	non	non	non	non	non
8								X	non	non	non	non	non	non	non	non
9	non	non	non	non	non	non	non	non	X							
10	non	non	non	non	non	non	non	non		X						
11	non	non	non	non	non	non	non	non			X					
12	non	non	non	non	non	non	non	non				X				
13	non	non	non	non	non	non	non	non					X			
14	non	non	non	non	non	non	non	non						X		
15	non	non	non	non	non	non	non	non							X	
16	non	non	non	non	non	non	non	non								x

Table 71: Results two sided t-tests for policy analysis set-up 14

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	X	0.9			0.186	0.21										
2	0.9	X			0.15	0.17										
3			X	0.46			0.22	0.25								
4			0.46	X			0.08	0.10								
5	0.186	0.15			X	0.93										
6	0.21	0.17			0.93	X										
7			0.22	0.08			X	0.94								
8			0.25	0.1			0.94	X								
9									X	0.36			0.47	0.91		
10									0.36	X			0.12	0.34	0.1	
11											X	0.44			0.437	0.82
12											0.44	X			0.14	0.35
13									0.47	0.12			X	0.58		
14									0.91	0.34			0.58	X		
15										0.1	0.437	0.14			X	0.61
16											0.82	0.35			0.61	X

OLS Linear Regression

All of the results of the OLS Linear Regression analysis can be seen in Table 72 and Table 73.

Table 72: OLS regression results Set-up 4

Dep. Variable	Infections scen
Model	OLS
Method	Least Squares
No. Observations	16
Df Residuals	11
Df Model	4
Covariance Type	nonrobust
R-squared	0.984
Adj. R-squared	0.978
F-statistic	164.8
Prob (F-statistic)	9.78e-10
Log-Likelihood	23.482
AIC	-36.96
BIC	-33.10

Table 73: Results of coefficient OLS Linear Regression analysis for set-up 4

	Coef	**std err**	**t**	**P>|t|**
const	1.0522	0.038	27.986	0.000
mask_wearing	-0.8517	0.034	-25.328	0.000
cleaning	-0.0060	0.034	-0.179	0.861
ventilation	-0.1346	0.034	-4.001	0.002
capacity office	-0.0425	0.034	-1.265	0.232

Set-up 14

For set-up 14, what NPIs present the policy numbers refer to can be seen in Table 74. These policies numbers are used in Table 75 and Table 76 to show the results of the Levene's tests and t-tests.

Table 74: Policy used for set-up 14

Policy nr	Wearing a mask	Cleaning rate	Ventilation rate	Number of people allowed	Occupancy rate
1	No mask	High	High	Maximum	No limit
2	No mask	Normal	High	Maximum	No limit
3	No mask	High	Low	Maximum	No limit
4	No mask	Normal	Low	Maximum	No limit
5	No mask	High	High	50 %	No limit
6	No mask	Normal	High	50 %	No limit
7	No mask	High	Low	50 %	No limit
8	No mask	Normal	Low	50 %	No limit
9	No mask	High	High	Maximum	Maximum of 50 %
10	No mask	Normal	High	Maximum	Maximum of 50 %
11	No mask	High	Low	Maximum	Maximum of 50 %
12	No mask	Normal	Low	Maximum	Maximum of 50 %
13	No mask	High	High	50 %	Maximum of 50 %
14	No mask	Normal	High	50 %	Maximum of 50 %
15	No mask	High	Low	50 %	Maximum of 50 %
16	No mask	Normal	Low	50 %	Maximum of 50 %
17	Mask while walking	High	High	Maximum	No limit
18	Mask while walking	Normal	High	Maximum	No limit
19	Mask while walking	High	Low	Maximum	No limit
20	Mask while walking	Normal	Low	Maximum	No limit
21	Mask while walking	High	High	50 %	No limit
22	Mask while walking	Normal	High	50 %	No limit
23	Mask while walking	High	Low	50 %	No limit
24	Mask while walking	Normal	Low	50 %	No limit
25	Mask while walking	High	High	Maximum	Maximum of 50 %
26	Mask while walking	Normal	High	Maximum	Maximum of 50 %
27	Mask while walking	High	Low	Maximum	Maximum of 50 %
28	Mask while walking	Normal	Low	Maximum	Maximum of 50 %
29	Mask while walking	High	High	50 %	Maximum of 50 %
30	Mask while walking	Normal	High	50 %	Maximum of 50 %
31	Mask while walking	High	Low	50 %	Maximum of 50 %
32	Mask while walking	Normal	Low	50 %	Maximum of 50 %

Levene's results

For the Levene's test, each cell in the matrix in Table 75 represents whether the set-up of the column and the set-up of the row are data samples for both set-ups are homogenous. In the states 'non', the combination is non homogenous, if the cell is empty the combination is homogenous.

For the matrix of the independent t-test, in Table 76, each cell represent whether the average number of infections of set-up of the column and the set-up of the row significantly differ from one another. When the cells says 'diff' the set-ups significantly differ from one another (the P value was lower than 0.05). In case the p value is lower, the difference cannot be proven and the p value of the t-test is given in the cell.

Table 75: Results Levene's test for policies set-up 14

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	51	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	X				non	non			non	non			non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
2		X			non	non							non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
3			X		non	non	non	non	non	non			non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
4				X	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
5					X												non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
6						X											non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
7							X										non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
8								X									non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
9									X								non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
10										X							non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
11											X						non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
12												X					non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
13													X				non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
14														X			non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
15															X		non	non	non	non	non	non	non	non	non	non	non	non	non	non	non	non
16																X																
17																	X															
18																		X														
19																			X													
20																				X												

21																					X											
22																						X										non
23																							X									
24																								X	non	non	non	non				
25																									X							
26																										X						
27																											X					
28																												X				
29																													X			
30																														X		
31																															X	
32																																x

All two sided t-tests

Table 76: Results two sided t-tests for policy analysis set-up 14

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	X	0.98																														
2	0.98	X																														
3			X	0.80	0.66	0.75																										
4			0.80	X	0.46	0.60																										
5			0.66	0.46	X	0.84																										
6			0.75	0.60	0.84	X																										
7							X	0.57																								
8							0.57	X																								
9									X	0.91	0.11	0.09	0.88	0.95																		
10									0.91	X	0.13	0.11	0.78	0.85	0.05																	
11									0.11	0.13	X	0.93	0.07	0.09	0.83	0.63																
12									0.09	0.11	0.93	X	0.05	0.07	0.91	0.71																
13									0.88	0.78	0.07	0.05	X	0.92																		
14									0.95	0.85	0.09	0.07	0.92	X																		
15										0.51	0.83	0.91			X	0.78																
16											0.63	0.71			0.78	X																
17																	X	0.			0.	0.					0.08	0.34				

																		75			59	75										
18																	0.75	X			0.37	0.50						0.24				
19																			X	0.79				0.65								
20																			0.79	x			1.00	0.88								
21																	0.59	0.37			X	0.79					0.12	0.52				
22																	0.75	0.50			0.79	X					0.09	0.43				
23																				1.00			X	0.86								
24																			0.65	0.88			0.86	X								
25																									X	0.42	0.12			0.10	0.99	0.64
26																									0.42	X	0.44	0.15			0.42	0.71
27																	0.08				0.12	0.09			0.12	0.44	X	0.50			0.12	0.25
28																	0.34	0.24			0.52	0.43				0.15	0.50	X				0.07
29																													X	0.66		
30																									0.10				0.66	X	0.08	
31																									0.99	0.42	0.12			0.08	X	0.63
32																									0.64	0.71	0.25	0.07			0.63	x

OLS Linear Regression

The results of the OLS Linear Regression for the policy analysis in for set-up 14 can be seen in Table 77 and Table 78.

Table 77: OLS regression results for policy analysis set-up 14

Dependent Variable	Infections scen 14
Model	OLS
Method	Least Squares
No. Observations	32
Df Residuals	26
Df Model	5
Covariance Type	Nonrobust
R-squared	0.852
Adj. R-squared	0.823
F-statistic	29.85
Prob (F-statistic)	5.47e-10
Log-Likelihood	-7.5534
AIC	27.11
BIC	35.90

Table 78: Results of coefficient OLS Linear Regression analysis for set-up 14

	Coefficient	**std err**	**T statistic**	**P value**
constant	1.8341	0.147	12.461	0.000
Mask wearing	-1.3037	0.120	-10.848	0.000
cleaning	-0.0071	0.120	-0.059	0.954
ventilation	-0.1968	0.120	-1.637	0.114
capacity office	0.0934	0.120	0.777	0.444
capacity meeting rooms	-0.6389	0.120	-5.317	0.000