Taking back control

Virtue ethics as an information basis on the implementation of a smoking-reducing chatbot.

AUTHOR:
Elizardo Laclé *(s1015215)*
elizardo.lacle@ru.nl

First Supervisor:
Dr. Anco Peeters
Donders Institute
anco.peeters@donders.ru.nl

Second Reader:
Lotte van Elteren MSc
AI Department
lotte.van.elteren@ru.nl

Radboud University

March 2, 2023

Abstract

Smoking is known to have a negative impact on one's health, but just like most other addictions, it is difficult to quit (Doll & Hill, 1950). For this, we have given a variety of design recommendations to implement a virtue-informed chatbot (VIC). The recommendations are based on Virtue Ethical and psychological research to ensure that the implementation is beneficial to its users. As complete cessation is often found difficult we have taken the approach of some intake reduction instead. For this, we have made use of the Virtue Temperance as this has to do with the moderation of pleasures. From this view we have also looked at ways we could then positively influence behaviour and try to combine these into design recommendations. These recommendations are based on both the research already done in this thesis, as well as existing aspects in existing technologies. The recommendations we give should help the future engineering of this VIC with general notes as well as specific examples.

Contents

Introduction

When thinking about addictions, most often substance addictions, such as to Methamphetamine (crystal meth), come to mind. Another form would be a behavioural addiction such as to gambling or shopping. This latter category is often less noticeable as it has solely to do with, as the name suggests, behaviour in its initial stages. However, there is a well-known addiction in modern-day society that seems to be generally accepted; an addiction to smoking.

This general acceptance is shocking when considering that in 1950 Richard Doll already showed that smoking can cause severe health damage (Doll & Hill, 1950). Since then multiple studies have come to the same conclusion, showing that it is a major cause for developing artery abnormalities (McBride, 1992).

These results would lead one to assume that both the public and the government start to realise that smoking should be refrained from. This thought process is not entirely incorrect but does contain some flaws. In more recent years the Dutch Government has been creating advertisements that explain both the health risks for those who smoke as well as those surrounding them. These advertisements were focused on aiding pregnant women in their journey on quitting smoking, where they mentioned 'Help a pregnant woman who stopped smoking by not smoking in their presence. This way you don't tempt her.' ("Stoppen met roken commercial #rookvrijopgroeien", 2017). However, this approach focuses on maintaining the cessation of smoking, compared to the prevention of it. In 2020 the Dutch Trimbos-institute (an independent knowledge institute for alcohol, tobacco, drugs, and mental health) published a report that showed that roughly 22% of Dutch citizens above 18 years old smoked in 2018 and these numbers have barely decreased since (Bommelé, 2020). In the same report, it is shown that more than 70% of the smokers from this research smoked on a daily basis meaning that the larger part of those who smoke do this very regularly.

This habituation is part of the addiction that they have developed. As described before smoking can have severe health consequences and it is thus best if people quit entirely, however since it is an addiction this is often easier said than done. An alternative to this problem could be reducing the (daily) smoke intake an individual takes on a regular basis.

Multiple studies have shown that smoke reduction can be beneficial to cessation as well a decrease in health risks (Godtfredsen, 2005; Hughes & Carpenter, 2006). However, these studies do also mention that cessation in and of itself is the best option to decrease these risks. This conclusion is logical but does not mean that cessation is easily achieved by many. If it was, why are there still so many smokers?

In this thesis, we will argue that smoke intake reduction is a more feasible approach for many current smokers. Our alternative approach would put the user's goals at its centre and try to guide them through this process. This would lead to setting individual end goals per person compared to cessation in which everyone has the same end goal, quitting smoking. This is to give the user a more personally tailored solution to their smoking addiction which in turn makes them feel more in control over it.

For this, we will make use of Virtue Ethics as it starts from the notion that human flourishing should be put at its centre. This is in line with our focus of putting the user's personal goals at centre stage. We want to ensure that this reduction is done ethically since we are dealing with complex and sensitive issues. Two alternative approaches that could be considered would be Deontology and Consequentialism but in our opinion these

lack in some way or form as explained in Chapter 1.3.

Lastly, in line with virtue ethics, we want to get as many people to reduce their smoke intake as possible. This can be achieved in a variety of ways but with the developments and improvements of technology comes the benefit that these are becoming widely accessible. Chatbots have already been shown to be able to help reduce depression symptoms in young adults, which indicates their effectiveness in supporting behavioural adjustments (Fitzpatrick et al., 2017).

To argue for the above-mentioned points it is important to note that this is an initial attempt to create a Virtue Informed Chatbot (VIC), not a final version. We will try to create such a VIC by first expanding our choice for Virtue Ethics by investigating stakeholders, different virtues, and reasoning from other normative ethical theories. Secondly, we will look at ways to intervene by making use of the information gathered in the previous chapter and combining it with studies from the field of psychology. This gives us a variety of ways that we could implement in our VIC. Lastly, we will combine the content of the previous two chapters and give design recommendations based on this. For this, we will use expertise from Human-Computer Interaction as we want to ensure that users feel comfortable making use of the VIC.

1 Virtue Ethics

Before we dive into developing a Virtue Informed Chatbot (VIC) we need to understand what Virtue ethics is, why we need it and why this it is best to use Virtue Ethics.

1.1 Stakeholders

When talking about smoking there are multiple parties that are involved in regard to the negative consequences. Throughout this thesis, we will refer to these parties as *stakeholders*. These stakeholders can be involved in differing ***amounts*** as well as at different points in time. In this case, the first, and most obvious one, is the group of people who smoke. These stakeholders are directly involved and have influence over the outcomes of their smoking habits. Secondly, I will highlight passive smokers. These stakeholders have little influence but are still exposed to smoke inhalation in many cases. Thirdly we will focus on health insurance companies. This group will bear the consequences of those who smoke in regard to the financial impact when they fall ill.

The first and main stakeholders will be Active Smokers. The group of Active Smokers contain individuals who actively increase their nicotine intake by for example smoking cigarettes. This definition will help us further identify sub-categories for our stakeholders and exclude other forms of smoking such as weed (and other drugs) as these might need other forms of support compared to nicotine intake. An important note to make is that this is a very broad stakeholder group. In this group fall a variety of people, such as people who smoke multiple cigarette packs a day, but also people who smoke very irregularly at for example parties. We will therefore define the subcategories for Active Smokers as *Heavy smokers*, *Moderate smokers*, and *Party smokers*.

Firstly lets discuss Heavy Smokers. Throughout a variety of studies the consistency for this category differ widely. Where one study may believe that more than 24 cigarettes a day would be a heavy smoker, others say this line is more towards 10 (Boulos et al., 2009; Wilson et al., 1992). This is most likely due to the difference in the moment of writing and publishing. From experience, we have seen that heavy smokers will easily smoke one package of cigarettes a day. Considering that a package often consists of 20 cigarettes, we will use this as our boundary for a heavy smoker (Blackwell et al., 2020).

When determining the boundary for the group of heavy smokers, the boundary for Moderate Smokers is logical; less than that. This means that this group will on average smoke less than 20 cigarettes a day. An important note to make here, however, is that this group smokes on a regular (daily) basis.

Lastly, we consider the group of Party Smokers. This group does not smoke on a daily basis and often also only smoke in social settings (such as parties). This group will not be taken into account when considering the reduction of smoke intake due to the inconsistency of their smoking 'habits'.

The second stakeholder that will be looked at is Passive Smokers. According to the Cambridge Dictionary, passive smoking is defined as "the unwanted breathing in of other people's cigarette smoke, especially by people who do not smoke"(Cambridge University Press., 2022a). For this thesis, we will specify this scope as "the act of breathing in smoke that contains nicotine from other people". The reason for this specificity comes from our definition of our first stakeholder.

Our third stakeholder would be Health Insurances. These insurances are in place to ensure that someone is taken care of when falling ill, without having to worry (too much) about the financial consequences. However, when someone smokes they increase their risk of health complications (Doll & Hill, 1950). In regard to these complications, health insurance companies will often have to cover the expenses made to treat the patient. This means that they are directly influenced by the first-mentioned stakeholder.

As for our main focus in this thesis, we will zoom in on the first mentioned stakeholder; Active Smokers. This group is chosen for a variety of reasons. The first and foremost reason has to do with the approach that will be taken to tackle the problem. Our approach will look at this problem from a Virtue Ethics perspective (see 1.3). Since Virtue Ethics puts human flourishing at its centre stage, it is only logical to take this as our main group. These stakeholders would gain the most in regard to their quality of life, as well as health. Next to that, they have influence over the other two mentioned stakeholders. Whereas neither passive smokers nor health insurances have this influence over smokers. The group of passive smokers is, as described, inhaling smoke that is produced by a smoker. This means that they have little influence on the intake apart from moving away from the smoke or smoker. For health insurance companies, this influence is even more thinly spread. They are not involved in the lives of smokers apart from paying for their treatments. A way they could influence this more is by making people who smoke pay more for their treatments, however, this would be considered discrimination. This leaves us with the first stakeholder as our best option. This is in line with our approach and seems to be the best to tackle before continuing to tackle other problems.

1.2 Three main theories in Ethics

Ethics is bluntly a branch of philosophy that is involved with researching the concepts of right and wrong (behaviour). Due to the nature of this field there is often not such a clear cut answer as with Mathematics or Science. However, there are multiple ways to look at certain problems. Within ethics there are three main theories; Virtue Ethics, Deontology, and Consequentialism.
These main theories have been around for quite some time, especially when noting that the origin of Virtue Ethics is often attributed to Aristotle. In Virtue Ethics the human character is put to centre stage. If a person would act in accordance with virtue they would do, and be, good. A virtue according to Aristotle is a trait of character that is optimal for someone to have. A virtue is a perfect middle ground between two extremes called vices which would be either an excess or a deficiency of said virtue (Hursthouse & Pettigrove, 2022). In order to reach virtue it is important to develop oneself. This development aids the flourishing of a person leading to eudaimonia. This approach is very different when compared to both Deontology and Consequentialism, which are both concerned with the actions of a person.
As the name suggests, the latter is looking at the consequences or outcomes of an action and if these could be considered good or correct. Often this means that all involved variables must be examined and evaluated. This is a good approach for cases where these outcomes and variables can easily be defined but it becomes more difficult when uncertainty is involved. In other words, the moment we are not able to predict a rough sketch of the future we have difficulty implementing utilitarianism to its best

Another approach would be Deontology. Deontology examines the intentions or motivations of an action. This means that if the intention of an action is morally correct, then the action is, irrespective of its outcome, good. In other words, an act may be considered right even if it produces a bad outcome, as long as it follows the moral law. These intentions or rules, however, do not necessarily prevent actions that would be considered morally wrong. A famous illustration of this would be the laws of robotics by Isaac Asimov, where he illustrates that despite having 'ethical rules' there could be disastrous consequences on the execution of this [1].

1.3 Why Virtue Ethics?

As explained above there are three main theories within the field of Ethics. In this thesis, we will make use of the theory of Virtue Ethics. Our main reason for this is due to it putting human flourishing at its centre which is essentially what one would want to achieve. Next to this, there are some critical issues when looking at Consequentialism and Deontology in combination with the uncertain future.
As explained for Consequentialism, making accurate decisions becomes difficult when variables are undefined or cant be assigned. As Shannon Vallor explains in her book 'Technology and the Virtues' it is not a problem of picking the right course of action, but a problem of discerning which possible courses there are (Vallor, 2018). According to Vallor, this discernment is practically incalculable as the potentials, especially those in technology, are both too numerous and too opaque when making an attempt of assigning reliable probabilities of specific potential outcomes.
For Deontology this is different. As explained earlier, Deontology has to do with following a set of rules. At its core Deontology says that one should not make special exceptions for ourselves, or others when making these moral decisions. Immanuel Kant envisioned this theory on a 'universal role' principle; would it be okay (or rational) if everyone in the world follows your rule in a similar situation? If the answer would be no, the rule is immoral, if it would be yes you could execute your action morally. This is a good intention and in many cases would be very helpful. However, in a world where technology is rapidly advancing and expanding such a model would not be practical. Over the last decades the usage of technology, especially in our daily lives, has exponentially grown and it doesn't seem to stop. With the current set of rules, we might not encompass the technological advances made, meaning we would always be trying to catch up compared to being able to adequately react to the situation.

1.4 Temperance as a Virtue

As mentioned before this thesis will focus on increasing one's Temperance to a virtuous level. To do this we first need to define what we believe Temperance is.
According to the Cambridge Dictionary, Temperance has two definitions. One being 'control of your own behaviour, such as not drinking or eating too much', and the other being 'the habit of not drinking alcohol because you believe it is dangerous or wrong' (Cambridge University Press., 2022b). For this thesis, we will be looking at the former definition.

[1] Isaac Asimov had created three rules for a short science-fiction story which gained popularity over the years. The rules boiling down to; A robot must not harm a human, by neither action or inaction, it must obey human commands and lastly it must protect itself from harm. Where latter rules must not conflict with former rules.

Looking at this definition we could come to the conclusion that Temperance could be the ultimate virtue. Other virtues have to do with the way a person acts but they are only virtuous when executed at the right moments in the right situations. This means that in many situations one would need to restrain themselves from saying or doing something, or on the other side of the spectrum accept that they need to loosen up a little for the benefit of, for example, the group atmosphere.
An important note is that one needs to be aware of their actions and emotions before being able to control them. Because of this, we will adjust our definition to 'Being aware of one's own behaviour (including actions, thoughts and emotions) and being able to control these in a desirable way'. One addition that can be noticed is the added phrase of 'in a desirable way'. The reason for this is due to the focus of human flourishing. A person might be aware of their emotions and control them in such a way that they will not be bothered by them. However, this is undesirable as this means that someone doesn't (allow themselves to) flourish. This latter statement could be investigated more, which is advisable for different theses but in this thesis, we will not go into it further.
Another difference that can be noticed here is the addition of 'thoughts'. A person's actions and emotions might be visible externally in many cases, but thoughts rarely are. If we look at the virtue of Temperance one would need to master all three in order to be virtuous. However, since this is very difficult, if not impossible, we will not look into becoming virtuous but we will focus on becoming *more* virtuous. For this reason, it is thus hard to focus on thoughts, but this is a topic that can be looked into further for becoming a virtuous human being.

It is important to note that up until this point, we have looked into solely Temperance as our virtue but this does not mean that this is the only virtue that could have been investigated. In the context of reducing smoke intake there are other virtues that could be interesting such as self-discipline and perseverance.
To start of with the former, self-discipline, it is important to lay the groundwork again. According to the Cambridge dictionary, self-discipline is the act of being able to execute tasks one is required to do despite not wanting to do them (Cambridge University Press., 2022c). This means that if, for example, someone has to do something such as the dishes, but does not want to do them it requires self-discipline. The main difference that can be made between self-discipline and Temperance is that Temperance requires someone to abstain from doing something, whilst self-discipline requires someone to actually do something. For the reduction of smoke intake, this would thus be less satisfactory and between these two virtues, Temperance fits better when looking at smoke intake reduction.

The other virtue that could have been interesting investigating would be Perseverance. As before we will first look at its definition in the dictionary; "continued effort to do or achieve something, even when this is difficult or takes a long time" (Cambridge University Press., 2022d). This seems to tie in closely with self-discipline which we have described before but there is an important distinction between the two. Perseverance is to continue on doing something despite (external) difficulty, self-discipline has to do with internal difficulties as there is little willingness. In other words, doing the dishes whilst not wanting to boils down to being self-disciplined. Whilst continuing doing the dishes when you notice that there are burnt stains on a pan that wont come off. In other words, you need to put in continuous effort to clean them.
In the case of reducing smoking intake this could be an interesting approach as the goal is

to have a user who persevere in their reduction. However, for this it is important to first start with this reduction and only at that point perseverance comes in. In other words, we would need to already have the user start on this before we can use our chatbot. When comparing this to our virtue of Temperance we can see that the latter can more easily be used at the start of this process.

But despite noticing that other virtues might not fit perfectly, what makes Temperance fit best on it own? To answer this it is important to look back at why we are working with Virtue Ethics. As explained before Virtue Ethics puts human flourishing and eudaimonia at its centre. In the case of smoking it would be optimal to have as little smoke intake as possible as this impacts one's health. However, this is not necessarily always possible and thus we need to moderate this. As explained previously Temperance is in essence a mean in regard to pleasures, i.e. the moderation of things that give pleasure. And as smoking is often considered a pleasure for smokers Temperance would be the perfect virtue to help reduce smoking intake.

1.5 The Extremes and ideal situation of Temperance

Now that we have explained that Temperance is the best fit virtue for our goal it is important to also support this more concretely. To show why we will focus on Temperance it is important to also show the vices that are involved in our case as well as what the ultimate virtuous goal would be.

To start off we will focus on the vice that has to do with too little Temperance; Self-indulgence (Aristotle, 2014). Self-indulgence is the act of allowing yourself to have or do anything that you enjoy (Cambridge University Press., 2022e). Aristotle goes further to explain that "The self-indulgent man, then, craves for all pleasant things or those that are most pleasant, and is led by his appetite to choose these at the cost of everything else" (Aristotle, 2014). In essence what he is describing is a severe addiction, choosing X over Y at any cost. Luckily in many cases people are not this severely addicted to a substance such as cigarettes, but regular smokers are often on this side of the scale resulting in their environment dealing with the consequences of their smoking habits.
The other side of this virtue scale would be having an excess of Temperance; Abstinence. This complete abstinence of pleasure can be seen as inhuman, as pleasure seems to often be a driving factor for people. But this abstinence would also be close to impossible in our current world where many things around us are shaped to give us some form of pleasure. Especially for those who are currently already smoking complete abstinence is difficult.

Our golden mean would be the perfect balance between the two; Virtuous Temperance. A way to moderate our pleasure seeking behaviour where we neither indulge, but also dont need to shy away from forms of pleasure. Being able to decide if you need to smoke at when you feel the urge to, and being able to moderate this could be considered more virtuous. In regards to human flourishing this would still not be optimal as smoking has been proven to cause health problems in the long run (McBride, 1992). However, in regards to more realistic goals this reduction can be considered realistically virtuous. Being able to be in charge of one-self and being able to moderate their smoke intake, and possibly reduce it, is a good step towards becoming more virtuous. Therefore becoming Realistically Virtuous can be considered a Silver Mean compared to being Generally Virtuous.

2 Self-control in Psychology

As explained previously self-control is 'Being aware of one's own actions, thoughts and emotions and being able to control these in a desirable way'. However, before we continue in the development of the Virtue-Informed Chatbot (VIC) we need to lay some ground in the field of psychology. This has to do with the fact that we will be working on supporting someone to change their behaviour. This is difficult to do from solely an ethics perspective and thus we require expertise from a field that works on this; psychology.

To start with this we will delve into the typical understanding of Temperance within psychology. After this, we will look into psychological mechanisms involved in Temperance and what obstructions there may be. Next we will explore ways that Temperance is learned and how it is often maintained. As we explore this we need to inspect ways to combine the previous two points and see how we can intervene to support the development of Temperance. After we have created a solid basis of psychological underpinnings we are going to look into the initial creation of a Virtue-Informed Chatbot, by making use of techniques from Human-Computer Interaction. For this, we will also look into the (possibly) already existing chatbots that make use of encouragement of the development of Temperance. Lastly, we will discuss the limitations and potential challenges of the currently existing psychological research.

Throughout this thesis we have been working on showing that Temperance is a virtue that is best suited for the reduction of smoking. However, in psychology Temperance is rarely studied as it is a virtue and not a psychological mechanism. Temperance is involved with multiple mechanisms, most notably self-control which is studied in a variety of ways. Self-control, as well as self-regulation, is often studied when looking to detect and control certain impulses. In other words, our virtue Temperance can be found as self-control and self-regulation in psychology research.

In psychology, self-control is often described as being able to choose between alternatives that arrive at different points in time (Rachlin, 1974). An example would simply be choosing between a tasty but unhealthy snack or a less tasty but healthier snack. The difference is in the moment of gratification. If one would eat the unhealthy snack they might get health problems later but have immediate gratification. As an alternative, they could choose to eat the healthier snack and have more gratification later (when they are healthy).

This effect has also been studied in a variety of ways, the most popular one being 'the Marshmallow Experiment' (Shoda et al., 1990). In this experiment, children were given a choice between receiving immediate gratification (e.g. one marshmallow) or delaying it and receiving something of larger gratification (e.g. two marshmallows). Before the study, the children were presented with a variety of items and were asked which one(s) they liked more so the researchers could pick items that could also indicate this gratification delay. As a result of this study, the researchers came to the conclusion that the children that were able to delay their gratification (and were thus better able to execute self-control) had other benefits compared to their peers as well such as coping with frustration and stress in adolescence (Shoda et al., 1990).

This delay in gratification is, as with many psychological mechanisms, believed to be trainable as if it were a muscle (Muraven, 2010). In their study, they mention that exercising self-control can have an impact if executed regularly, even if these executions

are small acts of self-control. This indicates that our VIC should help its users train their self-control to achieve the goal of reducing their smoking intake. This can be done in a variety of ways such as proposing that the user would drink a glass of water when they would want to grab a soda, and gradually increasing the difficulty of these acts.

2.1 The essence of psychological underpinnings

Now that we know how psychology experts define self-control it is important to look at what problems arise with it. The reason we explore it has to do with the mismatch between virtue and phenomena. In general, virtues are a goal that can not explicitly be defined in modern language. Psychology however works on the basis of models, theories and experimental data which can be expressed more easily in modern natural language. So to understand how self-control can be built we need to understand how it works psychologically and why it is the case that it can differ per person. Once this is clear we can further investigate how to stimulate the development of self-control.
Since self-control is such a broad mechanism there are multiple causes that can be identified to explain this difference. These can be split up into internal and external factors. One of the internal factors that influence one's ability to exercise self-control is the notion of ego-depletion. The theory of ego-depletion suggests that self-control can be depleted through usage as if it were only available in limited capacity (Baumeister, 2018). In other words, one is not able to execute more self-control if they don't have any left. Research supports this notion saying that once a person executes self-control they seem to have less self-control for subsequent tasks shortly after the initial one (Baumeister, 2003).

Another factor that influences ones ability to execute self-control well is motivation. This motivation can be attributed to both internal as well as external factors as it collaboration between the two. If someone is promised something very gratifying when they execute their self-control this could motivate them to do it, compared to being unsure if the gratification will happen. The motivation in this case would be the external promise for their internal gratification. However, a study done in 2003 showed that this motivation can also be complementary to self-control when it is completely external. The study showed that participants who believed that the task they did would help others, were able to execute more self-control compared to those who weren't told this (Muraven & Slessareva, 2003). In other words, the participants themselves had nothing to gain from this task themselves but were able to execute self-control.

An external factor that influences an individual's self-control is their social circle. Social pressure, and especially peer pressure, has been shown to affect one's self-control (Bonein & Denant-Boèmont, 2015). When someone is in a social setting with their peers they are often more inclined to try to fit into the group and behave differently then how they would in other circumstances. This can also be seen with self-control. For example, when someone is trying to drink less alcohol but everyone surrounding them is having a beer, there is a high chance they will grab one as well meaning that their self-control is impeded by their social surroundings.

A factor that might seem more hidden compared to others is planning. As explained previously self-control has to do with the ability to choose between different levels of gratification in different points in time. If someone has a lot of difficulty planning there is a high chance that their view of the future is not as accurate as they might hope. This inaccuracy has an influence on the amount of self-control someone can execute due to the

fact that one needs to be able to distinguish clearly between now and 'then' and between the levels of gratification between those periods, which requires planning. Looking at it from a practical standpoint if someone would want to lose weight they need to plan their diet and their work-out sessions, as you cant just hope for the best.

Another factor that is a combination of internal and external factors is decision fatigue. Decision fatigue is a phenomenon that occurs when a person's decision-making stamina pool is drained (Baumeister & Vohs, 2007). This means that someone has difficulty making decisions and thus might prefer to just pick a random option without (critically) thinking about it. In other words, one is too tired to make decisions that would lead to the execution of self-control.

2.2 How is self-control generally learned?

Now that we understand how Self-Control works it is important to understand how it is developed and how one learns to be in control. To investigate how self-control is learned we look at ways this happens in children. The reason for this is that children seem to be guided by impulses, playing with toy A at this moment and running to play with toy B the other moment. These impulsive behaviours imply that these children have difficulty executing self-control as they prefer to go for immediate gratification. However, at some point, they seem to be able to execute self-control and delay these gratifications but the difficulty is; how does this happen?

To start off we can look back at the Marshmallow experiment. These children were able to execute self-control and inhibit their desire to eat the marshmallow immediately (Shoda et al., 1990). This indicates that children between the ages of (roughly) three and a half, and five and a half have learned some form of self-control. This is in line with a study done in 2014 where they studied the development of self-control in children between the ages of three and nine (Tao et al., 2014). In this study, they came to the conclusion that growth of self-control improves between the ages of 5 and 6. This conclusion gives us an indication of when the peak of this growth would be. But knowing when this peak is is not enough as we need to distinguish the different ways that self-control can be developed.

One option that has been shown to have a positive effect is timely (consistent) reminders (Barker & Munakata, 2015). During their experiment, they instructed the children to only open boxes with a specific shape on top and tested if reminders would help with the execution of self-control within these children. They found that the children were better able to execute self-control when they were reminded to do so (e.g. "Make sure you wait to open the box until you see the right shape!"). Next to this they also experimented if an imposed delay would help, as a previous study had shown that imposing this delay had positive outcomes (Simpson et al., 2011). However, Barker and Munakata found little evidence to support that imposed delay improved self-control. They did find that a spontaneous delay, a delay imposed by the child themselves, did give a good indication of the execution of self-control. They reason that this is not a cause for better self-control but merely a symptom. "Children can take time to think, but giving them time does not guarantee that they will think." (Barker & Munakata, 2015).
These timely reminders could take the form of notifications for our Chatbot. Where the user would receive a short notification along the lines of "It is smart to skip a cigarette if you want to reduce your smoke intake." at random intervals. These phrases however

would need to be worked out better and should (if possible) be context-dependent. In other words, the user should not receive this message when they just smoked or were not planning on doing so.

Another way that children often learn to harness self-control is through a change of their environment (Duckworth et al., 2016). This might sound counterintuitive since a change of environment indicates that one does not have the self-control to succeed in the original environment. However, that is exactly the execution of self-control one needs. To change one's environment it is necessary to notice that the original environment is not fruitful. Changing this original setting thus requires self-control as it is simpler to not change anything and give in. This chosen change shows both self-awareness which is needed for good execution of self-control, as well as control over oneself, to be able to change it.

This also ties in nicely with another form of learning self-control; practice. This form of learning might sound obvious when reaching this point of this thesis, but it is an important, if not the most important, form of learning. In 2011 a study has shown that children that had trouble executing self-control had benefited significantly from practising on a regular basis (Tominey & McClelland, 2011). For someone who has little problems exerting self-control a change of environment might make it easier to resist temptation, but for those who already struggle it might be the only way to be able to resist the temptation at all. In other words, without that change of environment, they are not even able to practice since the temptation is too large for their current levels of self-control. However, with this change they are able to first practice with temptations that require less self-control and thus be able to practice more often.

Lastly an important factor to learn self-control is to also show gratification. An important note to make here is that this last factor has more to do with how to support learning of self-control. A study done in 2013 shows that children consider environmental reliability for their rational decision making process (Kidd et al., 2013). In the experiment Kidd et. al. performed they had reproduced the marshmallow experiment in either a reliable or unreliable environment. The difference between the two was that the promised gratification would always be given in the reliable condition whereas this was not the case for the unreliable condition. In the unreliable condition the researcher promised to return with a reward but failed to do so. This difference seemed to impact the child's ability to exert self-control, as the children who had the reliable condition waited significantly longer than those who had the unreliable condition. In other words, showing and giving the gratification are important to prevent the reduction of one's self-control. This prevention of reduction is important to not cancel out the growth that is made via the other aspects.
When looking at a real-world socioeconomic environment this also makes sense. When there is very little evidence to support that a promised reward is guaranteed, why would it be preferred over a guaranteed (but possibly smaller) reward. When a next meal for example is not guaranteed because you are living in poor circumstances it is not illogical to dive in on the first opportunity to have a proper meal. It is thus also important that our VIC circumvents this possibility of shrinkage of self-control. A way to do this is to have a form of self evaluation tracker that is linked to a form of gratification such as a form of plant the user can grow.

2.3 Temperance involved psychological mechanisms

Now that we have seen factors that affect self-control and ways that it can be learned and strengthened, it is time to combine the two. We will look into ways we can make use of these influences to our advantage by either strengthening or weakening them. As well as trying to optimise the ways of learning mentioned above.
To start off, there were a variety of factors that we noticed influence one's ability to exert self-control; Decision-fatigue, Ego-depletion, Lack of Planning, Motivation and Social influences.

Whilst many of these impede self-control motivation is an influence that has been proven to strengthen self-control. However, for our goal of smoke intake reduction, we assume that the user has chosen to use this method and is motivated to continue using it. This does not mean that no further attention will be spent on motivation but it does mean that for the scope of this thesis, we will not be delving into it in detail.

A factor that could heavily influence the user both positively and negatively is their social circle. Peer pressure has been shown to be able to influence one's behaviour on a longer persistent basis (Gallani, 2017). This indicates that making use of their social circle to cultivate self-control would be beneficial. However, from both experience and studies, it has been shown that people who smoke often come from social circles where multiple others smoke as well (Bewley et al., 1974). In other words, people in their social circles will more likely than not also smoke, which makes this effect less useful. A way to still make use of this would be for the user to inform their peers and ask them to help with their goal. The effects of this will most likely vary but this form of self-imposed peer-pressure can thus be chosen by the user and might help motivate them.
Another more drastic option would be to choose to not surround themselves by their current (smoking) social circle. This would be in line with the learning technique of changing one's environment, but this is a pretty drastic change. Since this change would be so drastic this is not something we will further investigate. Something that however is worth investigating and is in line with the technique of changing environment is changing one's individual environment. An example of this would be putting away certain temptations such as cigarettes and ash-trays. This change of environment supports their increase in self-control as described by Duckworth et. al. (Duckworth et al., 2016).

Something that our VIC might be able to influence is the training of self-control in a safe environment. As explained previously it has been shown that training self-control will have a beneficial effect on those who have difficulty making use of self-control. Since the users of this VIC are looking for ways to support their goal to reduce their smoke intake, they are implicitly looking for a way to improve their self-control. In other words, making use of this training will most likely be beneficial to the majority of users. There are a variety of ways this can be executed, but an important factor to take into account here is ego-depletion. When training self-control it is important that this training does not impede other executions of self-control that might be necessary soon after. Therefore in the initial stages, this training should not happen during moments when self-control is necessary for other tasks such as work or dinner. A way to then execute this would be to start these pieces of training in the evening after dinner but before the user goes to bed. An example would be brushing their teeth with their non-dominant hand until it becomes a habit. The reason for this is that people tend to fall back on habits when they have

difficulty making use of their executive functions (Foerde et al., 2006). In other words, when something becomes a habit you do not train self-control as intensively. After this training other forms could be tried such as meditation.

Another factor with which the VIC could help would be planning. As established by Duckworth et al. people are more likely to succeed in achieving their goal when thinking about possible obstacles and coming up with actions to tackle these before starting (Duckworth et al., 2018). In other words, if the user can reason with the VIC about possible obstacles when reducing their smoking might be difficult, together they can look into ways to counter them. These problems will differ per user and should thus also be adjusted per user. Some ways the VIC could help is by suggesting solutions that have been shown to work by research for cessation such as Nicotine Replacement Therapy (NRT) which gives you nicotine in forms other than that of smoke inhalation. NRT can help by relieving some of the physical aspects of withdrawal allowing the user to focus on more psychological symptoms. This is often used for complete cessation but can support the reduction of smoking as well as shown by an interview study conducted in 2011 (and published in *Nicotine and Tobacco Research* in 2012) (Beard et al., 2012).

2.4 Limitations and Challenges of current research

Currently, we have been focusing on the assumption that self-control can be seen as a limited resource and can thus be depleted resulting in ego-depletion. However, the field of psychology has not necessarily come to the conclusion that this assumption can be completely supported, but is merely a theory. This has to do with multiple reasons, the two main factors being that this theory is subject to the replication crisis and that the original researchers have been suspected of P-hacking.

The former problem, being the subject of the replication crisis, is something that many theories within psychology are subjected to. The replication crisis has to do with the fact that many psychological studies are methodically difficult, or even impossible, to reproduce. In 2012 a study has been done on this phenomenon where they came to the conclusion that it is likely that there are serious replication problems in the field of Psychology (Pashler & Harris, 2012). Since then this problem has received more attention and thus studies after this time period are often considered more credible as these have thus been under stricter supervision.

The other accusation that has risen on the theory of ego-depletion is the suspicion of P-hacking. P-hacking is known under a variety of names such as data dredging and data snooping, but they all boil down to misrepresenting results. This is done by using a variety of analyses methods and actively looking for those that give the desired results, and then reporting only these outcomes. As can be expected, this means that the results of a study can't be trusted if the analysis is not clearly defined. However, this does not mean that the results are false. This suspicion of P-hacking means that the results of this study need to be taken with a grain of salt and one should be sceptical. This scepticism also arose in a group of researchers in 2019 where Friese et al. investigated these claims (Friese et al., 2019). They investigated multiple studies that have been done on ego-depletion over the years to see if this theory could be supported. In the study, they tried to find evidence to convince those that were opposed to the theory that the theory was true, and those that supported the theory that it was false. In their conclusion they mention that they could find no clear evidence that was able to strongly support the theory of ego-depletion, but

also no clear evidence to *definitively dethrone* it (Friese et al., 2019).

Due to the scope of this thesis, we will not re-investigate this study but will work from the assumption that ego-depletion is an occurring phenomenon and thus that self-control is a limited resource. The reasoning for this has to do with a simple proposition. If we assume that self-control is limited, we must ensure that it never gets completely finished. If it were not limited this means that this emptying could not take place. In other words, the former is more restrictive on our implementation, meaning it should thus also work if the theory of ego-depletion is later shown to be incorrect.

2.5 Human-Computer Interaction techniques

After gathering the knowledge about how we can influence and harness Temperance, we need to look into ways to do this in regard to our chatbot. When designing there are often certain steps one takes which have been written down by the Design Council UK (Design Council, 2019). They define the four steps that are taken as; Discover, Define, Develop, and Deliver. The former two steps have already been taken by noticing that the cessation of smoking is difficult and refining it to how a VIC could help in this. The latter two (Develop & Deliver) will be investigated in this part of this thesis as well as in Chapter Three (Ch. 3). In these last two steps, it is especially important that users are involved. The reason for this is two-fold as explained by (Sharp et al., 2019) and can be summarised in; managing expectations and ownership.

The former boils down to ensuring that users have realistic expectations of the product that is built. The user should not encounter unexpected problems or negative surprises. If the user feels like the product they received is not the product they wanted, they are often displeased and will thus be less likely to use the product (Oliver, 1980). This can often be seen in advertisements where products are often shown from their best side and thus seem perfect but are in reality just ordinary products that weren't really needed by many.

The second point that Sharp et al. bring up is that of the users' feel of ownership. They discuss that users should be involved in the process as this will make them feel like they contributed to the product. This feeling can then grow into a sense of ownership towards the product and are more likely to support its use (Bano et al., 2016).

3 The Chatbot

Now that we have seen what Temperance is, why we use it, and how we can use its related psychological mechanisms to our advantage, we can start on the design of the chatbot. It is important that during this development we make use of the knowledge we have gathered but also incorporate already existing design approaches. For this thesis, we have taken an approach that focuses on the user and on the flourishing of this user. For this we have looked at ways to help cultivate Temperance in a way that is supported by psychological research. However, to incorporate this it is important to understand that not all design approaches start from such a point. Two approaches that do start from a core value/virtue are Value Sensitive Design, and Virtuous Practice Design. These two approaches differ in their focus and execution but both have interesting takes on how to start a design process from such a core value/virtue. We will explore both approaches and compare them, in the end, to understand which one is best suited for our project.

3.1 Value Sensitive Design

Firstly we will look into Value Sensitive Design, also known as VSD. This approach was coined around the 1990s with the goal of designing technology that takes human values into account throughout the design process (Friedman, 1998). In a later book, they mention that this process makes use of the tri-partite methodology. This means that there are three iterative steps that have to be taken to create the optimal outcome (Sears et al., 2007). These steps are referred to as Conceptual Investigations, Empirical Investigations, and Technical Investigations.

The first of the three investigations, the conceptual investigation, focuses on the way the system is set up and how it might influence direct and indirect stakeholders. These investigations generally consist of more high-level research, meaning often reasoning about certain topics. In many cases, an initial outline is made with help of these forms of investigation. Where Conceptual Investigation stops, Empirical Investigation continues. Empirical investigations are done on human activities that can be measured, documented, or simply observed. These forms of investigations look at the ways users interact with the system via both quantitative and qualitative methods. Next to these two investigations, there is the Technological Investigation. For this investigation, it is important to look into the most suitable technology as often already existing technologies are more suitable for specific activities and supporting values whilst having more trouble realising this for other activities and values. An important distinction between this investigation compared to the other two is that the Technical Investigation has two different forms. The first one focuses on how already existing (technological) properties and their underlying mechanisms can hinder or support human values. The second form turns this around and focuses on creating such mechanisms and properties anew.

These three investigations come together in the practical implementation where Friedman et al. give explicit suggestions. They break this down in eight written out steps. They also added two extra subsections regarding a list of often implicated Human Values, and heuristics to use for one's Technical Investigation. We will not go into detail about the list of human values as this is not relevant for the applicability of this implementation in our case. Firstly they suggest starting with a Value, Technology or Context of Use. The latter has to do with the environment in which the technology will be implemented. Their

suggested second step is to identify both the Direct and Indirect Stakeholders. After this identification, they emphasize that the benefits and harms should be identified for each stakeholder after which they should be mapped to corresponding values. For their fifth step, they suggest executing the conceptual investigation on the values that have been shown as essential, i.e. the key values. They mention that these values can be conflicting and that these potential conflicts should thus also be identified. As their seventh step, they suggest integrating these considerations into the organisational structure they are going to be implemented in. The reason for this is due to the possible collision between value importance, and economic or power objectives or other factors. Lastly, they advise interviewing stakeholders for the empirical investigation. This is to better understand what their thoughts are on the already existing technologies but also a possible proposed design.

As mentioned before they also give some heuristics that could be useful during the technical investigation. The first heuristic they bring up has to do with the mapping of trade-offs between values, especially when they are conflicting. The second heuristic is in regards to unanticipated values, where they mention that it is advised to preemptively implement flexibility into your project. Lastly, they mention that underlying protocols that are able to release data should have the option to be turned off.

A few years later, in 2013, Ibo van de Poel discussed one specific aspect of the steps mentioned above, the translation of Values into more concrete and tangible design requirements (van de Poel, 2013). This creation of tangibility is done by first translating the values into norms, and then translating those norms into design requirements. The reason for this is that the translation from Values to design requirements is made more systematic. Van de Poel adds that in this way value judgements are also made more explicit, transparent and debatable due to specific translations at the different levels of the hierarchy of a value. This transparency gives room for reflecting critically and opens the possibility to pinpoint where a disagreement on the design has occurred. A critical note Van de Poel makes here is that this approach is non-deductive and context-dependent. He notes that due to the nature of this approach value judgements arise but the way they are handled makes them more explicit.

3.2 Virtue Practice Design

In more recent years some criticism has arisen surrounding the use of Value Sensitive Design. These criticisms come from the point that VSD has main focus of intervention on the design process, compared to the environments and practices they will be used in (Reijers & Gordijn, 2019). A suggested alternative approach to this is Virtue Practice Design, which focuses on this latter point; the (technical) practices in which the technology is involved. They believe that focusing on just the end product is not enough as for (virtue) ethical technology is also formed by their environment. Reijers and Gordijn do mention that VSD does add applicability of ethical concerns, but that it is not complete as it is merely one approach that could be used for this integration. In addition, they argue that the list of possible values is inexhaustible and thus could never be considered complete. One last point that is brought up is the non-neutrality of the technologies that are designed using VSD. They argue that this non-neutrality can embody itself in for example racist values as noted in "Do Artifacts Have Politics?" (Winner, 1980). This non-neutrality can be derived from the unclear difference between 'values' and general 'preferences' as VSD defines

values as "what a person or group of people consider important in life" (Friedman, 1996). This unclarity could cause that 'values' that stakeholders might find important (their preferences) are deemed more important than moral values that can take other criteria into account.

To counteract these problems Reijers and Gordijn made a first attempt to develop, Virtue Practice Design. For this they use the work of Shannon Vallor, a philosopher of technology, and Alasdair MacIntyre, a moral and political philosopher (MacIntyre, 2007; Vallor, 2018)[2]. Their influences are visible throughout multiple parts of the approach Reijers and Gordijn take but these will be shortly highlighted here as well as this is important when we later compare both design approaches

The influence of MacIntyre is very visible in the main difference between VSD and VPD. Reijers and Gordijn mention that VPD is not the design of 'a thing' but rather the design of shaping a practice, as is MacIntyre's view on Virtue Ethics. MacIntyre mentions that human qualities, that are internal, need exercise to achieve excellence (MacIntyre, 2007). These internal factors can be better explained by showing the difference in excellence of two different activities; playing the piano and making a good trade. How well someone plays the piano can be determined by how well the piece is played, the judgement is contained within the activity. To put it simply, how well someone plays the piano can only be compared to someone else playing the piano. This is not the case with, for example, making a good trade as the determining factor for a good trade is often the amount of profit which is gained. The key difference between these two excellences is that the latter is dependent on an external factor, whereas the former is dependent on internal factors. This leads to MacIntyre saying that the standards of excellence will also depend on narratives. Being able to play the piano well could have entirely different standards when looking at it from a musician's standard compared to a listener's standard. This means that these standards can differ from person to person, meaning the system that is developed needs to be adaptive as well when considering using virtues. In other words, VPD puts the individual more towards its centre stage as it focused on internal factors, whereas VSD would put the majority's opinion in the spotlight. This thus also characterises their different approaches in designing a form of new technology.

As mentioned Reijers and Gordijn also take inspiration from Shannon Vallor. They do this by making use of the heuristics of virtues proposed by her (Vallor, 2018). Vallor has created an extensive list of virtues that have been around for a long time, but updating them to the current technological world refering to them as Technormal Virtues. And as these virtues are derived from the discussion on what is morally good and bad, these virtues are not arbitrary judgements as values are in VSD. The values in VSD could simply be decided by judgements of those who apply it, compared to being better morally based.

These criticisms and adjustments lead Reijers and Gordijn to propose the alternative tripartite method of VPD. They propose three alternative parts; Investigation of narratives, Reflection based on the virtues, and Prescriptions for technical practices.

For the first part of this tripartite process, the investigation of narratives, the aim is to understand relevant practices. This is done by gathering narratives and investigating them. These narratives can be considered the context in which the technology is supposed

[2]These sources have been republished at later points in time so the newest (and most accurate source) are currently cited. For the exact sources refer to their original article (Reijers & Gordijn, 2019).

to be used provided by those involved. They further elaborate on this by explaining that the goal of the technology should be to help the user cultivate a chosen virtue. And that it is important to first know how this can be done by and for the user, in other words, what the narrative is to hone this virtue.

The next phase is making the practice investigated previously more concrete. During this phase the previously gathered practices will be reflected on, and it will be investigated which virtue aligns best with the practices that might be useful. This can be done by making use of the virtue heuristics that were mentioned before (Vallor, 2018).

When considering the third and last step one might find many similarities between VSD and VPD, as also noted by the authors (Reijers & Gordijn, 2019). This step closely follows the earlier explained Technical Investigation by Friedman, but is significantly broader. This has to do with it focusing on the entirety of the practice, instead of solely focusing on the technical design. This in turn means that there are multiple aspects that can be important prescriptions for this stage such as training, education and regulation of the chosen practice. Reijers and Gordijn categorise these different types of prescriptions into design, regulation, and human development. These categorisations give way to developing a system that can support the cultivation of a virtue, by not solely focusing on the end goal, but also on the way to it. This approach is greatly similar to the suggestions we already made before, the (perfect) end-goal might not easily be reached but trying to get there is already a great step.

An important notion that needs to be made is that this approach is an initial attempt, not a perfected solution. Reijers and Gordijn mention that this approach also needs to be critically investigated as Virtue Ethics in and of itself can be critisised as well (Reijers & Gordijn, 2019). One of the main arguments made is by MacIntyre himself by mentioning the high correlation between tradition and virtue (MacIntyre, 2007). He explains that traditions are often considered good, or even virtuous, but due to this some traditions might uphold morally debatable virtues.

3.3 Our Design approach

During this thesis, we have taken many steps that correlate with the approach of Friedman; Value Sensitive Design (Friedman, 1998). But for the implementation itself, we haven't yet made use of the addition of Van de Poel to this approach. Comparing this to Virtue Practice Design, as proposed by Reijers and Gordijn has shown that for our project a variety of parts of this approach fit better with our current work (Reijers & Gordijn, 2019). A part of this choice has do with the fact that VSD makes use of values as their heuristic compared to making use of virtues, which we have done up until now. Next to this the implementation of the explained values that are used in VSD, can more easily be used in a deontological setting or a consequentialist one, compared to a Virtue Ethical setting. We will thus make use of VPD as our main design approach with taking parts of VSD where deemed fit, as a combination of these approaches can be fruitful. This has been supported by a reply from Steven Umbrello to VPD, where he mentions that VPD is currently still flawed but can be a useful method to add to VSD (Umbrello, 2020).

We have already gone through the first two steps of both approaches by investigating our stakeholders and investigating which virtue/value we believe is most important. As described in Chapter 1 we will focus on Moderate Active Smokers and try to support

their reduction via the cultivation of Temperance. For our last phase, we will be giving a variety of recommendations including but not limited to supporting the user with their development as well as trying to educate them on what they are doing and why this is beneficial. Next to this we will elaborate on already existing technologies, and investigate their strengths as well as explore current flaws that would need improvement.

3.4 Implementation Recommendations

Now that we have seen which paths are possible and have chosen to go for a combination between VSD and VDP, it is important to make our recommendations more concrete. As discussed in Chapter 2 (Ch. 2) there are two main aspects that we want to focus on; Planning and Training. These two find root in the literature mentioned previously but can be seen (partially) implemented in already existing technologies.

It is important to recognise that these recommendations are not complete, nor can they be. Creating such a virtue-supporting system design can never be complete, as can be seen from the two aforementioned methods. Therefore we will focus on the aspects we believe to be important for the creation of the Virtue Informed Chatbot we aim to establish. This also means that we will not discuss basic Human-Computer Interaction aspects, such as 'the system should be intuitive' or 'the text in the app must be easily readable', in great detail. These aspects are very important for the creation, but these are less important for the framework we are currently trying to create.

3.4.1 Planning

As discussed earlier multiple factors can influence one's temperance. One that can be easily supported by a chatbot or similar technology is planning. Two factors that we mentioned in chapter 2 that have to do with this aspect are timely reminders and future planning. For these aspects, it is important to look at already existing technologies that support this.

QuitNow! already does (part of) this from the moment you start using the app. It asks you to provide a quitting date that can both be in the past, but also in the future. For the latter option, it then gives a count-down timer of when you will have quit. This in and of itself, however, is not enough when looking to reduce future smoke intake. We want to support the user on their journey towards their reduction as well, not solely when they have already quit. In other words, we will need to look into how this planning can be optimised. This initial goal setting is a good start for this as we have already seen from previous studies (Gollwitzer, 1999). This means that we need to have an initial (realistic) goal date when the user has achieved their intended reduction. During this phase, it is important to discuss with the user how easy they think this journey would be. Do they believe that they can easily reduce their intake? Then this date can be relatively soon. Does the user feel like they will struggle a lot with this? Then this date should be postponed to optimize consistent results. The goal is that this date is final but should be revisable when the user shows great struggle with the reduction process.

The second aspect of planning ties in neatly with this as timely reminders can help with this guided reduction. After this date has been set timely reminders can support the user in remembering their goal. These reminders should focus on the user themselves, and how they are currently experiencing their journey. This can be done by for example letting the VIC ask the user "Hey, today you were planning on smoking X cigarettes. How is

that going right now?"[3]. This example should be given roughly around the late morning and possibly again in the afternoon, where the user has most likely already smoked some cigarettes but can still adjust their smoking behaviour for that day. Below you can see a table with example phrases and when these can be used (Table 1). These reminders should not be too often and especially in later phases it might can prove fruitful to reduce the number of reminders and users have grown accustomed to them and might find them obnoxious. This should be further investigated once implemented, by for example making use of AB Testing, and actively asking for feedback on this.

Moment of Usage	Example Phrase
Soon after the User wakes up	"Yesterday you were able to smoke just X cigarettes. Why dont you try to keep this up!"
	"For the past week you have only smoked X cigarettes a day. Let's try if you can go for $X - 1$."
After roughly 1/3 of the User's day has passed	"Hey, you were planning on smoking X cigarettes today. How is that going right now?"
	"Hopefully you feel like you can stick to your goal of X cigarettes today, if not let me know!"
After roughly 2/3 of the User's day has passed	"Your getting closer to your goal, are you still on track?"
	"This morning you decided your goal was X cigarettes, how is this going?"
Close to bedtime	"How did today go? Were you able to stick to your planning?"
	"Let's recap today. Was it easy to smoke just X cigarettes or was it a challenge?"

Table 1: *Table with example phrases the VIC could use as notifications to the user, where X represents a variable integer depending on user input.*

3.4.2 Training

As described previously to hone a virtue it is important to continuously train it. This can be done in a variety of ways as described in chapter 2, the three main aspects we will focus on in this part will be: Incremental Exercise, Resilience to Ego-depletion and Enhancing your Environment. These three aspects are closely tied to each other in regard to training temperance optimally in the context of reducing smoke intake.

When looking at apps that help a user train something, this is often done in multiple steps. Firstly, investigate the user's current level of the practice. This is because each user has a different starting *levels*, and can thus handle different amounts/intensities of exercise. For our purpose, we will be determining the entry level of the user on a self-based questionnaire. The user will thus be able to have a large amount of control over the execution of the process. These questions will not be extensive as to not scare users off, but will be rather short and simple. Questions that can be asked are for example;

- How many cigarettes do you currently smoke per day? If you don't smoke cigarettes, what would your estimated daily nicotine intake be?

[3]Where X would be replaced by the number of cigarettes the user planned to maximally smoke that day.

- Have you tried reducing your smoking before using this VIC?
- How confident are you in your journey to reduction?
- Do you feel like you are heavily dependent on your cigarettes?
- etc.

These questions would give an initial overview of how easy the user thinks their reduction process will be, and thus give an indication on how much they might be able to handle. This however does not mean that a user will be thrown in the deep when they feel confident. The goal is merely to pre-process how swift the journey might be. To give this some more context we can think of someone who feels like they are really addicted to their nicotine and will thus not be able to easily reduce their intake. For this user, it is important to very slowly increment the difficulty, and first suggest leaving only one cigarette a day and trying this for two weeks. If they succeed, try to leave two cigarettes and try it for two weeks, etc. For someone who would be more confident in their reduction, this process can be sped up more, by for example suggesting to leave two cigarettes from the start compared to only one.

After determining the starting 'difficulty' it is also important to train this outside of just the context of smoking. As studies have shown, practising Self-Control (the psychological term for Temperance) in any context improves one's overall Self-Control (Muraven, 2010). A drawback of exercising Self-Control, however, is the occurance of Ego-Depletion, where someone has difficulty exercising Self-Control multiple times in close succesion. This means that these training moments should not occur at times when it is important to exert Self-Control for other activities such as work. For this reason, we suggest that the VIC recommends exercises to enhance Self-Control between the users Dinner and Bed-time. This way the user should have some time to exercise without having too many negative drawbacks from Ego-Depletion. Exercises that can be suggested are for example:[4]

- Brushing one's teeth with their non-dominant hand.
- Meditating for an increasing number of minutes.
- Doing Yoga for an increasing amount of time.
- Writing with their non-dominant hand.
- Focus on correcting one's posture.
- Try not to use curse/swear words.

[4]This list is far from complete as there are many ways to exercise Self-Control, but this is to have a better overview of what the possibilities are.

3.4.3 Motivation

After reading this thesis, this subsection might seem odd compared to the previous two. In Chapter 2 we explained as we believe that users will already be motivated to reduce their smoking intake but this does not mean that this motivation can be neglected. In many smoking apps such as *QuitNow!* and *Smoke Free*, rely often on external motivation. A study done in 2014 has shown that extrinsic goals, such as money, are dominantly present as goal content for cessation apps (Choi et al., 2014). Both these apps will give facts on how much money the user has saved and how long they have now abstained from smoking, but neither shows indication to support the journey to this cessation. This is entirely perpendicular to our goal of self-enrichment, as we want to hone an intrinsic virtue that will support an extrinsic goal. For this reason, it is important that we look at ways to support this stance. Ways that have been shown to support this inner change are discussing

3.4.4 Other Factors

As mentioned at the start of this subsection, we will not be able to cover all aspects that are needed for an optimal VIC. In the subsections above we have shown aspects that are generally not covered, and are specific to our system.

But there are also other factors, such as engagement with the system, that are not covered by our psychological research. This aspect of engagement with the system can already be found in a variety of existing systems. In gaming apps, this is for example done by giving the user a reward for opening the app each day where the rewards scale for each consecutive day. Another app that uses a form of external motivation is called 'Forest'. This app helps a user use their phone *less*, and motivates them by slowly growing a tree over time for the duration that the user does not interact with their phone. If a user would use their phone before the 'end time' has passed their tree will stop growing and die. A main similarity that is often seen between these apps and approaches is that the user receives something from the system, be it a direct reward or something that grows over time. This is to keep the users engaged and as studies have shown that engagement increases the likelihood that someone quits smoking which we assume will work similarly beneficial for reduction (Regmi et al., 2017). Therefore it is crucial that we look for something that keeps the users engaged with the system without them becoming dependent on the app, as this would be replacing one bad habit with another.

4 Conclusion

In this thesis, we have made an initial attempt to design a Virtue Informed Chatbot that supported the reduction of smoke intake for individual users.

We have investigated different stakeholders that are in some way influenced by smoking. Since our goal is to help the reduction of smoking itself we further investigated Active smokers. Following this, we have shown why we believe Temperance is the most fit virtue to achieve this goal and in which ways this can be beneficial. To provide more concrete methods of intervention we researched how Temperance is defined and investigated in the field of Psychology. This was necessary to ensure that our recommended adjustments in behaviour were found on a solid basis, as well as to be able to figure out which flaws and strengths we could use to our advantage. Lastly, we combined these findings into recommendations for a possible VIC design. This design makes use of technologies that already exist and tries to use the strengths of one to negate the flaws of another, creating a more complete system.

Despite our best efforts, we must note that this design is not complete as there are many more aspects to a nicotine/smoke addiction than we were able to cover. This has been both due to the limited amount of time as well as it being too broad for the scope of this thesis. It is important to investigate this in more detail when trying to develop this design further. Next to this, these recommendations are only recommendations, not implementations. We have given many handles as to how to possibly implement this and further test this, but we have not been able to implement this ourselves. The reason for this is two-fold as we wanted to ensure our recommendations were sound and found basis in already existing systems, as well as it being difficult to test. The time scope of our thesis is roughly six months and this in combination with aiming to cultivate a virtue makes it nearly impossible to draw conclusions from an initial (short) testing phase.

References

Aristotle. (2014). *Cambridge texts in the history of philosophy: Aristotle: Nicomachean ethics* (R. Crisp, Ed.; 2nd ed.) [Book 3 (III), Chapter 10 (X).]. Cambridge University Press.

Bano, M., Zowghi, D., & da Rimini, F. (2016). User satisfaction and system success: An empirical exploration of user involvement in software development. *Empirical Software Engineering*, *22*(5), 2339–2372. https://doi.org/10.1007/s10664-016-9465-1

Barker, J. E., & Munakata, Y. (2015). Time isn't of the essence. *Psychological Science*, *26*(12), 1898–1908. https://doi.org/10.1177/0956797615604625

Baumeister, R. F. (2003). Ego depletion and self-regulation failure: A resource model of self-control. *Alcoholism: Clinical and Experimental Research*, *27*(2), 281–284. https://doi.org/10.1097/01.alc.0000060879.61384.a4

Baumeister, R. F. (2018). Self-regulation and self-control. https://doi.org/10.4324/9781315175775

Baumeister, R. F., & Vohs, K. D. (2007). Self-regulation, ego depletion, and motivation. *Social and Personality Psychology Compass*, *1*(1), 115–128. https://doi.org/10.1111/j.1751-9004.2007.00001.x

Beard, E., Vangeli, E., Michie, S., & West, R. (2012). The use of nicotine replacement therapy for smoking reduction and temporary abstinence: An interview study. *Nicotine and Tobacco Research*, *14*(7), 849–856. https://doi.org/10.1093/ntr/ntr297

Bewley, B. R., Bland, J. M., & Harris, R. (1974). Factors associated with the starting of cigarette smoking by primary school children. *Journal of Epidemiology and Community Health*, *28*(1), 37–44. https://doi.org/10.1136/jech.28.1.37

Blackwell, A. K., Lee, I., Scollo, M., Wakefield, M., Munafò, M. R., & Marteau, T. (2020). Size matters but when, why and for whom?

Bommelé, M., Jeroen; Willemsen. (2020). Kerncijfers roken 2018. *Nationaal Expertisecentrum Takabsontmoediging*. https://www.trimbos.nl/docs/f9f84642-6f50-4c57-9392-019525f18ef2.pdf

Bonein, A., & Denant-Boèmont, L. (2015). Self-control, commitment and peer pressure: A laboratory experiment. *Experimental Economics*, *18*(4), 543–568. https://doi.org/10.1007/s10683-014-9419-7

Boulos, D. N., Loffredo, C. A., El Setouhy, M., Abdel-Aziz, F., Israel, E., & Mohamed, M. K. (2009). Nondaily, light daily, and moderate-to-heavy cigarette smokers in a rural area of egypt: A population-based survey. *Nicotine and Tobacco Research*, *11*(2), 134–138. https://doi.org/10.1093/ntr/ntp016

Cambridge University Press. (2022a). Cambridge dictionary [Passive smoking; Accessed: 06-12-2022, Snapshot: 04-05-2022]. http://web.archive.org/web/20220504073917/https://dictionary.cambridge.org/dictionary/english/passive-smoking

Cambridge University Press. (2022b). Cambridge dictionary [Temperance; Accessed: 06-12-2022, Snapshot: 04-05-2022]. http://web.archive.org/web/20220504045449/https://dictionary.cambridge.org/dictionary/english/temperance

Cambridge University Press. (2022c). Cambridge dictionary [Self-Discipline; Accessed: 06-12-2022, Snapshot: 04-05-2022]. http://web.archive.org/web/20220504065541/https://dictionary.cambridge.org/dictionary/english/self-discipline

Cambridge University Press. (2022d). Cambridge dictionary [Perseverance; Accessed: 02-01-2023, Snapshot: 27-03-2022]. http://web.archive.org/web/20220327210211/https://dictionary.cambridge.org/dictionary/english/perseverance

Cambridge University Press. (2022e). Cambridge dictionary [Self-Indulgence; Accessed: 02-01-2023, Snapshot: 04-05-2022]. http://web.archive.org/web/20220504062247/http://dictionary.cambridge.org/dictionary/english/self-indulgence

Choi, J., Noh, G.-Y., & Park, D.-J. (2014). Smoking cessation apps for smartphones: Content analysis with the self-determination theory. *J Med Internet Res*, *16*(2), e44. https://doi.org/10.2196/jmir.3061

Design Council. (2019). The double diamond: A universally accepted depiction of the design process [Accessed: 21-01-2022, Snapshot: 17-01-2023]. http://web.archive.org/web/20230117052808/https://www.designcouncil.org.uk/our-work/news-opinion/double-diamond-universally-accepted-depiction-design-process

Doll, R., & Hill, A. B. (1950). Smoking and carcinoma of the lung. *BMJ*, *2*(4682), 739–748. https://doi.org/10.1136/bmj.2.4682.739

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2016). Situational strategies for self-control. *Perspectives on Psychological Science*, *11*(1), 35–55. https://doi.org/10.1177/1745691615623247

Duckworth, A. L., Milkman, K. L., & Laibson, D. (2018). Beyond willpower: Strategies for reducing failures of self-control. *Psychological Science in the Public Interest*, *19*(3), 102–129. https://doi.org/10.1177/1529100618821893

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial. *JMIR Ment Health*, *4*(2), e19. https://doi.org/10.2196/mental.7785

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences*, *103*(31), 11778–11783. https://doi.org/10.1073/pnas.0602659103

Friedman, B. (1996). Value-sensitive design. *Interactions*, *3*(6), 16–23. https://doi.org/10.1145/242485.242493

Friedman, B. (1998). *Human values and the design of computer technology*. Center for the Study of Linguistics; Information.

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2019). Is ego depletion real? an analysis of arguments [PMID: 29591537]. *Personality and Social Psychology Review*, *23*(2), 107–131. https://doi.org/10.1177/1088868318762183

Gallani, S. (2017). Incentives, peer pressure, and behavior persistence. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2897572

Godtfredsen, N. S. (2005). Effect of smoking reduction on lung cancer risk. *JAMA*, *294*(12), 1505. https://doi.org/10.1001/jama.294.12.1505

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*(7), 493–503. https://doi.org/10.1037/0003-066x.54.7.493

Hughes, J., & Carpenter, M. (2006). Does smoking reduction increase future cessation and decrease disease risk? a qualitative review. *Nicotine and Tobacco Research*, *8*(6), 739–749. https://doi.org/10.1080/14622200600789726

Hursthouse, R., & Pettigrove, G. (2022). Virtue Ethics. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Winter 2022). Metaphysics Research Lab, Stanford University.

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, *126*(1), 109–114. https://doi.org/10.1016/j.cognition.2012.08.004

MacIntyre, A. (2007). *After virtue: A study in moral theory, third edition*. University of Notre Dame Press.

McBride, P. E. (1992). The health consequences of smoking: Cardiovascular diseases. *Medical Clinics of North America*, *76*(2), 333–353. https://doi.org/10.1016/s0025-7125(16)30356-x

Muraven, M. (2010). Building self-control strength: Practicing self-control leads to improved self-control performance. *Journal of Experimental Social Psychology*, *46*(2), 465–468. https://doi.org/https://doi.org/10.1016/j.jesp.2009.12.011

Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources [PMID: 15018677]. *Personality and Social Psychology Bulletin*, *29*(7), 894–906. https://doi.org/10.1177/0146167203029007008

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, *17*(4), 460. https://doi.org/10.2307/3150499

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? three arguments examined. *Perspectives on Psychological Science*, *7*(6), 531–536. https://doi.org/10.1177/1745691612463401

Rachlin, H. (1974). Self-control. *Behaviorism*, *2*(1), 94–107. Retrieved January 11, 2023, from http://www.jstor.org/stable/27758811

Regmi, K., Kassim, N., Ahmad, N., & Tuah, N. (2017). Effectiveness of mobile apps for smoking cessation: A review. *Tobacco Prevention and Cessation*, *3*(April). https://doi.org/10.18332/tpc/70088

Reijers, W., & Gordijn, B. (2019). Moving from value sensitive design to virtuous practice design. *Journal of Information, Communication and Ethics in Society*, *17*(2), 196–209. https://doi.org/10.1108/jices-10-2018-0080

Sears, A., Jacko, J. A., & Jacko, J. A. (2007). The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications, second edition. https://doi.org/10.1201/9781410615862

Sharp, H., Preece, J., & Rogers, Y. (2019). 2.2.2 the importance of involving users. In *Interaction design: Beyond human-computer interaction* (pp. 43–45). Wiley.

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology*, *26*(6), 978–986. https://doi.org/10.1037/0012-1649.26.6.978

Simpson, A., Riggs, K. J., Beck, S. R., Gorniak, S. L., Wu, Y., Abbott, D., & Diamond, A. (2011). Refining the understanding of inhibitory processes: How response prepotency is created and overcome. *Developmental Science*, *15*(1), 62–73. https://doi.org/10.1111/j.1467-7687.2011.01105.x

Stoppen met roken commercial #rookvrijopgroeien. (2017). https://www.youtube.com/watch?v=HvGD4vz2oZM

Tao, T., Wang, L., Fan, C., & Gao, W. (2014). Development of self-control in children aged 3 to 9 years: Perspective from a dual-systems model. *Scientific Reports*, *4*(1). https://doi.org/10.1038/srep07272

Tominey, S. L., & McClelland, M. M. (2011). Red light, purple light: Findings from a randomized trial using circle time games to improve behavioral self-regulation in preschool. *Early Education and Development*, *22*(3), 489–519. https://doi.org/10.1080/10409289.2011.574258

Umbrello, S. (2020). Combinatory and complementary practices of values and virtues in design: A reply to reijers and gordijn. *Filosofia*, (65), 107–121. https://doi.org/10.13135/2704-8195/5236

Vallor, S. (2018). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.

van de Poel, I. (2013). Translating values into design requirements. *Philosophy and Engineering: Reflections on Practice, Principles and Process*, 253–266. https://doi.org/10.1007/978-94-007-7762-0_20

Wilson, D., Wakefield, M., Owen, N., & Roberts, L. (1992). Characteristics of heavy smokers. *Preventive Medicine*, *21*(3), 311–319. https://doi.org/https://doi.org/10.1016/0091-7435(92)90030-L

Winner, L. (1980). Do artifacts have politics? *Daedalus*, *109*(1), 121–136. Retrieved February 7, 2023, from http://www.jstor.org/stable/20024652