

BACHELOR THESIS  
ARTIFICIAL INTELLIGENCE

**Radboud University**



---

# Falsification machines in medical decision making

---

*Author:*

Ralph van Eerd  
S4800702

*First supervisor:*

Dr. W.F.G. Haselager  
Artificial Intelligence, Donders Institute  
w.haselager@donders.ru.nl

*Second supervisor:*

Dr. H.K. Schraffenberger  
Donders Institute  
h.schraffenberger@ru.nl



27 January 2021

## **Abstract**

Decision Support Systems are systems that assist users during complex decision making processes. However, in the medical field is the risk that the use of these support systems can lead to increased numbers of erroneous diagnoses. In most of these cases, it could be complicated to tell who is responsible for these errors. A proposed solution to increase the feeling of responsibility of the user, also known as the sense of agency, is a Falsification Machine in addition to a Decision Support System. A Falsification Machine asks questions to check the final decision made by the user. To investigate if a Falsification Machine is a suitable solution, an experiment was conducted with students to show if there is an increase in the sense of agency when a Falsification Machine, in combination with a Decision Support System, was used.

# Table of Contents

<b>1. Introduction.....</b>	<b>4</b>
<b>2. Theoretical approach.....</b>	<b>7</b>
2.1. <i>Risks of Decision Support System.....</i>	7
2.2. <i>Sense of Agency.....</i>	8
2.3. <i>Meaningful human control.....</i>	10
2.4. <i>Overreliance.....</i>	11
2.5. <i>Linking of concepts.....</i>	12
2.6. <i>Eight modules for implementing a FM.....</i>	13
<b>3. Methods.....</b>	<b>15</b>
3.1. <i>Design of the research.....</i>	15
3.2. <i>Operationalization.....</i>	15
3.3. <i>Data collection method.....</i>	16
3.4. <i>Construction of the cases.....</i>	17
3.5. <i>Construction of FM questions and statements.....</i>	20
3.6. <i>Structure of the questionnaire.....</i>	22
3.7. <i>Data analysis.....</i>	23
<b>4. Results.....</b>	<b>26</b>
4.1. <i>Medical cases.....</i>	29
4.2. <i>Law cases.....</i>	29
<b>5. Conclusion.....</b>	<b>30</b>
<b>6. Discussion.....</b>	<b>33</b>
6.1. <i>Further research.....</i>	34
<b>7. Bibliography.....</b>	<b>35</b>
<b>8. Appendix.....</b>	<b>38</b>
8.1. <i>Table to create medical cases.....</i>	38
8.2. <i>Conversations.....</i>	38
8.2.1. <i>Conversation with medical student.....</i>	38
8.2.2. <i>Conversation with law student.....</i>	39
8.3. <i>Cases.....</i>	40
8.3.1. <i>Casus 1: Ongewild gewichtsverlies.....</i>	40
8.3.2. <i>Casus 2: droge hoest.....</i>	42
8.3.3. <i>Casus 3: Bijlmer Noodweer.....</i>	44
8.3.4. <i>Casus 4: Steen van een viaduct.....</i>	46
8.4. <i>Case results.....</i>	48
8.4.1. <i>Case 1.....</i>	48
8.4.2. <i>Case 2.....</i>	48
8.4.3. <i>Case 3.....</i>	49
8.4.4. <i>Case 4.....</i>	49
8.4.5. <i>Sense of Agency participants.....</i>	49

# 1. Introduction

Diagnosing patients is an important task in the medical field. When diagnosing patients, health workers, such as doctors, nurses, and general practitioners, can be aided by electronic devices, such as a computer(program). These programs aid during the decision-making process and they are called clinical decision support systems, and they will be referred to as CDSS. The function of a decision support system (DSS) in general, is to gather and analyse data, compiling the data into useful information and provide advice to the user during complex decision problems (Arnott & Pervan, 2015). A CDSS is a type of DSS, but it is used in the medical field.

In addition to a CDSS, there are other types of DSS in various fields, such as management, ICT, and finance (Eom & Kim, 2006), but all these types of DSS share the same fundamental characteristics. These characteristics include a logic model, which consists of the relations between variables (or rules) used in the environment; a data model, which combines the data and rules and generates the output; a data collection, which offers information to the data model; and lastly, a user interface, which communicates with the user and provides the output in the form of recommendations (Finlay, 1990).

An example of a CDSS used in the Netherlands is Gaston. Gaston consists of two components, namely a guideline editor and a decision support module (DSM). With the guideline editor, the user is able to describe when and how Gaston has to act. The DSM collects data from multiple information systems, such as the hospital information system (HIS) and the patient's file. The DSM will provide the user with warnings, advice or help if a situation, which is stored in the guideline editor, occurs. (Korsten, Grouls, & Minderman, 2008)

However, the use of a CDSS involves some potential risks. These risks include negative impact on user skills, alert fatigue, system & content maintenance, user distrust, inaccurate or poor-quality data (Sutton, et al., 2020), automation bias and automation. The risks that will mainly be focused on in this thesis is the automation bias and automation complacency, which is the overreliance on the performance of the CDSS. (Goddard, Roudsari, & Wyatt, 2011)

Automation bias on CDSS can lead to errors in diagnosing patients due to incorrect advice from the CDSS. Research from Friedman et al. (1999) has shown that in some cases doctors substitute their correct diagnosis for an erroneous diagnosis due to flawed advice from the CDSS. In 12.0% of the cases a positive change, an erroneous diagnosis corrected by the CDSS into a correct diagnosis, was observed. However, in 6.0% of the cases a negative change was observed. This

negative change is caused by overreliance by the user on the CDSS. This results in an overall gain of 6% when using the CDSS. However, the negative change in diagnosis is not acceptable, because automation failure in the medical field can result in fatalities and shows that improvements when using a CDSS are needed. The effects of overreliance are still being researched and a concrete solution has not been found yet. A proposed solution to the problem of overreliance is to continuously monitor the CDSS, however, this solution has not been investigated yet (Sutton, et al., 2020).

An alternative possible solution is to continuously monitor not only the CDSS, but the entire decision process, including the health worker. This can be done with a Falsification Machine (FM). A FM tries to help the user during or after the decision-making process by asking questions or pinpointing important factors. Consequently, the machine can check whether the user made a conscious decision, held into account all variables and, therefore, was more in control.

A problem that can arise because of the dependency on the support systems is the lack of human control in the decision-making process, resulting in the person not feeling responsible for the outcome of the decision. There are two components that are important when discussing human control. The first is the sense of agency of the human, which is the feeling of how much control the user has over their actions and their consequences (Moore, 2016). The second component of human control is the meaningfulness of human control. Not computers and their algorithms, but humans should be in control, and thus responsible, for relevant decisions. Therefore, meaningful human control is present if the user is in control in the decision-making process. (van den Hoven & Santoni de Sio, 2018)

When a doctor makes a diagnosis, they make the decision. As a consequence, the doctor is responsible for the decision, whether it is a correct or a faulty diagnosis. Because the doctor is responsible for the decision made, and not the CDSS, there is meaningful human control. However, if the doctor changes his correct diagnosis to a faulty one because of the advice the CDSS gave him, there is still meaningful human control, but overreliance as well. This is because the doctor's confidence in the CDSS is higher than in his own knowledge (Goddard, Roudsari, & Wyatt, 2011). Consequently, the doctor might feel as if he is not responsible for the decision, as it was the CDSS's advice that caused the faulty diagnosis, resulting in a lack of sense of agency from the doctor. A FM is designed to increase the doctor's trust in his own knowledge, resulting in an increase in the sense of agency for the doctor.

The goal of this research is to analyse a FM as a possible solution to the problem of lack of sense of agency for a doctor when there is overreliance on a CDSS. Therefore, the following research question has been made:

“What is the influence of a falsification machine on the sense of agency for medical students when making a decision in the medical field in comparison with non-medical students making a decision in the medical field with a falsification machine?”

It is expected that there will be a bigger increase in the sense of agency for medical students when they make a decision in the medical field with the use of a falsification machine than for non-medical students when making a decision in the medical field with the use of a falsification machine. This expectation is due to the expertise of the medical students in comparison with non-medical students, because of the difference in education. This is because experts rely less on support systems in comparison to non-experts in the field of the expert (Goddard, Roudsari, & Wyatt, 2012). In this research, the medical students will be used as representatives of the medical field. This is because they are more easily reached than doctors or other medical specialists.

In the following chapters, an approach is proposed in order to find an answer to the research question. Chapter two will cover the theoretical approach, by explaining the central concepts. In chapter three, the methods will be explained. This will include how the experiment is constructed and performed. In chapter four, results will be analysed, followed by a conclusion in chapter five. Chapter six contains a discussion, stating possible future research and improvements for the current research.

## 2. Theoretical approach

In this chapter, the risks of the decision support system will be clarified. Furthermore, the concepts of sense of agency, overreliance, meaningful human control and their relationship will be explained and analysed. Moreover, the multiple levels of implementing the FM and how the FM creates the questions will be discussed. This explanation of the levels of implementation of the falsification machine might contribute to the relevancy in the Artificial Intelligence field. This research can act as a blueprint for implementing the possible parts of a FM and show the results of using a FM alongside a DSS.

### 2.1. Risks of Decision Support System

There are numerous risks when a DSS is used. Some of these risks will be explained in this part. The first risk is a negative impact on the user's skill level. When a person performs a task, they are responsible for that task. However, with the use of a DSS an impression can be created that checking the correctness of the task is no longer necessary. Another pitfall is that the user's skills are not used as often as before the DSS was applied. This creates a dependency for the user on the DSS for the specific task. Consequently, the user trusts the DSS more than their own knowledge, resulting in an overreliance on the DSS. (Sutton, et al., 2020)

The second risk of using a DSS is alert fatigueness. This risk emerges if the DSS provides too many insignificant alerts or suggestions, which leads to habituation for the user and the user distrusting the advice from the DSS. As a consequence, the DSS will be used less often. (Sutton, et al., 2020)

A third risk relates to the system and content maintenance of the DSS. System and content maintenance is the difficulty of keeping the DSS operating and up to date. The challenge of maintaining the content used by the DSS is that the content is rapidly changing and the environment in which it operates is complex. If the data that the DSS uses is incomplete or even incorrect the recommendation or advice given by the DSS will be incomplete or incorrect, which will result in more automation failures. (Sutton, et al., 2020)

User distrust is the fourth risk of using a DSS. This means that the user of the DSS does not agree with the way the DSS operates or is implemented. As a consequence, the DSS is used less often. (Sutton, et al., 2020)

The fifth risk is inaccurate or poor-quality data. A DSS uses data to produce advice or recommendations. However, an environment, in which the DSS operates, can contain multiple data sources, which have to be combined into one dataset for the DSS. These sources can contain duplicate or incorrect data. This data will then be used in the decision making process, which can lead to errors. (Sutton, et al., 2020)

Goddard, Roudsari and Wyatt mention in their paper (2011) that there is another potential risk when using a DSS, consisting of two components, namely automation bias and automation complacency. The term automation bias entails that a user prefers to rely on information provided by the DSS instead of information provided by a human, either the user itself or another human. This dependency is present even if the advice is incorrect, which results in the user overrelying on the DSS. Automation complacency is similar. However, the difference is that automation complacency focuses on the user overrelying on the DSS's accuracy (Grissinger, 2019), whereas automation bias focuses on how much the user relies on a DSS compared to human knowledge (Goddard, Roudsari, & Wyatt, 2011). The automation complacency results in less surveillance of the system, because the user trusts the DSS to give the correct advice or recommendation. Therefore, the user over relies on the DSS. It is important to note that both experts and non-experts can have automation bias and automation complacency (Grissinger, 2019).

There are three reasons for automation bias and automation complacency to happen. The first reason for relying more on a DSS instead of own knowledge is the excessive trust in the accuracy of a DSS. The second reason is that it is easier for a human to trust a machine than to think for itself. Lastly, it is caused if a machine is doing the same task as the human. The user will hand over their responsibility to the machine or the user will reduce their effort because the machine will perform the task. (Grissinger, 2019)

## 2.2. Sense of Agency

The concept of sense of agency refers to the feeling of how much control someone has over their actions and consequences. Sense of agency or “experience of conscious will” (Wegner & Wheatley, 1999) occurs if a person intends to perform an action, after which the action happens, and this intention was the only possible explanation why the action happened (Moore, 2016). It is a main assumption of this thesis that increasing the sense of agency during the final decision making of a human user of a CDSS could reduce the effects of automation bias and automation complacency. The underlying idea is that the more a user is actively, consciously, involved in the

final decision making, the less chance there is for complacency to set in. This assumption should be investigated further, but that lies outside of the scope of the current thesis.

Important characteristics of sense of agency are that it is not objective and that it is difficult to measure. As stated by Moore (2016), sense of agency is not a reproduction of objective reality, because it refers to a person's view of their own feelings and intentions. Feelings and intentions can be distinctively different from reality, which makes the sense of agency possibly differ from reality as well. The difficulty to measure sense of agency lies in the subjectivity of the content that needs to be measured, namely one's feelings and intentions.

Even though sense of agency is difficult to measure, two approaches for measuring this concept have been developed, namely a direct approach and an indirect approach. When using a direct approach, statements, questionnaires or rating scales need to be filled in by a person for a specific event. These questions are about the extent to which the person feels as if they are the agent of their action or the effect of their action. (Tapal, Oren, Dar, & Eitam, 2017)

In the paper of Tapal, Oren, Dar and Eitam (2017), 13 statements are given to measure the sense of agency, which can be found in Figure 1. It is important that some of these statements question the same items, but then in opposite ways, meaning one question is formulated negatively and another question is formulated positively. The reason for this is that a bias can be created when only one formulation is used (Sauro, 2011). In addition to this, using the exact words of responsibility and sense of agency may influence the user, resulting in biased answers. Therefore, this direct approach implicitly asks the user about their sense of agency. (Tapal, Oren, Dar, & Eitam, 2017)

The indirect approach of measuring sense of agency includes the analysis of two concepts, namely sensory attenuation and intentional binding. Sensory attenuation entails that people tend to perceive less

intense sensory consequences when the action is their own in contrast to movements from others, as shown by Blakemore et al. (1998; 1999). If sensory attenuation is present, then there is a higher sense of agency. This is because the person feels as if they did the action, even if they are not consciously aware of doing the action.

1. I am in full control of what I do
2. I am just an instrument in the hands of somebody or something else
3. My actions just happen without my intention
4. I am the author of my actions
5. The consequences of my actions feel like they don't logically follow my actions
6. My movements are automatic—my body simply makes them
7. The outcomes of my actions generally surprise me
8. Things I do are subject only to my free will
9. The decision whether and when to act is within my hands
10. Nothing I do is actually voluntary
11. While I am in action, I feel like I am a remote controlled robot
12. My behavior is planned by me from the very beginning to the very end
13. I am completely responsible for everything that results from my actions

Figure 1. Statements for measuring the sense of agency

The intentional binding as a measurement examines the time perception of the person. As shown by Haggard, Clark and Kalogeras (2002), people tend to narrow the time interval between their voluntary action and their perceived effect. Relating this to sense of agency, when there is a higher intentional binding, then there is a higher sense of agency. This is because if the perceived time interval is smaller than the actual time interval, people unconsciously claim responsibility for that action and its effect.

### 2.3. Meaningful human control

According to van den Hoven & Santoni de Sio (2018) meaningful human control suggests that humans, not computers and their algorithms, should ultimately remain in control of, and thus morally responsible for, relevant decisions. People have to be held accountable for their actions and decisions, because computer(programs) cannot be held accountable for the, possibly incorrect, advice they give. The decision, which can be based on the advice of the computer(program), is ultimately made by the user.

With this definition, it is difficult to determine how meaningful human control can be designed and achieved. There are two conditions that have to be met in order for meaningful human control to exist. These conditions are the tracking and the tracing condition. In the first condition, tracking, a DSS, or any autonomous decision-making system, should respond to the human moral reasons relevant in the situation, limited to the rules in which the machine operates. In the second condition, tracing, decisions or operations that the autonomous system, in this case a DSS, should be leading back to at least one human. The reason for this is because only humans are able to have a proper moral understanding. The decision of the DSS, or any autonomous system, should be based on this moral understanding. In addition to this, the tracing condition requires human operators who are capable and motivated to actively prevent unwanted outcomes to occur in the first place. (van den Hoven & Santoni de Sio, 2018)

Meaningful human control is difficult to measure. However, the tracking and tracing conditions need to be met in order for meaningful human control to be present. In the current literature, there has not been given any approach to measure the tracking and the tracing conditions. This difficulty in measuring the two conditions is because a person cannot objectively say that they are in control, and thus responsible, because a person's feelings and emotions are in the way. A possible approach to determine who is responsible, and if meaningful human control is present, is an objective third party, such as a judge, who has to determine who can be held accountable.

It is a main assumption of this thesis that the sense of agency increases when the tracking and tracing conditions are met. The underlying idea is that the user is more actively and consciously involved in the decision making process when a support system adheres to the morals of the user and the decisions of the support systems can be traced back to the user. This assumption should be investigated further, but that lies outside of the scope of the current thesis.

## 2.4. Overreliance

A CDSS was first implemented to assist doctors to aid during the complex decision making process of diagnosing patients. With the help of a CDSS, a doctor should reach a higher number of correct diagnoses. This statement is supported by a study by Friedman et al. (1999), which shows that without the use of a CDSS subjects reached a 39.5% (764 cases) success rate of diagnosing patients. With the use of a CDSS a 45.4% (879 cases) level of correct diagnoses occurred, which is a 6% increase. However, this increase consisted of a 12% increase of positive cases, where an erroneous answer was altered to a correct diagnosis, and a 6% increase of negative cases, where a correct diagnosis was altered into an erroneous answer. This last statistic indicates an overreliance on the CDSS. Overreliance is a state in which there is an excessive trust or dependency on something or someone.

Overreliance can be found in both experts and non-experts (Parasuraman & Manzey, 2010). However, as stated by Goddard, Roudsari and Wyatt (2012), experts, who have more trust in their own capabilities, rely less on support systems in comparison with non-experts in the field of the expert. Nevertheless, when a user trusts in a support system, there is an increase in the reliance of the user on the system. This holds for both experts and non-experts. In addition to this, the reliance on the support system increases when experts are uncertain about their decision and if there is an extensive collaboration. (Goddard, Roudsari, & Wyatt, 2012)

Overreliance in itself is difficult to measure. Therefore, situations need to be analysed in which overreliance occurs. Automation bias and automation complacency are such situations. The overreliance in the situation of automation bias can be seen in the number of correct diagnoses changed to faulty ones, which is called negative consultation. Automation complacency results in less identified automation failures, which is a result of overreliance, because the users trust the accuracy of the support system. An indication for automation complacency is, thus, the number of unobserved automation failures. (Goddard, Roudsari, & Wyatt, 2011; Grissinger, 2019)

Automation bias and automation complacency are difficult to measure as well. However, research by Goddard, Roudsari and Wyatt (2012) has shown one approach by which this can be done. This approach is to compare the performance of doctors when they have to make decisions either with or without the aid of a CDSS.

## 2.5. Linking of concepts

In this part, the previously explained concepts will be linked to each other. Moreover, the relationships between the concepts will be defined. In order to link these concepts, the concept of responsibility will be used.

Responsibility means that people can be held accountable for what they do, which, in turn, allows people to be held accountable and be punished or rewarded accordingly (Moore, 2016). However, the possible risk of using a DSS is the shift in responsibility from the user to the DSS by the user when performing a task. This shift in responsibility has an effect on the sense of agency of the user.

When a user shifts the responsibility of the task to the CDSS, there is a decrease in their sense of agency. This is because they feel as if it is not solely their task and decision anymore, because the CDSS is part of the decision making process as well. This is an example of how sense of agency can be used to imply a division of responsibility to various actors of the decision making process (Moore, 2016). When there is an increase in automation, meaning that support systems are used more often, there is a reduction in the sense of agency (Berberian, Sarrazin, Le Blaye, & Haggard, 2012).

As has been mentioned before, humans should be in control, and therefore, responsible for the relevant decisions, not computers and their algorithms (van den Hoven & Santoni de Sio, 2018). Therefore, the division of responsibility by the sense of agency cannot be a legal division in responsibility. The user can only feel as if the responsibility is not entirely their own anymore, because the user mentally divides the responsibility between himself and other actors within the decision making process.

There is another way in which responsibility and control are linked. According to Fischer and Ravizza (1998), meaningful human control and responsibility are linked if a person possesses guidance control over the action they perform. Thus, a person is responsible for an action if there is guidance control.

Guidance control consists of two conditions. The first condition is that a DSS should be moderately reason-responsive. This concept is similar to the tracking condition, explained in Chapter 2.3. Moderately reason-responsiveness means that a DSS must be responsive to the human moral reasons relevant in the situation. It is important that the human is able to consciously make a decision, where they can recognize and bring themselves to (not) perform an action, even if the advice from the DSS might differ from their own choice of action. As a result, the human is accountable, and responsible, for their actions, and not the DSS. (van den Hoven & Santoni de Sio, 2018)

The second condition is that the decision should be the user's own. This is similar to the tracing condition, explained in Chapter 2.3. It means that the user has to have enough knowledge to have a proper moral understanding of the decision and the problem. When a user does not have enough knowledge to understand a problem, they give the advice of a DSS more weight, and, thus, rely more on the DSS. When the user does not have a proper moral understanding, the necessary condition for being morally responsible is not met. (van den Hoven & Santoni de Sio, 2018)

Health practitioners have legal and moral responsibility for the decisions they make. Overreliance on the CDSS may disrupt this balance. The sense of agency of the user will have a negative effect on overreliance on the CDSS; the higher the sense of agency, the lower the overreliance of the CDSS. When the user has a bigger sense of agency, they will feel more moral responsibility for the task. As a result, if they have enough confidence in their knowledge, they will depend on their own ability to do the task. Therefore, the user will not depend on the CDSS too much. If the user does not have enough knowledge to feel confident in making the decision, they will seek the help of the support systems.

Automation bias, which is a type of overreliance, can result in negative consultations and automation complacency can result in unnoticed failure of the CDSS. This results in a difficulty in assigning responsibility in the decision making process, because the doctor is the one who is responsible, because he is the one making the decision. However, when there is overreliance, the doctor gives too much weight to the advice of the CDSS. Therefore, the doctor does not make the decision anymore, because it is influenced too much by the advice of the CDSS.

## 2.6. Eight modules for implementing a FM

A Falsification Machine (FM) should help increase the sense of agency of the user. If the FM asks a question, a user should feel more sense of agency, because they cannot claim that they

only listened to the CDSS recommendation to make the decision. Therefore, the user gains more control in the decision making process.

A FM tries to help the user during or after the decision-making process by asking questions or pinpointing important factors. These questions will be based on the diagnosis of the user, output of the CDSS, symptoms of the patient and the dataset that the CDSS has used to produce the output. In the best scenario, an extensive dataset is used that contains most known diseases, their symptoms, their possible treatments and success rate. However, it is difficult to have an up-to-date database that contains accurate and good-quality data, as has been explained in chapter 1.1.

In his thesis, Niels (Cornelissen, 2021) explains eight modules as design requirements for implementing a FM. Niels only uses the first three modules for his implementation, because of the scope of his thesis, but here all modules will be briefly explained. These modules are designed for a FM that critiques the decision-making process, which uses a CDSS.

The first module is a design requirement that states that the FM should check if there are major differences between the diagnosis of the DSS and the user. This design requirement is called the correctness critic. The second module is a design requirement that states that the FM should check if there are major differences between the diagnosis of the user and the symptoms of the patient. This design requirement is called the consistency critic. The next module is the differential critic. This design requirement states that the FM should form its own diagnosis, and ask a question to guide the user to the diagnosis of the FM. The fourth module states that the FM should check for red flags regarding prescription combinations, which is another consistency critic. The fifth module is a design principle which states that the FM should provide which supplementary tests should be run to be able to improve the diagnosis. This is called a tool critic. The sixth module considers the use of preceding cases and their diagnoses as supplementary information in the decision making process. The seventh module is a design requirement for the FM, which states that the FM should list the risks, possible side effects and success rate of the suggested diagnosis of the user. The two modules are both experiential critics. The last module is an alternative critic, which states that the FM should recommend alternative diagnoses.

## 3. Methods

### 3.1. Design of the research

For this investigation, a quantitative approach is used, or more specifically a survey research. The quantitative approach helps with explaining social phenomena with facts and figures. McLeod (2019) states that the data which is collected should be measurable and analysed through numerical comparison. This means that the concepts, which are of interest to the research, should be operationalized to obtain the data and results for the research.

The quantitative approach helps to quantify different concepts and to compare results (McLeod, 2019). In order to reach an answer to the research question, the concept of sense of agency need to be quantified. For this research, it is essential to be able to compare the answers to the statements of sense of agency between the subjects, as there has to be a comparison between medical and non-medical students.

In regard to the COVID-19 measures, under the current conditions the optimal way to increase the generalizability is conducting an online questionnaire. The reason for this is because more people can be reached, resulting in a larger sample population. (McLeod, 2019)

### 3.2. Operationalization

Overreliance on a support system is major problem in the decision making process (Friedman, et al., 1999). This problem will be analysed by looking at three dimensions, namely sense of agency, (over)reliance and meaningful human control. The central dimension in this analysis will be sense of agency, as the research question is roughly stated as: “What is the influence of a FM on the sense of agency of medical students in the medical field in comparison with non-medical students.”

The first dimension, sense of agency, will be measured with two aspects of the participant’s experience, namely control and conscious awareness of their actions (Tapal, Oren, Dar, & Eitam, 2017). Control is a part of the experience that considers if the participant thinks they are in control. Conscious awareness indicates if the participant is consciously aware of themselves and their surroundings.

The second dimension that will be measured is the interaction with the environment. Interaction with the environment considers the interaction with the different actors of the decision making

process and how these actors influence the decision of the participant. This will measure the reliance on the support systems during the decision making process.

The third dimension is meaningful human control. This can be measured by the tracking and the tracing condition. These concepts have been explained in the theoretical approach.

<b>Concept</b>	<b>Dimension</b>	<b>Item</b>
Overreliance	Sense of agency	Control
		Conscious awareness of actions
	Reliance	Interaction with the environment
	Meaningful human control	Tracking
		Tracing

Table 1. Operationalization

### 3.3. Data collection method

For this research, the main source for the data collection is a questionnaire. Multiple pilots have been conducted before the final version was sent to the participants. The first pilot had the aim of determining the estimated time participants would take to finish the cases. The questionnaire should not be too long, because when a questionnaire is too long, there is a reduction in people finishing the questionnaire. Therefore, a questionnaire with four cases has been chosen. The goal of the second pilot was to determine the language in which the questionnaire should be conducted, either English or Dutch. It has been found that more people can be reached when the questionnaire is in Dutch. In addition to this, medical and law terms are more easily understood by the participants when they are in their mother language. Therefore, the language in which the questionnaire is held is Dutch. The final pilot was conducted to determine how the statements should be formulated. The statements at the end of each case should be clear, otherwise the statements could be misinterpreted, and the data gathered could be inaccurate.

The four cases consist of a medical case with a Decision Support System (DSS) and a Falsification Machine (FM) question; a medical case with a DSS, but without a FM; a law case with a with a DSS and a FM question and lastly a law case with a DSS, but without a FM question (see table 2 and 3). All four cases have a different subject matter, meaning that the information given in one case is only applicable to that one case, not in all other cases. Two populations of students will participate in the data collection, one group of medical students, and another group

of non-medical students. In the tables below the experimental design is shown. It should be noted that case 1 and 5 are the exact same, the same holds for case 2 and 6, 3 and 7, and 4 and 8.

	DSSyes FMyes	DSSyes FMno
Medical	1	2
Law	3	4

Table 2. Medical students

	DSSyes FMyes	DSSyes FMno
Medical	5	6
Law	7	8

Table 3. Non-medical students

The medical cases are used, because the medical students are experts in this field. The law cases will be used as a base case, because every participant has roughly the same amount of knowledge in this field. If a law student makes the questionnaire the answers of the participant will not be used as data for the results, because law students are experts in the base case, which would lead to biased results.

To show if the hypothesis holds, on average the gain in sense of agency of the medical students making a decision with a FM has to be higher than non-medical students using a FM in the decision making process. Relating this to the table 2 and table 3, the difference in sense of agency between 1 and 2 needs to be in favour of 1. This difference needs to be compared to the difference in sense of agency for the non-medical students. This will be done by measuring the difference in sense of agency between 5 and 6, which should be in favour of 5. If the difference is in favour of medical students, then this means that the FM increases the sense of agency of the medical students more than that the FM increases the sense of agency for non-medical students. The following formula can be made for this hypothesis:

$$(1 - 2) > (5 - 6)$$

### 3.4. Construction of the cases

As stated before, the questionnaire consists of four cases. Two medical cases and two law cases. All cases start with a description of the case that explains the situation and states a problem. The

case descriptions are followed by advice from a DSS. This DSS will give a recommendation to the user about the most probable answer, given the situation.

After the DSS's recommendation, four possible answers are stated. If there are difficult terms or concepts within the answers or case description that a non-expert is unlikely to know, there are information buttons added, which the participant can click to get some extra information. Otherwise, non-experts are unable to answer the question, due to lack of knowledge to give a well substantiated answer.

To construct the description of the situation for the medical cases, a table was used created by Van der Stigchel (2021). This table, which can be found in appendix 8.1, contains important categories of information about the patient that should be in the description. The two cases are built around one main symptom, which can be caused by four different illnesses (Huisarts DD, n.d.). For the first medical case, this symptom is "involuntary weight loss" and the main symptom of the second medical case is "coughing". The illnesses related to the symptom of involuntary weight loss are colon cancer, depression, hyperthyroid, diabetes mellitus type 2. The illnesses related to the symptom of coughing are lung cancer, upper respiratory tract infection, side effect of lisinopril and gastroesophageal reflux disease. These illnesses will be the four possible answers to their corresponding cases. Then, the cases are expanded with information from the different categories from the table. At the end, all the information is combined in a case presentation (Budgell, 2008). Non-medical students cannot be expected to know all illnesses and their corresponding symptoms. Therefore, they can use the information buttons for extra information if necessary.

To create the two medical case presentations and their answers, a collaboration with a fifth-year medical student was necessary (see appendix 2). This collaboration was required to ensure that the cases were complex enough to be challenging for medical students and that the cases contain enough valuable information to give an answer to the question.

The recommendation by the DSS for the medical cases is made at random beforehand, meaning that the recommendation by the DSS does not differ per participant. In addition to this, there is no correct answer. All answers are possible, because symptoms connected with every answer are present in the case presentation. The symptoms every answer is connected to can be seen in case 1 and case 2 in Appendix 8.3.1 and 8.3.2.

To construct the law cases two real cases were used. Using the entire real case would not be possible, because the cases are too extensive to use as a case presentation in a questionnaire. Thus, only a simplified version of the real case is used. The first law case of the questionnaire is based on the Bijlmer case (Irfan, 2016). This case involves a woman who was being robbed by two men. She used a (illegal) gun to defend herself, because of a traumatic experience in the past. She fatally wounded one man and injured the other man. The woman is now being prosecuted. The participant in the questionnaire is asked to give a sentence to the woman.

For the second case, a simplified version of case number 09/900065-05 is used (Het Rechtenstudentje, 2005). This case concerns a group of men, who were causing trouble by putting branches and street signs in the middle of the road. Afterwards, they threw a large piece of brick off a bridge on to the highway. At the same time, a woman was driving in her car on the highway and she was fatally hit by the piece of brick. Two of the four men remained in the car at all times, while the other two were causing all the trouble. All four of them are now being prosecuted. The participant's task is to sentence the men.

To create the law cases for the questionnaire the case presentation with information about the situation (syllabus) is used. Afterwards, a demand from the court for the suspect is set forth. A DSS recommendation states which sentence, given the situation, would be the best solution to the problem. Finally, the participant has to answer what they think the best sanction would be for the suspect(s) if they were the judge in these cases. The part of the opinion, which is the decision of the judge (University of Maine School of Law, 2019), is not given to the participant, but it is used to formulate one of the answers for the case. Additionally, there are information buttons to explain the complex concepts used in the law cases.

To adapt the original law cases into a form suitable for the use in the questionnaire, there was a collaboration with a third-year law student. (see appendix 2). This collaboration was required to ensure that the cases used the appropriate law terminology and were complex enough to simulate a real law case. When sections of the real law cases were excluded to create the law cases used in the questionnaire, the remaining case still had to resemble a law case to give the participant the feeling they had to solve a real case. The law cases could not be too complex or contain too much law jargon, because the participants are not familiar with these concepts.

The recommendation by the DSS for the law cases is made at random beforehand, meaning that the recommendation does not differ per participant. After the recommendation by the DSS, the participant can choose their answer. One of these answers is based on the conviction made by the

judge in the real case. The other answers were made in collaboration with the law student to guarantee realistic alternative answers.

The first medical case and the first law case have a FM question, the other two do not. This structure has been chosen, because with this method the effect of the FM can be measured in both medical and law cases for the medical and non-medical students. With this information, the research question can be answered, because the results can be compared. All cases have multiple statements to measure the sense of agency. The statements are adjusted to the differences in the cases, such as use of FM or law or medical case. How these statements were created will be explained in the next part of the method, part 3.5. In part 3.6., the structure of the questionnaire will be given in the form of a flowchart.

### 3.5. Construction of FM questions and statements

In two cases, the FM asks the participant a question. The FM question is asked after the participants fill in their answer. After the FM question is asked, the participant is given the opportunity to switch their answer. The questions which the FM can ask are premade, because the FM was not implemented. Every answer of the case with the FM has one premade FM question. The questions are created with the use of the eight modules explained in the theoretical approach.

For the medical cases, the FM question will ask about a symptom in the case presentation, which is not a symptom of the illness the participant has given. The FM will try to increase the reflection of the users about the quality of their answers. This is similar to the third module that Niels Cornelissen proposes, namely the FM as a differential critic. As stated before, with this module the FM tries to provide an alternative solution to the answer given by the user.

The FM questions used in the law cases use the same principle as the FM question of the medical cases. However, the FM question in the law cases tries to guide the participant to another answer by counter thinking the belief in their current answer. The FM assumes that all other answers are better than the participant's current answer.

At the end of the case, there are statements that measure sense of agency and reliance on external systems. These statements are based on the article by Tapal, Oren, Dar, and Eitam (2017) and use the direct approach of measuring sense of agency. The statements have to be answered on a Likert scale from 1 (totally don't agree) to 7 (totally agree) and measure the different aspects of

sense of agency of the participant (Lauermann & Karabenick, 2013). The validity increases because of the use of five statements, because it covers most of the concept of sense of agency and overreliance.

There are four or five statements depending on the use of a FM in the case. When the FM is used in the medical case the statements are “I made my diagnosis completely independently“, “The diagnosis was a conscious decision”, “The diagnosis support system has influenced my answer”, “The question from the falsification machine has influenced my answer”, “I am responsible for what my actions might bring about”. If there is not a FM question the fourth statement is not used. The law cases with a FM question are “I made my decision completely independent“, “The decision was a conscious decision”, “The decision support system has influenced my answer”, “The question from the falsification machine has influenced my answer”, “I am responsible for what my actions might bring about”. If there is not a FM question the fourth statement is not used. An overview of the structure of the statements can be found in flowchart 2 (Chapter 3.6).

The statements can be linked to the items made in the operationalization. To measure the item of control, the statement “I made my diagnosis completely independently” is made. The statements “The diagnosis was a conscious decision” and “I am responsible for what my actions might bring about” measure the item of conscious awareness. The statements “The diagnosis/decision support system has influenced my answer” and “The question from the falsification machine has influenced my answer” measure the interaction with the environment.

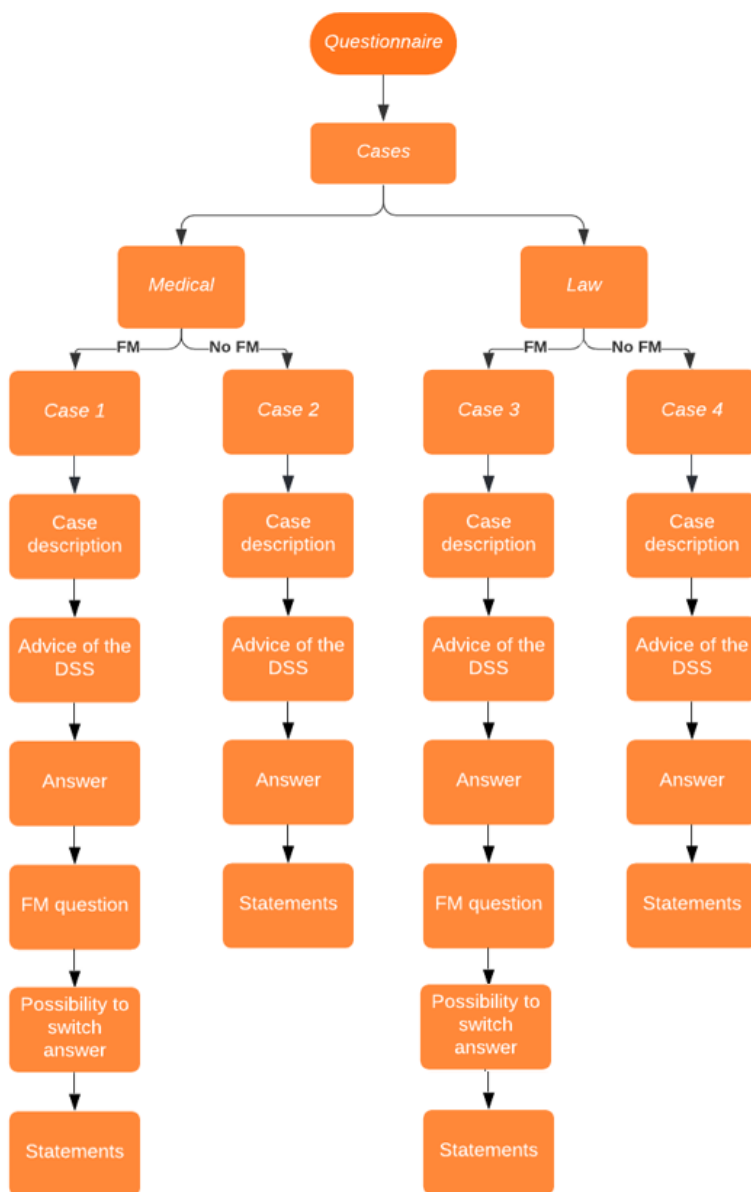
In the operationalization, it is stated that meaningful human control can be measured by two items, namely with the tracking and tracing conditions. However, to measure if there is a difference in the sense of agency for the participant, the meaningful human control dimension needs to be a fixed state. In this case, the conditions of tracking and tracing need to be met in order for meaningful human control to be present. The tracking condition is met because the participant is able to consciously make a decision, and they are able to change their answer, thus perform a different action than their first action. The tracing condition states that the user should have enough knowledge to have a proper moral understanding of the problem. This condition is met in the medical cases because, for the medical students, it is expected they have enough knowledge of medicine to answer the question, and for the non-medical students, and medical students as well, there are information buttons to give them enough information to be able to answer the question while understanding the possible answers. Because it was expected that no participant has sufficient knowledge to answer the questions relating to the law cases, information

buttons were added to give them enough information to have a proper moral understanding of the terms used in the case and its answers.

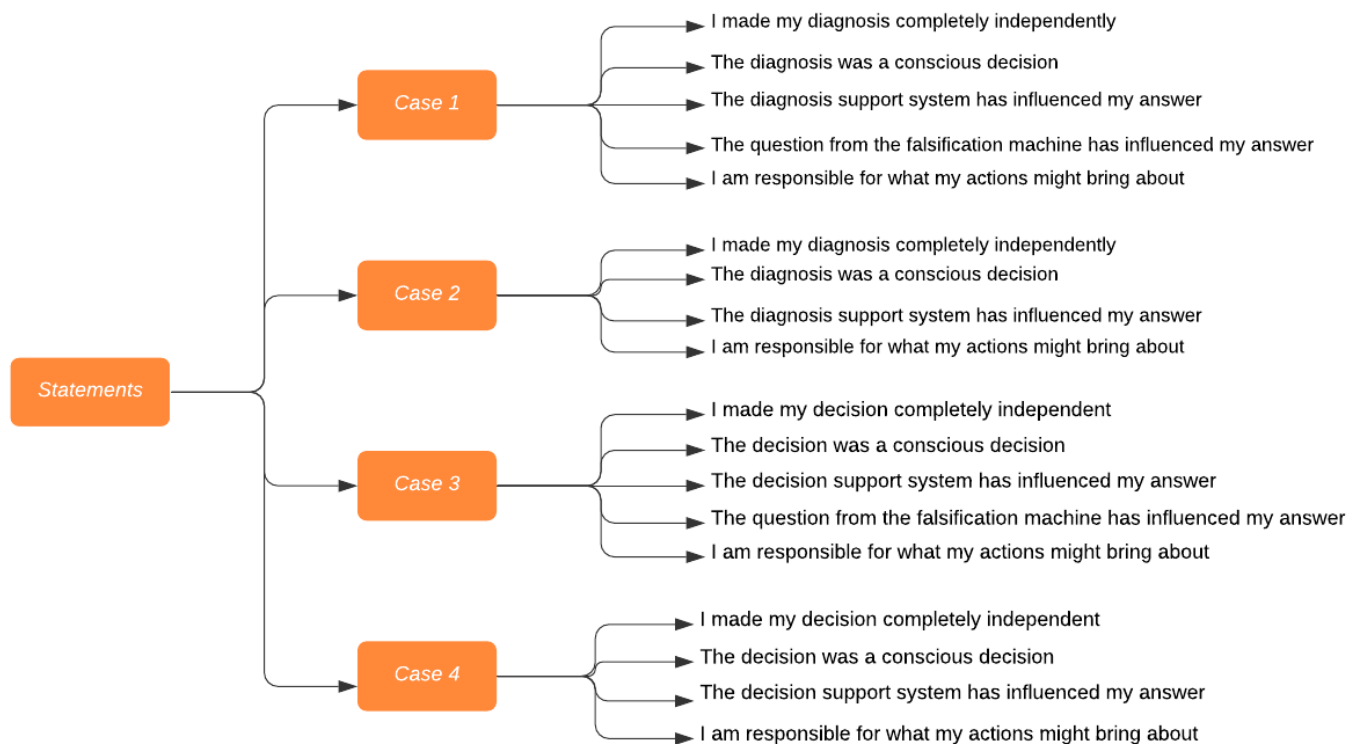
### 3.6. Structure of the questionnaire

To give a better overview of the structure of the questionnaire, two flowcharts have been made. The first flowchart shows the general structure of the questionnaire. In the second flowchart, the cases with their respective statements can be found.

Case 1 in the flowcharts below is equal to 1 and 5 from table 2 and 3. Case 2 in the flowcharts below is equal to 2 and 6 from table 2 and 3. Case 3 in the flowcharts below is equal to 3 and 7 from table 2 and 3. Case 4 in the flowcharts below is equal to 4 and 8 from table 2 and 3.



Flowchart 1. Structure of questionnaire



Flowchart 2. Structure cases with statements

### 3.7. Data analysis

When analysing the data, each possible answer of the statements will be given a number in order to compare them. The answer of “I completely don’t agree” will be given a value of 1 and the answer of “I completely agree” will be given a value of 7. For each statement, a separate average will be calculated for the answers given the medical students and the non-medical students. In addition to this, another attribute will be collected, which identifies if a participant has changed their answer after the FM question in case 1 and case 3.

The scores from the statements will result in a score on the items made in the operationalization. These will be used to answer the research question. To analyse the sense of agency the item control is measured with the statement “I made my diagnosis completely independently” and conscious awareness of actions is measured with the statements “The diagnosis was a conscious decision” and “I am responsible for what my actions might bring about”. To measure if there is reliance on support systems, interaction with the environment is analysed with the statements “The diagnosis support system has influenced my answer” and “The question from the falsification machine has influenced my answer”. If there is no statement about the FM (case 2

and 4), interaction with the environment is analysed with only the statement “The diagnosis support system has influenced my answer”.

A Likert scale will be used for the answers for the statements. When using a Likert scale, the subjectively experienced difference between the options need be equal to that between other options. This means that the difference between, for instance, option one and two might be experienced as bigger than the difference in option four and five. In this thesis, however, it is assumed that the differences between the options are all equal. With this assumption, a numerical analysis of the answers for each statement for the different cases can be made. Further investigation is required to establish whether or to what extent this assumption is justified, but that lies outside of the scope of this thesis.

To quantify the concept of sense of agency (SoA), the items of control and conscious awareness first have to be calculated and put into a formula. This will be done for the two groups, medical and non-medical, separately. This formula is the following:

$$Value\ SoA = \frac{C + CA}{2}$$

The item of control (C) is calculated by the mean score for the statement “I made my diagnosis completely independently” answered by one group. The item conscious awareness (CA) is calculated by adding the mean scores of the two statements of “The diagnosis was a conscious decision” and “I am responsible for what my actions might bring about” together. After adding the two statements of CA together, this score will be divided by two. This will be done to get separate scores for C and CA with an equal weight. Together, they will form the SoA. The scores for C and CA are then added together and divided by two to get the value of the SoA. This will be done for both groups separately.

The interaction of the environment will consist of two separate attributes, namely the statement of “The diagnosis support system has influenced my answer” and the statement “The question from the falsification machine has influenced my answer”. There will not be one score for the interaction with the environment, because it is important to identify the different influences of the DSS and the FM.

The scores of the interaction with the environment will be used to analyse the value of SoA. When there is a high mean score on the statement of “The diagnosis support system has influenced my answer”, then there is a higher possibility of overreliance, which reduces the sense

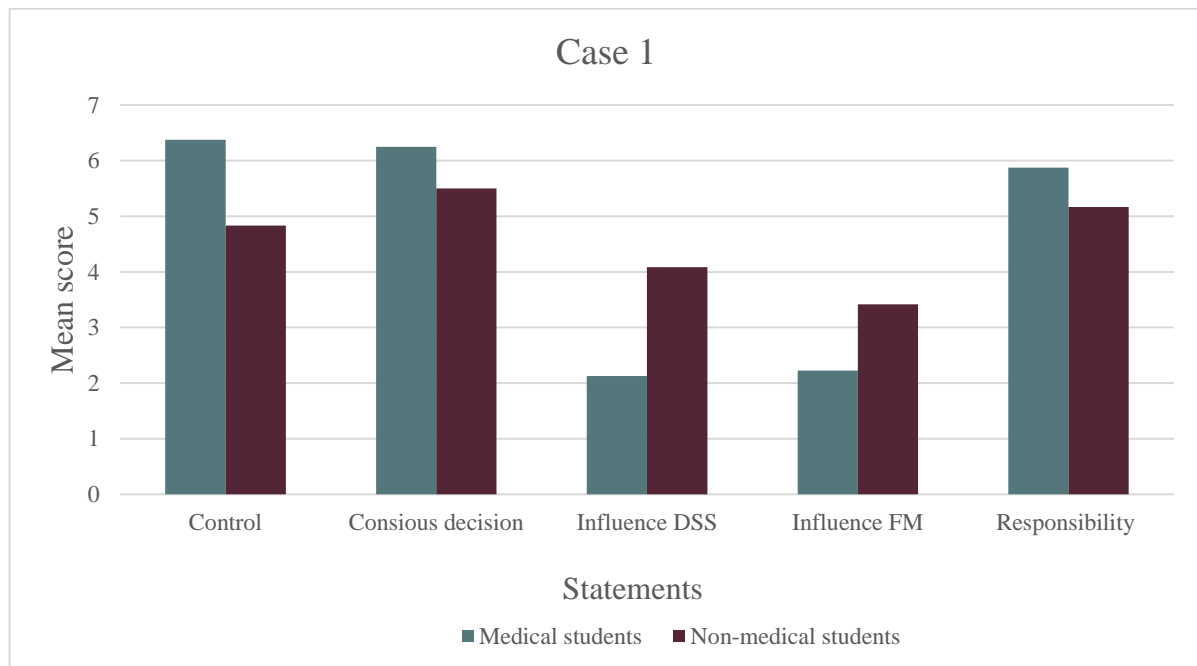
of agency of the user (Friedman, et al., 1999). Therefore, a low mean score on this statement means that there is a lower possibility of overreliance, resulting in a higher sense of agency. A high mean score on the statement of “The question from the falsification machine has influenced my answer” means that the possibility of overreliance, on a DSS or the user’s own knowledge, is reduced. This is because the user has to restart the decision making process due to the question of the FM. As a consequence, there is a higher probability that the user changes their answer, because the user takes into account the information that the FM provides. Therefore, it is important to identify if the participants used the option to change their answer after the FM asks its question. To conclude, the FM should have a positive influence on the SoA, meaning that a high score on the statement of “The question from the falsification machine has influenced my answer” results in a higher score on SoA. It is important that these two statements will be analysed separately, because the influence of the FM in the decision making process on sense of agency, which is reduced by overreliance on a DSS, has to be identified in order to answer the research question.

To conclude, the output of the cases will consist of the statements for each case and if a participant changed answers. The output of the statements will consist of the SoA, which is calculated with C and CA, and the interaction with environment, which is the interaction with the DSS and the FM.

## 4. Results

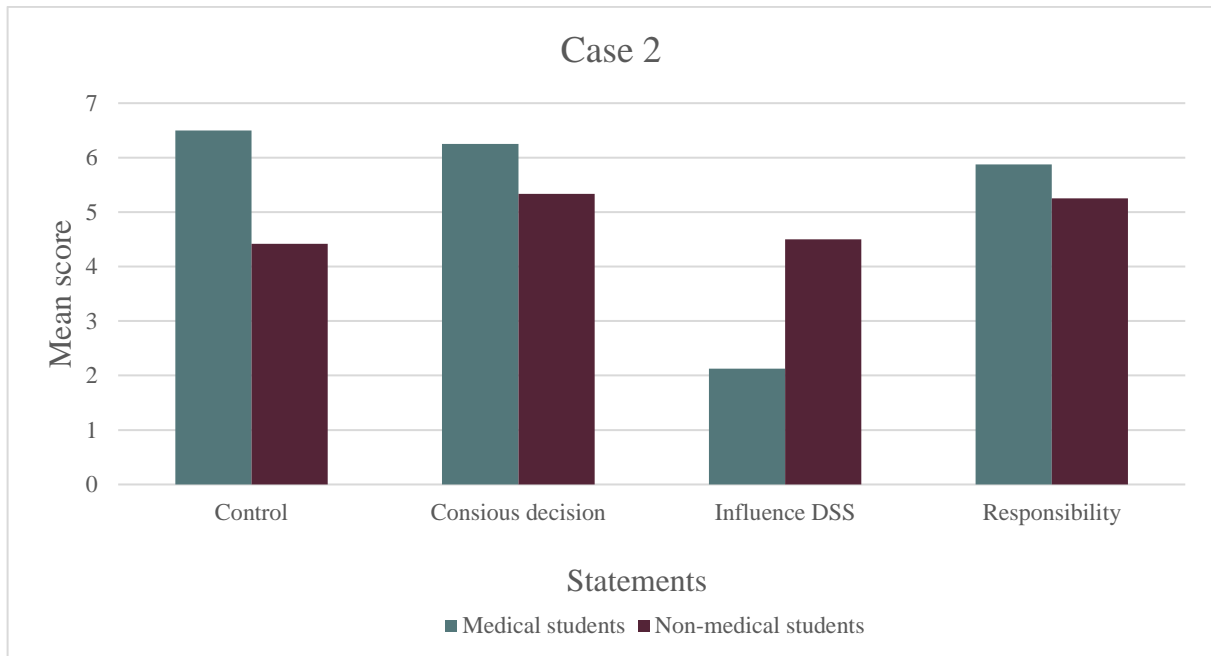
There have been 42 responses collected, however, only twenty participants completed the questionnaire. For the 22 people who did not complete the questionnaire, six responses were registered as started, but the participants did not enter any input, which means that there is not any information about these responses. The other sixteen responders stopped after the first case, about half before answering the statements and the other half after answering the statements. This could be explained by the participants expectation about the length of the questionnaire or the difficulty of the case. For the results, the responses of the twenty participants, who finished the questionnaire, will be used. Of these twenty responses, eight were filled in by medical students and twelve by non-medical students. Due to the low number of responses, the results cannot be regarded as significant, because a minimal sample size of a large population is around 100 (Alshibly, 2018). This means that there is a minimum of 200 respondents in order for the results of this research to be significant, because a minimum of 100 respondents need to represent the medical students and another 100 respondents need to represent the non-medical group.

In the following graphs, the mean scores of the answers from each statement per case is shown for both the medical and the non-medical students. The exact numbers and the range of the answers can be found in Appendix 8.4. After the graphs, the results will be analysed in more detail.



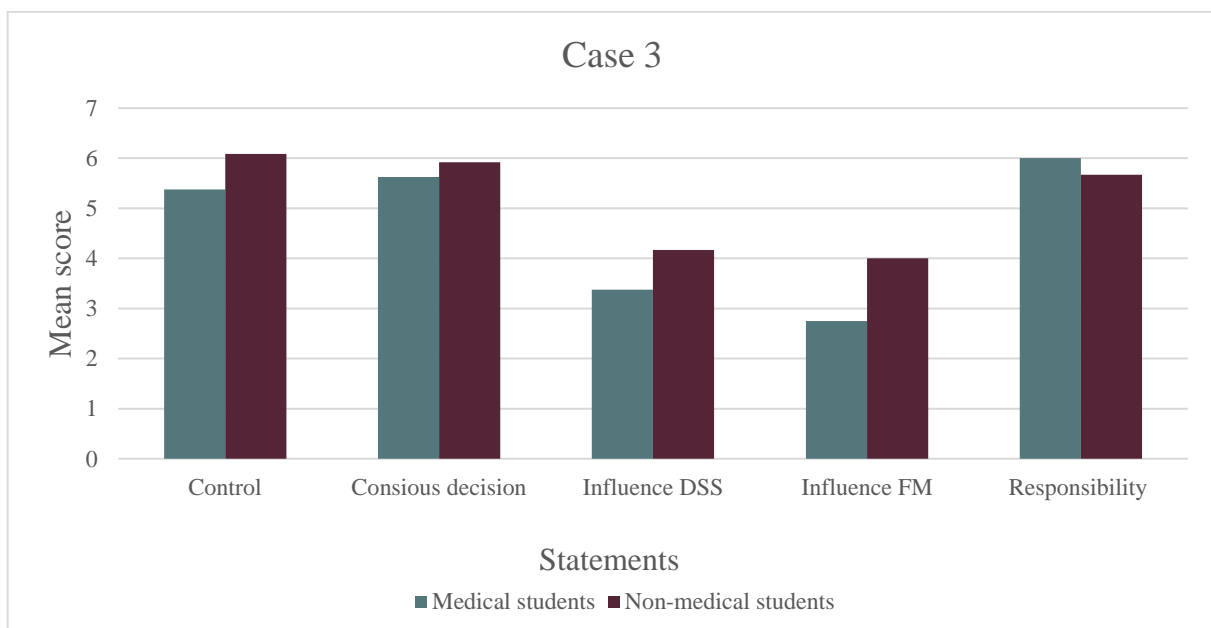
Graph 1. Results case 1

For case 1, the score for the item control (C) for medical students is 6.375. The score for the item conscious awareness (CA) for medical students is 6.025. The score of CA is calculated by adding the mean scores of 6.25 and 5.875 together and then dividing by two. The score of C for non-medical students is 4.8333, and the score of CA for non-medical students is 5.333.



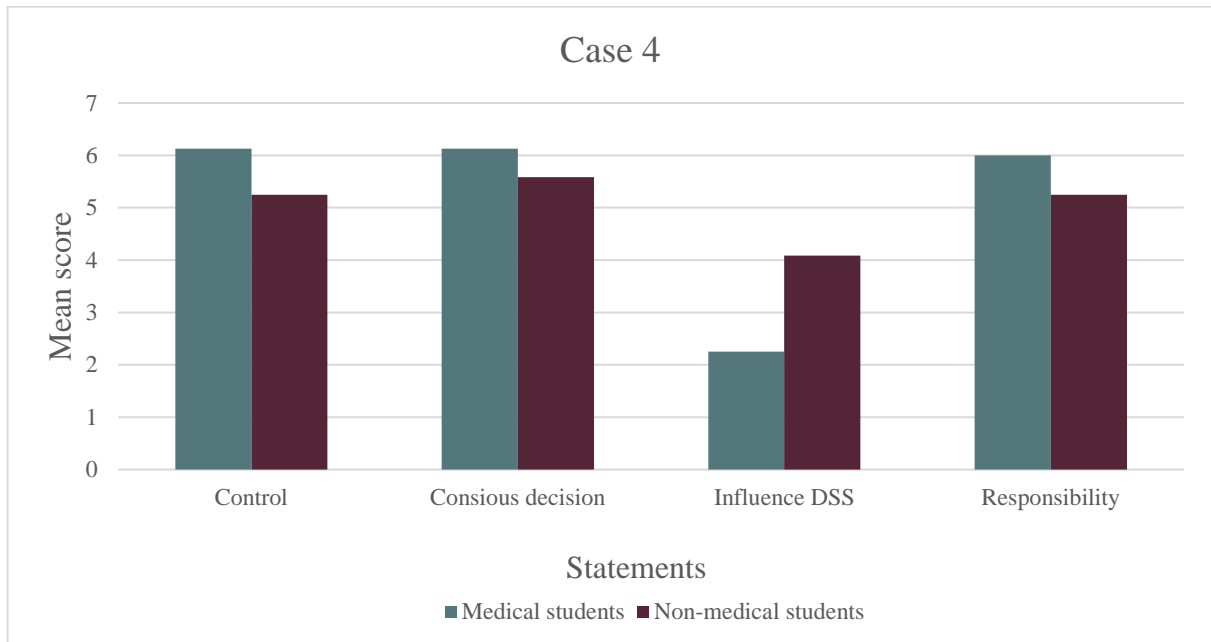
Graph 2. Results case 2

For case 2, the score of C for medical students is 6.5. The score of CA for medical students is 6.025. The score of C for non-medical students is 4.417, and the score of CA for non-medical students is 5.291.



Graph 3. Results case 3

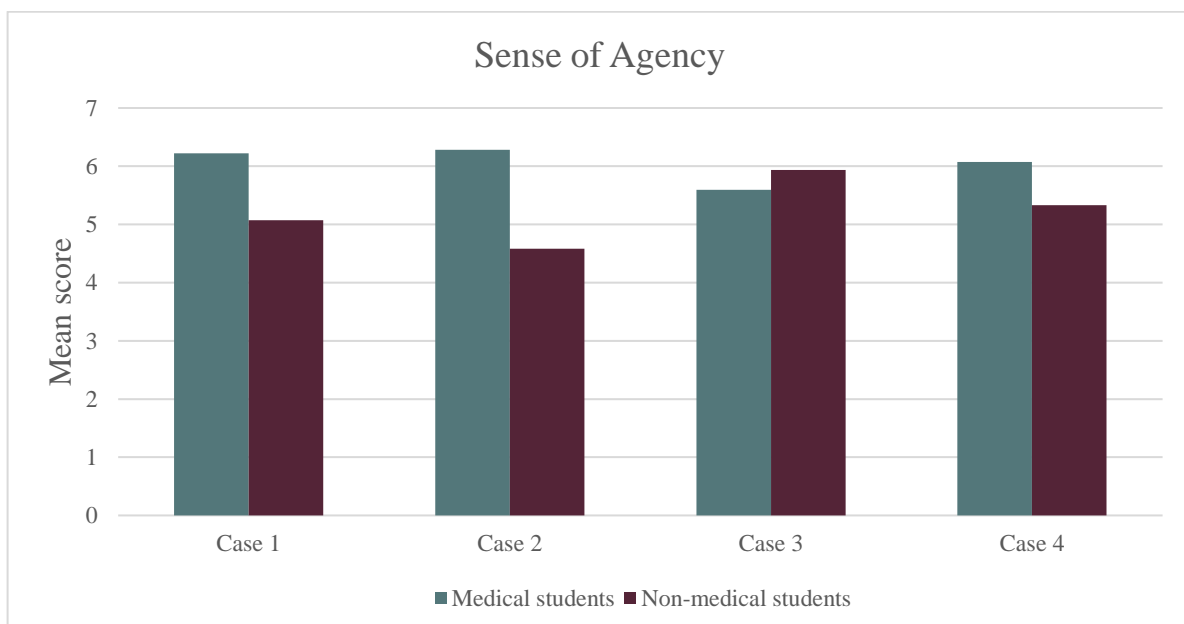
For case 3, the score of C for medical students is 5.375. The score of CA for medical students is 5.8125. The score of C for non-medical students is 6.083, and the CA for non-medical students is 5.792.



Graph 4. Results case 4

For case 4, the score of C for medical students is 6.125. The score of CA for medical students is 6.025. The score of C for non-medical students is 5.25, and the score of CA for non-medical students is 5.4165.

To calculate the Sense of Agency (SoA), these scores are added together and divided by two as explained in the formula in chapter 3.6 Data analysis. The results can be found in the table below.



Graph 5. Results Sense of Agency

## 4.1. Medical cases

The difference in SoA for the medical students between the two medical cases is 0.062. The SoA is higher in case 2, where no FM is used. The mean score for the statement “The diagnosis support system has influenced my answer” is the same in case 1 and case 2 for the medical students, namely 2.125. In case 1 the mean score is 2.225 for the statement “The question from the falsification machine has influenced my answer”. This can be seen in graph 5 and Appendix 8.4.5. There was no changing of answers by the medical students in both case 1 and case 2.

The difference in SoA for the non-medical students between the two medical cases is 0.488. The SoA is higher in case 1, where a FM is used. The mean score for the statement “The diagnosis support system has influenced my answer” is in the first case 4.083 and in the second case 4.5. In the first case, the mean score for the statement “The question from the falsification machine has influenced my answer” is 3.416. This can be seen in graph 5 and Appendix 8.4.5. Two of the non-medical students used the option to change their answer after the FM asked its question.

## 4.2. Law cases

The difference in SoA for the medical students between the two law cases is 0.482. The SoA is higher in case 4, where no FM is used. The mean score for the statement “The diagnosis support system has influenced my answer” is in the third case 3.375 and in the fourth case 2.25. In case 3 the mean score is 2.75 for the statement “The question from the falsification machine has influenced my answer”. This can be seen in graph 5 and Appendix 8.4.5. There was no changing of answers by the medical students in both case 3 and case 4.

The difference in SoA for the non-medical students between the two law cases is 0.604. The SoA is higher in case 3, where a FM is used. The mean score for the statement “The diagnosis support system has influenced my answer” is in the third case 4.167 and in the fourth case 4.083. In the third case, the mean score for the statement “The question from the falsification machine has influenced my answer” is 4.0. This can be seen in graph 5 and Appendix 8.4.5. One of the non-medical students used the option to change their answer after the FM asked its question.

## 5. Conclusion

To answer the research question, multiple sub-conclusions will be drawn from the results. These sub-conclusions will be combined to make a conclusion regarding the hypothesis and the research question. It should be emphasized that the number of subjects is too low to treat the results as being significant or indicative. Strictly speaking, nothing can be concluded from these results due to low  $n$ . However, the results can be taken as demonstrations of the viability of the experimental setup, in the sense that the data can, in principle, be used to reflect upon the research question. Hence, in the remainder of the conclusion this basis will be proceeded on. For the sake of clarity and brevity, ‘conclusions’ and ‘results’ will be spoken of, even though these are, at best, merely indicative of potential trends in the data.

The sub-conclusions are that medical students have a higher sense of agency (SoA) when not using a FM, non-medical students have a higher SoA when using a FM, medical students always have a higher SoA in medical cases than in law cases, and non-medical students always have a higher SoA in the law cases than in the medical cases. The main conclusion is that the influence of the FM on the SoA of the medical students is less than the influence of the FM on the SoA of the non-medical students in the medical field.

The first sub-conclusion that can be drawn is that medical students have a higher SoA with cases that do not use a FM. The difference in the SoA of the medical students for the two medical cases is 0.064 in favour of case 2, where there was not a FM used. This holds for the law cases as well. The difference in SoA of the medical students for the two law cases is 0.482 in favour of case 4. Therefore, the medical students have a higher SoA when they make a decision without a FM.

The second sub-conclusion that can be drawn is that non-medical students have a higher SoA with cases that use a FM. The difference in SoA of the non-medical students for the two medical cases is 0.488 in favour of case 1, where the FM is used. This holds for the law cases as well. The difference in SoA of the non-medical students for the two law cases is 0.604 in favour of case 3. Therefore, the non-medical students have a higher SoA when they make a decision with a FM.

As can be seen in Graph 5, the medical students have more SoA in the medical cases than in the law cases. A possible explanation for this is that using a DSS or FM in their field of knowledge would increase the SoA instead of a field that the medical students do not have expertise in. However, as shown by the mean score on the statement “The diagnosis support system has influenced my answer”, there is no higher score on the influence of the DSS in the decision

making process in the medical cases than in the law cases. On the contrary, the highest mean score on this statement can be found in case 3. However, these mean scores are still below neutral, which is the score of 4 (neither agree nor disagree). These numbers can be found in the corresponding graphs in Chapter 4. Results. This low score on the statement of “The diagnosis support system has influenced my answer” means that the DSS does not have, or is not experienced as having, an influence on the medical students in the decision making process. The same holds for the use of a FM during this process. This can be seen by two aspects. The first aspect is the first conclusion that was drawn, namely that medical students always have a higher SoA when no FM is used. The second aspect is that the mean score on the statement of “The falsification machine has influenced my answer” is 2.225 in the medical case and 2.75 in the law case. These scores are below the neutral value of 4, which indicates that the medical students are not influenced, or do not experience being influenced, by the FM during the decision making process. Because all values related to the interaction with the environment are low, the score on the item interaction with the environment is low as well. To conclude, medical students are not affected by either a DSS or a FM.

Another sub-conclusion that can be drawn from the results is that non-medical students have a higher SoA in the law cases than in the medical cases. A possible explanation for this result is that non-medical students are less certain about their decision in the medical cases than in the law cases. This uncertainty would decrease the mean scores on the statements of “I made my diagnosis completely independently”, “The diagnosis was a conscious decision” and “I am responsible for what my actions might bring about” in the medical cases. In addition to this, the scores on the statements of “The diagnosis support system has influenced my answer” and “The falsification machine has influenced my answer” should, therefore, be higher, as the non-medical students would rely more on the DSS and the FM when they are less certain about their decision in medical cases than in the law cases. However, the results show that the mean scores on the statement “The diagnosis support system has influenced my answer” in the medical cases are similar to the mean scores on this statement in the law cases. These mean scores are all close to 4, which is the neutral value. In addition, the mean score on the statement of FM is higher in the law case than in the medical case. For the law case, this value is exactly 4, and for the medical case this is 3.416. These numbers can be found in the corresponding graphs in Chapter 4. Results. It should be noted that the SoA is higher in the law case. The values for the item interaction with the environment for the non-medical students are slightly higher than the values for the medical students. This indicates a higher interaction with the environment in comparison with the medical

students. This could indicate that the FM increases the SoA for the non-medical student. However, more research is needed to provide support for this statement.

When combining these conclusions, an answer to the research question can be made. The research question asks what the influence is of a FM on the sense of agency for medical students when making a decision in the medical field in comparison with non-medical students making a decision in the medical field with a falsification machine. As has been explained, the medical students are not influenced by the FM in the decision making process, either in the law or in the medical cases. However, the non-medical students are influenced by the FM in the decision making process, both in the medical and in the law cases. Therefore, it can be concluded that the influence of FM on the SoA of the medical students is less than the influence of the FM on the SoA of the non-medical students in the medical field. This results in the hypothesis not being supported, because the data from the medical students do not show an increase of SoA when using a FM.

## 6. Discussion

There are multiple aspects of this research that could be improved and there are various future research topics in the field of decision support systems (DSS) and falsification machines (FM).

The main problem that occurred during this research is lack of literature regarding the falsification machine. The concept of FM is new and therefore, I could not rely on other papers in regard to the effects of FM use in the medical field for example. Furthermore, the results of this research could not be compared with previous results from other papers, which makes it hard to be sure if the results are reliable. Another issue regarding the DSS are the concepts linked to the use of a DSS. These concepts, such as reliance, sense of agency and meaningful human control, are hard to define and measure. In addition to this, the literature there was no consensus about how to measure or how to define certain topics. There are some papers about these concepts individually, but these concepts are never all linked together. Therefore, their effect on each other is never investigated nor identified. This means that this study was fully explorative in nature, and that, given the low number of participants as well, results are only suggestive of possible further directions of exploration.

The cases used in this research were created with the help of a medical and law student and were tested by doing several pilots. However, it was difficult to measure if the cases were not too easy, complex or vague. If the cases were not up to par, this would influence the results. This is because if the cases were too easy, participants would not need a DSS or FM to make their decision. This would increase the difficulty of measuring the influence of the support systems. Moreover, the DSS and FM questions were made by me, meaning that there was no real implementation of the DSS and the FM. The DSS could be implemented and tell the participant why a certain answer would be recommended, because this is how a real DSS works (reference to Gaston). The FM questions were made according to the modules of the paper from Niels Cornelissen. However, only two modules were used. If a more complete FM is used, with more modules implemented, a different outcome might be measured.

The medical students and medical cases were used as experts performing diagnosis in their field of knowledge. However, differences in knowledge levels between medical students were not accounted for. This holds for the non-medical students and law cases as well. The level of knowledge of each participant is different. This is countered by the information buttons, but participants with a lesser amount of knowledge could have been more uncertain about their

decision, leading to a higher possibility of overreliance on the DSS and the FM. As a consequence, the results can be influenced by this.

As there were only twenty questionnaires fully completed, the results cannot be generalized. A reason for the limited amount of completed questionnaires could be the length of the questionnaire. To get the results to answer the research question, four cases had to be included in the questionnaire, otherwise no comparison could be made between medical and law cases. There had to be two cases for each field, because in one case, the FM needed to be included and in the other one it did not. Therefore, the minimum number of cases was 4. Nevertheless, the time needed to fill in the questionnaire could not be too long, as this would reduce the number of completed questionnaires. Therefore, only five statements to measure the reliance and the sense of agency were chosen.

The results showed that the medical students did not interact with their environment, namely the DSS and the FM. Therefore, the high sense of agency that the medical students had could not be attributed to the use of either a DSS or FM. This resulted in not being able to compare the high sense of agency by the medical students with the sense of agency by the non-medical students, resulting in the research question not being fully answered.

## 6.1. Further research

As there is currently no literature and information about the FM and its potential, more research could be done, not only in the medical field, but all fields where a DSS could be implemented. Furthermore, research can be done with a fully implemented DSS and FM, which would increase the accuracy of the results of the research. There was no link found between the medical students and increasing the sense of agency because of a FM. Nonetheless, this link might exist when the number of participants is increased.

However, even if the FM solves the problem of overreliance on a DSS, there is the potential of an increase in the possibility of other risks associated with a DSS happening. This needs to be taken into account when doing further research on the use of a FM.

## 7. Bibliography

- Alshibly, H. (2018). Re: How to determine sample size in quantitative research? Retrieved from [https://www.researchgate.net/post/How\\_to\\_determine\\_sample\\_size\\_in\\_quantitative\\_research/5a4b2f553d7f4b89d837da40/citation/download](https://www.researchgate.net/post/How_to_determine_sample_size_in_quantitative_research/5a4b2f553d7f4b89d837da40/citation/download).
- Arnott, D., & Pervan, G. (2015). A critical analysis of decision support systems research. London: Palgrave Macmillan. doi:10.1057/9781137509888\_5
- Berberian, B., Sarrazin, J. C., Le Blaye, P., & Haggard, P. (2012). Automation technology and sense of control: a window on human agency. *PloS one*, 7(3), e34075. doi:10.1371/journal.pone.0034075
- Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1999). The cerebellum contributes to somatosensory cortical activity during self-produced tactile stimulation. *NeuroImage*, 10, 448-459. doi:10.1006/nimg.1999.0478
- Budgell, B. (2008). Guidelines to writing case studies. *The Journal of the Canadian Chiropractic Association*, 52(4), 199-204.
- Campbell, E. M., Sittig, D. F., Guappone, K. P., Dykstra, R. H., & Ash, J. S. (2007). Overdependence on technology: an unintended adverse consequence of computerized provider order entry. *AMIA Symposium*, 2007, 94-98.
- Cornelissen, N. (2021). Falsification Machines: a modular approach to Falsifying Decision Support. Bachelor thesis dept. of AI, Radboud University.
- Eom, S., & Kim, E. (2006). A survey of decision support systems applications (1995-2001). *Journal of the Operational Research Society*, 57, 1264-1278. doi:10.1057/palgrave.jors.2602140
- Finlay, P. (1990). Decision support systems and expert systems: A comparison of their components and design methodologies. *Computers and Operations Research*, 17(6), 535-543.
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Friedman, C. P., Elstein, A. S., Wolf, F. M., Murphy, G. C., Franz, T. M., Heckerling, P. S., . . . Abraham, V. (1999). Enhancement of Clinicians's Diagnostic Reasoning by Computer-Based Consultation: A Multisite Study of 2 Systems. *JAMA*, 282(19), 1851-1856. doi:10.1001/jama.282.19.1851

- Frith, C. D. (2014). Action, agency and responsibility. *Neuropsychologia*, 55, 137-142.  
doi:10.1016/j.neuropsychologia.2013.09.007
- Goddard, K., Roudsari, A., & Wyatt, J. (2011, January). Automation bias - A hidden issue for clinical decision support system use. *Studies in Health Technology and Informatics*, 164, 17-22. doi:10.3233/978-1-60750-709-3-17
- Goddard, K., Roudsari, A., & Wyatt, J. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19, 121-127. doi:10.1136/amiajnl-2011-000089
- Goodman, K. W. (1998). *Ethics, computing, and medicine: informatics and the transformation of health care*. Cambridge: Cambridge University Press. Retrieved November 27, 2020, from  
[https://books.google.nl/books?hl=nl&lr=&id=Wb6apuz7\\_ocC&oi=fnd&pg=PA24&dq=ethics+computing+and+medicine&ots=D4-7sAws6U&sig=JPNUBGF4qQTOmzFwU-3mShAEHKc#v=onepage&q&f=true](https://books.google.nl/books?hl=nl&lr=&id=Wb6apuz7_ocC&oi=fnd&pg=PA24&dq=ethics+computing+and+medicine&ots=D4-7sAws6U&sig=JPNUBGF4qQTOmzFwU-3mShAEHKc#v=onepage&q&f=true)
- Grissinger, M. (2019). Understanding Human Over-Reliance On Technology. *P&T: a peer-reviewed journal for formulary management*, 44(6), 320-375.
- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5, 382-385. doi:10.1038/nn827
- Het Rechtenstudentje. (2005, June 29). *ECLI:NL:RBSGR:2005:AT8463*. Retrieved from Het Rechtenstudentje:  
<https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBSGR:2005:AT8463>
- Huisarts DD. (n.d.). Retrieved November 15, 2020, from <http://www.huisartsdd.nl/>
- Irfan, J. (2016, Februari 6). *ECLI:NL:HR:1984:AC8567 (Bijlmer Noodweer-arrest)*. Retrieved from Het Rechtenstudentje:  
<https://www.hetrechtenstudentje.nl/jurisprudentie/eclinlhr1984ac8567-bijlmer-noodweer-arrest/>
- Jones, S. P. (2016, July). *How to write a great research paper*. Retrieved from Microsoft:  
<https://www.microsoft.com/en-us/research/academic-program/write-great-research-paper/>
- Korsten, H. H., Grouls, R. J., & Minderman, N. (2008, March 4). *De sluitsteen van de zorg*. Retrieved December 28, 2020, from Medisch Contact:  
<https://www.medischcontact.nl/nieuws/laatste-nieuws/artikel/de-sluitsteen-van-de-zorg.htm>

- Lauermann, F., & Karabenick, S. A. (2013). The meaning and measure of teachers' sense of responsibility for educational outcomes. *Elsevier*. doi:10.1016/j.tate.2012.10.001
- McLeod, S. A. (2019, July 30). *Qualitative vs. Quantitative research*. Retrieved December 15, 2020, from Simply Psychology: <https://simplypsychology.org/qualitative-quantitative.html>
- Moore, J. W. (2016). What is the Sense of Agency and Why Does It Matter? *Frontiers in Psychology*, 7, 1272. doi:10.3389/fpsy.2016.01272
- Parasuraman, R., & Manzey, D. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors*, 52, 381-410. doi:10.1177/0018720810376055
- Sauro, J. (2011, April 26). *Are both positive and negative items necessary in questionnaires?* Retrieved January 5, 2021, from Measuring U: <https://measuringu.com/positive-negative/#:~:text=By%20including%20a%20mix%20of,which%20should%20reduce%20these%20biases.>
- Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ digital medicine*, 3, 17. doi:10.1038/s41746-020-0221-y
- Tapal, A., Oren, E., Dar, R., & Eitam, B. (2017). The Sense of Agency Scale: A Measure of Consciously Perceived Control over One's Mind, Body, and the Immediate Environment. *Frontiers in Psychology*, 8, 1552. doi:10.3389/fpsyg.2017.01552
- University of Maine School of Law. (2019, November 7). *Legal Research*. Retrieved from Lawguides: <https://lawguides.maine.gov/LR/readingcases>
- van den Hoven, J., & Santoni de Sio, F. (2018, February 28). Meaningful Human Control over Autonomous Systems: A Philosophical Account. *Frontiers in Robotics and AI*, 14. doi:10.3389/frobt.2018.00015
- van der Stigchel, B. (2021). Resilience Towards Bias in Artificial Agents: Human-Agent Team Performing a Triage Task. Master thesis dpt. of AI, Radboud University.
- Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, 54(7), 480-492. doi:10.1037/0003-066X.54.7.480

## 8. Appendix

### 8.1. Table to create medical cases

	Opties voor deze categorie				
Informatie in verhaaltje					
Leeftijd					
Sociale status	Geen kinderen en geen partner	Partner zonder kinderen	Heeft kinderen	Heeft kleinkinderen	
Fitheid	Chronisch ziek/verzorgingstehuis	Ouderdomsklachten	Neutraal	Fit/bezig bijtje	Zeer sportief
Beroep	vrijwillig	geen	betaald		
Hoe ziek geworden	Eigen schuld	Neutraal	Heel veel pech		
Hoe naar ziekenhuis gekomen	Zelf	gebracht	ambulance		
familie situatie	net gescheiden	net familielid overleden	familielid ernstig ziek	Zwangere vrouw/dochter	
Rol in zorg	Geen rol	Discutabele rol	Duidelijke ziekenhuisrol		
Snelheid slechter geworden	Aan begin van de ziekte	Wordt langzaam slechter	Verslechtert in rap tempo		
Naam	Nederlands	Onbekend	Buitenlands		
Soort klachten	Hoesten	Benauwdheid	Koorts	Droge keel	Futloos

This table has been made by Van der Stigchel (2021) in her master thesis to create medical cases.

## 8.2. Conversations

### 8.2.1. Conversation with medical student

This conversation shows the process of how the medical cases were made.

R.v.E. Hoe kun je het beste een casus opbouwen?

R.V. Niet te moeilijk taalgebruik, maar wel uitgebreide informatie geven.

R.v.E. Welke informatie moet er zoal instaan?

R.V. Hoe vaak een klacht voorkomt, de medische geschiedenis van de patiënt, en veel achtergrondinformatie (bijv. werk, gezinssituatie, etc.)

R.v.E. Hoe kunnen we het beste beginnen met een casus schrijven waarbij vier mogelijkheden zijn?

R.V. Ik zal even naar huisartsdd.nl gaan en dan kunnen we wat simpele symptomen zoeken, met verschillende ziektebeelden erbij. Hoesten is bijvoorbeeld een mogelijkheid. Er zijn namelijk meerdere ziektebeelden waarbij hoest een symptoom is.

R.v.E. Welke zijn dat bijvoorbeeld?

R.V. Longkanker, bovenste luchtweginfectie, bijwerking van medicatie, en gastro-oesofageale reflux. Deze zou je kunnen gebruiken voor je eerste casus. De andere symptomen van deze ziektebeelden kun je op internet vinden.

R.v.E. Heb je toevallig nog zo'n soort symptoom?

R.V. Ja hoor, ik zal even gaan zoeken. Ongewild gewichtsverlies kan wel een optie zijn. Dit kan komen door darmkanker, depressie, diabetes type 2, hyperthyreoïdie. Hierbij kun je ook weer andere symptomen zoeken op het internet.

R.v.E. Top, dankjewel voor je hulp!

R.V. Ik zal de casussen nog reviseren voor je als je ze klaar hebt.

### 8.2.2. Conversation with law student

This conversation shows the process of how the law cases were made.

R.v.E. Is het mogelijk om twee goede, korte rechtencasussen te maken?

J.d.V. Dat is wel lastig, want voor een casus is altijd wel een gedeelte van het wetboek nodig, waardoor casussen vrij complex en lang zijn, met veel rechten-termen die leken niet zullen snappen.

R.v.E. Wat adviseer je om nu te doen? Ik heb ze wel nodig namelijk.

J.d.V. Als je bij deze twee termen een casus zoekt, namelijk "noodweer" en "in vereniging handelen", dan deze in het kort opschrijven, stuur deze dan naar mij zodat ik het kan controleren of het alle basisinformatie bevat, maar het nog steeds redelijk goede casussen zijn.

R.v.E. Oke, dit zal ik doen. Dankjewel voor je hulp alvast!

The conversation ended.

A few days later, there was another conversation with the same persons.

R.v.E. Ik heb twee casussen gevonden en die om proberen te schrijven naar kortere casussen met vier verschillende antwoordmogelijkheden. Alleen ben ik niet zeker of ze nog wel kloppen volgens de normen van een rechtencasus.

J.d.V. Ik zal er even naar kijken en ze aanpassen zodat ze kloppen volgens een rechtencasus en waar nodig wat rechtenjargon toevoegen.

R.v.E. Dankjewel, als het je gelukt is zou je ze dan weer naar mij willen sturen.

## 8.3. Cases

### 8.3.1. Casus 1: Ongewild gewichtsverlies

Mevrouw Janssen komt op uw spreekuur met de klacht van ongewild gewichtsverlies. Ze is 53 jaar oud en werkt als filiaalmanager bij de Etos in het dorp. U kent haar vanwege een urineweginfectie 4 jaar geleden. Ook is haar man vorig jaar overleden door een auto-ongeluk, **waarvoor ze tweemaal bij u is geweest om over te praten**. Verder heeft ze 3 kinderen die uit huis wonen en heeft zelf een kat.

Ze vertelt dat ze de afgelopen maanden heeft gemerkt dat ze is afgevallen. **Het gaat om 6 kilogram in de afgelopen 2 maanden**. Hoewel ze er zelf blij mee is, vraagt ze zich toch af hoe het komt omdat ze niet actief bezig is om af te vallen. Als je doorvraagt geeft ze aan dat **haar eetlust nog steeds slecht** is sinds haar man overleden is, maar dat dit al het gehele jaar zo is. Drinken doet ze naar eigen zeggen wel genoeg, **want ze heeft bijna altijd wel dorst**. Ze heeft niks veranderd in haar dieet en probeert net als vroeger nog wel lekker voor zichzelf te koken. **Hier heeft ze dan ook nog wel plezier in**, zeker als haar kinderen mee-eten. Wel is ze **vaak moe** als ze na werk thuiskomt en moet heel soms een afspraak 's avonds afzeggen, zeker als het op haar werk weer zo druk is geweest. Ze heeft soms last van lastige klanten en er gebeuren soms vervelende dingen, wat ze dan mee naar huis neemt en waar ze dan over **piekert**. Daar kan ze zich dan erg over **opwinden en dan wordt ze er helemaal warm en naar van**.

Haar ontlasting is “net als anders”. Bruin van kleur, **wat aan de dunne kant** maar geen diarree. Ze hoeft nooit te persen. Ze gaat 2 keer per dag en **weet niet of er slijm/bloed bij de ontlasting zit**, want ze “kijkt nooit naar wat ze achterlaat”. Met plassen heeft ze geen problemen, behoudens ooit de urineweginfectie waar ze sindsdien geen last meer van heeft gehad.

Haar tante is op 86-jarige leeftijd overleden aan darmkanker. Mevrouw Janssen gebruikt geen medicijnen. Er zijn geen allergieën bekend. Ze gebruikt ongeveer **4 eenheden alcohol per week, vroeger wel veel meer samen met haar man**. Ze drinkt 4 kopjes koffie per dag, rookt sinds vorig jaar weer vast twee pakjes per week, maar daarvoor heeft ze 30 jaar niet meer gerookt. Drugs gebruikt ze niet.

Wat is naar aanleiding van het bovenstaande verhaal de meest waarschijnlijke oorzaak van het ongewild gewichtsverlies van Mevrouw Janssen?

**Op basis van de gegeven symptomen en de dataset van de voorafgaande casus adviseert het diagnose ondersteuningssysteem dat u gebruikt antwoord C als meest waarschijnlijke diagnose.**

Antwoordmogelijkheden:

A. **Darmkanker** (Info knop: De klachten van darmkanker zijn veelal aspecifiek. Alarmsymptomen die mogelijk op darmkanker kunnen wijzen zijn een veranderd ontlastingspatroon, bloedarmoede en winderigheid. Een slechte levensstijl en een hoge alcoholconsumptie kan de kans op darmkanker vergroten.)

B. **Depressie** (Info knop: Depressie is een stemmingsstoornis die zich kenmerkt door een verlies van levenslust of zware neerslachtigheid. Symptomen zijn onder andere een sombere stemming, verlies van interesse of plezier en vermoeidheid.)

C. **Hyperthyreoïdie** (Info knop: Hypothyreoïdie oftewel een te snel werkende schildklier kenmerkt zich onder andere door vermoeidheid, bleke huid en traagheid in denken en handelen.)

D. **Diabetes mellitus type 2** (Info knop: Oftewel suikerziekte heeft als meest voorkomende symptomen veel plassen en drinken, vaak last van blaasontsteking en vermoeidheid.)

Antwoord participant:

Vraag van de Falsificatie machine op antwoord A: Kan darmkanker de meeste symptomen van Mevrouw Janssen, zoals altijd dorst hebben, verklaren?

Vraag van de Falsificatie machine op antwoord B: Kan een depressie de oorzaak zijn van zo'n groot gewichtsverlies?

Vraag van de Falsificatie machine op antwoord C: Heeft u eraan gedacht dat hyperthyreoïdie een vrij zeldzame aandoening is en voorkomt bij 2 op de 1000 mensen?

Vraag van de Falsificatie machine op antwoord D: Kan diabetes mellitus type 2 de oorzaak zijn van de verminderde eetlust bij Mevrouw Janssen?

Mogelijkheid voor de participant om van antwoord te wisselen:

Nu volgen een aantal stellingen over de door u gemaakte beslissing. De stellingen gaan over hoe je tot je beslissing gekomen bent.

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
Ik ben volledig zelfstandig tot mijn diagnose gekomen	0	0	0	0	0	0	0
De diagnose was mijn bewuste beslissing	0	0	0	0	0	0	0
Het diagnose ondersteuning systeem heeft mijn antwoord beïnvloed	0	0	0	0	0	0	0
De vraag van de falsificatie machine heeft mijn antwoord beïnvloed	0	0	0	0	0	0	0
Ik ben verantwoordelijk voor wat mijn acties teweegbrengen	0	0	0	0	0	0	0

### 8.3.2. Casus 2: droge hoest

Meneer Hendriksen komt op uw spreekuur met klachten van hoest. Hij is 68 jaar oud, is met vervroegd pensioen gegaan na een leven als ambtenaar bij de gemeente en woont samen met zijn vrouw met wie hij twee zonen heeft. Ook heeft hij inmiddels een kleindochter. Hij is in de praktijk bekend met hoge bloeddruk, waarvoor u 2 maanden geleden bij hem lisinopril bent gestart. Verder is zijn voorgeschiedenis blanco.

Sinds anderhalve maand heeft hij last van een hoest waarbij hij geen slijm opgeeft. Het is af en toe aanwezig, maar hij kan niet goed aangeven hoe vaak. Bij doorvragen is het “een paar keer per dag” waarop hij een flinke hoestbui heeft. Dan prikkelen zijn longen de hele tijd, waardoor hij moet hoesten. Soms wordt hij 's nachts ook wakker om te hoesten. Hij neemt dan wat water om de hoestkriebel weg te krijgen, wat soms werkt.

Hij is de afgelopen tijd niet ziek geweest en heeft geen koorts gehad, maar heeft een tijd geleden wel last gehad van een verkoudheid.

Hij heeft heel soms iets last van benauwdheid maar niet echt tijdens inspanning. Echt sporten doet hij niet, want “de trap oplopen is al beweging genoeg”. Eten gaat zoals gewoon, al heeft hij soms last van een oprisping als hij te veel gegeten heeft. Hij eet soms een klein beetje minder en is de afgelopen tijd een “paar pondjes afgevallen” wat hij prima vindt.

In de familie komen geen ziektes voor. Hij gebruikt alleen lisinopril en af en toe een paracetamolletje voor de hoofdpijn. Hij is niet bekend met allergieën. Alcohol gebruikt hij sporadisch op een feestje. Roken doet hij al 45 jaar, waarbij hij de laatste 20 jaar wel minder is gaan roken en nu nog maar op een half pakje per dag zit.

Wat is naar aanleiding van het bovenstaande verhaal de meest waarschijnlijke diagnose voor Meneer Hendriksen?

**Op basis van de gegeven symptomen en de dataset van de voorafgaande casus adviseert het diagnose ondersteuningssysteem dat u gebruikt antwoord B als meest waarschijnlijke diagnose.**

Antwoordmogelijkheden:

A. **Longkanker** (Info button: Veelvoorkomende symptomen zijn hoesten, ophoesten van bloed en kortademigheid. Vaak gaan er ook gewichtsverlies en vermoeidheid mee gepaard.)

B. **Bovenste luchtweginfectie** (Info button: Is een infectie in de neusholte, keelholte of longblaasjes. Symptomen zijn onder andere hoesten, niezen, kortademigheid en verkoudheid.)

C. **Bijwerking medicatie** (Info button: Lisinopril verlaagt de bloeddruk en verbetert de pompkracht van het hart. Bijwerkingen kunnen zijn hoesten, duizeligheid en overgevoeligheid. Dit komt bij ongeveer 10 procent van de gevallen voor.)

D. **Gastro-oesofagale reflux** (Info button: Oftewel brandend maagzuur. Symptomen kunnen zijn het proeven van maagzuur in de mond, pijn achter het borstbeen en slikproblemen. Komt vaak voor tijdens de nacht wat tot hoesten en vermoeidheid kan leiden.)

Antwoord participant:

Nu volgen een aantal stellingen over de door u gemaakte beslissing. De stellingen gaan over hoe je tot je beslissing gekomen bent.

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
Ik ben volledig zelfstandig tot mijn diagnose gekomen	0	0	0	0	0	0	0
De diagnose was mijn bewuste beslissing	0	0	0	0	0	0	0
Het diagnose ondersteuning systeem heeft mijn antwoord beïnvloed	0	0	0	0	0	0	0
Ik ben verantwoordelijk voor wat mijn acties teweegbrengen	0	0	0	0	0	0	0

### 8.3.3. Casus 3: Bijlmer Noodweer

U bent rechter in de zaak van Mevrouw de Jong. In de Amsterdamse woonwijk Bijlmermeer vinden de laatste tijd veel tasjesroven plaats die vaak op dezelfde wijze door twee jongemannen worden gepleegd. Omdat Mevrouw de Jong een paar maanden geleden al eens eerder van haar tasje was beroofd, had zij illegaal een pistool gekocht en was ze lid geworden van een schietvereniging. In de hal van haar flatgebouw in de Bijlmer stond ze met een geladen pistool in haar jaszak op de lift te wachten. Daar werd ze overvallen door twee jongens (R. en G.) die haar tas wilden stelen. Ze werd brutaal vastgegrepen en met een mes bedreigd. Omdat een waarschuwingsschot niet hielp, heeft ze nogmaals geschoten. R. werd dodelijk in de borst geraakt en strompelde weg. Toen G. terugkwam om alsnog haar tas te roven, lag ze nog op de grond en heeft ze nogmaals het vuur geopend, waarbij G. zware verwondingen opliep.

De vrouw wordt vervolgd door de officier van justitie. Ze doet een beroep op noodweer (Infoknop: Noodweer of legitieme verdediging is geweld dat ongestraft gebruikt mag worden om aanranding jegens eigen of anders lijf tegen te gaan.) en subsidiair noodweerexcès (Infoknop: Noodweerexcès is de overschrijding van de grenzen van de noodzakelijke verdediging, indien zij het onmiddellijk gevolg is geweest van een hevige gemoedsbeweging, door de aanranding veroorzaakt.). Verdachte wordt (1) doodslag, (Infoknop: Doodslag is het opzettelijk, maar niet met voorbedachten rade, beroven van het leven van een ander. Wanneer er wel sprake is van voorbedachten rade spreekt men hier van moord. Voor doodslag kan maximaal 15 jaar cel geëist worden.) (2) poging tot doodslag (Infoknop: Poging tot doodslag is het proberen iemand van het leven te beroven. Voor poging tot doodslag kan maximaal 10 jaar cel geëist worden) en (3) verboden vuurwapenbezit (Infoknop: Voor verboden wapenbezit van een pistool kan maximaal drie maanden cel geëist worden.) ten laste gelegd.

Kan er een geslaagd beroep op noodweer dan wel noodweerexcès worden gedaan, ondanks het feit dat verdachte een verboden vuurwapen heeft gebruikt? Het gerechtshof stelde dat de verdachte gezien haar ervaring en de korte afstand waarop de verdachten zich bevonden, in staat had moeten zijn geweest om hen op minder vitale plekken te raken. Mevrouw de Jong zegt echter dat het schieten op vitale plekken een onmiddellijk gevolg was van een hevige gemoedsbeweging die door de bedreiging met het mes werd veroorzaakt.

Hoe zou u Mevrouw de Jong, naar aanleiding van bovenstaand verhaal, berechten?

**Op basis van de gegeven feiten en de dataset van de voorafgaande casus adviseert het beslissing ondersteuningssysteem dat u gebruikt antwoord D als beste vonnis.**

Antwoordmogelijkheden:

- A. Mevrouw de Jong moet ontslagen worden van alle rechtsvervolging, want er is sprake van noodweer en het vuurwapenbezit was nooit ontdekt als zij niet overvallen was.
- B. Mevrouw de Jong moet alleen veroordeeld worden voor het bezitten van een illegaal vuurwapen, maar niet voor punt 1 of 2, want er is sprake van noodweer.
- C. Mevrouw de Jong moet alleen veroordeeld worden voor het bezitten van een illegaal vuurwapen, maar niet voor punt 1 of 2, want er is sprake van noodweerexcès.

D. Mevrouw de Jong moet voor alle te laste gelegde daden veroordeeld worden. Het voorhanden hebben van een verboden vuurwapen maakt een beroep op noodweer(exces) onmogelijk.

Antwoord participant:

Vraag van de Falsificatie machine op antwoord A: Heeft u het feit dat Mevrouw de Jong het wapen doorgeladen, dus klaar voor gebruik, bij zich droeg genoeg gewicht gegeven in uw beslissing?

Vraag van de Falsificatie machine op antwoord B: Is de reactie van Mevrouw de Jong niet te heftig gezien de omstandigheden waarin ze zich bevond?

Vraag van de Falsificatie machine op antwoord C: Kan Mevrouw Janssen zich zo bedreigd gevoeld hebben dat iemand neerschieten en iemand ernstig verwonden als noodweereces gezien kan worden?

Vraag van de Falsificatie machine op antwoord D: Is Mevrouw de Jong een gevaar voor de samenleving met een kans op herhaling?

Mogelijkheid voor de participant om van antwoord te wisselen:

Nu volgen een aantal stellingen over de door u gemaakte beslissing. De stellingen gaan over hoe je tot je beslissing gekomen bent.

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
Ik ben volledig zelfstandig tot de uitspraak gekomen	0	0	0	0	0	0	0
De uitspraak mijn bewuste beslissing	0	0	0	0	0	0	0
Het beslissing ondersteuning systeem heeft mijn antwoord beïnvloed	0	0	0	0	0	0	0
De vraag van de falsificatie machine heeft mijn antwoord beïnvloed	0	0	0	0	0	0	0
Ik ben verantwoordelijk voor wat mijn acties teweegbrengen	0	0	0	0	0	0	0

### 8.3.4. Casus 4: Steen van een viaduct

Verdachten Jeroen van G. (22), Davy V. (18), Stephan H. (19) en Frederik B. (18) verveelden zich op een avond en besloten om af te spreken en rond te rijden in de omgeving. Na een tijd kwamen de jongens bij een viaduct aan en daar besloten van G. en V. takken op de snelweg te gooien. Vervolgens besloten zij een zak puin van het viaduct te gooien en de betonnen voet van een verkeersbord de weg op te slepen. Wonderbaarlijk genoeg is hierbij niemand gewond geraakt.

Na verder rondgereden te hebben, stopten de verdachten op een viaduct over de A4. Frederik en Stephan horen Jeroen en Davy zeggen dat zij een stoeptegel van het viaduct gaan gooien, maar hebben zich hier niet van gedistantieerd. Vervolgens voegden Jeroen en Davy de daad bij het woord. Het slachtoffer reed op dat moment met haar auto onder het viaduct door. De tegel is door de voorruit gegaan en heeft het slachtoffer op het hoofd geraakt, hetgeen haar onmiddellijke dood tot gevolg had. Stephan en Frederik zaten tijdens dit alles in de auto te wachten.

Jeroen en Davy zagen de tegel door de ruit van de auto gaan en de vier vluchtten vervolgens weg in de auto. Later lazen de verdachten in de krant dat de 30-jarige vrouw is komen te overlijden, maar spraken af zichzelf niet aan te geven. Een week later werden ze door de politie opgepakt.

De verdachten komen allen uit stabiele gezinnen en hebben geen strafblad. Het OM legt de verdachten (1) medeplegen van moord (Infoknop: Er is sprake van medeplegen van moord als er geen verschil is in rol en bijdrage van de verschillende plegers.) en (2) medeplegen van poging tot moord (Infoknop: Er is sprake van medeplegen van poging tot moord als er geen verschil is in de rol en bijdrage van de verschillende plegers en er moet een opzet zijn tot het plegen van een misdrijf.) ten laste. Tegen Jeroen van G., vader van een zontje van 1, en Davy V. wordt 8 jaar geëist. Tegen Stephan H. en Frederik B. wordt beiden vier jaar geëist.

U bent rechter in deze zaak. Hoe zou uw vonnis luiden?

**Op basis van de gegeven feiten en de dataset van de voorafgaande casus adviseert het beslissing ondersteuningssysteem dat u gebruikt antwoord A als beste vonnis.**

Antwoordmogelijkheden:

- A. De verdachten moeten allen 6 jaar krijgen, want de verdachten zijn allen medeplegers aan de misdaad. Ze hebben in vereniging gehandeld en hierdoor moeten zij allen dezelfde straf krijgen.
- B. Jeroen van G. en Davy G. moeten veroordeeld worden voor de geëiste straf; Stephan en Frederik moeten worden vrijgesproken, omdat zij de steen niet gegooid hebben en ook geen opzet hadden op het sterven van de vrouw.
- C. De verdachten moeten allen de geëiste straf krijgen, want de verdachten hebben tezamen en in vereniging met elkaar opzettelijk en met voorbedachte rade de vrouw van het leven beroofd.

D. De verdachten moeten niet voor moord maar voor doodslag veroordeeld worden, omdat zij niet van tevoren besloten hebben iemand van het leven te beroven. Het gooien van de stoeptegel gebeurde in een opwelling. Dit heeft als gevolg dat alle de helft van hun geëiste straf krijgen.

Antwoord participant:

Nu volgen een aantal stellingen over de door u gemaakte beslissing. De stellingen gaan over hoe je tot je beslissing gekomen bent.

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
Ik ben volledig zelfstandig tot de uitspraak gekomen	0	0	0	0	0	0	0
De uitspraak was mijn bewuste beslissing	0	0	0	0	0	0	0
Het beslissing ondersteuning systeem heeft mijn antwoord beïnvloed	0	0	0	0	0	0	0
Ik ben verantwoordelijk voor wat mijn acties teweegbrengen	0	0	0	0	0	0	0

## 8.4. Case results

### 8.4.1. Case 1

<i>Case 1</i>	<b>Medical students</b>	<b>Non-medical students</b>
<i>I made my diagnosis completely independently</i>	Mean: 6.375 Range: 5-7	Mean: 4.833 Range: 2-7
<i>The diagnosis was a conscious decision</i>	Mean: 6.25 Range: 5-7	Mean: 5.5 Range: 3-7
<i>The diagnosis support system has influenced my answer</i>	Mean: 2.125 Range: 1-5	Mean: 4.083 Range: 2-6
<i>The question from the falsification machine has influenced my answer</i>	Mean: 2.225 Range: 1-5	Mean: 3.416 Range: 2-7
<i>I am responsible for what my actions might bring about</i>	Mean: 5.875 Range: 4-7	Mean: 5.167 Range: 2-7

### 8.4.2. Case 2

<i>Case 2</i>	<b>Medical students</b>	<b>Non-medical students</b>
<i>I made my diagnosis completely independently</i>	Mean: 6.5 Range: 6-7	Mean: 4.417 Range: 2-7
<i>The diagnosis was a conscious decision</i>	Mean: 6.25 Range: 6-7	Mean: 5.333 Range: 3-7
<i>The diagnosis support system has influenced my answer</i>	Mean: 2.125 Range: 1-5	Mean: 4.5 Range: 1-7
<i>I am responsible for what my actions might bring about</i>	Mean: 5.875 Range: 4-7	Mean: 5.25 Range: 2-7

### 8.4.3. Case 3

<b>Case 3</b>	<b>Medical students</b>	<b>Non-medical students</b>
<i>I made my diagnosis completely independently</i>	Mean: 5.375 Range: 3-7	Mean: 6.083 Range: 2-7
<i>The diagnosis was a conscious decision</i>	Mean: 5.625 Range: 5-6	Mean: 5.917 Range: 4-7
<i>The diagnosis support system has influenced my answer</i>	Mean: 3.375 Range: 2-5	Mean: 4.167 Range: 1-6
<i>The question from the falsification machine has influenced my answer</i>	Mean: 2.75 Range: 2-5	Mean: 4 Range: 1-6
<i>I am responsible for what my actions might bring about</i>	Mean: 6 Range: 5-7	Mean: 5.667 Range: 2-7

### 8.4.4. Case 4

<b>Case 4</b>	<b>Medical students</b>	<b>Non-medical students</b>
<i>I made my diagnosis completely independently</i>	Mean: 6.125 Range: 6-7	Mean: 5.25 Range: 2-7
<i>The diagnosis was a conscious decision</i>	Mean: 6.125 Range: 6-7	Mean: 5.583 Range: 4-7
<i>The diagnosis support system has influenced my answer</i>	Mean: 2.25 Range: 1-4	Mean: 4.083 Range: 2-6
<i>I am responsible for what my actions might bring about</i>	Mean: 6 Range: 5-7	Mean: 5.25 Range: 2-6

### 8.4.5. Sense of Agency participants

<b>Sense of Agency</b>	<b>Medical students</b>	<b>Non-medical students</b>
<i>Case 1</i>	6.219	5.073
<i>Case 2</i>	6.281	4.585
<i>Case 3</i>	5.593	5.937
<i>Case 4</i>	6.075	5.333