LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Department "Institut für Informatik"
Chair of Human Computer Interaction
Prof. Dr. Andreas Butz

Master Thesis

Automation supporting Sports Trainers

Julia Pühl
jcpuehl@gmail.com

Commenced: 13. 2. 2023
Completed: 14. 08. 2023
Supervisor: Dennis Dietz, Florian Bemmann
Examiner: Prof. Dr. Andreas Butz

Abstract

In this thesis, the use of automation in fitness training is explored, particularly its effects on trainers' expertise and user experience. The research involves implementing an automated training plan feature to support trainers. The goal is to investigate whether such automation could compromise trainers' expertise or bring other negative consequences. A user study was conducted, tracking the trainers' interactions with the automated feature to identify benefits and potential drawbacks. Results indicate gains in the trainers' efficiency as the time they needed to create training plans was significantly reduced using the automated tool.

However, assessing negative impacts, particularly the loss of expertise, brought difficulties. Trainers maintained extensive engagement, frequently adjusting automated output. This active involvement suggests that they were able to integrate their expertise despite the use of automation. Yet, understanding the motivations behind these adjustments remained challenging due to limited qualitative data on the trainers' perspectives. This hindered insights into whether adjustments were caused by usability concerns or by accuracy issues.

In conclusion, embedding diverse interaction options within automated features can mitigate the risk to restrict human expertise. On the positive side, automation accelerates processes, but to prevent negative impacts, user engagement needs to be ensured. The study requires additional qualitative insights to better comprehend automation-user dynamics.

I hereby declare that I carried out the present work independently, marked all quotations as such as well as all sources and aids used.

München, August 14, 2023

Contents

1 Introduction

1.1 Problem

Automated processes and support by AI have become a self-evident part of our daily lives: automatic confirmation emails, driverless cars, automatically corrected grammar errors, or texts generated entirely artificially. Automation and AI make our lives much easier in many ways because they provide quick and easy access to exactly the information we need or show us suggestions that match our interests [24]. Especially in the working environment, these technological enhancements open up a lot of opportunities that would be impossible to achieve by human labor, such as the analysis of big data or rapidly executing repetitive tasks. Moreover, in many areas, human workforce is scarce and expensive [25]; automating processes might not only be cheaper but also more reliable because human faults can be avoided [23].

These technologies keep being enhanced and distributed rapidly with the focus on performance and improvements, often disregarding the impacts they have on us and our current ways of living. There are all kinds of opinions and speculations on what the future with these tools will look like, but the majority agrees that their potential risks should be taken seriously. Sam Altman, the CEO of OpenAI, recently even compared the risk of unchecked AI development to threats like nuclear wars or pandemics [25].

Technological changes and inventions are part of the passage of time. The question is, why the current one is so remarkable. One reason is the complexity and thus unpredictability of the recently developed tools, like ChatGPT or other AI models[11]. It is therefore easy to lose overview over what they do and where they play a role. Another reason is that they affect almost everyone: it could be a 3D printer for craftsmen, robotic surgery, or code pilot generating an app - automation and AI are applied in a very broad range of fields and professions.

Subsequently, for many, this comes with the fear that new technologies will take over their job and make them superfluous. Ten years ago, Frey and Osborn [19] estimated, that 47% of all jobs in the U.S. are likely to be computerized. Arntz et al. [11] questioned this result and took a different approach to determining how many and which jobs are likely to be affected by automation and digitalization. They also took into consideration the individual tasks that a job includes. This way, they came to the conclusion that only 9% of jobs in the U.S. 'face high automatability' [11]. A finding they both stated is that lower educational jobs requiring less expertise are more endangered to be automated. However, regardless of whether a job is taken over fully or partly by recent technological advances, it changes common ways of working.

Another potential challenge that comes with automation and AI is the loss of control: using it comes with the requirement to trust it, even without understanding the underlying logic. While this causes some to be skeptical, others tend to over-rely on technology and do not question their outcome [35]. To prevent either of these reactions, more research is needed on how to ensure that new technologies are integrated into our lives in a way that is beneficial for us.

In summary, the problem is that recent advances in technology are developing and spreading very quickly affecting us in various ways, which makes it difficult to track, understand and control these changes. Therefore, more research is required firstly, on the impacts that using automation and AI can have and how these technologies can be applied to create an added value.

1.2 Idea

The answer of ChatGPT to the question, of whether automation and AI will make humans redundant, is: the maximal benefit comes from the combination of both, human qualities and the power of AI and automation. This is also a much-stated conclusion in literature [54, 11]. Technologies should be used as a tool that amplifies human skills instead of diminishing them. This is the idea behind this thesis: to combine human expertise and automation. The goal is to analyze, whether this combination can bring added value. This will be determined by identifying potential positive and negative impacts of using automation.

RQ1: What impacts can using automation have on expert users?

RQ2: What added value does automation bring expert users?

The impacts and added value of automation will be investigated with the focus on expert users to ensure that the users have expertise. The reason for this specification is that the main goal of this paper is to find out whether a balance between automation and human skills can be achieved or if one excludes the other. This will be investigated by applying automation in a real example, an app for personal trainers to coach their trainees. With the app, trainers can track the athletes' data, give feedback and generate training plans for them. Creating these plans requires knowledge and experience as there are many factors to modify in order to achieve a certain training goal. On the other hand, it is a repetitive task because the plans usually follow a certain pattern. In practice, trainers are often held up by manually typing in the same values multiple times. In addition to that, a trainer typically has multiple trainees who all expect to be supported individually - an issue known as the one-to-many problem. Automation can provide solutions to these challenges. Therefore, the app is used as an exemplary use case, where automation can be integrated, while expertise still plays an important role. The benefits and impacts of automation will be investigated with the example of a tool that automatically generates training plans.

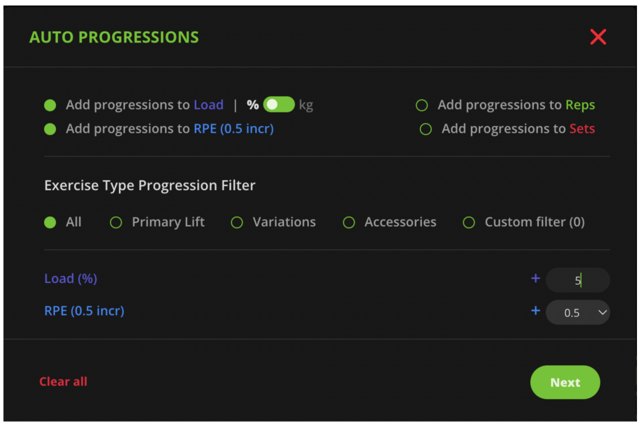


Figure 1.1: The Auto Progression Feature of MyStrengthBook [4]

An algorithm that sets up a training plan automatically has already been implemented in various existing apps and programs. One of them is AlphaProgression [2], an app that first asks the user to specify initial settings, and then generates a training plan accordingly. The app guides the user through the plan and all its exercises, indicating when and which exercises should be trained for how long or in how many repetitions. The user can type in whether they managed to perform the exercise the way it was defined by the app or not. The app includes a logic that modifies the exercises of the following week according to the performance of the user. MyFitCoach [8] is very similar to AlphaProgression; the user flow through both apps is compared in Figure 1.2.

Other examples of apps that create training plans are JuggernautAI [5] or enduco [6]. JuggernautAI is, like AlphaProgression and MyFitCoach, focused on powerlifting and bodybuilding.

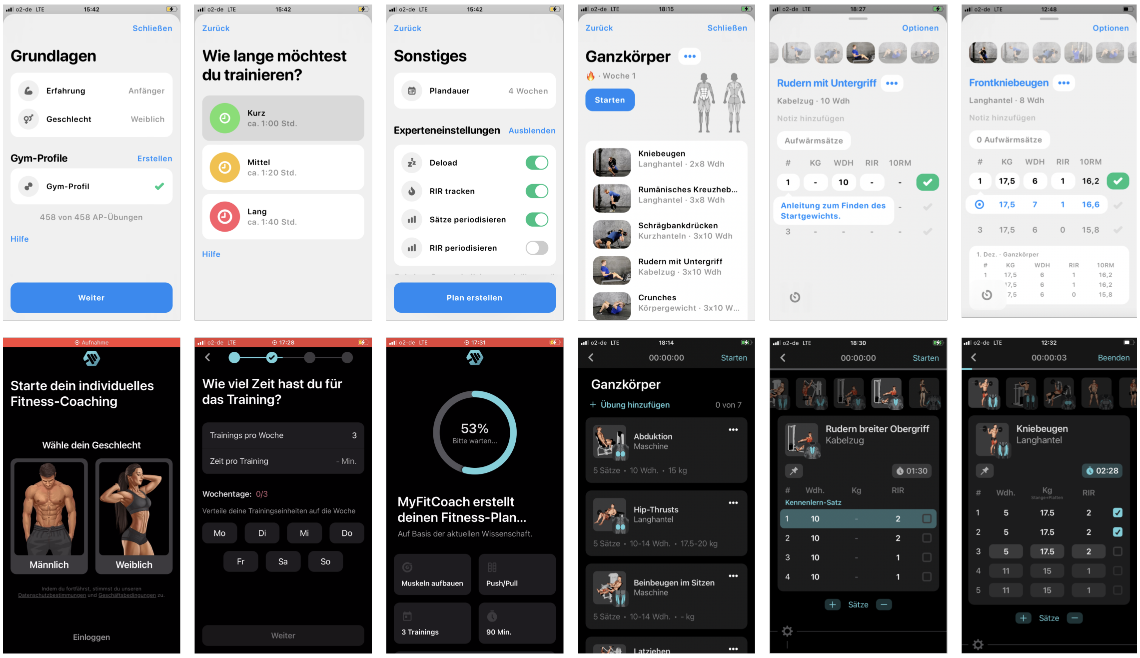


Figure 1.2: User Flow through AlphaProgression [2] (top) and MyFitCoach [8] (bottom)

Enduco on the other side was developed mainly for cardio training for example for runners or cyclists. The so far mentioned apps have in common that they use some sort of AI that creates the training plans and guides the user through their training so that human trainers are no longer needed. Opposed to that, the idea of this thesis is to support trainers with AI, not to replace them. There are also apps following this approach, for example by MyStrengthBook [4]. In the MyStrengthBook app, trainers can create training plans for their athletes and track their progress. However, MyStrengthBook does not include any kind of logic or intelligence. It offers the feature 'Auto Progression', but this requires the trainer to give input and does not automatically come up with a new training plan (see Figure 1.1).

In order to answer the research questions, a tool will be built that is able to generate a training plan automatically. It involves a logic that can tune the plan according to the athlete's individual performance. However, the trainer is still required to review the automated plan. The trainer can accept it or modify it. This training plan tool will be integrated into the existing Traindoo app, where it will be tested by its users. The tool is meant to exemplarily represent a possible combination between human expertise and automation. By testing it, its potential added value and impacts can be found and the following hypotheses can be verified or rejected:

Hypotheses on positive impacts of automation

I. Automation helps trainers to save time.

II. Automation increases the trainers' productivity and quality of work.

Hypotheses on negative impacts of automation

III. Automation restricts the integration of human expertise.

IV. Automation creates stress or skepticism among the trainers.

V. Automation causes a lack of motivation and attentiveness among the trainers.

1.3 Outline

In order to answer the research questions, the following steps will be taken. First, literature research is done to clarify, what automation means, in what contexts it is used in combination with human expertise, and what challenges come with it. Then, the focus is put on the user side to be able to characterize and understand the trainers, their ways of working, and their limitations better. Data on this is collected via a pre-study survey sent to the trainers. Furthermore, the sports training context and especially the logic behind training plans are researched to be able to develop a tool that automatically generates training plans.

Knowing the context and users for which automation will be applied, the main topics will be addressed, which are the impacts and added value of automation. In order to analyze them, an automated training plan tool is built and integrated with the existing training app. This process consists of the common steps of a software development cycle: first, the requirements are determined according to the outcome of the user survey and research. Then, the new system is designed and implemented in React Typescript. To identify a potential added value, a user study is conducted. For this, user data is tracked on how trainers usually create new training plans using the existing methods and on how they use the new automated tool. Both data sets are compared in order to determine possible impacts and whether using the new tool has added value for the trainers. The trainers are asked to give feedback via a questionnaire during and after the user study on their experience with the tool and its usability.

2 Background

As the research question will be investigated by testing an automated tool applied in the sports context, some basic terms and concepts of sports training and training planning will be explained first.

2.1 Online Sports Training

Just like many other fields, sports has become a very versatile setting that no longer only takes place in real life or in person. At the latest since COVID-19, online fitness plan providers or coaches have experienced a massive increase in demand. The World Economic Forum [22] in September 2020 announced that the number of fitness and health app downloads has increased globally by 46%. Reasons for this growth are not only the pandemic, as a market analysis published by Globe Newswire shows [7]. The high demand in online fitness training is due to technological advances, like AI, apps and wearables, people's 'on-the-go' lifestyles and their wish for personalized programs. This demand is expected to grow significantly in the upcoming ten years.

These changes do not only affect athletes, but also trainers. Apps and online coaching opens up many new opportunities; tracking and supervising the training does not have to happen on paper and in real-time anymore. This is the idea behind the Traindoo App: an app where athletes and trainers can simultaneously enter their training progress or feedback, instead of exchanging fixed tables or plans of training data in real life.

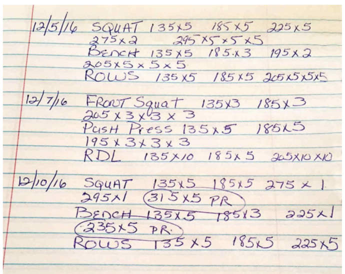


Figure 2.1: Training Log by Glenn Pendlay [1]

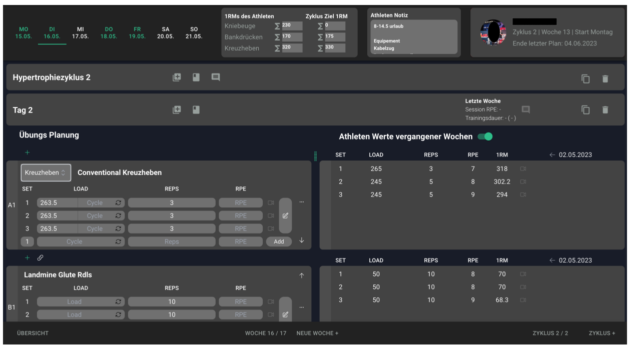


Figure 2.2: Training Plan in the Traindoo App

Figure 2.1 shows a training plan by Glenn Pendlay, a famous American Olympic weightlifting coach [1]. To him, the key to success in weightlifting was keeping a training log. A training log contains all details about the executed exercises and determines the parameters of the upcoming exercises. Tracking the training progress that way is especially common in power- and weightlifing, bodybuilding and calisthenics [43]. Therefore, the Traindoo App is mainly focused on these types of sports. Figure 2.2 shows what a training plan looks like in the app. On the left half of the app screen, the trainer can set the values for the exercises and on the right half the athlete can enter what they achieved. The exercises in powerlifting and bodybuilding can be similar, however, the training goals are different. Powerlifting is about reaching a maximal weight in the three main lifts squat, deadlift and benchpress [30]. In bodybuilding, the goal is to grow as much muscle mass as possible via strength training with weights [10]. In calisthenics, the basic idea is to mainly train with body weight instead of additional weights [47].

2.2 Planning a Training

As mentioned in 1.2, creating a training plan requires knowledge and experience because it involves many parameters that can be tuned individually depending on the training goal. There are three main goals: endurance, hypertrophy and strength [30]. Hypertrophy means making a muscle increase in size by causing muscle fibers to grow. Strength is the ability of a muscle to withstand certain stress in form of pushing, pulling, lifting etc., independent of its size [3]. Endurance is the ability of a muscle to take that stress for a certain amount of time, independent of the magnitude of the stress [42].

Factors Determining a Training Plan

When setting up a training plan for a new athlete with unknown training history, first, some general information about the athlete and their life style needs to be collected; for example, age, diet or sleeping rhythms. Furthermore, it is essential to specify the athlete's goal, their motivation (e.g. fun or competitions) and their availability [39]. To create a training plan for a known athlete with known training history, the first factor to determine is the phase they are currently in: do they have a competition ahead, are they in recovery, also called deload phase, or do they follow general, unspecific training [39]. Apart from the phase, in most cases also the training cycle is relevant. The concept behind training cycles is to split the training into periods of overload and recovery. The periodization causes the body to constantly adapt to the changes and this way build up fitness faster than chronologically constructed training schedules [30]. A periodized training plan consists of macro-, meso- and microcycles, whereof microcycles are the smallest and macrocycles the biggest unit [13]. Furthermore, there are different types of periodization, the linear, the undulating, the reverse and the block periodization [30]. The differences between these types will become apparent in the following.

The remaining factors to determine for a new training plan are the exercises it should contain and what values their parameters should be set to. To decide on suitable exercises, it is necessary to know the way in which they affect the muscle. A muscle has three working modes, concentric (positive-dynamic), eccentric (negative-dynamic) and isometric (static) [30]. The first part of a pull-up for example is concentric and the second part eccentric. A plank or a wall sit would be isometric. The main parameters of an exercise are the load, with which it should be executed, and the number of repetitions indicating how many times it should be repeated. There are more parameters, like the number of sets, the RIR (reps in reserve) or RPE (rating of perceived exertion), that can all be tuned according to the desired training outcome [13]. The so far described process of setting up a training plan is summarized in Figure 2.5.

Training Plan Theories and Examples

In general, creating a training plan is about deciding which factors are relevant and how they should be modified. The modification of parameters from training plan to training plan is called progression, because it is a gradual adaption of values to make the exercises more challenging [16]. There are many different theories on how to modify the above described factors and parameters when planning a training. The basic factors, that influence the remaining parameters are experience level and goal of the athlete [13]. The experience level of an athlete impacts a big part of the training, for example its duration and frequency, but also the number of exercises and sets and the load and repetitions per exercise (see Figure 2.3). Load and repetitions also depend on the athlete's goal. Someone, whose goal is to build maximal strength, should train with high loads and little repetitions. The opposite applies when training endurance (see Figure 2.4).

ILB-Grobraster

Leistungsstufe	Zeitstufe in Monaten	Trainingssystem (Organisationsform)	Trainingshäufigkeit pro Woche	Übungen pro Muskelgruppe	Sätze pro Übung	Intensität[1] (in % ILB)
Orientierungsstufe	0 - 1,5	Ganzkörper	2	1 - 2	1 - 2	gering[2]
Beginner	1,5 - 6	Ganzkörper	2	1 - 2	1 - 2	50 - 70
Geübter	6 - 12	Ganzkörper	2 - 3	1 - 2	2	60 - 80
Fortgeschrittener	> 12	Ganzkörper / Split	3 - 4	1 - 3	2 - 3	70 - 90
Leistungstrainierender	> 36	Ganzkörper / Split	3 - 6	1 - 4	2 - 4	80 - 100

[1] Der Referenzwert für die prozentualen Intensitätsangaben ist das im ILB-Max-Test ermittelte Gewicht
[2] In der Orientierungsstufe wird noch kein ILB-Max-Test durchgeführt. Training mit geringen Intensitäten, orientiert am "subjektiven Belastungsempfinden".

Figure 2.3: Different experience levels and how they influence the setup of a training [13]

Abgrenzung Trainingsziele

Trainingsziel	Anzahl Wdh	alternativ TUT[1]
Kraftausdauer	15 - 20	> 50 Sek
Muskelhypertrophie	8 - 15	20 - 50 Sek
Maximalkraft	5 - 8	< 20 Sek

[1] Die "time under tension" beschreibt die Dauer einer Wdh oder in diesem Falle die Dauer eines Satzes.

Figure 2.4: Different training goals and how they influence training parameters [13]

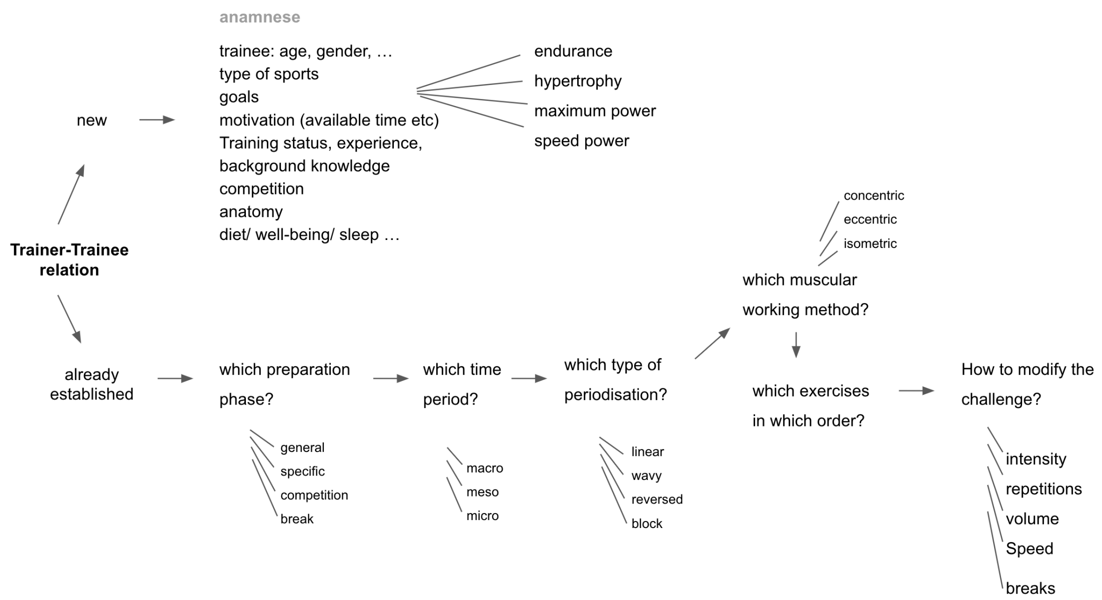


Figure 2.5: Variable parameters within the process of creating a training plan

It gets more complex when considering the different types of alternation between exercise sets and between training weeks. Modifying the parameters is important because that way, the body is forced to continuously adapt to the changes, in other words, training progress can be achieved [30]. There is a number of patterns for the alternation between sets and weeks. The standard pattern is to increase the load from week to week with constant repetitions and constant sets. Another pattern is the pyramid method, where load and repetition are changed between sets: it starts off with high repetitions and a small load with decreasing repetitions and increasing load towards the last set [13]. Furthermore, there is also the 5x5 pattern and the 5/3/1 method. In the 5x5 pattern, an exercise is repeated five times in five sets [14]. The 5/3/1 method refers to lowering the number of repetitions from five in the first set to one in the last [45]. The type of pattern that is chosen again depends on the overall training goal or sometimes also on personal training philosophies.

3 Related Work

In the course of this work so far, the terms automation and AI were often used interchangeably. To shed more light on recent technological advances, in the following chapter, the differences between them, their use cases and challenges that come with them, will be pointed out. When reading about automation, it is many times mentioned together with AI. Because of the recent release of the AI chat bot Chat GPT, the more dominant part is often AI. The idea of some powerful computer intelligence comes to one's mind without questioning further, what is really meant by that. In this chapter, automation is differentiated from AI and other recent technologies to be able to better place the focus of this work within the range of other technologies.

3.1 Automation and Other Recent Technologies

Artificial Intelligence is the simulation of the human intelligence, while automation means letting technology perform a task with little or no human intervention [18]. This is also, what AI is doing, but with the ability to interact, summarize and learn. Automated jobs typically do not involve such skills. These 'smart' capabilities are what makes, for example Chat GPT, special compared to already existing technologies, like search engines. At first sight, both do the same, giving answers to our questions, but search engines simply output, what they found on the internet, while intelligent chatbots generate their own response based on their learned knowledge and on our input [17].

Intelligent Chatbots belong to a form of AI called Natural Language Processing, which is only one of many other types of AI, like computer vision, robotics or predictive analytics [53]. The field of AI is huge and will not be discussed further here, since the point of this chapter is only to give on overview over recent technologies. As opposed to AI, automation is typically used for tasks that humans would also be capable of fulfulling, only needing much more time [33]. This is the case for repetitive jobs, for example, or for processes that involve large amounts of data. Frey and Osborne [19] estimated in 2013, that employments which are at high risk of automation are mostly office and administrative support, service and sales. Jobs within education, community service, arts and media are at a lower risk because they do not involve as many repetitive tasks. Instead, they require qualities like creativity, social intelligence and manual dexterity [19]. Frey and Osborne [19] considered these capabilities to be non-automatable. Today, with the progress of AI, this has changed and it has become more difficult to find a task, that cannot be automated, especially in combination with AI. Still, it has become apparent that even with these advances, new tasks come up that require human intervention, like reacting to failure or unforeseen challenges [11].

Figure 3.1: IUI with a Virtual Companion [36]

Automation and AI are not just used as standalone tools, but they are often combined with existing systems. One example is the integration of AI into user interfaces, known as intelligent user interfaces (IUIs) [27]. With recent advances, technologies become smarter and more automated, but still, some sort of interface between them and us as users almost always remains. Therefore, the idea is to optimize interfaces using AI. The ideal interface is one that knows the user and is able to understand them just from natural human interaction methods. This can be achieved using AI methods, like speech recognition or computer vision [27]. IUIs can interpret the user's desires and adapt to them and the situation individually; for example, they change the visualization space or content structure accordingly. IUIs know about the user and the situation because they use various models containing information about the user, the task they want to achieve and the context [28]. These models are continuously updated to ensure that the interface reacts correctly. Noh et al. [36] built an IUI with a virtual companion to assist elderly using smart phones. They displayed the companion as 3D character that reacts to the user via speech and activity recognition and shows the user the required information as animation. When the phone receives a call, for example, instead of displaying the call, the companion lets the user know with the corresponding gesture (see Figure 3.1).

3.2 Automated and Intelligent Support in Different Contexts

The idea of this thesis is to support professionals by automating parts of their job without losing their expertise. The concept of such an automated support system can be found for all kinds of professions since experts across various fields often have similar struggles: they need to handle a lot of data in a limited amount of time and face the one-to-many problem [54, 37]. A short overview over automated support systems used in other areas is given in the following.

Healthcare and Medicine

The one-to-many problem is quite severe in healthcare and medicine because of the shortage of staff. This has the consequence, that sometimes staff is employed that has not finished their education or has not had enough training yet [23]. According to Harleem et al. [23], many patients suffer or even die from faulty treatment by insufficiently trained doctors. Medical training can be very complex because often the real-life scenario, for example surgery, is complex to recreate in a practice setup. Automation and AI can improve this situation by automatically evaluating the trainees' performance and giving them direct feedback. Furthermore, it can provide very accurate simulations of complex real-life scenarios in hospitals [48].

An example for such an automated training tool was developed by Tiellet et al. [48]. They investigated how well veterinary students are able to practice surgery using only hypervideos (videos with enhanced features). They found that students who did not train with these hypervideos performed worse in the surgery than the ones who did use them. This specific example does not include AI or automation, but still serves as representative example since the hypervideos could also be replaced by AI simulations. Another experiment, that does include automation and AI, was conducted by Ahmad et al. [9]. They developed an algorithm that is able to produce feedback on a trainee's surgery performance. Then they investigated, how accurate this automatically generated feedback was compared to feedback by human experts. In the experiment, the trainees were asked to perform two different tasks. For one task, the automated feedback differed only 1% from the expert feedback, while for the second task it differed by 21%. This shows that the evaluation skills of an AI might not easily be able to keep up with human skills.

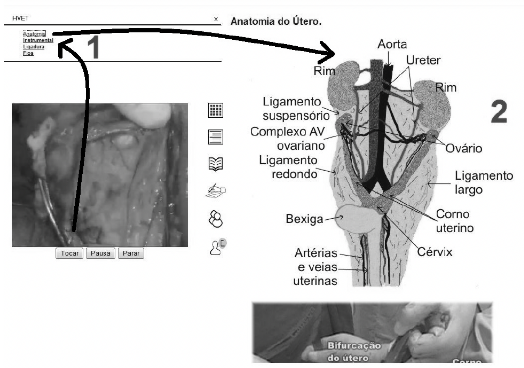


Figure 3.2: Hypervideoes giving automated feedback to veterinary students [48]

Education

Within the field of education, there is broad range of potential use cases for automated support systems. Pera and Ng [38] developed a recommender system that gives teachers book suggestions that match their students' state of knowledge and interests. They found, that the automatically suggested books were more accurate than the suggestions given by the teachers. This could be due to the fact that technology is much better at getting an overview over a lot of data and outputting the required info. However, they also point out that the recommender system still requires input by the teacher to provide some context and boundary information.

Another example of automation used in education is the feedback system built by Goncalves et al. [21]. In their setup, the students had to work on remote laboratory exercises. They developed a system developed that reported the errors the students made. This way, they got immediate feedback without any intervention by the teacher. This can be very beneficial, especially in an online learning setup. Moreover, the system might also reveal some of the students' misunderstandings or deficiencies, that would not have become apparent to the teacher otherwise. However, Goncalves et al. [21] also point out that the system lacked to ability to identify the type of error. In their experiment, 68% of the students' errors were caused by previous errors. The system did not recognize this correlation and thus gave a less precise feedback than a teacher would have.

The idea of automated feedback systems for students is not new - already in 2001 Odekirk-Hash and Zachary [37] published a study on a tool that automatically assesses students' performance in programming tasks. Their goal was to address the unbalanced teacher to students ratio. Interestingly, in their experiment, Odekirk-Hash and Zachary [37] found that students with access to the tool performed equally as the students without, but asked the teacher less frequently. Therefore, they concluded their automated feedback tool can help teachers overcome the one-to-many challenge and works well as supplementary support.

With the most recent developments, especially with the release of Chat Gpt, now support by AI is available for everyone, for teachers as well as for students. Besides, the opportunities and benefits, this also brings a lot of new question, for example how to assess work that students produced with the help of AI.

Sports

Similarly as automated feedback is used in medical training and education, it is also applied in sports. Trajkova [50] recently started a research on different AI-based video tools that evaluate the performance of ballet dancers. Interestingly, her goal is not only to support ballet trainers, as it was the case for most applications in medicine and education, but primarily to make ballet training more accessible. Ballet teachers can be expensive and their expertise was often exclusively passed on to them by previous generations. In her research project, Trajkova [50] wants to develop a tool that evaluates videos of ballet dancers. With this tool, everyone would be able to learn ballet, independent of whether they can afford or have access to a teacher, or not. Digitalizing and storing ballet knowledge this way also opens up many opportunities because everyone can contribute to it or reuse it. Nevertheless, the question remains whether feedback by such a tool is as precise and effective as by a teacher.

Figure 3.3: Automated feedback system for Ballet dancers [50]

Lee et al. [31] implemented a feedback automation system for therapists to give feedback on their patients' exercises. They chose a different approach as their idea is to get the best from both, automated and human knowledge. They combined a data-driven model with a model based on the therapist's feedback. By including the therapist's knowledge into the model, it is able give more personalized results. Lee et al. [31] found that this kind of hybrid model performs better than the data-driven model alone. This shows the importance of human knowledge despite the abilities of automation. Furthermore, they emphasize that the hybrid approach does not only improve the model's performance but also enables the therapist to understand and control the model.

The challenge is not only to make automated feedback as high-quality as human feedback, but also to bring it across in a comprehensible way. Tang et al. [46] explore different visualization options to present feedback on physiotherapy exercises. They reported that one of the main difficulties was to provide enough guidance while not overloading the user with too many details. Suggested solutions are to only show parts of the body or to use scaffolding. Besides giving understandable feedback, another challenge is to capture the patients as they are doing the exercises. Main issues with this are camera placement and depth capture. In short, when it comes to conveying feedback, human communication skills are difficult to replicate. This leads to the assumption, that even if feedback was not fully produced by humans, it might be beneficial to let them transmit the message.

Summary

The ways automation is applied for expert users in other fields has been researched because experts from different fields often face the same issues that automation can provide solutions to, for example the shortage of resources like time and staff or a restricted accessibility [54]. It has been shown that the use of automation to support experts was partly successful and partly revealed shortcomings. For instance, the automated system used by teachers that selects matching books for their students gave more accurate results than what the teachers would have been able to come up with [38]; the automated feedback for ballet dancers relieves ballet teachers and makes ballet dancing more accessible [50]. On the other side, an automated error report system for students was found to be less helpful than the teacher's feedback because it was not able to identify the type of errors [21]. Automated feedback systems for patients doing physiotherapy exercises did not perform as well as human therapists either because they lacked soft skills, like comprehensible communication [46], or differed a lot from the therapists' feedback [31]. Hence, from the investigated use cases of automation for experts, it is unclear whether automation brings an added value or not. Therefore, a separate use case, the automated training plan tool, will be investigated throughout this paper to be able to better specify the positive and negative impacts of automation and with this, its added value.

3.3 Interacting with Automation

Independent of the field automation is used in, it always changes our ways of working. Tasks that we used to perform ourselves, applying our skills and knowledge, are completed automatically. In a few cases, this means that we do not have to intervene anymore at all, but most of the times new tasks come up that require human qualities, such as controlling automated tools or giving them the right input [11]. Interacting with these tools can have various impacts: it can trigger emotions like fear or excitement [35] or it can have long-term influences, like the decrease of conscientiousness or expertise [54]. In the following, ways of interacting with automation is explored and what possible consequences for users, in particular for professionals, this can have.

Ways of Using Automation and AI

Wang et al. [54] did a research on what scientists think about integrating automated AI in their work process. They identified three ways, automated AI is considered as among the scientists: AutoAI can be a collaborator, a teacher or a standalone expert. In the first case, AutoAI takes over trivial parts of the expert's job, points out possible weaknesses of the expert's work or inspires them to take another approach. In the other two scenarios, the user does not have enough knowledge, so that AutoAI takes over the job partly, while teaching the user, or fully as standalone expert. Then, the problem is that users might not be able to interpret the result of the AutoAI or deploy it correctly in different situation. Furthermore, they would also not be able to recognize when the AutoAI produced nonsense. From this, Wang et al. [54] conclude that it is crucial to consider technology as augmentation of human skills, not as replacement.

Consequences and Reactions caused by Automation and AI

When users have enough expertise, they can better control and understand automated system, but they still face other challenges. Experts can lose some of their flexibility for example because they have to adhere to some predefined method [54]. Furthermore, even if only insignificant tasks are executed by automation, dealing with these tasks sometimes helps experts to get a deeper understanding of the task or they accidentally discover hidden correlations on the way.

The scientists' opinions on using automation and AI, that Wang et al. [54] stated in their paper are mixed. On the one side, they saw technology as chance to increase the expert's productivity because automation can be faster and more accurate. It can also be cheaper and beneficial for an expert team in general, because less specific expertise is required and a broader group of people can contribute to a task with the help of AI. On the other side, some expressed the concern that using these tools makes complex jobs accessible to anyone, also to those who are not qualified for it. Both sides agreed that human expertise is still needed to check and interpret the outcome of automated processes.

There are not only concerns about the impact on the outcome of the job, but also about the impact on the users themselves. In general, users tend to evaluate potential risks or threats before interacting with new technologies [51]. Automation and AI can be threatening in particular because of its complexity, scope and use of a lot of personal data [35]. In addition to that, there are the already described changes caused by automation and AI of unknown extent, like loss of control or even jobs or new working structures [20]. Common reactions to this are fear, stress or insecurity [35].

Preventing Negative Impacts of Automation and AI

In literature, there is a lot of research on factors that influence users to engage with a technology or not, known as the behavioural intentions to use' (BIU) [35]. These factors are essential for developing beneficial, human-centered technologies. Beneficial AI is a much used term in literature [12, 40] which means to put human safety and well-being over technological advances when developing new tools [41]. There are many approaches to identifying the factors influencing user behavior, for example the Technology Acceptance Model (TAM) [51] or the Unified Theory of Acceptance and Use of Technology (UTAUT) [52].

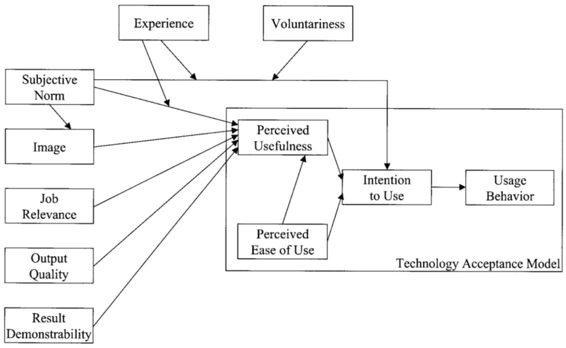


Figure 3.4: TAM model [51]

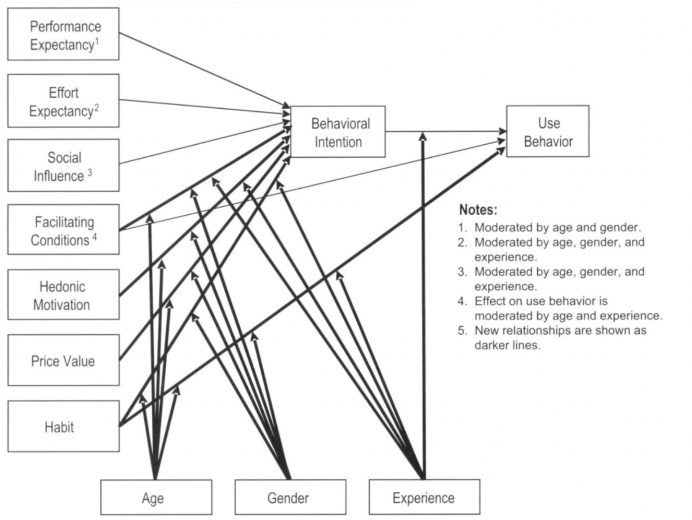


Figure 3.5: UTAUT model [52]

Both models include theories on when consumers are willing to use a technology or not. According to the TAM [51], this decision is based on the ease of use and usefulness of a system. The UTAUT model [52] has been developed more recently and involves additional aspects such as the consumer's social context or their affinity to innovation. If these factors are handled correctly, negative impacts on users by technology are limited and thus users are willing to engage with it.

Meyer-Waarden et al. [35] expand these models and also consider cognitive aspects (trust, security and privacy), affective aspects (social recognition, hedonism) and well-being. Interestingly, they found that users prefer secure, controllable technologies over highly advanced and fully automated systems. From this, they concluded that trust is one of the most important factors for users to be willing to accept new technologies. Also Wang et al. [54] state that trust is an important prerequisite for the interaction with automated systems. This brings up the question: when do users trust technology?

Trust in Automation and AI

According to Meyer-Waarden et al. [35], trust is established by three factors: transparency of how the system operates, a competent performance of the system and the ability to regain control over the system anytime. Transparency as essential factor is also emphasized in Wang et al.'s research [54] on the collaboration between scientists and AI. They concluded that for automated systems to be used successfully, they have to be transparent and explainable. Explainability as crucial requirement for the usage of automation and AI is reoccurring in many other sources [49, 44, 26]. The European Comission released guidelines for requirements of trustworthy AI where explainability is one of them [44]. Jacovi et al. [26] state explainability as cause of intrinsic trust, meaning that the user trusts a system because they can understand how it works. Explainability is also one of the four criteria that Toreini et al. [49] identified for technologies to be trustworthy. Their remaining three criteria are fairness (to avoid biased systems), auditability (to be able to monitor the operation of the system) and safety (to ensure that the system is robust against attacks by hackers).

The importance of safety to the user, is also highlighted by Meyer-Waarden [35]. They state that 93% of all european citizens have concerns about data theft or fraud. Accordingly, if users do not feel that their data is safe with an automated system, they do not trust it. Also among the guidelines by the EU commission for trustworthy AI, requirements like technical robustness, data governance and privacy are listed [44]. Mehrotra et al. [34] take a different approach to investigating what it takes to develop trustworthy technology. In their experiment, they let participants work with AI agents that acted according to different values. The result was that agents whose values were similar to the participants' ones were trusted more. This makes value similarity also a factor to consider when establishing trust in technology. A summary of all factors that can create trust among users is shown in Figure 3.6.

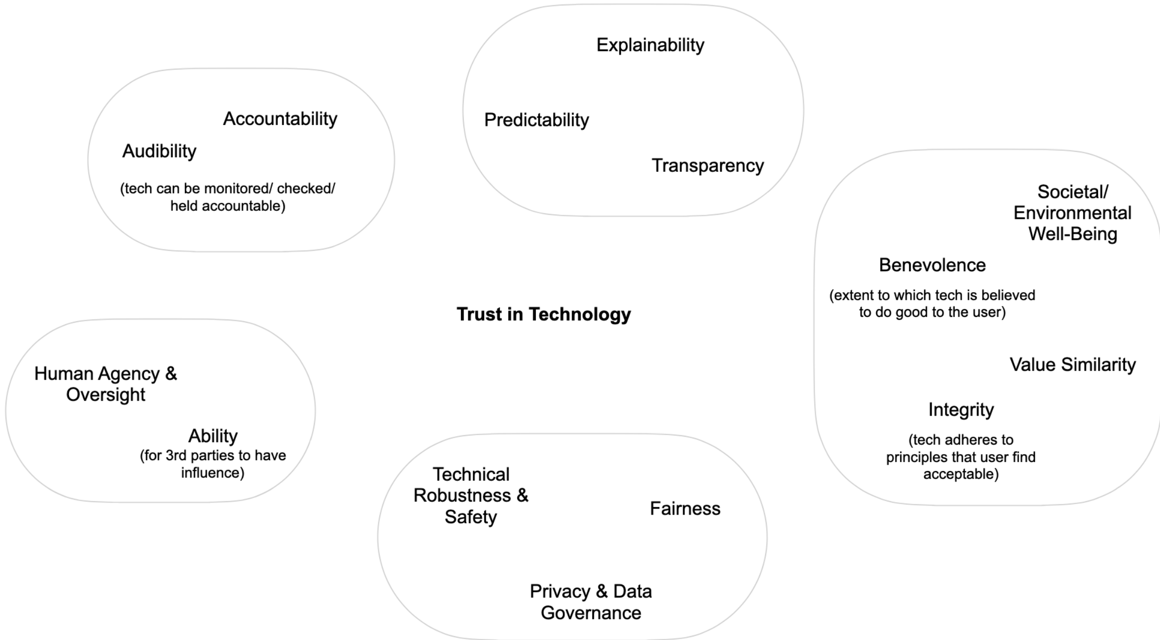


Figure 3.6: Factors establishing Trust in Technology [44, 34, 49, 26]

4 Concept

Before implementing a tool that automatically generates training plans, a short overview will be given of the status of the app into which the tool will be integrated. Furthermore, the context of the app will be specified by analyzing its users. From this, requirements for the tool will be derived which will guide the implementation process.

4.1 The Current System

4.1.1 Traindoo

The Traindoo app originated from the idea to provide a platform for trainers and their trainees to exchange training data and feedback quickly. Digitalizing coaching this way makes it more flexible, since it is time and location independent, and opens up additional possibilities, such as data tracking and analysis. The Traindoo app includes three main functionalities: planning a training week, tracking health and training data, and giving feedback. The app exists as a web version for trainers and as a mobile version for athletes.

In order to create a new training plan, the trainer can assemble it by adding or modifying exercises for each day of the upcoming week. The athlete sees the created plan in their mobile version. When performing the specified exercises, the athlete can enter the corresponding values, like load and repetitions, and add a comment or video. The trainer then checks the athlete's entries and gives feedback. There are two options for the trainer to set up a training plan: by starting from scratch with an empty plan or by copying and pasting the plan from the previous week. Either way, the trainer has to go through every training day and add or modify each set of each exercise manually. Parts of this process are shown in Figure 4.1.

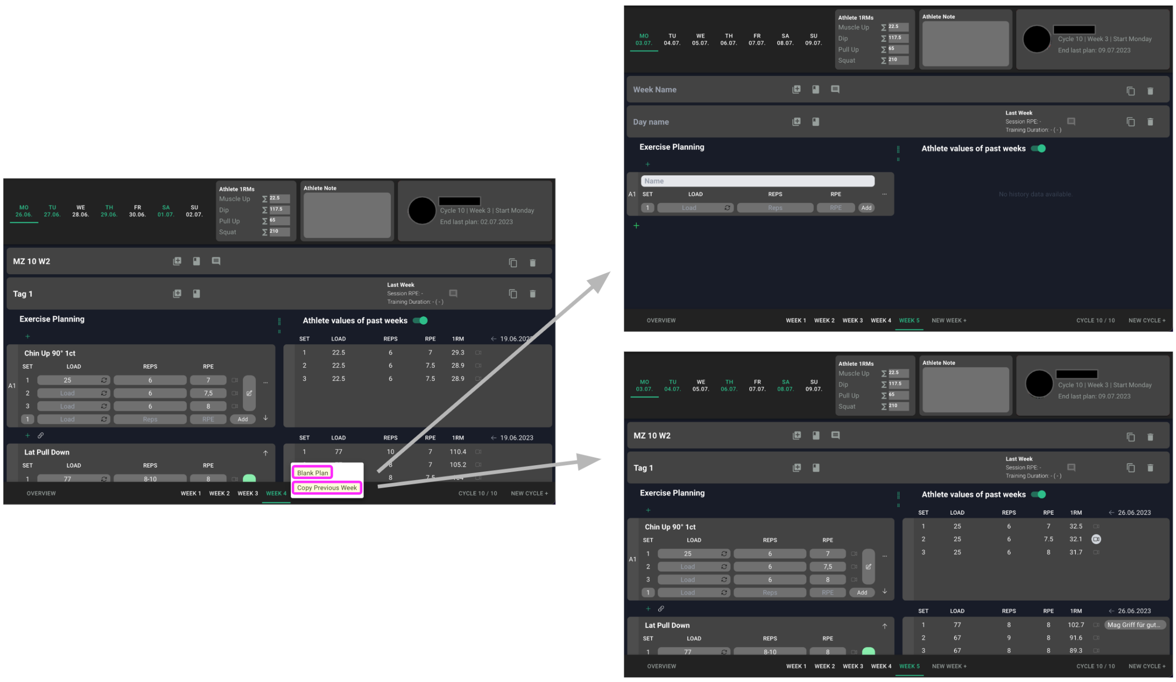


Figure 4.1: Creating a new training plan via the existing options 'Blank Plan' and 'Copy Previous Week'

4.1.2 User Analysis

The above described ways of creating a training plan can be very repetitive, especially since training plans often follow a certain pattern (see Chapter 2.2). Therefore, the idea is to automate the planning process; thereby, automation needs to be integrated in such a way, that it brings added value for the trainers. As concluded from literature, to achieve the maximal benefit, automation should be used as a tool to amplify human skills, rather than to replace them. Furthermore, it has been found, that the negative impacts can be reduced by developing human-centered technology. Thus, the user, in this case, the trainer, has to be put in focus. All users of the Traindoo app have been asked to answer an online survey aimed at collecting data on their background, ways of working, and challenges. This way, a better understanding of the user group can be obtained, which will guide the development of the automated training tool. The survey is divided into a general part about the trainers' background and a more specific part investigating their goals, habits, and challenges. The last part is aimed at determining the technology affinity among the trainers.

Background Information on the Trainers
In total, 48 trainers completed the survey. Most of them (one-third) are active in powerlifting, followed by bodybuilding and general fitness. The minority coaches calisthenics or works as physiotherapists. For around 17% of the participants, the type of sport is unknown (see Figure 4.2). Most trainers do not have more than ten trainees. Only three of the participants coach over 20 athletes. The rest teaches between 10 and 20 athletes or has an unknown number of trainees (see Figure 4.3).

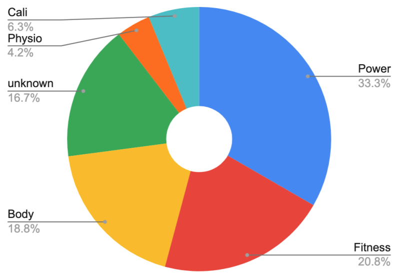


Figure 4.2: Distribution of the sports types among the trainers

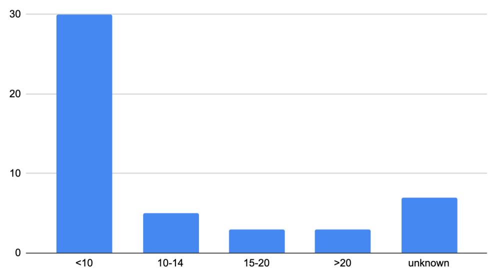


Figure 4.3: Number of trainees

The number of years, the participants have been active as trainers is pretty equally distributed between less than a year and more than five years (see Figure 4.4). This implies that the participants have various levels of experience. Also, the ways they got to this knowledge and experience are quite diverse. 50% acquired their trainer knowledge on their own initiative via self-study and research. Only one quarter followed a study course or similar education program in order to become a trainer. The remaining quarter learned from practical experience or role models (see Figure 4.5). What most have in common, however, is that they are trainers as part-time job. Only roughly 30% are full-time trainers.

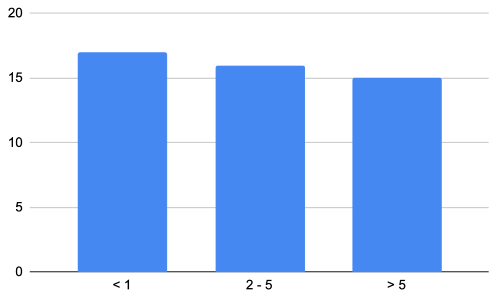


Figure 4.4: Number of years the participants have been active as trainer

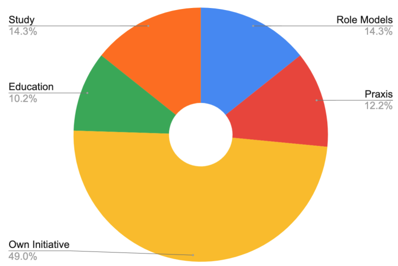


Figure 4.5: Sources where trainers got their knowledge from

The Trainers' Goals, Working Habits and Challenges

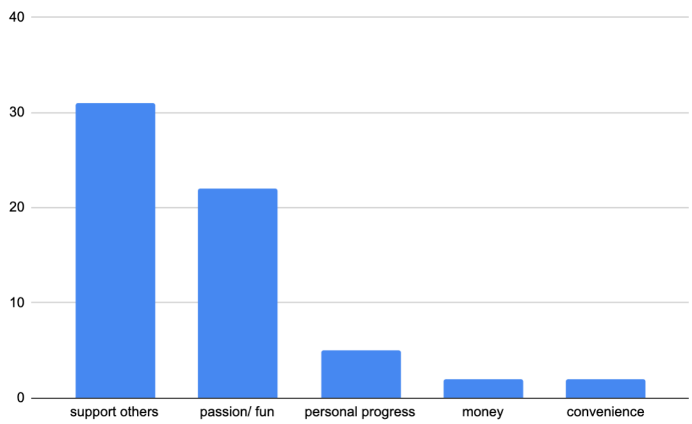


Figure 4.6: Reasons why the participants have become trainers

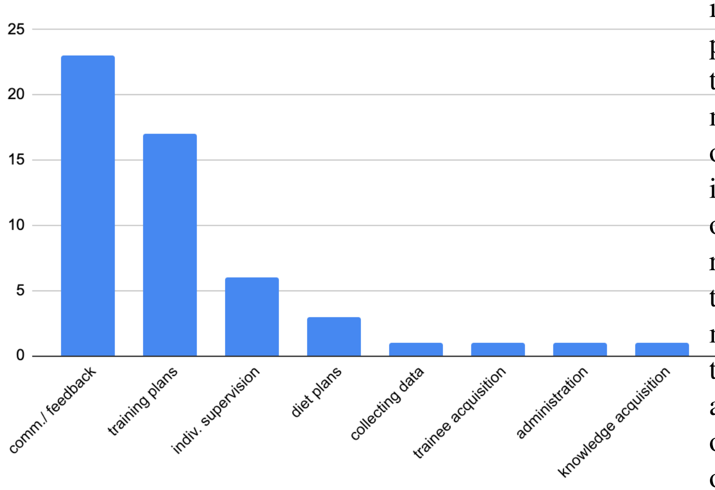


Figure 4.7: Tasks that trainers stated to be particularly time-demanding

To be able to better understand and analyze the trainers' user behavior, the survey also included more personal questions about their work as trainers. One question asked for the reasons for becoming a trainer (see Figure 4.6). Most participants answered that they want to support others and help them in reaching their goals or in discovering the benefits of sports. The second most common reason was the trainers' passion for the sport. A few also stated that being a trainer enables them to become more successful in their sport themselves and to personally grow. It is remarkable that only two trainers mentioned money as the main reason. Another two became trainers because it went well with their careers or existing job. Overall, the results show, that the majority of the participants are intrinsically motivated to be trainers and follow their job by conviction. This lets assume that they tend to put a lot of effort into it. A consequence of intrinsic motivation can be that people invest too many of their resources and end up running out of time or being overworked. To find out if this is the case for the trainers, they were asked what tasks of their job they spend a lot of time on. The answers show that the number one of the most time-consuming tasks is the communication and feedback between trainer and trainee. Number two is creating training plans. Training plans being the second most time-consuming factor leads to the assumption that automating the planning process could free up some of the trainers' time. The remaining tasks categorized by the trainers as particularly time-consuming are shown in Figure 4.7.

To find out more about the trainers' relation to their trainees, the survey also contained a question of whether trainers frequently interact with their trainees one-to-one. 94% answered yes; two of the ones who said no gave a shortage of time as a reason, and one said that there is no need for a one-to-one exchange. This result matches well with the fact that most trainers are intrinsically motivated and spend most time on communication and feedback, which has been found in the above-described answers.

Lastly, the survey was meant to find out what functionalities the trainers would like to have added to the Traindoo app. The answers include a variety of suggestions. Many of them concern the creation and visualization of training plans. For the visualization, a much-mentioned request is to have a clear and comprehensive overview of the training plan and related information, such as the training data of the previous week for example. When creating a new training plan, many would find it helpful to have template plans as orientation for adapting the training parameters. Additionally, multiple participants requested a feature showing progression patterns or prognoses when setting up a new plan. Such progression patterns seem to be relevant, especially for the load parameter: one participant would like the app to be able to calculate the precise load value automatically based on some basic input by the trainer, for example. Another one wrote that the app should recognize when the load has been changed by the athlete and take over the modified value. Besides templates and progression patterns, also the option to pre-plan a training was among the suggestions. In the following, some quotes from the trainers' answers are given:

> "I would like to have a view that not only shows me the values for the current week in the second half of the screen when creating a new training day, but also a variable selectable training unit from the past." (ID: eylPu)

> "If the athlete changes their weights, [...] that it [the app] takes over directly for the next training session." (ID: A62xH)

> "automatic set weight (coach enters percent of 1RM, app generates weight)" (ID: TNtRy)

Technology Affinity among the Trainers

Next to the trainers' background and working habits, another factor that influences the user behavior is their general opinion on and openness towards technology, known as technology affinity (TA). To measure the technology affinity, a method by Karrer et al. was used [29]. They identified four categories, each containing various TA statements, that the user is asked to confirm or deny: enthusiasm, competence with technological devices, and positive and negative attitude towards technology.

Based on Karrer et al.'s method, the survey included the following six phrases that the trainers were asked to disagree or agree with on a scale from 1 to 5 (1 being complete disagreement, 5 being complete agreement).

1. I am enthusiastic about new technology trends. [Enthusiasm]
2. I know most functions of my digital devices. [Competence]

3. Digitalization and automation facilitate our daily lives.

4. Digitalization and automation help to retrieve information. [Positive Attitude]

5. Digitalization and automation put many jobs at risk.

6. Digitalization and automation lead to intellectual impoverishment. [Negative Attitude]

The amount of (dis-)agreement among the trainers for each of the TA statements is shown in the following figures. Interestingly, the statements that are positively associated with technology got overall approval (Figures 4.10 and 4.11). For the negative statements, the opinions are more spread (Figures 4.12 and 4.13): most participants were unsure whether digitalization and automation put our jobs at risk, and the ones who confirmed this statement were as many as the ones who denied it. This uncertainty also becomes visible in the answers to the claim that digitalization and automation cause intellectual impoverishment; however, this statement was slightly more rejected than agreed with. The different answer distributions for the positive and negative statements are remarkable because, with the uniform agreement to the positive phrases, one could expect an equally uniform disagreement with the negative phrases, which is not the case. This could imply insecurity and a lack of knowledge regarding technology among the trainers. The fact that the trainers do not feel completely confident using technology is also shown in the answers to the competence statement. Here, the approval is slightly more moderate than for the enthusiasm about technology. Despite this insecurity, the trainers' enthusiasm for technology is high, as Figure 4.8 shows.

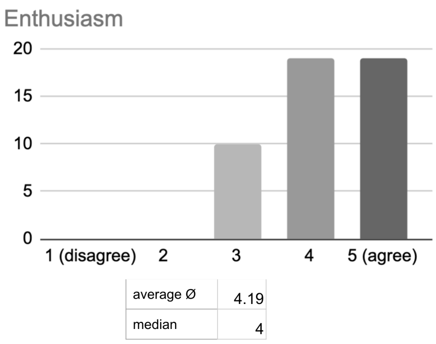


Figure 4.8: Answer Distribution of the 1st TA statement

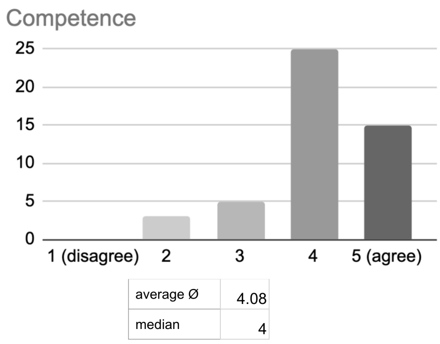


Figure 4.9: Answer Distribution of the 2nd TA statement

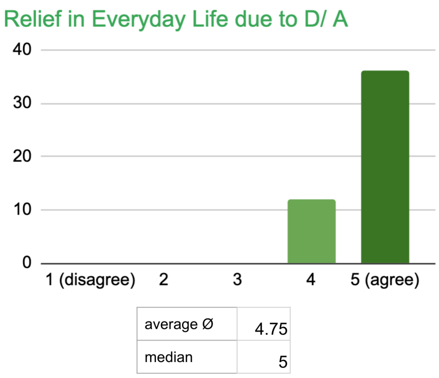


Figure 4.10: Answer Distribution of the 3rd TA statement

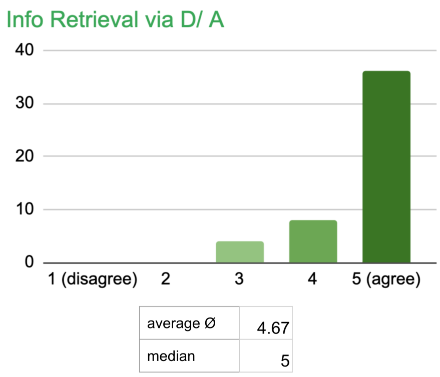


Figure 4.11: Answer Distribution of the 4th TA statement

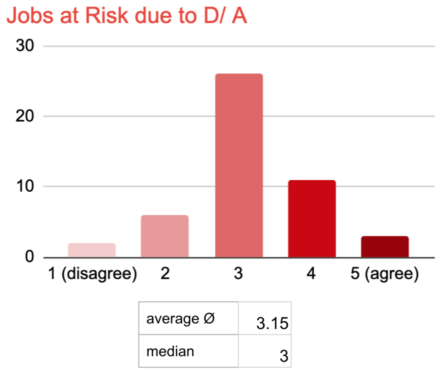


Figure 4.12: Answer Distribution of the 5th TA statement

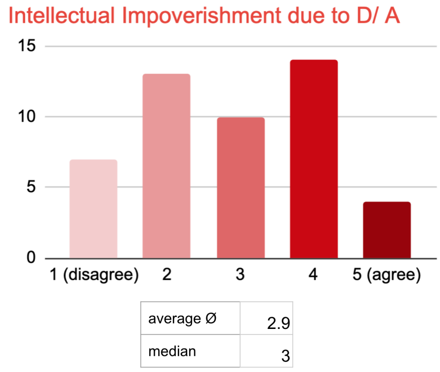


Figure 4.13: Answer Distribution of the 6th TA statement

4.1.3 Requirements

The usefulness of an automated training plan tool has been validated by the trainers' answers to the online survey. They show that most trainers are intrinsically motivated to follow their job and hence are likely to invest a lot of time and effort into it. The tasks that cost them the most time are communication and feedback and planning their athletes' training schedules. Automating the creation of training plans could speed up the process. Furthermore, trainers explicitly demanded more support in creating new training plans. According to their requests, an automated training plan tool should fulfill the following requirements:

R1: show a clear and complete overview of all relevant information
The data for the new training plan and all required data from previous plans is shown collectively in a structured way.

R2: adapt variable parameters of the new training plan automatically
Variable parameters of the new training plan, such as the load value, are adjusted automatically according to the trainer's or athlete's input.

R3: minimize the number of steps or clicks the trainer has to follow for creating a new training plan in order to speed up the process and make the trainer more productive

Other requirements that an automated training plan tool should fulfill can be derived from the research results of related work described in Chapter 3. The difficulty in developing a good automated system lies in the trade-off between simplifying and speeding up a process without losing quality, flexibility, or any of the trainer's valuable skills. Creating a new training plan requires knowledge, which is why it is important that the trainer maintain the ability to include their expertise and to check and adapt the output of the system. The other challenge is to counteract ungrounded negative bias towards automated systems among trainers. Many negative reactions can be prevented by creating trust in automated systems. Systems can be trusted if they fulfill certain criteria, such as transparency, explainability, safety, etc. From these research findings, the following additional requirements for an automated training plan tool can be derived:

R4: maintain the trainer's authority and flexibility
The trainer has control over the automatically generated training plan and can check and change the tool's output according to their expertise.

R5: maintain the quality of the trainer's work
Automatically generated training plans are as qualitative and accurate is manually created plans.

R6: make the output of the tool transparent, explainable, robust and conforming to the trainers' norm
The tool is comprehensive and fulfills all other criteria of a trustworthy system to establish trust and that way prevent negative impacts on the trainer.

4.2 The Proposed System

4.2.1 An Automated Training Plan Generator

The previously described findings from literature, related work, and the survey determine, what a system, that automatically generates training plans should look like. They boil down to two main requirements: the system should facilitate and speed up the process of creating training plans and the trainer should at the same time maintain the authority over the process. The need to support the planning process was identified from the results of the survey which showed that training planning is among the most time-consuming tasks of a trainer. Furthermore, it is a repetitive task, which automation is well-suited for. In order to provide this support, the system should be able to adapt the training parameters of the new plan automatically. As explained in Chapter 2.2, the modification of parameters from plan to plan is called progression. Therefore, the system will be called Progression Tool. To facilitate the planning process further, the number of clicks and user actions, the trainer has to perform to achieve a new training plan, should be reduced.

Besides the optimization of the planning process, the second main requirement for the Progression Tool is that despite the use of automation, the trainer maintains their authority and flexibility. This is important for multiple reasons; as found in literature, it not only ensures a better quality of the outcome but is also necessary for the user to establish trust in the system. This aspect is especially relevant in this context because as found in the survey, most trainers follow their job out of passion and with the purpose to help others. Therefore, they are unlikely to use a system, that they do not fully approve and understand. Accordingly, the Progression Tool needs to have a clear interface with input fields for the trainer to apply their expertise and possibly make changes to the tool's output.

Before designing the interface, the functionality of the tool has to be determined. The Progression Tool needs a logic, based on which it is able to automatically generate a training plan. To ensure that the generated plan is individually adapted to the trainee it is made for, several factors have to be taken into consideration, such as the trainee's training history and demographic data.

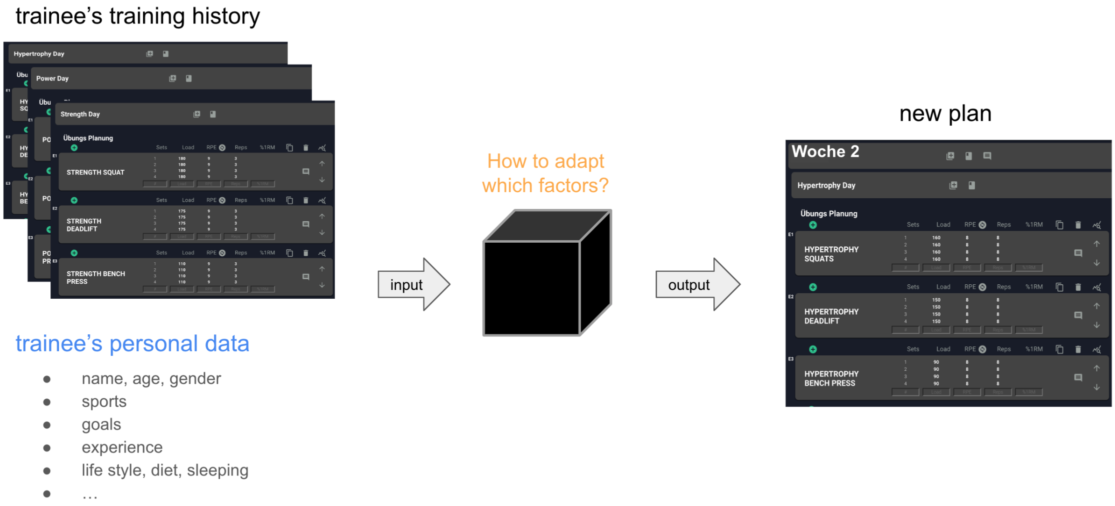


Figure 4.14: Input information relevant for the planning logic

With this information as input, the new training plan has to be calculated. Calculating a new plan means adjusting all training parameters so that the trainee is most likely to achieve their training goal. There is a large number of parameters that could potentially be calculated (highlighted in Figure 4.15).

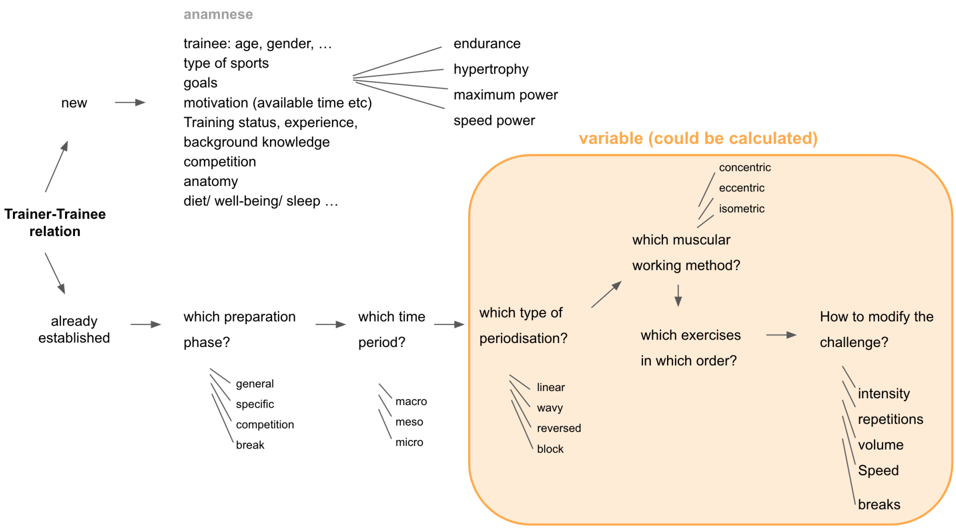


Figure 4.15: Variable parameters within the process of creating a training plan

For better feasibility, only specific variables are considered and the remaining conditions are taken as predefined. As found during the analysis of existing planning logic and in the answers of the trainers to the survey, one of the most frequently modified parameters are the intensity, determined by the load moved during an exercise, and the volume, which is defined by the number of repetitions of an exercise. Therefore, the logic of the Progression Tool is focused on adapting the load of the exercises of the new training plan, corresponding to the trainee's individual training history. Consequently, if there is no training history available, the logic cannot be applied. A more detailed description of the logic behind the Progression Tool and how the new loads are calculated is given in Chapter 5.3.

The Progression Tool is not only required to generate the new plan automatically but also to leave the trainer enough opportunity to give their own input. At the same time, the number of clicks and user interactions should be kept minimal to make the flow through the planning process as fast and intuitive as possible. Based on these aspects, the following user interface was designed for the Progression Tool:

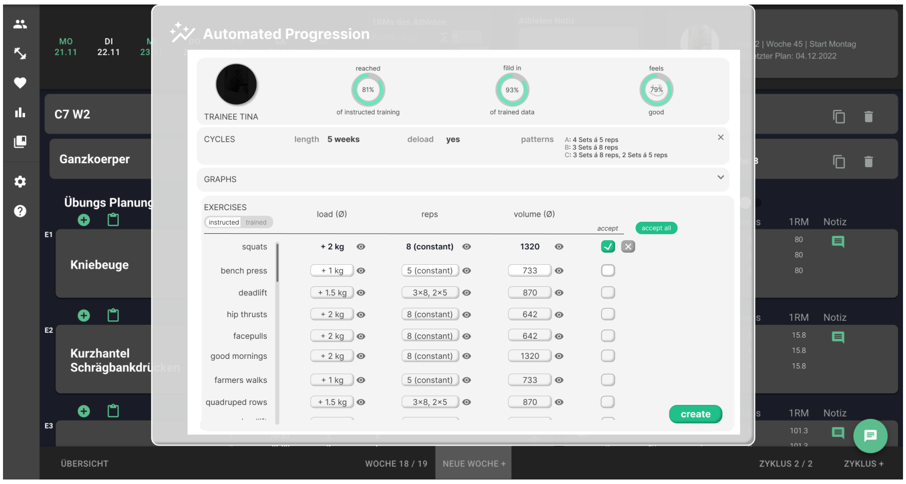


Figure 4.16: Prototype of the user interface of the Progression Tool

For every exercise, the value by which each parameter should be changed is shown. In the first row, for example, the tool suggests increasing the load of the 'squat' exercise by two kilograms and keeping the number of repetitions at eight. If the trainer approves this suggestion, they can click the check box on the right. Otherwise, they can click on the corresponding field and change the values. The trainer is only able to create a new plan, after all exercises have been approved, i.e. after all boxes have been checked. There is also an 'accept all' button that marks all exercises as approved at once. This option is meant to speed up the process on the one side and on the other side will give insight later into how fast trainers are to accept and trust the tool's suggestions. The purpose of the checkboxes is to require the trainers' expertise and approval so that the tool cannot produce an output that has not been reviewed by the trainer. This guarantees that the trainer remains in control over the system as this has been described earlier being one of the criteria that makes an automated system user-friendly and trustworthy. In addition to that, the trainer needs to be able to change the tool's output, which is also given via the input fields. With the list structure that shows all exercises on top of each other, the trainer can view everything at once without needing to click anywhere else (as was the case previously, where exercises were not shown as a list, but split up between days). The goal of this is to provide a clear overview and reduce the number of clicks. Above the list of exercises, there are graphics showing additional information shown about the trainee and their progress over time, which is meant to quickly give the trainer an idea of the status of the trainee.

4.2.2 Hypotheses

The above-proposed tool, representing an example of how automation can be adapted, will be used to investigate the following hypotheses.

Hypotheses on positive impacts of automation

I. Automation helps trainers to save time.

II. Automation increases the trainers' productivity and quality of work.

Hypotheses on negative impacts of automation

III. Automation restricts the integration of human expertise.

IV. Automation creates stress or skepticism among the trainers.

V. Automation causes a lack of motivation and attentiveness among the trainers.

5 System Design

In the previous chapter, the focus was on the functionality of the Progression Tool itself and the requirements it should fulfill. This chapter is about describing how the tool is applied in practice. Therefore, an exemplary scenario is presented of how the tool is intended to be used. This scenario is then generalized in a use case diagram.

5.1 Scenarios and Use Cases

Exemplary Use Case Scenario

Tina is a trainer in powerlifting and supervises fourteen athletes. She uses an app called Traindoo to track her trainees' progress and to create training plans for them. Every trainee needs a new plan every week, which is why on Sundays, Tina usually spends the whole evening setting up plans for all her athletes for the new week. In the Traindoo app, Tina can create a training plan from scratch or using the plan of the previous week as a template. This means, she either has to completely re-fill the plan with the correct exercises and values or she has to adapt the old exercise values from the previous week. This costs a lot of time and additionally, is very repetitive since the adaptions follow a certain progression pattern. Furthermore, she has to click back and forth a lot in order to recall each trainee's individual training performance of the previous weeks to determine whether to increase or decrease the intensity of an exercise and by how much. Traindoo recently released a tool that automatically suggests a plan for the upcoming week. Instead of modifying every value of every exercise by hand, Tina can look over the tool's suggested exercise values and accept them individually or collectively, by checking the single accept boxes or by clicking the accept all button. If, for a specific exercise, Tina thinks that the tool's suggestion is not adequate, she can modify it manually. Besides, the tool shows the athlete's progress status on each of the exercises so that Tina does not have to click through passed training weeks anymore to review their performance.

The process of using the Progression Tool, which is exemplarily described in the above scenario, is generalized in the following use case diagram (see Figure 5.1). The use case diagram is meant to clarify all possible actions the trainer can take when using the tool and what they will trigger within the functionality of the tool. Furthermore, it points out how its functionality fulfills which of the requirements listed in Chapter 4.1.3.

After the trainer has started the planning process with the Progression Tool, it pulls the trainee's training data from all previous weeks. Based on this data, the tool calculates how much each exercise of the plan should be changed for the upcoming week. The exercises with the adapted values and by how much they have been changed are then shown in the interface of the tool (R2, R1), where the trainer can check them and approve or modify them (R4). Depending on the trainer's interactions, the calculated data and the status of the plan are updated. The tool needs to keep track of the approval status of the plan because it can only allow the trainer to finalize the plan generation after he or she has verified all exercises (R4). Once this is the case, the trainer can trigger the generation of the plan meaning that the tool takes all calculated and modified values and puts them into a new training plan, which is then saved to the database and displayed to the trainer.

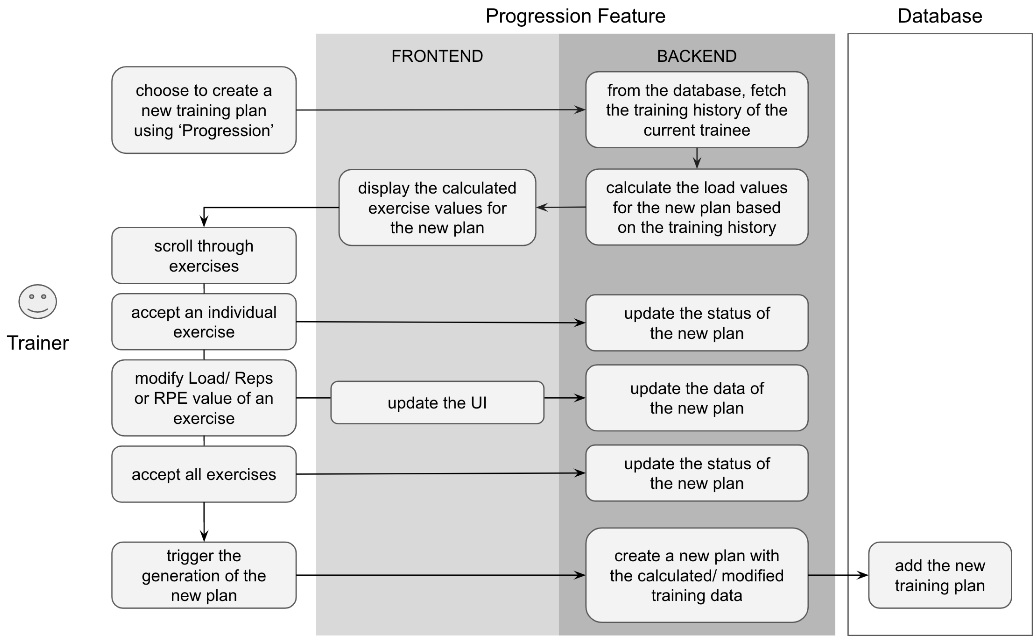


Figure 5.1: Use Case Diagram of the Progression Tool

5.2 UI and User Flow Model

The user flow model presents the above-described user interactions and corresponding functionalities of the tool from a user perspective. This means it shows the UI of the tool and how the user navigates through different views of the UI (see Figure 5.2).

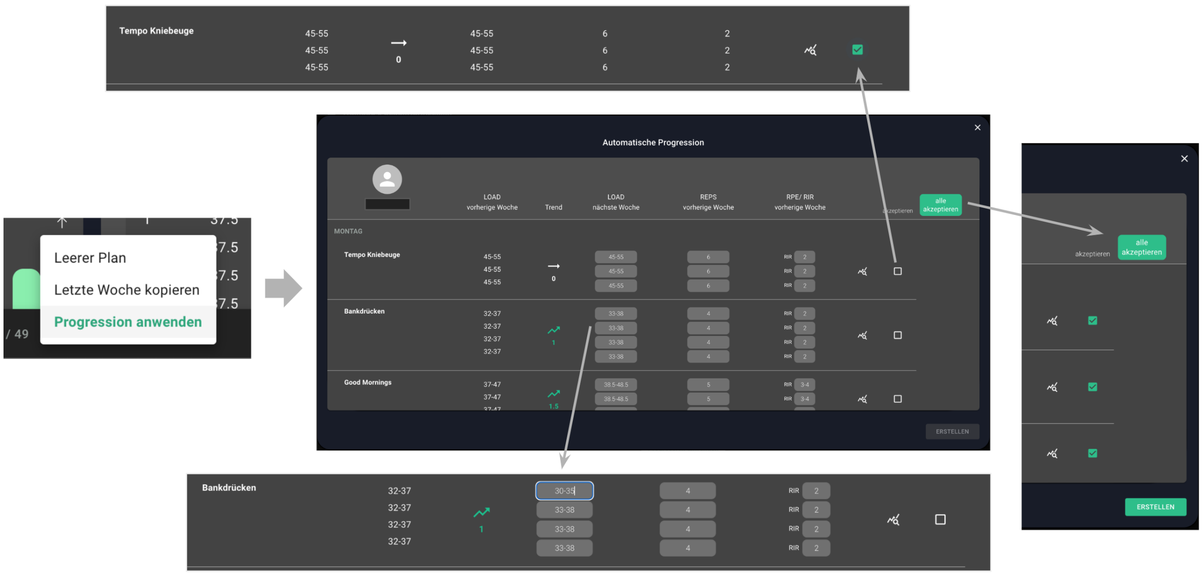


Figure 5.2: User flow through the Progression Tool

5.3 Implementation

Before implementing the Progression Tool, the prototype shown in Figure 4.16 was discussed with some trainers using the Traindoo app. Based on their input, some changes were made to the prototype. One of their feedback was to also show the exercise values of the previous week, not only by how much they are suggested to be changed. Furthermore, the trainers found it more useful to have an indication of the athlete's progress for each exercise separately instead of a general progress overview as it was planned for the top part of the interface. A comparison between the interface draft and the implemented interface can be seen in Figure 5.3.

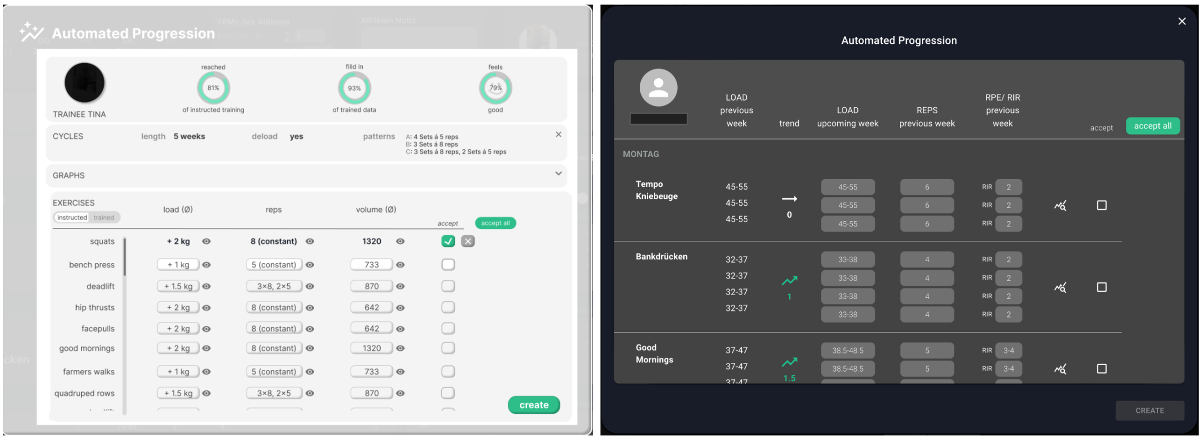


Figure 5.3: Prototype vs. Implementation of the Progression Tool

The Progression Tool was programmed in React Typescript and integrated into the existing Traindoo web application. The app's data is stored in a firestore database, which the tool accesses to calculate the progression data and generate a new training plan. More particularly, it draws all training weeks of the selected athlete and extracts the data on how they performed on each exercise, i.e. the amount of weight (load) they used for each exercise. In order to determine the load values, the athlete should train within the upcoming week, the average is calculated by which the athlete has increased the weight from week to week for each exercise. For this, first, the training weeks which deviate from the overall training progress, such as weeks in which the athlete was sick, on holiday, or on a break, are identified as outliers and removed from the calculation. Then, the load increase from week to week is extracted and summed up to be divided by the total number of training weeks. This is done for each exercise and the resulting increase values are the amounts by which the loads of the previous week should change. In case the previous week does not include training data, the last week for which training data is available is identified. The increases are added to the loads of the most recent training and displayed as suggestions in the tool's interface (see Figure 5.4).

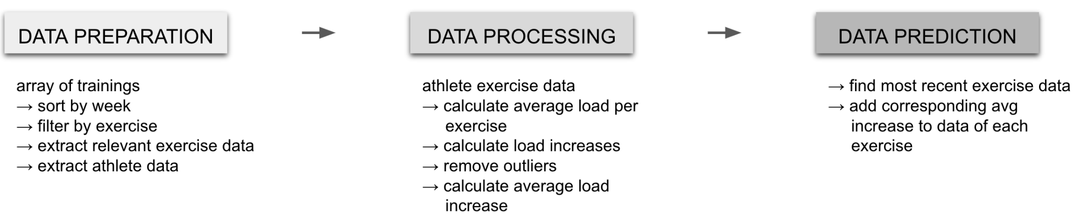


Figure 5.4: Steps through the calculation of the exercise values for the new training plan

6 Study

The purpose of the study is to analyze the trainers' user behavior with the Progression Tool to determine the possible effects of automation. The results of the study can then be used to answer the research questions and verify or reject the hypotheses. As stated in the hypotheses, the effects of automation on the trainers can be potential benefits, such as a reduced amount of time and effort. On the other side, trainers can also be affected negatively by using the Progression Tool; it might lower their level of accuracy and hence the quality of their work or it could affect the trainers themselves by making them feel more stressed or critical for example. Tracking the trainers' user behavior gives insights into which of these assumptions are most likely to be true.

6.1 Study Method

6.1.1 Procedure

For the study, the Progression Tool was unlocked for selected trainers. These trainers have been introduced to the tool and the study in beforehand via an app notification. They were explained how the Progression Tool works in a short click-through tutorial. Afterward, they were asked if they are willing to test the tool and agree to the conditions of the study. The conditions of the study were the following: over the course of six weeks, data on the usage of the Progression Tool is tracked anonymously; halfway through the testing phase, the trainers are asked to give a quick update on their experience with the tool, by giving a rating and open feedback. At the end of the six weeks, they are given a questionnaire about the usability of the tool [15]. Upon the trainers' approval of these conditions, the Progression Tool was unlocked and the tracking started. In summary, the study consisted of the following parts:

1. Participant acquisition via app notifications
2. Tracking of the trainers' usage of the tool
3. Interim rating and feedback
4. Questionnaire on the trainers' user experience with the tool

6.1.2 Participants

Among all trainers using the Traindoo app, the 48 participants of the survey were qualified to take part in the study because only for them the necessary survey data was available. The survey data includes information about the trainers' educational background, ways of working, sports field, and technical affinity as described in chapter 4.1.2. This data is needed to be able to analyze the data tracked during the study. Hence, 48 trainers received a notification via the Traindoo app, that introduced them to the Progression Tool and the study. 26 of them agreed to take part in it. The majority of them (35%) were trainers in powerlifting, 23% were bodybuilders, and 15% fitness coaches. A very small percentage had a calisthenics background and also very few were active as physiotherapists. The remaining trainers' sports background was unknown. Regarding their educational background, most of them acquired their trainer knowledge via their own research, practical experience, or role models (77%). 15% studied and 8% followed some other form of education. Furthermore, most had been a trainer for 2 to 5 years (46%). The number of trainers who had more than five years of experience was 31% and the rest had only recently become a trainer. The majority was trainer as half-time job (65%) and had less than ten trainees (69%). Only 12% of the participating trainers were female.

6.2 Data Tracking

In order to track the trainers' user behavior with the Progression Tool, time stamps were saved in the database of the app, whenever they performed a specific user action. For user actions that included user input, the value of that input was also saved. To be able to interpret the usage of the Progression Tool better, both the usage of the tool and the usage of existing methods to create a training plan were tracked (Figures 6.3 and 6.4).

tracked user actions	saved data	→ information required for user behavior analysis
1) start session 'empty week'/ 'copy week'	timestamp	How long does it take to create a training plan?
2) 3) register action 'added exercise'/ 'entered exercise title'	timestamp	How many/ what type of user actions did trainers perform during the planning?
4) register input 'modified load/ reps/ RIR'	timestamp, value	How many/ what type of user actions did trainers perform during the planning?
5) 6) register action 'added new set'/ 'added comment'	timestamp	How many/ what type of user actions did trainers perform during the planning?

Figure 6.1: Saved data for each tracked user action with the existing options (see Figure 6.3) and what info it is meant to provide

tracked user actions	saved data	→ information required for user behavior analysis
1) start session 'progression'	timestamp	How long does it take to create a training plan?
2) register input 'modified load/ reps/ RIR'	timestamp, old value, new value	How often/ by how much are the suggested values modified?
3) 4) register action 'accept clicked'/ 'accept all clicked'	timestamp	How often/ when are the accept buttons clicked?
5) register action 'inactive'	timestamp, value	For how long did trainers actively use the tool?
6) end session 'create'/ 'close'	timestamp	How often did trainers finish a plan? How often did they drop out?

Figure 6.2: Saved data for each tracked user action with the Progression Tool (see Figure 6.4) and what info it is meant to provide

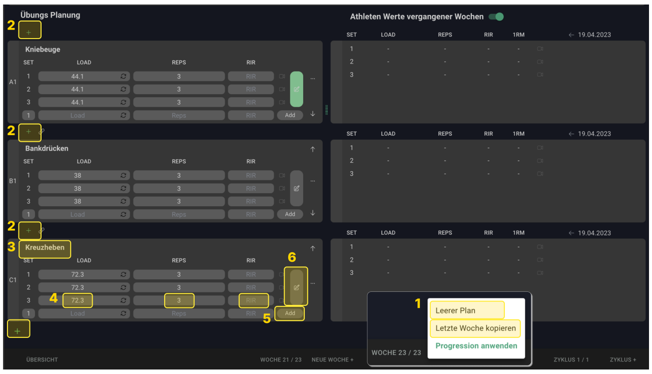


Figure 6.3: Tracked user actions for the existing options

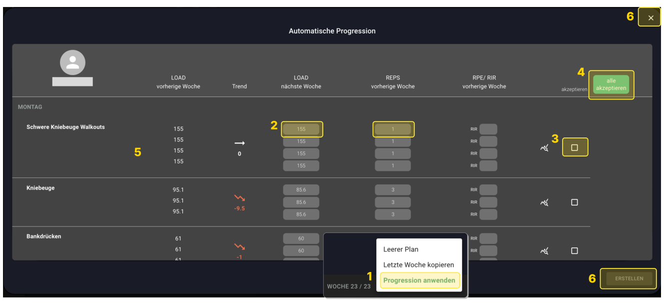


Figure 6.4: Tracked user actions for the Progression Tool

Only user actions relevant to the research questions and hypotheses were tracked. The first hypothesis states that by using the Progression Tool, trainers are more productive and save time. This can be verified or denied by comparing the time trainers need to set up a training plan using the tool to the time they would usually need without it. The second hypothesis assumes that automation does not have any negative impacts, neither on the trainers themselves nor on their performance. A possible negative effect of automation is that instead of supporting them, it makes users more stressed because they have the feeling that they need to double-check everything, in other words, they do not trust it. This assumption mostly requires qualitative data. However, there are also quantitative data that could provide useful information, such as the number of times the automated values have been adapted by the trainers and by how much. A large number of manual changes might imply that the tool does not offer noticeable support. Furthermore, again the time needed to create a plan with the tool could indicate whether trainers are more likely to feel stressed and skeptical about using it. For a more accurate representation of the trainers' emotions and attitudes toward the tool, qualitative data from the Usability Survey and the trainers' tech affinity score will be considered.

Another potential negative effect of automation on trainers is that it makes them less attentive or motivated which causes the quality of their work to go down. This is again a quite subjective assumption. Still, a trend can be detected by analyzing the trainers' interactions with the tool. The 'accept all' button, for example, can be an indicator for how carefully trainers check the tool's suggestions. More specifically, the 'accept all' button being clicked shortly after starting the planning process, without any value adaptions or single 'accept' buttons clicks makes it likely that the trainer approved the tool's output without double-checking it. This in return, could be a sign of a lack of attentiveness or motivation. Based on these considerations, it was determined which user actions should be tracked. An overview of the tracked data and which information it is meant to provide for the user behavior analysis is given in Tables 6.1 and 6.2. Each user action in the tables has a number that corresponds to the numbers depicted in Figure 6.3 and 6.4 to clarify where the user actions take place in the interface.

6.3 Interim Feedback and Questionnaire on Usability and User Experience

During the study phase, the participants received an app notification that asked them about their opinion so far on the Progression Tool. They could rate the tool with stars, from 'not helpful at all' (one star) to 'very helpful' (five stars). Afterward, they were asked to explain their rating and also had the option to give additional feedback in an open text field (see Figure 6.5).

After the six weeks of testing the Progression Tool, the trainers were asked to fill out a questionnaire on their user experience with the tool. The questionnaire was sent to them via an app notification and contained ten statements on various usability aspects. The statements were based on the System Usability Score (SUS) Questionnaire, which the trainers could disagree or agree with on a scale from one to five. This way, a score was obtained that indicates the degree of usability of the tool. Furthermore, the questionnaire included four open questions about the trainers' opinions on the tool. With the results of the questionnaire, the goal was to detect facts and correlations that could not be caught by quantitative data tracking. These results play an important role because they give insight into the trainers' subjective perception of the tool, which is needed to answer the research questions.

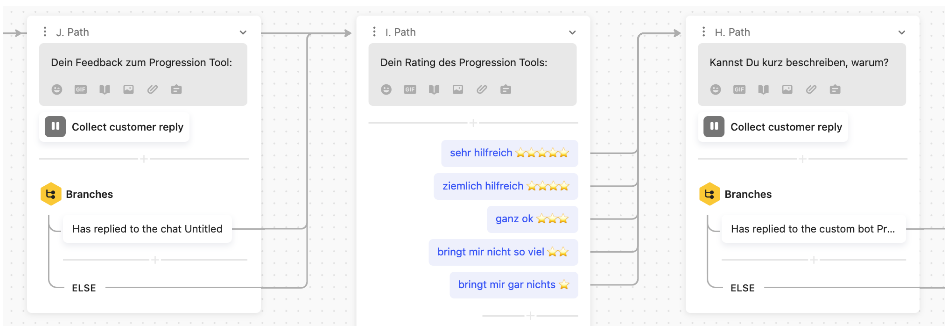


Figure 6.5: Feedback collection via app notifications

7 Results

7.1 Quantitative Results

In the following, it will be described how the data that has been tracked during the study of six weeks was processed and visualized.

How was the tool used over time?

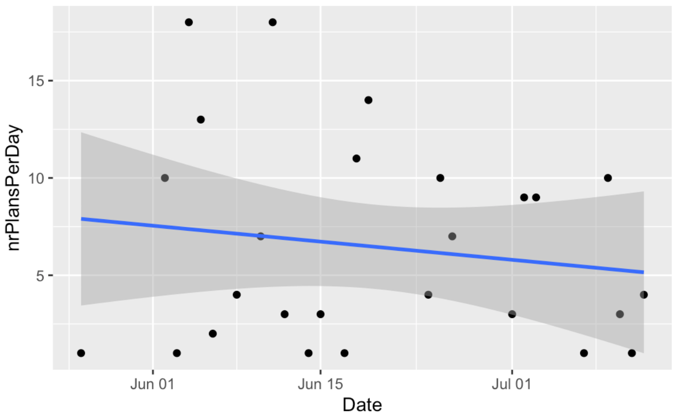


Figure 7.1: Usage of the Progression Tool over time

The first aspect, that has been considered was, in general, how often the trainers used the Progression throughout the study phase. These basic findings help to put the following results into the right context. Figure 7.1 shows, how frequently the tool was used on each day. One point represents the number of training plans that all trainers created on that day, using the Progression Tool. This number varies a lot from day to day, but overall the trend is decreasing as the blue line shows. Reasons for this could be the novelty effect of the tool in the beginning or the trainers being drawn back to the old methods of creating a training plan or that they were simply not happy with the tool.

However, when looking at the usage of the tool for each trainer individually, it becomes apparent that some trainers actually used the tool more frequently over time (see Figure 7.2). This suggests that the overall decrease does not come from a general dissatisfaction with the tool, but that the tool was more useful in some circumstances than in others. These circumstances will be identified in the following.

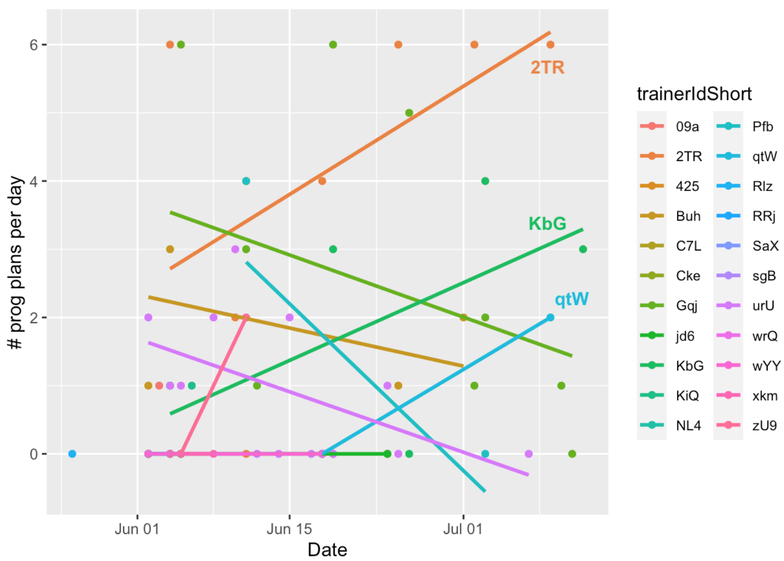


Figure 7.2: Usage of the Progression Tool over time per trainer

How much time and how many user actions did trainers spend on the tool?

As described in Chapter 4.2.1, the Progression Tool suggests how the exercises of the plan for the upcoming week should be adapted. The trainer has the option to accept all these suggestions at once, to accept each suggestion individually, or to modify the suggested values. Each modification and click on an 'accept' or 'accept all' button was tracked as user event. For each trainer, the number of user events that occurred during the usage of the tool was counted and divided by the number of progression training plans they created. This gave the average number of actions each trainer performed when using the tool, shown in light green in Figure 7.4. It changes a lot from trainer to trainer, ranging from almost no interactions to more the 120 actions on average per plan.

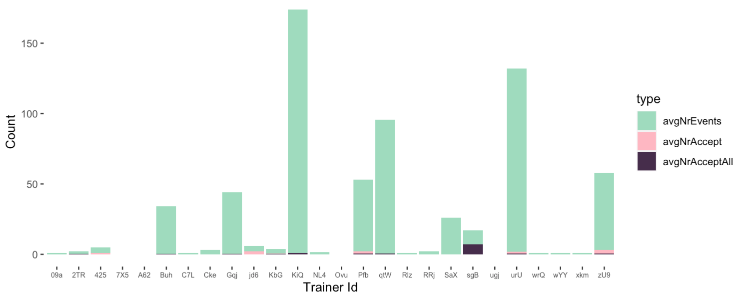


Figure 7.4: Average frequency of the different types of user actions with the progression tool per trainer

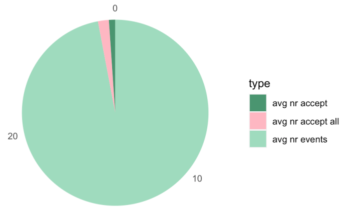


Figure 7.3: Average frequency of the different types of user actions with the progression tool overall

The amount of time spent by each trainer using the tool is, similar to the number of actions, quite inconsistent. This becomes apparent when comparing the bars of Figures 7.4 and 7.5, which shows the median amount of time each trainer spent on the tool per plan. However, the time values deviate less from trainer to trainer (standard deviation = 1.55) than the number of user actions (standard deviation = 47.09). Accordingly, some trainers had remarkably many more user actions than others while the time they spent on the tool did not differ as much. Also, the trainer who spent the most time on the tool is not the trainer with the maximum user actions. This observation needs to be analysed further to find possible reasons why the number of user actions of some trainers is so high (see 8.1). Moreover, it is interesting to compare the number of actions and the time each trainer spent on the tool to how frequently they used it (Figure 7.6). Here, the relation is the other way around: the trainer who used the tool most frequently has spent very little time and user actions on it. And the trainer who performed most actions used

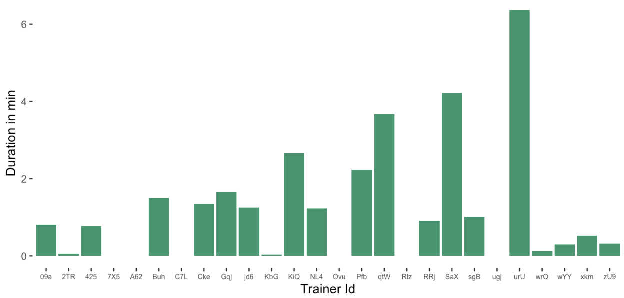


Figure 7.5: Median time each trainer spent on using the Progression Tool (per plan in minutes)

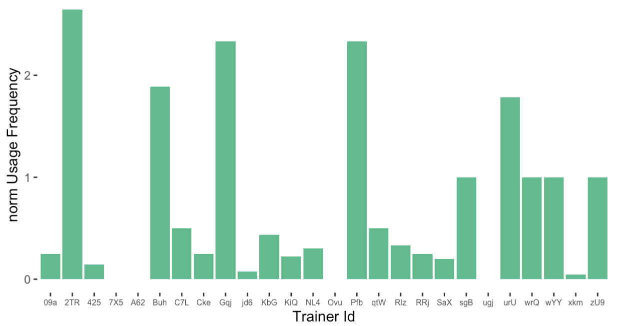


Figure 7.6: Overall number of times each trainer used the Progression Tool (normalized by their number of trainees

it rather rarely. However, this relation cannot be generalized as there are also trainers with both, a high frequency of usage and a high number of user actions and amount of time. The t-tests confirm this observation (see Table 7.1). They were done splitting the trainers into a group with large and a group with low time values (number of user actions) [32]. The trainers were split at the median. Then, a two-sample t-test was run in R, to check whether the trainers who spent less time (user actions), used the tool significantly more. The results are presented in Table 7.1.

H 0	p-value	outcome of t-test
Trainers who spent less time on the tool, used it more.	0.3125	p > alpha (0.05) –> reject H 0, meaning: The time trainers spent on the tool does not indicate whether the trainers used the tool more or less.
Trainers who spent less user actions on the tool, used it more.	0.1644	p > alpha (0.05) –> reject H 0, meaning: The number of user actions trainers spent on the tool does not indicate whether the trainers used the tool more or less.

Table 7.1: Significance of the difference in time and number of user actions between trainers who used the tool frequently and those who did not

Next to comparing the number of user actions to the time and frequency of usage, the user actions themselves can be analyzed in more detail. The user actions can be differentiated between different types of actions, like button clicks or user inputs. The darker green and pink in Figures 7.3 and 7.4 indicate, how many of the total number of user events were 'accept' or 'accept all' button clicks. For almost all the trainers, the 'accept' button clicks are hardly visible; thus, almost nobody used them. Instead, by far the most common user action was to modify the suggested values (Figure 7.3). The number of 'accept all' button clicks being low is not surprising as it was intended to be clicked only once during the planning process. However, the number of 'accept' button clicks to be so low while manual adaptions by the trainers occurred a lot was not expected and brings up further questions. Why did the trainers make a lot of manual changes? Were the suggestions of the tool not accurate enough or were the trainers hesitant towards accepting the automated suggestions because of trust issues or because of different reasons?

To investigate these questions, additional metrics have to be considered. One of them is the difference between the values the automated tool suggested and the values the trainers changed them into. A large gap between them would indicate an insufficient accuracy of the tool. In this case, the reason for the large number of modifications is likely to be the functionality of the tool itself. On the other side, if the changes made by the trainers differ only slightly from the suggested values, it might be more due to the trainers' attitudes and habits that they modified many values. Subtracting all values the tool suggested from the values the trainer entered and taking the mean of these differences gives -2.47. However, some trainer inputs should not be taken into consideration because they were a mistake or the trainer got distracted and did not complete what they were typing. To filter out these inputs, the outliers were removed. Re-computing the mean of the differences without outliers gives -0.61. Hence, when trainers modified the tool's suggestions, it was often only by just a very small amount.

Another metric that can give insight into the trainers' attitude towards the tool, is at what point they clicked the 'accept all' button. If they clicked it directly after starting to use the tool, they cannot have had much time to look at the tool's suggestions meaning that they accepted them without much hesitation. To determine this, the time when the trainer opened the tool has been subtracted from the time the trainer clicked the 'accept all' button. On average, one minute and five seconds have passed before trainers clicked on 'accept all'. To interpret this value, it has to be put into relation to the total amount of time that the trainers used the tool to create a training plan. As described earlier, the maximum median amount of time that a trainer spent with the tool for one training plan is roughly six minutes (see Figure 7.5). Across all trainers, the average duration of a planning session with the tool is one minute and 14 seconds. Therefore, the time of one minute and five seconds, which it takes trainers on average to click on 'accept all' is relatively long compared to the total amount of time trainers use the tool.

How did trainers use the tool compared to the existing options?

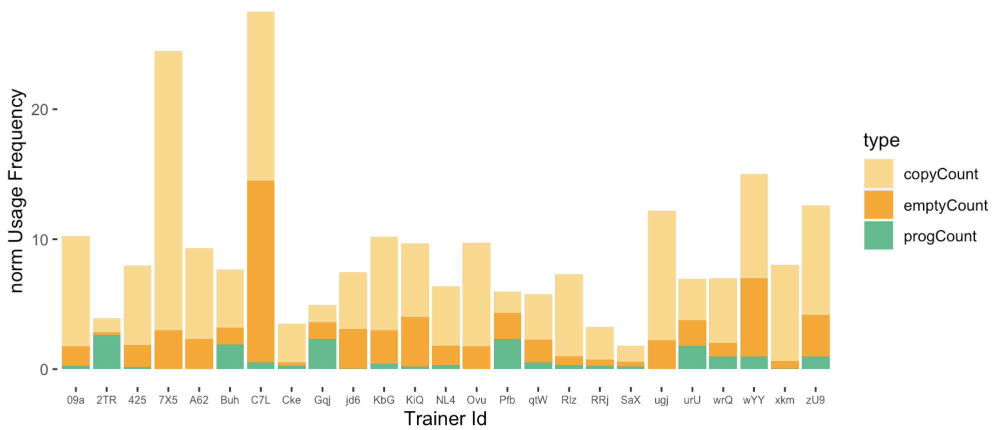


Figure 7.7: Number of times each trainer used the Progression Tool 'copy plan' and 'empty week' method to create a training plan (normalized by the number of trainees)

One of the main goals of this study is to identify a potential added value of the automated Progression Tool over the two already existing manual planning options. To achieve this, the usage of the tool is compared to how trainers used the already existing methods. One of these methods is to build a new plan from scratch called the 'empty week' method. The other method is to start planning a new training by copying and pasting the plan of the old week and adapting it, called the 'copy week' method. When the trainers wanted to create a new training plan, they could choose between these two methods and the automated tool. Figure 7.7 shows how many times each trainer used which option, normalized by the number of their trainees (light orange: 'copy week', orange: 'empty week', green: tool). Overall, the new planning tool clearly is the least used option. However, looking at the trainers individually, some used the tool more than the old methods.

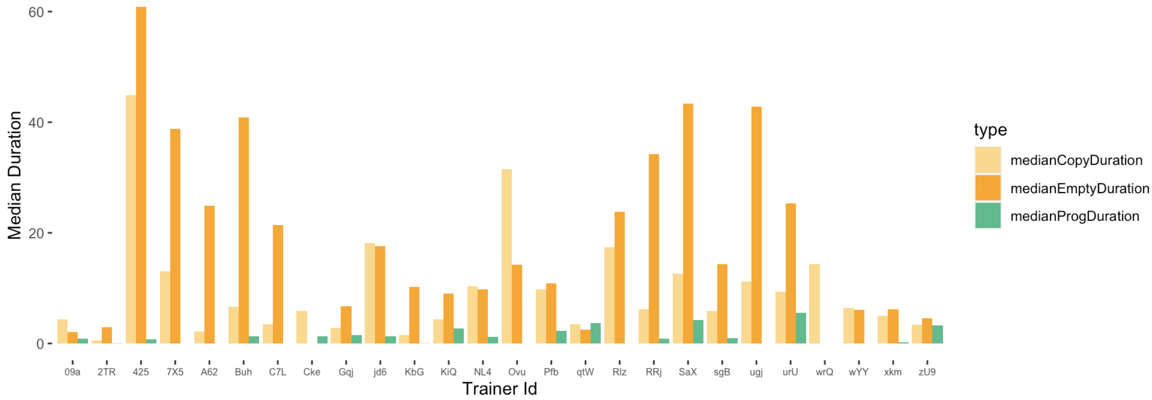


Figure 7.8: Median amount of time (in min) each trainer spent on each planning method (Progression Tool = green, 'copy plan' = light orange, 'empty week' = dark orange)

To determine whether the Progression Tool could facilitate and speed up the planning process, the average times required for setting up a plan with and without the tool are compared. Figure 7.10 visualizes the average duration of creating a training plan for each method. It shows that it took the trainers much less time to make a training plan using the automated tool than using the old methods. Setting up a plan from scratch via the 'empty week' method took 34 minutes on average, using the 'copy week' option 13 and with the tool only one minute. When looking at the trainers individually, it becomes apparent that the majority of them need most time for the planning option 'empty plan'. The amount of time needed for the 'copy week' method is also higher than the times needed with the Progression Tool, except for trainer 'qtW' and 'zU9'.

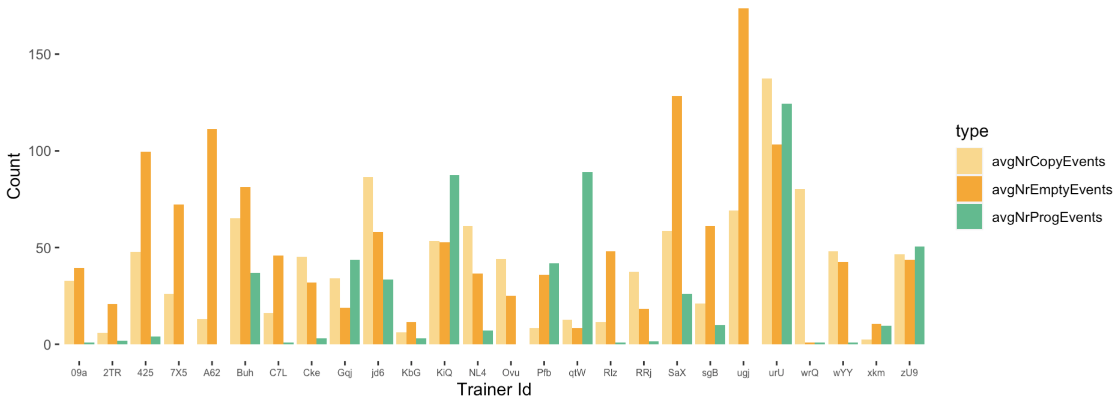


Figure 7.9: Average number of user actions each trainer spent on each planning method (Progression Tool = green, 'copy plan' = light orange, 'empty week' = dark orange)

This suggests, that the automated tool indeed speeds up the planning process and saves the trainers time. However, to verify this, it has to be ensured that in the measured time, the trainers really successfully used the tool and did not drop out halfway. This can be determined by analyzing the 'create' and 'close' click rates. The 'close' button stops the planning process and closes the tool. The 'create' button triggers the configuration of the plan, meaning the tool was used successfully. For the majority of the trainers, no 'close' button clicks were recorded. On average, the 'close' button was clicked 1.04 times through the study phase and the 'create' button 34.12 times. Hence, trainers mostly used the tool as intended and did not drop out while using it.

Since for the 'empty week' method the training plan had to be set up from scratch and also the 'copy week' method required a lot of adaptions, it seems plausible that using the Progression Tool takes less time. Accordingly, it was expected that the number of actions performed during the usage of the Progression Tool is also smaller. However, the difference is not as remarkable as for the time, as can be seen in Figure 7.11. The average number of user events per plan recorded for the Progression Tool is at 22. Using the ’copy plan’ and ’empty plan’ methods, trainers performed on average 41 and 53 user actions, respectively. This means that generally, trainers needed a lot less time for setting up a training plan using the Progression Tool, but still interact with the tool comparatively much. Figure 7.9 visualizes the number of user actions per planning method for each trainer individually. It reveals that two trainers even had the most actions when using the Progression Tool. Still, they needed a lot less time with the tool than using the ’copy plan’ or ’empty plan’ method. Possible reasons for this will be explained in the discussion.

In order to ensure, that both, the amount of time and the number of user actions spent on the tool, were significantly lower than for the existing methods, t-tests were performed; data set one of the two-sided t-test contained all time (user action) values recorded for the Progression Tool and data set two all values recorded for the ’copy plan’ and ’empty week’ method, respectively. The obtained p values confirmed that the time and number of user actions were significantly lower when trainers used the Progression Tool (see Table 7.2).

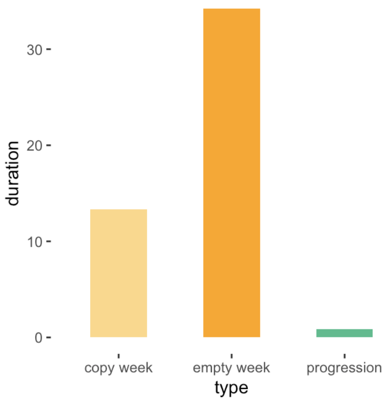


Figure 7.10: Average amount of time spent on each planning option (across all trainers)

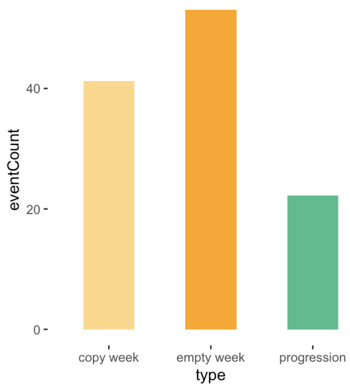


Figure 7.11: Average number of user actions spent on each planning option (across all trainers)

H 0	p-value	outcome of t-test
Trainers needed less time to create a training plan using the tool than using the ’copy plan’ (’empty week’) method.	0.0001639 (0.00001816)	p < alpha (0.05) –> accept H 0, meaning: The time trainers needed to create a training plan using the tool was significantly lower than using the ’copy plan’ (’empty week’) method.
Trainers had to perform less user actions to create a training plan using the tool than using the ’copy plan’ (’empty week’) method.	0.03895 (0.004982)	p < alpha (0.05) –> accept H 0, meaning: The number of actions trainers performed to create a training plan using the tool was significantly lower than using the ’copy plan’ (’empty week’) method.

Table 7.2: Significance of the difference in time and user actions needed to create a new plan using the different planning methods

Who used the tool the most?

The trainers who used the Progression Tool frequently will now be analyzed in more detail. For this, the trainers are sorted by usage frequency and split at the median usage frequency. This gives a 'low frequency' and 'high frequency' user group. Then, the types of sports of both groups, their educational background, and the scores on their technological affinity are considered. The left part of Figure 7.12 shows what sports the trainers do that used the tool frequently: one-third of them are Fitness trainers, and a quarter is active in powerlifting. Around 17% are bodybuilders and 8% is active as physiotherapist. On the other side, the division of sports among the trainers who did not use the tool frequently is less diverse. Almost half of them come from powerlifting and roughly one-third from bodybuilding. This brings up the assumption, that the tool was used less among trainers from powerlifting. To investigate this, t-tests were run on the sports types of the 'high frequency' and the sports types of the 'low frequency' user group. Thereby, there was only a differentiation made between 'powerlifting' and 'no powerlifting'. The outcome of the t-test did not confirm the assumption that powerlifting trainers used the tool less (see Table 7.3).

Next to the sports type, also the educational level of the trainers who used the Progression Tool frequently will be analyzed. To recall, in the pre-study survey, the trainers were asked where they got their knowledge as trainers. The answering options were: 'from practical experience', 'via role models', 'through my own research', 'via an apprenticeship', 'at university'. The first three options were combined into one group (category 1) and 'via an apprenticeship' and 'at university' were separate categories (category 2 and 3). Figure 7.13 shows the distribution of education categories among the trainers with ascending usage frequency. It suggests that trainers who studied used the tool tendentially more. Performing a t-test on this supposition yields that the p-value is indeed smaller than 0.05, meaning that it can be accepted (see Table 7.4). The t-test was made, as before, by splitting the trainers at the median usage frequency into two groups, a low- and a high-frequency group. Then, the education categories of the first group are tested against those of the second group.

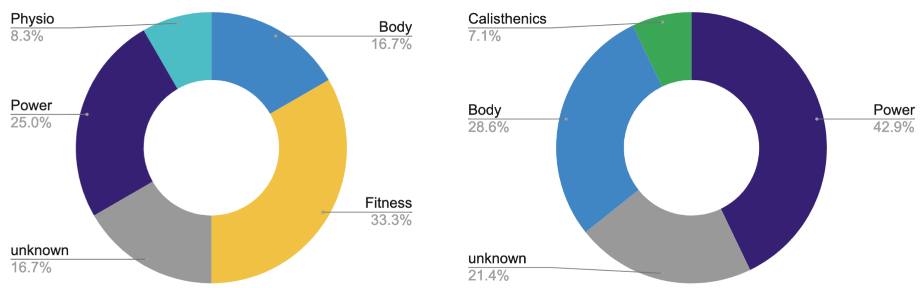


Figure 7.12: Sports types of trainers who did use the Progression Tool frequently (left) and sports types of those who did not (right)

H 0	p-value	outcome of t-test
The tool was used less among trainers who did powerlifting.	0.2329	p > alpha (0.05) –> reject H 0, meaning: trainers from powerlifting did not use the tool significantly less

Table 7.3: Significance of the difference in sports types between trainers who used the tool a lot and those who did not

Furthermore, the technological affinity of trainers was analyzed in relation to the frequency of usage. The technological affinity was measured by presenting statements to the trainers from four categories, enthusiasm, competence, and positive and negative attitude with respect to technology. The trainers had to agree or disagree with these statements on a scale from 1 to 5. Based on their answers, a score for each category can be calculated. Figures 7.14 and 7.15 show the scores for the categories ’positive attitude’ and ’negative attitude’ of the trainers in relation to their frequency of usage. Figure 7.14 shows that no matter if trainers used the Progression Tool a lot or not, they all agreed almost equally strong to the positive statements about technology. On the other side, for the negative statements, there seems to be a relation between usage frequency and the amount of disagreement. Trainers who used to tool more also disagreed more with the negative statements about technology. Hence it could be argued, that trainers with a higher usage frequency have a less negative attitude towards technology. However, running a t-test on this in the same way as it was described for the educational background shows that for trainers who used the tool a lot, neither their positive nor their negative attitude towards technology is significantly different (see Table 7.4).

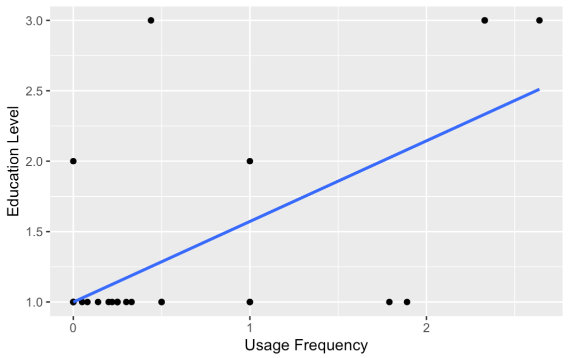


Figure 7.13: Correlation between frequency of usage and education

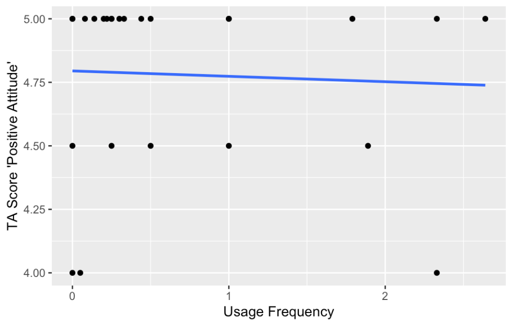


Figure 7.14: Correlation between frequency of usage and positive TA score

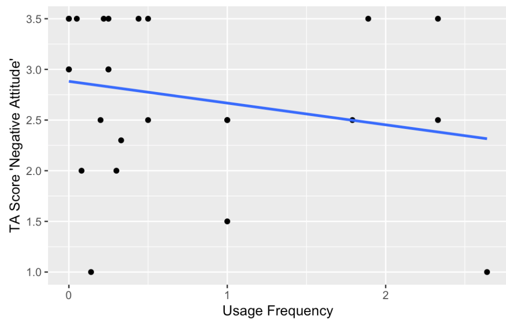


Figure 7.15: Correlation between frequency of usage and negative TA score

For the above correlations, all participating trainers have been considered. Now, the three trainers with the highest recorded usage frequency will be analyzed in more detail (see Figure 7.6). Two of them are active in Bodybuilding, and the other one is Fitness trainer. All three of them have studied and have between two to five and more than five years of experience. As already described earlier, the trainer who used the tool most, spent very little time and user actions on it. The other two also needed relatively little time when using the Progression Tool, but had an average number of interactions with the tool. Interestingly, these trainers used the tool even more than the two existing methods (see Figure 7.7). This brings up the question, of what benefits these trainers saw in the tool that motivated them to prefer it over the existing options. To investigate this, the time and actions they needed for creating a plan without the tool have to be considered. Figure 7.9 shows that the three trainers indeed could save a lot of time using the tool compared to the methods ’copy plan’ and ’empty plan’. However, two of them performed more user actions with the tool than without it. Possible explanations for these observations are given in the discussion.

H 0	p-value	outcome of t-test
The tool was used more by trainers who studied.	0.04117	p < alpha (0.05) –> accept H 0, meaning: among the trainers who used the tool a lot, the percentage of trainers who studied is significantly higher
The tool was used more by trainers who agree with positive consequences of technology more.	0.7866	p > alpha (0.05) –> reject H 0, meaning: the positive attitude towards technology of trainers who used the tool a lot was not significantly higher than of those who did not
The tool was used more by trainers who refuse negative consequences of technology more .	0.6879	p > alpha (0.05) –> reject H 0, meaning: the negative attitude of trainers who used the tool a lot was not significantly lower than of those who did not

Table 7.4: Significance of the difference in education and technology affinity between trainers who used the tool a lot and those who did not

As a last remark about the frequent users, a closer look is taken at the trainers with the IDs '2TR', 'KbG', and 'qtW', who used the Progression Tool increasingly over time (see figure 7.2). One of them is fitness trainer, one is bodybuilder and the third's sport type is unknown. Two of them have studied and one is trainer as full-time job. The fitness and bodybuilding trainers are active as trainers for already more than five years. Furthermore, their number of trainees is relatively high, between 8 and 39. Their TA scores lie within the average TA scores.

Who did not use the tool at all?

There were four trainers who did not use the tool at all. Two of them were powerlifters, the other two were bodybuilders. They got their trainer knowledge from their own research, practical experience, or role models, except one who did an apprenticeship. With this, and also with their TA scores, they fit in the correlation schemes shown in Figures 7.13, 7.14, and 7.15. One suggestion could be that they did not use the tool because they were inactive in general, but this is not the case as Figure 7.7 shows. They all created quite a lot of training plans, mostly using the 'copy week' option. They also spent quite a lot of time and actions using this method and the 'empty week' option, some even more than average. Hence, a reason for their rejection of the tool does not become apparent from the tracked data. Qualitative data is needed to be able to identify possible reasons.

7.2 Qualitative Results

Besides the tracked data, information on the trainers' user behavior was obtained from their feedback during the study and from the usability survey sent at the end of the study (SUS) [15]. Unfortunately, the participation rate was too low to be able to calculate a usability score. Nevertheless, the trainers' answers and feedback gave insight into their opinions about the tool. Some of them are quoted in the following:

Interim Feedback

> "Not very useful for me as the suggestions did not match the exercises." (ID: RRj63)
>
> "The better comparability to the previous week makes programming easier." (ID: Gqjfy)
>
> "weekly adjustment is faster because you don't have to switch back and forth - automatic adaption of the rpe value would also be nice (from week to week, 0.5 or 1

higher) [...] to make it perfect, a small icon should be added to adjust the notes to an exercise (after excepting the suggestions)" (ID: urUES)

Final Survey (SUS

"I would continue using the tool because it means less work and allows me to work more efficiently. [...] Using the Progression Tool has made training planning faster and more efficient." (ID: 2TREr)

8 Discussion

8.1 Interpretation of the Study Results

In the following, the study results will be discussed. It will be described how they indicate whether to accept the hypothesis or counter-hypothesis. To recall, the hypotheses and counter-hypotheses are the following:

Hypotheses on positive impacts of automation

I. Automation helps trainers to save time.

II. Automation increases the trainers' productivity and quality of work.

Hypotheses on negative impacts of automation

III. Automation restricts the integration of human expertise.

IV. Automation creates stress or skepticism among the trainers.

V. Automation causes a lack of motivation and attentiveness among the trainers.

The quantitative results showed that the time trainers needed to create a training plan using the Progression Tool is significantly lower than when using the existing methods. Both, the average required time among all trainers and the median time of each individual trainer were remarkably lower for the Progression Tool. The average or median values ensure that the time values are not distorted by the fact that the Progression Tool was used much less frequently than the other two options. Furthermore, the 'create' button clicks were considered in order to guarantee that the time values really represent the duration of a complete training plan creation, rather than of an unfinished process. Hence, the decrease in time required for a training plan with the Progression Tool implies that the tool does save the trainers time. This finding is also confirmed by the qualitative results obtained from the survey and the trainers' feedback. One trainer mentioned that using the Progression Tool the "weekly adjustment is faster [...]" (ID: urUES) and another one stated that "[...] the Progression Tool has made training planning faster and more efficient" (ID: 2TREr). Therefore, Hypothesis I can be approved. These statements, together with the decrease in time suggest that Hypothesis II is also true. Additionally, feedback like using the tool means "[...] less work and allows me to work more efficiently [...]" (ID: 2TREr) supports the hypothesis further. However, to be able to truly verify it, feedback data of more trainers is needed and also more data on the trainers' activities outside the training planning, like the interactions with their trainees.

All actions related to the planning process were tracked extensively, though, as presented in the previous chapter. The tracked data showed that the number of actions performed during a planning process with the tool was on average lower than for the other two options. However, some trainers had most actions when using the Progression Tool. This means, that the number of actions could not always be reduced with the tool, as opposed to the time. The question arises, why trainers on average needed less time while still performing relatively many user actions. One possible explanation is that the tool enables a more efficient workflow so that trainers can perform the required actions for creating a training plan in a shorter amount of time. The tool shows all necessary information and the suggested value adaptions at once. This way, the trainers do not have to click through previous plans anymore to recall the trainee's performance and figure out what to change the exercise values into, as it is the case for the existing methods. Given this explanation, Hypothesis II can be accepted. It can additionally be supported by the trainers' feedback and survey answers, which contained for example, that "the better comparability to the previous week makes programming easier" (ID: Gqjfy) and that "weekly adjustment is faster because you don't have to switch back and forth [...]" (ID: urUES). Furthermore, the planning with the Progression Tool was considered to be "[...] more efficient" (ID: 2TREr). However, as stated earlier, to be able to verify this hypothesis more reliably, more qualitative data and more tracking data of the trainers' actions outside the training process are needed.

The relatively high number of user interactions with the Progression Tool by some trainers needs further explanation as the tool was originally meant to reduce it. The tracked data revealed that the big majority of the actions were manual modifications of the suggested values and only a very small percentage were 'accept' button clicks. Furthermore, when trainers modified the tool's suggestions, it was often only by just a very small amount. This could either be due to the insufficient accuracy of the tool, which was also given as feedback by one of the trainers. Another possibility is that it was also the trainers' skeptical or hesitant attitude towards tool that caused them to change a lot of the suggested values. Given the second interpretation, Hypothesis IV is likely to be true. For a more unambiguous interpretation, more qualitative data is required about the trainers' opinion on the tool. Apart from the reasons behind the number of interactions with the tool, it is interesting, that neither the number of actions nor the amount of time needed for using the tool seemed to have significantly influenced the trainers' decision on whether they used the tool a lot or not (Table 7.1).

Furthermore, it has been found that the time it took trainers on average to click on 'accept all' was considerably long compared to the total amount of time trainers use the tool for. This observation implies that the trainers invested time in double-checking the tool's suggestions, leading to the assumption that trainers had a rather critical attitude towards the tool. This additionally supports Hypothesis IV, but at the same time contradicts Hypothesis V since the trainers still paid attention and did not just blindly rely on the tool. This finding together with the relative high number of manual modifications shows that trainers were still able to give input and integrate their knowledge; this contradicts Hypothesis III, which therefore can be rejected. Hence, it was possible for trainers to apply their expertise despite the use of automation.

Nevertheless, there are also other explanations why trainers tent to change the tool's suggestions a lot. A possible reason could be the novelty effect; this means that the trainers interacted with the tool a lot in the beginning to get to know its functionality. The period of time over which data has been tracked is not long enough to be able to verify this. Next to the novelty effect, another possible reason for the relatively high number of manual modifications is the trainers' habit. From the 'copy plan' and 'empty week' method, trainers were used to having to do a lot of manual changes. Hence they might have continued working in this way out of habit even when using the Progression Tool. To be able to make a valid statement on the trainers' habits, also tracking data over a longer period of time would be required.

8.2 Outcome

The Progression Tool has been used by some trainers more than by others. Therefore, there seem to be certain circumstances under which the tool was more useful. These circumstances can be identified by looking at the characteristics of trainers who used the tool frequently.

As presented in the results, the trainers with frequent usage had a more diverse sports background, including fitness and physio therapy, while the trainers who did not use the tool mostly came from powerlifting and body building. In these sports, training plans are often created based on a very specific theory developed by the trainer. This could explain why powerlifting and bodybuilding trainers were not as open towards using the tool as trainers from sports, where standard planning theories are more common. On the other hand, this cannot be generalized because as the t-test showed, there is no significant difference between the sports types of the trainers with frequent usage and those with infrequent usage.

Another category described in the results was the educational background of the trainers. Among the trainers who used the Progression Tool frequently, the portion of trainers who have studied is higher than among those who did not. This tendency was shown to be significant by the results of the two sided t-test. In addition to that, trainers with a high frequency usage had been trainers between two to five or for more than five years. For these trainers, it can be assumed that they have expertise and that applying their expertise is important to them. Hence, if these trainers did use the tool, it is likely that it did allow the user to integrate their knowledge. This implies that Hypothesis III can be rejected. On the other side, this implication is questionable since the assumption that trainers with more years of experience and a university degree have more expertise is debatable.

An additional characteristic about trainers investigated in the results was their affinity to technology. It was found that everyone equally agreed with positive statements about technology, but the trainers who used the tool more seemed to have less negative opinions about technology. However, the t-tests showed that this relation is not significant. Hence, no relation could be confirmed between usage frequency and the trainers' attitude towards technology.

Lastly, it was investigated whether the number of trainees of the trainers had an influence on their usage frequency. It was found that the trainers who used the tool increasingly over time, had more trainees than average. For these trainers, the tool's benefit of being able to save time is especially relevant. Therefore, it is possible that the tool was especially useful for trainers with many trainees. In this context, it is also relevant to recall that almost all trainers want to provide personal support to their trainees. Providing personal support can be considered a typical human skill; thus, the trainers are unlikely to use an automated tool that does not allow them to apply this skill. If trainers used the tool frequently, this could indicate that it does enable the integration of human skills, i.e. that Hypothesis III can be rejected. However, there is not enough user data available to be able to verify this statement.

8.3 Limitations and Future Work

As mentioned in the discussion, many of the results from the tracked data cannot be generalized because of the low number of trainers who agreed on the data tracking and the relatively short time period of the data collection. Data from more participants over a longer study phase could have enabled a more reliable categorization of who did and did not use the tool; also the decreasing trend of the usage frequency of the tool (see Figure 7.1) could have been analyzed further with more data. Moreover, a limitation of the tracked data was that it only represented the trainers' activities that were related to the planning of trainings. With data on other activities, such as how much the trainers communicated with their trainees, more insights on the impacts of automation could have been collected. Additionally, it would have been also interesting to capture the trainees' perspective by tracking how their actions and experience changes with the use of automation.

The more severe limitation, however, was the extremely low participation rate in the usability survey at the end of the study. Therefore, a lot of qualitative data was missing that would have been needed to be able to understand the trainers' user behavior better. The results from the quantitative data sometimes left multiple interpretation options, as it was the case for the high number of user actions that had been tracked for the tool. More qualitative data would have endabled a better understanding whether this was indeed due to the trainers' skeptical attitude or because of the tool's insufficient accuracy. Especially for better capturing the trainers' attitude towards the tool and their feelings while using it, more qualitative data would have been needed. With this, in particular the negative impacts of the tool on the user, such as stress or fear, could have been determined in more detail. Hence, the findings of this paper might be distorted, representing the positive impacts more than the negative ones. This could have been also caused by the selection bias as it is more likely that trainers who liked the tool took part in the survey than trainers who did not. However, there were also several positive aspects that partly made up for these limitations, such as the unbiased representation of the user behavior obtained from the tracked quantitative data and relatively high participation rate in the first survey.

In future work, these limitations can be tackled by collecting more qualitative data on the user experience with automation and by collecting data from all affected users, in this case, not only from the trainers, but also their trainees. Furthermore, more quantitative data should be tracked that also includes information on the user behavior that is not directly related to the interaction with automation.

9 Conclusion

Many tasks that used to require human skills can now be easily completed by automation; however, it is widely agreed that the best results can be achieved when automation is used to complement human skills, instead of substituting them. Analyzing use cases in which this combination has been implemented revealed that it is not always clear whether automation does bring added value. Thereby, the focus was on scenarios in which the users had certain expertise, such as doctors, teachers, or therapists, to ensure the relevance of human skills. Despite the enthusiasm for the benefits of automation, its consequences on the users are often forgotten. The goal of this paper was to shed more light on the impacts by identifying possible drawbacks and benefits of automation on expert users. Therefore, a tool was developed that automatically generates training plans. This tool was meant to stand exemplary for the use of automation.

Before the tool was implemented, research has been conducted to find out what criteria an automated tool should fulfill to limit negative impacts. Possible negative impacts of automation are the loss of expertise, over-reliance, a decrease in attention and self-confidence or fear, stress, and skepticism. Positive effects, on the other side, are an increase in productivity and efficiency since fewer resources, like time and staff, are required and higher accuracy and accessibility. It has been found that negative impacts mostly arise when automation is used as a replacement instead of augmentation of human skills or when the user context has not been taken into account properly. A factor within the user context that has been identified to be very relevant for the human-centered use of automation was trust. In order for the user to trust automated tools they have to fulfill certain criteria, such as authority by the user, explainability, and capability. The implementation of the training plan tool was based on these findings. Additionally, a user analysis was done before the implementation to ensure a human-centered approach. It has shown that the majority of the user group are trainers in powerlifting and have less than ten trainees. The levels of experience among the trainers are evenly spread out between less than one year and more than five years of experience, but most trainers have acquired their knowledge via their own research, role models, or practical experience. Furthermore, the user analysis revealed that most trainers are intrinsically motivated, meaning that they became trainers out of passion or to help others. Accordingly, it is important for them to provide personal support to their trainees. It was also found that most trainers have a high affinity towards technology.

With the automated training plan tool, a user study has been conducted to collect usage data on the trainers' interactions with the tool. This data was analyzed to identify possible impacts of the use of automation on expert users. The main takeaway from the tracked usage data was, that trainers needed much less time to set up a training plan using the tool. The number of required user actions for creating a plan also decreased with the tool, but not as much and as consistently as the time. Some trainers even had the highest number of user actions when using the tool. This brought up the question, how the tool allowed trainers to save time while they still performed relatively many user actions. One possible explanation is that the tool enabled the trainers to work more efficiently. Moreover, the tracked data revealed that most of the user actions with the tool were manual modifications of the automated values. Furthermore, the trainers accepted the suggested values after a relatively long time. With these findings, the first research question can be answered as follows:

An increase in productivity can be verified as a positive impact of automation on the user as the tracked user data showed that trainers saved a considerable amount of time. Furthermore, the users' feedback contained that trainers were able to work more efficiently using the tool. The user being restricted in applying their expertise and skills as a negative impact of automation cannot be confirmed for multiple reasons. The trainers changed the tool's suggested values a lot, meaning they did integrate their expertise; accordingly, they also did not accept all of the automated output immediately implying that they reviewed them before approving them. Moreover, the trainers who used the tool most had a degree, which is remarkable because this was the case for only 14% of all trainers. The high usage frequency of the tool by knowledgeable trainers indicates that it allows for the integration of human expertise. In addition to that, the user analysis showed that almost all trainers found it important to provide personal support to their trainees. This implies that they would not have used the tool if it did not allow them to integrate human skills, like the ability to keep personal contact with their trainees. However, to be able to generalize these findings, tracking data of more users over a longer period of time would have been required. In particular, the lack of qualitative data made it difficult to identify valid findings on less measurable negative impacts, such as reactions and feeling like stress, boredom, or fear among the trainers caused by automation. Nevertheless, the tracked quantitative data showed that trainers were more critical than indifferent toward automation.

Overall, it can be concluded that automation can provide added value to expert users because it enables them to work faster and more efficiently. However, this is only possible under certain conditions: automation has to give users enough flexibility and input options and has to be tailored to the users' context. Hereby, the study results showed that the affinity of users towards technology is less relevant than the specific circumstances and challenges of the user of automation.

References

[1] Glenn pendlay: The training log. URL https://www.otpbooks.com/glenn-pendlay-training-log/.

[2] Alpha progression. Alpha Progression - Schnabel und Sinapius GbR. 2.25.1. Available at: https://alphaprogression.com/en, 2020.

[3] Hypertrophie: Muskelaufbau verstehen und richtig trainieren, Dec 2021. URL https://maxinutrition.de/maxi-blog/post/hypertrophie-muskelaufbau-verstehen-und-richtig-trainieren/.

[4] Mystrengthbook. My Strength Book Inc.. 2.0.240. Available at: https://mystrengthbook.com/, 2021.

[5] Juggernautai. Juggernaut Apps, LLC. 1.6.2. Available at: https://www.juggernautai.app/, 2022.

[6] enduco: Radfahren lauf app. Andre Siegl, Philipp Staudt, Gerrit Schelter, Benedict Strube, Lennard Schafer. 5.0.6. Available at: https://enduco.app/, 2023.

[7] Virtual fitness market, Apr 2023. URL https://www.globenewswire.com/news-release/2023/04/20/2650716/0/en/Virtual-Fitness-Market-Expected-to-Surpass-US-187-8-Billion-by-2032-Growing-at-a-CAGR-of-31-2.html.

[8] Myfitcoach. MyFitCoach GmbH. 1.13.4. Available at: https://www.myfitcoach.de/de/, 2023.

[9] Malik Anas Ahmad, Shamyl Bin Mansoor, Zohaib Amjad Khan, Waqar Ahmad Aqeel, and Shahroze Humayun Kabir. Benchmarking expert surgeons' path for evaluating a trainee surgeon's performance. In *Proceedings of the 12th acm siggraph international conference on virtual-reality continuum and its applications in industry*, pages 57–62, 2013.

[10] Ragami C Alves, Jonato Prestes, Alysson Enes, Wilson MA de Moraes, Thiago B Trindade, Belmiro F de Salles, Alan A Aragon, and Tacito P Souza-Junior. Training programs designed for muscle hypertrophy in bodybuilders: a narrative review. *Sports*, 8(11):149, 2020.

[11] Melanie Arntz, Terry Gregory, and Ulrich Zierahn. The risk of automation for jobs in oecd countries: A comparative analysis. 2016.

[12] Seth D Baum. On the promotion of safe and socially beneficial artificial intelligence. *Ai & Society*, 32(4):543–551, 2017.

[13] Martin Becker. Martin becker - trainingsplanung, 2023. URL https://www.martinbecker-trainingsplanung.de/Training-Trainingsmethoden.html.

[14] Patrick Berndt. Angewandte trainingswissenschaften – die 5x5-methode, Mar 2021. URL https://www.fitnessmanagement.de/fitness/fachartikel-angewandte-trainingswissenschaft-5x5-methode.

[15] John Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.

[16] Cardiac College. Progressing your exercise. URL https://www.healtheuniversity.ca/EN/CardiacCollege/Active/Progressing_Your_Exercise.

[17] Claire D. Costa. Chat gpt vs google: The seo outlook, Feb 2023. URL https://medium.com/geekculture/chat-gpt-vs-google-the-seo-outlook-eaf51410b20b.

[18] aiplusinfo editorial team. Automation vs ai: What is the difference, why is it important?, Sept 2022. URL https://www.aiplusinfo.com/blog/automation-vs-ai-what-is-the-difference-why-is-it-important/.

[19] Carl Benedikt Frey and Michael A Osborne. The future of employment: How susceptible are jobs to computerisation? *Technological forecasting and social change*, 114:254–280, 2017.

[20] B GE and Y CHOMBOS. Inside risks risks of automation: A cautionary total-system perspective of our cyberfuture. *COMMUNICATIONS OF THE ACM*, 59(10), 2016.

[21] Alexandre L Gonçalves, Gustavo R Alves, Lucas M Carlos, Juarez B da Silva, and João B da M Alves. Remote experimentation supported by learning analytics and recommender systems. In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*, pages 313–319, 2018.

[22] Report ID: GVR-3-68038-538-0. Fitness app market size, share trends analysis report by type (exercise weight loss, diet nutrition, activity tracking), by platform (android, ios), by device, by region, and segment forecasts, 2023 - 2030, 2022. URL https://www.grandviewresearch.com/industry-analysis/fitness-app-market.

[23] Abid Haleem, Mohd Javaid, and Ibrahim Haleem Khan. Current status and applications of artificial intelligence (ai) in medical field: An overview. *Current Medicine Research and Practice*, 9(6):231–237, 2019.

[24] Alex Hern. My week with chatgpt: can it make me a healthier, happier, more productive person?, Apr 2023. URL https://www.theguardian.com/technology/2023/apr/06/my-week-with-chatgpt-can-it-make-me-a-healthier-happier-more-productive-person.

[25] Alex Hern. What should the limits be? the father of chatgpt on whether ai will save humanity – or destroy it, Jun 2023. URL https://www.theguardian.com/technology/2023/jun/07/what-should-the-limits-be-the-father-of-chatgpt-on-whether-ai-will-save-humanity-or-destroy-it.

[26] Alon Jacovi, Ana Marasović, Tim Miller, and Yoav Goldberg. Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in ai. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pages 624–635, 2021.

[27] Nauman Jalil. Introduction to intelligent user interfaces (iuis). In *Software Usability*. IntechOpen, 2021.

[28] V López Jaquero, F Montero, JP Molina, and P Gonz´ lez. Intelligent user interfaces: Past, present and future. In *Engineering the User Interface: From Research to Practice*, pages 1–12. Springer, 2008.

[29] Katja Karrer, Charlotte Glaser, Caroline Clemens, and Carmen Bruder. Technikaffinität erfassen–der fragebogen ta-eg. *Der Mensch im Mittelpunkt technischer Systeme*, 8(2009): 196–201, 2009.

[30] Paolo Keßler. *Trainingsplanung im Kraftdreikampf - ein wissenschaftlicher Ansatz*. PhD thesis, 05 2020.

[31] Min Hun Lee, Daniel P Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez i Badia. Interactive hybrid approach to combine machine and human intelligence for personalized rehabilitation assessment. In *Proceedings of the ACM Conference on Health, Inference, and Learning*, pages 160–169, 2020.

[32] Evie McCrum-Gardner. Which is the correct statistical test to use? *British Journal of Oral and Maxillofacial Surgery*, 46(1):38–41, 2008.

[33] meghasaki. Difference between artificial intelligence and automation, Jul 2022. URL https://www.geeksforgeeks.org/difference-between-artificial-intelligence-and-automation/.

[34] Siddharth Mehrotra, Catholijn M Jonker, and Myrthe L Tielman. More similar values, more trust?-the effect of value similarity on trust in human-agent interaction. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pages 777–783, 2021.

[35] Lars Meyer-Waarden and Julien Cloarec. "baby, you can drive my car": Psychological antecedents that drive consumers' adoption of ai-powered autonomous vehicles. *Technovation*, 109:102348, 2022.

[36] Sangwoo Noh, Juyeon Han, Jinwoo Jo, and Ahyoung Choi. Virtual companion based mobile user interface: an intelligent and simplified mobile user interface for the elderly users. In *2017 International Symposium on Ubiquitous Virtual Reality (ISUVR)*, pages 8–9. IEEE, 2017.

[37] Elizabeth Odekirk-Hash and Joseph L Zachary. Automated feedback on programs means students need less help from teachers. In *Proceedings of the thirty-second SIGCSE technical symposium on Computer Science Education*, pages 55–59, 2001.

[38] Maria Soledad Pera and Yiu Kai Ng. How can we help our k-12 teachers?: using a recommender to make personalized book suggestions. In *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, volume 2, pages 335–342. IEEE, 2014.

[39] Maik Pister. Informelles interview zum ablauf einer trainingsplanung, Nov 2022.

[40] Stuart Russell, Daniel Dewey, and Max Tegmark. Research priorities for robust and beneficial artificial intelligence. *AI magazine*, 36(4):105–114, 2015.

[41] Prof. Dr. Albrecht Schmidt. Mensch maschine interaktion 2. Human Computer Interaction, April 2021. LMU.

[42] Stephen Seiler. What is best practice for training intensity and duration distribution in endurance athletes? *International journal of sports physiology and performance*, 5(3):276–291, 2010.

[43] Aditya Singh. Bodybuilding vs crossfit vs powerlifting vs calisthenics, Mar 2020. URL https://adityafitness.health.blog/2020/03/13/bodybuilding-vs-crossfit-vs-powerlifting-vs-calisthenics/.

[44] Nathalie A Smuha. The eu approach to ethics guidelines for trustworthy artificial intelligence. *Computer Law Review International*, 20(4):97–106, 2019.

[45] Johannes Steinhart. 5/3/1 wendler trainingsplan, Nov 2019. URL https://science-fitness.de/trainingsplaene/531-wendler-trainingsplan.

[46] Richard Tang, Xing-Dong Yang, Scott Bateman, Joaquim Jorge, and Anthony Tang. Physio@ home: Exploring visual guidance and feedback techniques for physiotherapy exercises. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 4123–4132, 2015.

[47] Ewan Thomas, Antonino Bianco, Esamuela P Mancuso, Antonino Patti, Garden Tabacchi, Antonio Paoli, Giuseppe Messina, and Antonio Palma. The effects of a calisthenics training intervention on posture, strength and body composition. *Isokinetics and exercise science*, 25 (3):215–222, 2017.

[48] Claudio AB Tiellet, André Grahl Pereira, Eliseo Berni Reategui, José Valdeni Lima, and Teresa Chambel. Design and evaluation of a hypervideo environment to support veterinary surgery learning. In *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia*, pages 213–222, 2010.

[49] Ehsan Toreini, Mhairi Aitken, Kovila Coopamootoo, Karen Elliott, Carlos Gonzalez Zelaya, and Aad Van Moorsel. The relationship between trust in ai and trustworthy machine learning technologies. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 272–283, 2020.

[50] Milka Trajkova. Designing ai-based feedback for ballet learning. pages 1–9, 2020.

[51] Viswanath Venkatesh and Fred D Davis. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 46(2):186–204, 2000.

[52] Viswanath Venkatesh, James YL Thong, and Xin Xu. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly*, pages 157–178, 2012.

[53] Leandro Vian and Marcelo de Gomensoro Malheiros. Analysis of the application of reinforcement learning algorithms on the starcraft ii video game. *Revista Destaques Acadêmicos*, 11(4), 2019.

[54] Dakuo Wang, Justin D Weisz, Michael Muller, Parikshit Ram, Werner Geyer, Casey Dugan, Yla Tausczik, Horst Samulowitz, and Alexander Gray. Human-ai collaboration in data science: Exploring data scientists' perceptions of automated ai. *Proceedings of the ACM on human-computer interaction*, 3(CSCW):1–24, 2019.

A Appendix

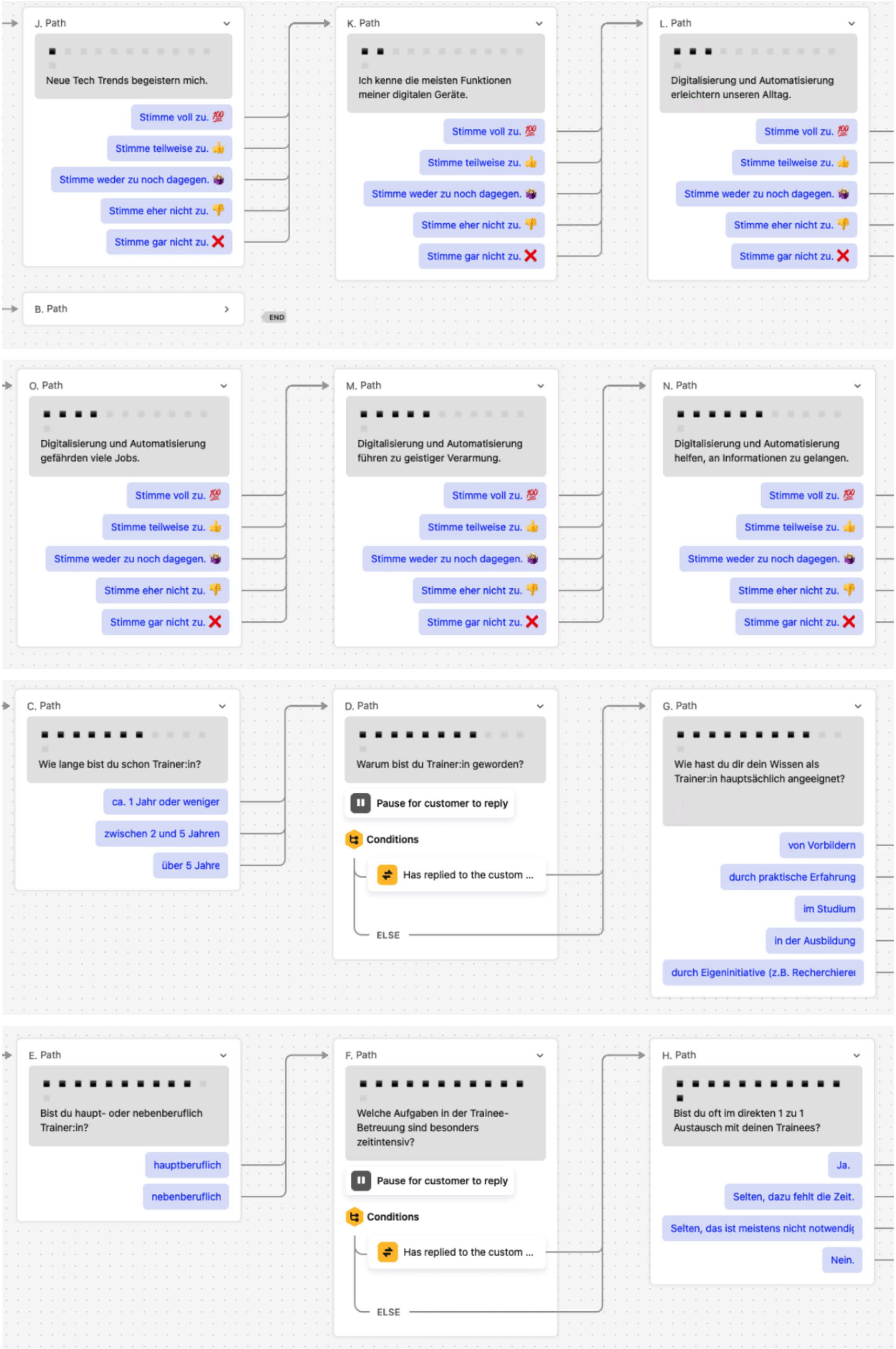


Figure A.1: Pre-Study Survey

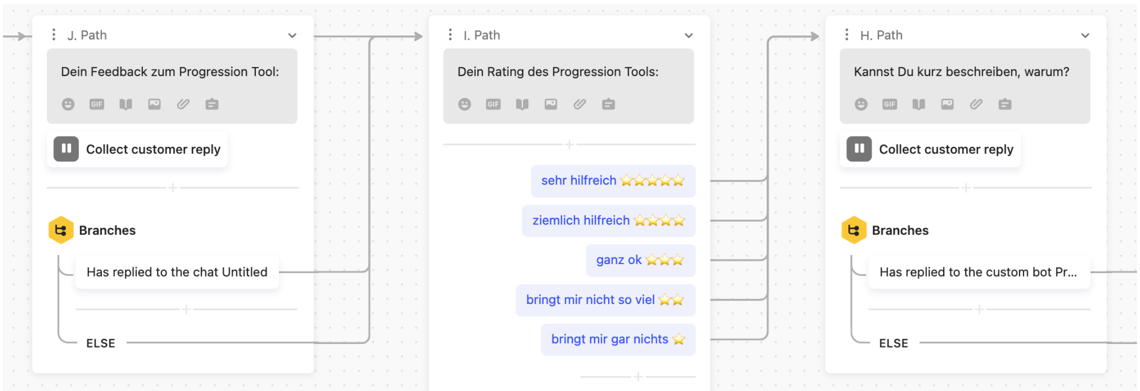
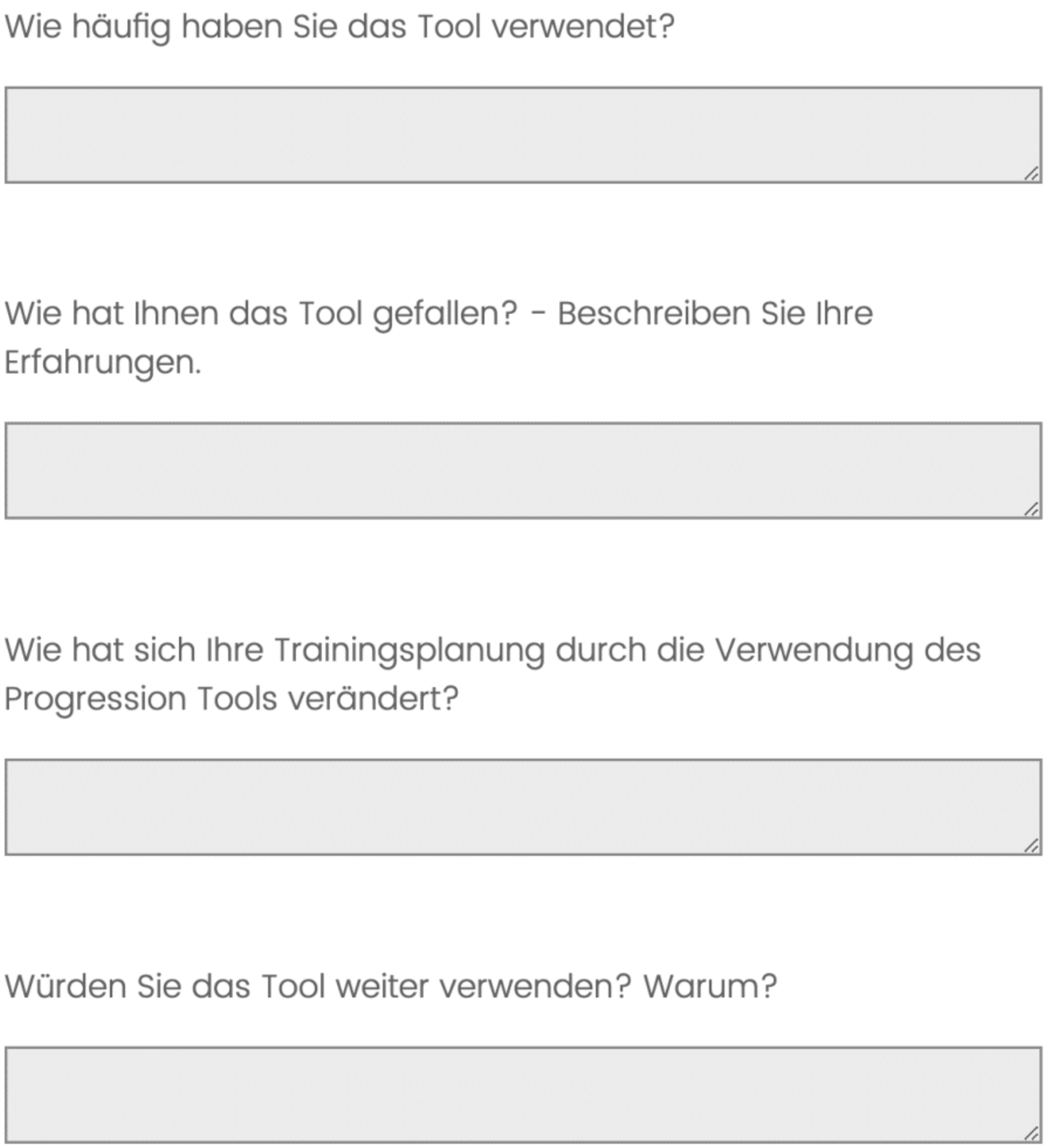


Figure A.2: Interim Feedback

Bitte beantworten Sie die folgenden Fragen im Bezug auf die Nutzung des Pogression Tools. Geben Sie dabei an, wie sehr Sie den folgenden Fragen zustimmen.

	Stimme überhaupt nicht zu				Stimme voll zu
Ich denke, dass ich das System gerne häufig benutzen würde.					
Ich fand das System unnötig komplex.					
Ich fand das System einfach zu benutzen.					
Ich glaube, ich würde die Hilfe einer technisch versierten Person benötigen, um das System benutzen zu können.					
Ich fand, die verschiedenen Funktionen in diesem System waren gut integriert.					
Ich denke, das System enthielt zu viele Inkonsistenzen.					
Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen.					
Ich fand das System sehr umständlich zu nutzen.					
Ich fühlte mich bei der Benutzung des Systems sehr sicher.					
Ich musste eine Menge lernen, bevor ich anfangen konnte das System zu verwenden.					

Figure A.3: SUS Survey Part 1



Wie häufig haben Sie das Tool verwendet?

Wie hat Ihnen das Tool gefallen? - Beschreiben Sie Ihre Erfahrungen.

Wie hat sich Ihre Trainingsplanung durch die Verwendung des Progression Tools verändert?

Würden Sie das Tool weiter verwenden? Warum?

Figure A.4: SUS Survey Part 2