Utrecht University

Game and Media Technology MSc thesis

Lostness in a building construction simulation

Klaas Hidde van den Berg, 3768309

PRO4ALL BUILDING TOGETHER

Utrecht University

supervised by
Dr. Herre van Oostendorp

Abstract

This research paper studies the observed lostness in a building construction project simulation. The simulation has been built from scratch, to ensure the availability of all relevant information. Lostness is measured by comparing the choices participants make to the optimal choices available. While users play the simulation, their lostness is monitored and when participants (employees in the construction branch) in the study group consistently score low on a specific category of questions, such as Innovation, the simulation is adapted: they are presented with an additional crisis situation which is meant to raise awareness, improve their performance and decrease their lostness. However, the used threshold for the adaptivity appeared not to fit the skill deficiency measured, leading to a low number of users presented with crisis situations late in the simulation. This led to a lack of informative results for applying dynamic adaptivity. Different thresholds, or simply placing the crisis situations earlier, should yield a more optimal environment. This could be tested in follow-up research. However, interesting correlations have been found: for instance, a significantly positive correlation between Education and the Quality and Risk score, and a negative correlation between Education and Age. Finally, this thesis could function as a starting point for future research, providing a basis from which adaptations could be made.

Contents

Chapter 1

Introduction

A man purposefully walks down a street, only to hesitate when arriving at an intersection. After consulting his phone, he sighs in relief and sets off to the right. A girl is making a drawing, but doubts her current progress. Checking the instructions, she erases a previous section and makes a second attempt. Two construction workers are arguing about the placement of a doorway, only to find out their blueprints differ in version and the supervisor is required to make the final call. What is the common factor in these examples?

These examples can all be attributed to “lostness”, the situation where a person or group of people is unsure as to how to proceed. As can be seen here, lostness occurs in any situation, not only in finding your directions. This thesis will explore lostness in the construction branch, and aim to find adaptivity that reduces this lostness and assists users in finding the optimal choices at any point in time. The lostness expertise will be provided by dr. H. van Oostendorp from (Utrecht, 2018), while the construction experience is available at (Pro4all, 2018).

First, related work will be examined and researched in order to establish a foundation from which to continue research. Afterwards, the research questions for this paper will be formulated, followed by a short terminology chapter. Next, the method will be elaborated upon, including the presentation of the information to the participants, calculation of lostness and participant procedure. Furthermore, the results and conclusion will be presented, followed by the discussion, future work and appendix.

Chapter 2

Related work

Related work can be derived from two main subjects for this research, being earlier research into Lostness and adaptivity, and research into the Construction branch.

2.1 Lostness

Lostness in computer applications has been observed as early as 1985, when these applications became more common and the increase in features and content necessitated a more navigable interface. (Elm and Woods, 1985) first examined participants' habits when working with analogous data when working with books, such as reading ahead, marking sections with shoestrings or physically marking a "trail" through the procedures. Thereafter, the same participants were given a crisis situation with a digital workstation (computer), and observations regarding their work methods were made. The findings included that aiding the user and not making him rely only on his memory greatly improved the user's navigation speed and understanding, and as such, that data availability design in itself is insufficient; the user should be kept informed of issues in parallel to his own execution, and an automatic indication of any relevant "Notes and Cautions" should be provided at all times. An important point this paper brings up is that lostness should be rated "...by a decrease in the ability to extract the information needed to successfully perform domain tasks, rather than by subjective feelings of being lost". This quote is a great starting point for the lostness research further down the line, aiding in developing the initial lostness measure to be used.

Over the next thirty years, although lostness has been researched to a degree, it has initially been seen as a less-than-crucial factor. This should not come as a surprise; a website or application requires content and features first, and when these are established, keeping them up-to-date can push a good interface into being a mere afterthought. In 1996, an attempt to actually measure or quantify lostness has been made. (Smith, 1996) recognises the importance of Elm's and Woods's research, stating objective measurements should outweigh users' feelings. The study has participants perform searches for specific information in a limited search space and analyses their behaviours; users inspecting many items might be simply exploring, or they might be lost. Apart from lostness, different measures are proposed, being "confidence" and "efficiency". Both of these are related, but are separately scored in the research. Six different values are measured, including total number of nodes accessed, number of different nodes accessed and total number of nodes visited whilst searching, and finally, these three measures are calculated. Ultimately, the paper proposes using path measures to measure lostness in hypertext systems, rather than using the traditional human-computer interaction measures of time and errors when assessing the usability of hypertext systems. Although this paper is extremely interesting and informative, much of the work and results are incompatible with this study, since no hypertext system as used there is being used here, with paths of possibly different lengths; as such, most of the findings can be used as inspiration, but nothing more.

The same, sadly, is true for many other works, including (Gwizdka and Spence, 2007), which abstracted lostness even further, attempting to retrieve structural and temporal measures that would accurately predict lostness. Their main research goal was to investigate "...how navigational efficiency and success may be assessed by considering the clickstream alone, without reference to the meaning or content of the web pages and links". After reviewing and selecting appropriate structural and temporal measures (including path measures, revisits, graph metrics and the similarity to the optimal navigation path) and increasing understanding of the commonalities and differences between these measures in order to find navigation styles, the research continues by studying which measures were the best predictors of either lostness or task success. After both a talk-aloud and a time-limit study, the results show similarity to the optimal path, navigation graph linearity and compactness, along with two second-order measures, can be useful diagnostics of user web behaviour.

Although not directly useful for this study, (Jeng, 2008) has included lostness as an aspect of the survey to rate the New Jersey Digital Highway. While this is in direct contrast to (Elm and Woods, 1985), since this work's main goal was not lostness, it is understandable the specifics of the field are not being adhered to. Unsurprisingly, lostness was found to negatively correlate with navigation; users who had no trouble navigating reported a low lostness, and vice versa.

Possibly more applicable here, (van Oostendorp and Karanam, 2013) has done research specifically into spatial lostness, with users finding their way through 3D virtual environments, and investigated the impact of providing model-generated navigation support in these environments. The support, mainly used in web environments, was argued to work in 3D environments as well, since many similarities between these types of environments exist. The model highlighted important aspects of the environment, such as information boards and directional arrows, to simplify path finding for the participants. The model turned out to accurately predict users' navigation behaviour, and, perhaps more importantly, managed to greatly assist the support group, while the control group without the support displayed significantly more lostness. Perhaps these methods can also be applied in this study, although the impact of their methods in this type of work might prove to remove too much challenge by clearly pointing out the preferred answers.

An extensive survey of the use of gamification has been performed by (Seaborn and Fels, 2015), studying many situations where game elements could enhance current systems. Education is a major field where gamification could assist greatly. The link between lostness and education is a simple one, since getting educated should reduce lostness. As such, aiding users with lostness issues in educational systems could increase their learning rate and, in turn, reduces the future lostness displayed. Gamification has been shown to motivate participation, increase the number of answers submitted and the duration of engagement, all reducing future lostness. Achievements, badges, feedback, graphical differences and other cosmetic changes could serve a twofold purpose: on the one hand, the user would be more motivated to continue, increasing long-term gains; on the other hand, these changes in themselves could already provide extra help with difficult problems, such as differently coloured buttons or achievements hinting at possible progress, to add another level of lostness reduction. These gamification elements could possibly also be applied in this study, increasing immersion or stimulating users to continue, while subtly assisting as well.

Very recently, (Shute et al., 2017) has provided insight into stealth assessment, retrieving information on the user's performance through information directly available, without having to ask the user for his feedback, and as such, without interrupting the user's workflow; this once again cycles back to (Elm and Woods, 1985), making sure objective observations are taken to measure lostness, instead of depending on subjective user feedback. While this study implemented many stealth assessment factors that could determine users' performance, and most of those were applicable in the game used, these factors are much more difficult to determine in a more focused program such as the one this study is developing. As such, sadly, many of the suggestions in this paper end up remaining unused.

2.2 Adaptivity

Once lostness is determined, however, the next step is adapting to the observed behaviour. This is called adaptivity, and can manifest in many shapes and forms. For instance, (Lieberman and Selker, 2000) advocates for using context to adapt to situations and previous information, even though most computer developers at that time were following context independence; abstraction from context ensures functions work the same regardless of the situation they are being performed in, and any consideration for the context might invalidate this certainty. Despite that, the more general availability of sensors and perceptual technologies, coupled with the rise of software agents for tools like mouses and keyboards, provides ideal situations to develop context-based adaptivity. This study examines context dependent systems and how these adapt to their users' habits. Although this study is quite specific in its research, focusing on a system called the COgnitive Adaptive Computer Help, or COACH, the underlying message remains; simply providing information context-free might be preferable for mathematical calculations, but for human-computer interaction, adapting to the user has great advantages. However, since the nature of these adaptations is not elaborated on much, the study is interesting, but not necessarily of great use to this work.

More general, and, as such, generally applicable, is (Jameson, 2003), a study into systems that adapt to their users in general. This study defines a very simple, yet very convincing, schema for processing in user-adaptive systems. The schema identifies five stages, with steps between them: the information initially known, acquisition of a user model from this information, the user model, application of the user model, and, finally, predictions or decisions about the user. This can be applied to virtually any user-adaptive system; for instance, a map app has information on previous use by the user, his detours and path following skills - this particular user might like paths through parks and tends to avoid busy streets. This information can be compiled to create a user model, where the app recognises these behaviours and generates a model of a quiet, relaxed user who does not mind taking a longer, but more pleasant journey. These parameters can in turn influence the planning, display or general path planning for the app, suggesting routes which appeal to the user right away. Although the original focus of this study was on suggestions and related searches, the applicability of this scheme makes for a great basis to work from.

Two interesting, but sadly not immediately useful, papers coincidentally both concern libraries. (Simeoni et al., 2009) studies how adaptivity can be implemented in gCube, a system for libraries to work with virtual research environments. Although adaptivity is the main focus of this work, the primary issue addressed is how to get adaptivity to work in the gCube environment, without going into much depth regarding the proposed adaptivity at all. (Frias-Martinez et al., 2009) evaluates adaptivity vs. adaptability for personalised digital libraries, but once again refrains from detailing the adaptivity actually being implemented; although concluding automatic adaptivity makes users perform better than manual adaptability does, the nature of the adaptivity being performed is not elaborated upon.

(Lopes and Bidarra, 2011) did a survey into adaptivity challenges in games and simulations, and gives quite a complete overview of different methods being used at that time. This study focuses on what and how adaptation and generation engines can or could adapt. It contains four sections, one of which will be discussed now while the other three will follow slightly later. The section to be discussed here is the Offline Adaptivity section, or making adjustments "considering player-dependent data, but prior to initiating any gameplay". As such, the study states it is mainly useful for content generation, and automatic content generation could be of great value for this. World generation is the main field where offline adaptivity is used, changing the world to be made depending on player actions or preferences. This is probably not of the most use for this study, but the other sections could prove to be more valuable.

Finally, in a different take on adaptivity, (Pujol et al., 2013) studies adaptivity in a storytelling application for the Acropolis Museum. The basis for the adaptivity is relatively simple, being confined to a different style of storytelling, focus on different areas and omission or inclusion of aspects relevant to the user profile, which is determined on entry into the museum by means of a short questionnaire. However, the results prove the concept is a promising one, being able to place users in categories and providing information accordingly. This is increased by a second layer of adaptivity; when users skip sections of the guide, this is a high certainty negative feedback. When users finish sections without interruption, this is a low certainty positive feedback. These types of feedback are then used to change the user profile and provide the most relevant information for the user. This type of interaction and on-the-go adaptivity is called online or dynamic adaptivity, and is what we are looking to use as well.

2.3 Dynamic Adaptivity

As seen before, (Lopes and Bidarra, 2011) contains four sections, but only one of them was discussed in the previous section. The other three all concern dynamic adaptivity, or adaptivity during gameplay. Firstly, the study distinguishes Steering Adaptivity, the purpose behind adaptivity, which can be used to better suit the game to a dynamic element such as player skills or team size. For entertainment games, this should aim at keeping the game fun, preventing "too easy" or "too hard" situations. Serious games can have different purposes than fun, though; they may aim at providing educational or training experiences, and in these cases, the adaptivity should steer the user into the direction of more effective knowledge transfer. While Steering Adaptivity is the purpose of the adaptivity, the target recipients are, for us, more important. Gameplay mechanics (such as aim assist in shooter games), the AI domain (like rubber band AI for racing games, or testing the player's weakness in sports games) and game narratives (altering the narrative due to the player's decisions, such as changing story elements or even enemy spawning based on challenge provided to the player) are the most often targeted areas, while game worlds (such as simpler levels for underachieving players) and scenarios (adding, removing or replacing events, for instance) are still lacking in broad, consolidated and integrated research focus.

This leads into the previously discussed section concerning offline adaptivity for content generation, but then also includes online adaptivity, or dynamic adaptivity: while playing, a player model is being kept track of, and every adaptation that is introduced is closely studied by the framework, where the player's reactions to the adaptation could lead to either a new adaptation or an update of the player model. At that time, research showed that online adaptivity was mainly concerned with adjusting challenge levels of NPCs and game AI (evaluating their current situation and adapting to rectify earlier mistakes, or predicting a player's actions and making decisions based on this information).

However, other directions have also seen some online adaptivity applications, including game levels (creating tracks for racing games during the race, based on players' characteristics, or using search-based procedural content generation to create personalised platform levels for fighting games) and even quest generation (using quests with skill tests matching the skill proficiencies in the player model). Sadly, most adaptivity was still predictable to a certain degree since all possible variations were required to be created a priori. The methods discussed here, including other methods still under development, could greatly assist in the methods being considered in this study, and the succinct overview of techniques should aid tremendously.

(Kickmeier-Rust et al., 2007) also utilises dynamic adaptivity, studying ELEKTRA in a learning environment and looking for possible ways to introduce adaptivity in a narrative. However, since the narrative is closely coupled with learning tasks, adaptively ordering and presenting learning tasks also result in a nonsensical narrative that is implausible. For this study, the solution was to introduce non-invasive microadaptivity; the adaptation system provides recommendations to ELEKTRA, but leaves the ultimate decision to the game engine as to whether or not to enact a recommendation, in order to ensure the changes are not detrimental to the learner's gaming experience. More information about the specific framework behind the adaptivity, and the narrative itself, is lacking, but the message to refrain from introducing elements detrimental to the user's experience is of great value for this study.

The final study of importance is (van Oostendorp et al., 2014), displaying very clearly how dynamic adaptivity can tailor to users' experience and assist in learning rates. This study presents a calamity situation where the user is required to rate the urgency of victims' medical needs, using the serious game Code Red Triage. The program then rates the user based on the correctness of their evaluations and their adhering to the steps from the official guidelines. As users score higher, the easier cases get removed from the game; this keeps the game from becoming too boring for more experienced users, but keeps the easier cases for weaker users to practice on. Although the level of engagement was not confirmed to be higher due to these changes, the learning efficiency was proven to be higher with these adaptive changes, showing that removing redundant situations could assist in improving knowledge transfer. This could also be of use for this thesis, and will definitely be taken into consideration.

2.4 Construction

This specific research will delve into the physical building construction branch, specifically looking into lostness in this sector. A simulation of a full construction project will be realised, making the user play every role in the process. For this, research into the construction branch has been done. Although information on topics such as software support or simulation assistance is available ((Flemming and Woodbury, 1995) & (Reichard and Papamichael, 2005)), these supply little information regarding the phases in construction and what scenarios would be preferable. More relevant are the researches into interesting scenario events, finding that a lack of knowledge and, specifically, motivation are the prominent causes of defects in construction ((Josephson and Hammarlund, 1999)), while financing is found to be the factor most responsible for delays in large building construction projects ((Assaf and Al-Hejji, 2006)). This information could be used to provide meaningful problems to users during the simulation.

Chapter 3

Research Questions

This thesis will have three research questions.

3.1 Lostness

First of all, the research will attempt to determine how lost users are in our simulation. This does not include lostness within the physical part of the simulation; the buttons should be easy to find and menus navigable. The research will purely focus on the decision making, assessing how well users can define and follow the optimal path at each point in time.

3.2 Reflection

Secondly, the process of measuring lostness will be critically studied. A lostness measure has been decided on, but how well does this measurement coincide with lostness or the lack thereof? Could other measurements, options or methods potentially improve upon the methods used? Results should be able to throw more light on this matter.

3.3 Dynamic Adaptivity

Thirdly, dynamic adaptivity will be investigated; is it possible to dynamically adapt the simulation to the displayed level of the user's lostness? This involves increasing difficulty when a user manages to follow the optimal path closely (van Oostendorp et al., 2014), or narrowing the choices when a user deviates from the optimal path, but the intention is to focus on the latter only.

The difference between regular adaptivity and dynamic adaptivity is based on the moment changes happen. Regular adaptivity takes data from the past, and applies it to new situations. For instance, when a new user starts the simulation, and the previous 10 participants had trouble with a specific section, the simulation changes that element. Dynamic adaptivity reacts to situations as they happen; a participant is having trouble with a specific group of questions, which makes the program reduce the difficulty of such questions later on, or in this case, introduce extra scenes to aid the user. In short, adaptivity involves changing a program depending on earlier behaviour, but it only makes changes in-between sessions, while dynamic adaptivity changes a program while it is being executed due to observed behaviour for this specific user. More information regarding this difference is clearly structured in (Lopes and Bidarra, 2011).

With this dynamic adaptivity responding to users' displayed lostness, this research aims to measure the difference between answers of users who got extra situations, and users who did not (the control group). Half of the users are placed in the control group, running the exact same program but without the extra situations when their scores are too low, while the other half gets extra situations when they score badly. Since the timing of users' answers is being kept track of, potential changes in answering patterns after encountering the crisis situation, compared to the answers given by the control group, can be analysed.

Ultimately, the goal of this thesis is to explore lostness in a simulation of a complete construction project. What important factors can be found, how can the degree of lostness be accurately measured, and is it possible to dynamically adapt the difficulty of a simulation to the user's displayed lostness? Hypothesised is that a performance measure should provide an accurate lostness value (Elm and Woods, 1985), which means concrete performance should also be measurable. In addition, dynamic adaptivity should be possible, but the exact manner is still to be determined.

Chapter 4

Terminology used

"Lostness" is the central concept for this research. As described earlier, it is the situation where a person or group of people is unsure as to how to proceed. This is quite a narrow definition, though, and other factors do come into play here. A better description for being lost for this research would be "to lack knowledge or insight into the importance of choosing a certain path, or not realising the drawbacks a method can bring".

Dynamic Adaptivity is another concept crucial for this research. Adaptivity in itself is adjusting something to something else, in this case adjusting the program depending on the input the user gives. Our goal is to aid users when they stray from the optimal path, which means dynamic adaptivity is required; adjusting the program while it is running, instead of in-between sessions (also online versus offline adaptivity). Users are expected to do only a single run through the program, after all. Applying dynamic adaptivity means their actions can change elements of the program, where the changes aim to help users who might be lost in the current situation.

Chapter 5

Method

5.1 Simulation construction

To make this project available for users, a number of factors were considered. While Pro4all already had several products in place, any of which could be altered and adapted to our needs, these products were specialised towards their own specific goals. Altering one of them to also measure lostness would not only be time-consuming and complex - keeping the current structure while adding functionality for measurement, detection of lost behaviour and keeping track of users' decisions - but would also make the program more confusing; its current goals and the new ones would inevitably clash, the existing products being complex in and of themselves, making lost behaviour difficult to classify and measure. This led us to the conclusion that building our own simulation, with the exact parameters chosen to introduce, should yield more optimal conditions and, consequently, results. The simulation would be created to be web-based, making access easy for users. Making a simulation, however, included some build time to get it running, after which a fitting scenario could be written to work through the simulation.

The simulation display is very straightforward, and the flow of the complete program can be seen in Figure 5.3. After a simple home screen containing a sole Start button, the user is shown the Introduction screen (text is provided in Appendix C). The user then reaches the Scene page, as can be seen in Figure 5.1. The upper text field sets the stage for the current scene. Images to the left, or occasionally articles which can be opened in an overlay, give additional information or clarification, visible in Figure 5.2. The bottom text field asks a question with regards to the situation, and a number of options are given. The option buttons then take the user to the next scene. The progress of the user is shown in the bottom, by means of both a progress bar, and a truck making its way along the bottom of the screen. Once all the scenes have been passed, the user is taken to a final survey, asking for some core details including education, job and years of work experience. All information, including choices made, exact

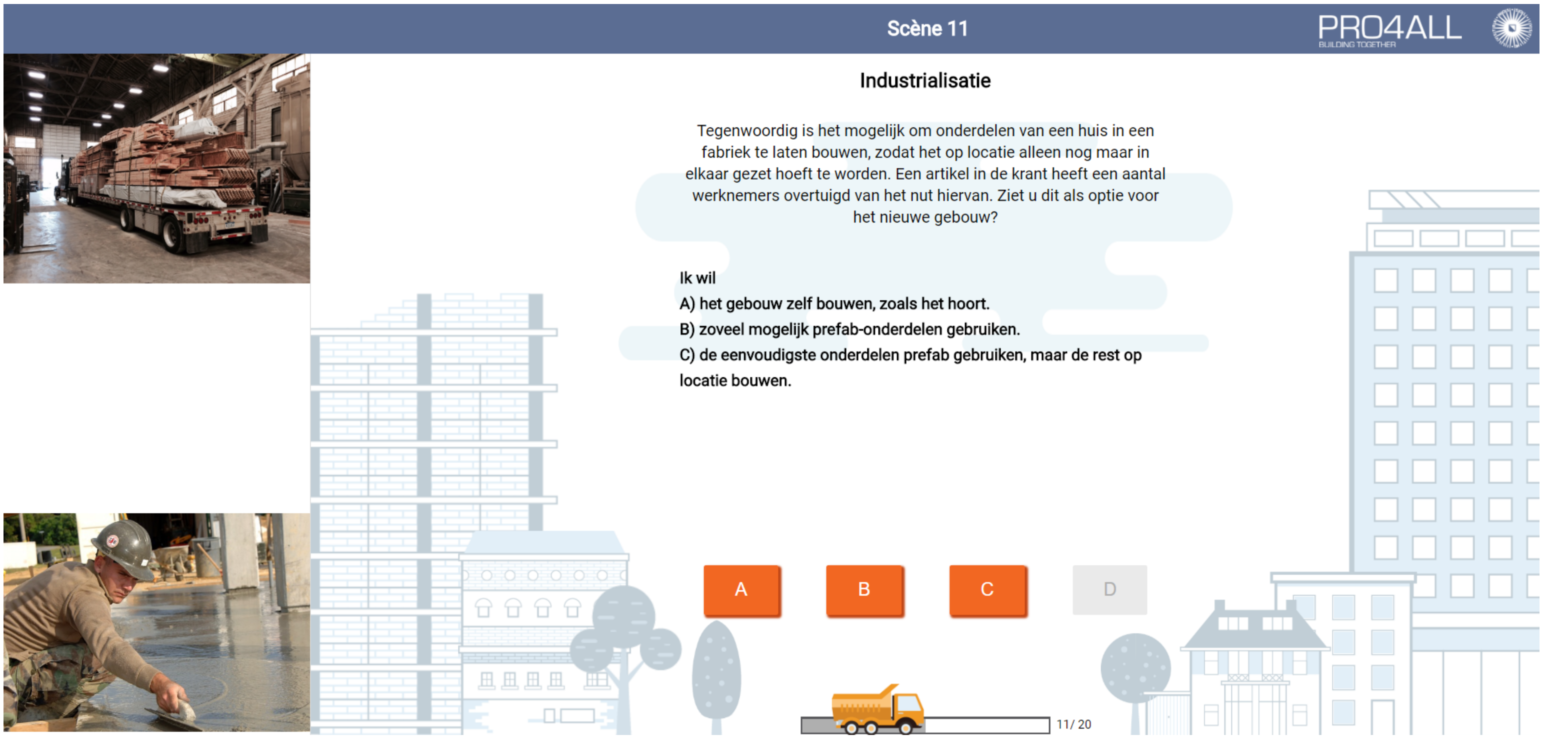

Figure 5.1: A scene in the simulation

- BREAKING NEWS -

Bouwvakkers vast in bouwput; opzichter zakt door de grond van schaamte

Twee bouwvakkers, opgesloten in een bouwput door een inschattingsfout van de opzichter! Het lijkt onwaarschijnlijk met alle voorschriften van tegenwoordig, maar dat is de situatie die plaats vond in Kleinhuizen, aan de rand van de stad.

Vanochtend om half tien was een team van vier bouwvakkers bezig funderingen aan te leggen. Hiervoor werd een bouwput uitgegraven, maar de opzichter, M. van den Brink, had één probleem over het hoofd gezien. "De grond is hier veel losser dan dat ik me realiseerde, en ging dus ook schuiven toen het gat te diep werd. Er lagen al bouwmaterialen op, die schuin het gat in vielen en nu de reddingsploegen tegenwerken."

Binnen tien minuten waren hulpdiensten ter plekke, maar de bouwmaterialen zijn ondertussen stevig in de grond vast komen te zitten. Op dit moment wordt er een tweede gat gegraven waardoor de bouwvakkers zo snel mogelijk kunnen worden geholpen. Eén van hen lijkt een kneuzing opgelopen te hebben, de ander maakt het naar omstandigheden goed. "Ik had nooit verwacht dat dit zou gebeuren, ik kan wel door de grond zakken van schaamte!", aldus Van den Brink.

Figure 5.2: An example of a newspaper article

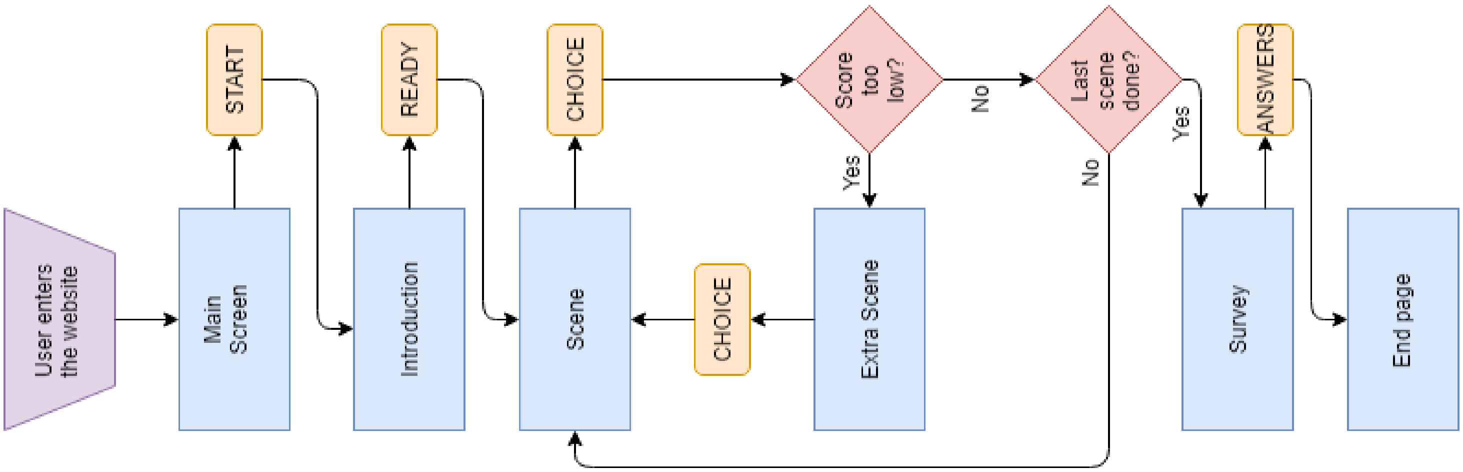

Figure 5.3: The flow of the program

time of choice and current score, are then saved in a database.

5.2 Scenario and scenes

Having decided on a general topic, and with the simulation program in development, an in-depth scenario was written. In collaboration with a former construction employee with more than 15 years of experience in the industry, an initial proposal was drafted. The scenario consisted of 20 scenes, each one presenting the user with a situation requiring an action. Answer options were provided, ranging from reserved responses ("I do not change this", "I keep working as we did" or "I do not tell my employees this") to significant redesigns of the work methods ("I move all our data to the cloud", "I aim for zero-emission" or "I incorporate this new method of working in our workflow"). The complete scenario covered most of a building project, from the conceptual design up to the finishing of a building. The scenes mostly focused on the executive aspect of a building project, from measuring errors and miscalculations to comparing contract prices and change orders. A number of scenes were also crafted in accordance with issues raised in (Josephson and Hammarlund, 1999) and (Assaf and Al-Hejji, 2006), such as material choice, funding and design decisions.

Discussion with our CEO resulted in a complete overhaul of the topics included, though; the focus shifted from the executive aspect to a more conceptual view, instead inquiring about innovation and digitalisation. These are more relevant to Pro4all, and since for the simulation this should not be a problem - the main requirement is the existence of an optimal path, which is still guaranteed - this was implemented instead. The new situations concern less physical, more intellectual problems, from investing in zero-emission and using digital blueprints to risk calculation and worker safety. The full script is included in Appendix A.

The skills and knowledge tested will be focused on conceptual knowledge, the ability to recognise optimal choices on the topics of digitalisation, innovation, quality and risk assessment. A good participant will recognise the optimal choices, will digitalise and innovate where beneficial and will take emission, worker safety and durability into consideration. Participants who make more conservative choices run unnecessary risks: a lack of digitalisation results in organisational difficulties for large projects, refusing to innovate can lead to outdated methods and impair the results, and the drawbacks of a lack of worker safety should be self-explanatory. Users who choose sub-optimal results are, as such, lost in the process of making decisions in their own field of expertise, while making the correct decisions has a huge impact.

5.3 Lostness calculation

To include dynamic adaptivity, three aspects need to be considered: monitoring an arbitrary measure to determine lostness, supplying possible adaptations, and then ensuring users who score low on the monitored measure are assigned one of these adaptations.

5.3.1 Monitoring

For the monitoring, following (Elm and Woods, 1985), a performance measure was decided upon; after all, objective performance measures should correspond with lostness better than users' feelings of being lost. Each scene was assigned to a category related to the construction industry:

- Digitalisation, for transforming analogous data into digital information;
- Innovation, for embracing new techniques and incorporating beneficial technological advancements;
- Quality and Risk, for durability, quality of life and worker safety

The scene distribution can be seen in Appendix D. With each scene having been assigned a category, each choice can now be assigned a score. For instance, the first question has three options; the optimal option will reward 10 points, the least optimal option will reward 0 points, and the in-the-middle option will reward 5 points. For different numbers of answers to a scene, the optimal and sub-optimal answers remain at 10 and 0 points, but the scores in between differ in scoring, as can be seen in Table 5.4. After all, in contrast to most earlier studies (such as (Elm and Woods, 1985), (Smith, 1996), (Gwizdka and Spence, 2007) and (van Oostendorp and Karanam, 2013)) which scored users on the number of nodes visited, this study investigates lostness displayed in domain issues. As such, the performance measure will still rate the distance from the optimal path, but instead of following a strictly different path (as would happen in a hyperspace study), the paths are the same, but the specifics change. This includes altering question text depending on earlier answers, or changing some information around based on choices made. Although initially, images were supposed to open on click, allowing tracking of the number of opening and closing of images as an extra lostness measure, this was later removed in favour of ease of use.

Number of answers	Sub-optimal	In-between			Optimal
2	0				10
3	0		5		10
4	0	3		6	10

Table 5.4: The scoring for individual scenes

These two pieces of information, the scoring per scene and the category distribution, lead us to the maximum scores; for digitalisation, with 10 scenes, this is 100 points; for innovation, with 5 scenes, this is 50 points; and for quality and risk, with only 4 scenes, this ends up being 40 points.

5.3.2 Adaptations

Once the monitoring of users' skills is in place, and with scores being kept track of, the next step is to supply possible adaptations. Initially, adding or removing answers to remove more difficult choices was planned, but with the advancement of the scenario and individual scenes, this seemed the less useful method. As such, the adaptation comes in the form of extra scenes, one for each category; these scenes are the last three in Appendix A, with the category also being the name. Contrary to (van Oostendorp et al., 2014) where easier cases are removed when users display experience and a lack of lostness, these scenes are crisis scenarios, being presented when a user displays lostness and a lack of awareness or knowledge. Issues arise and the user can decide to continue his current course (which led him to this situation), adapt and find a middle road between his current behaviour and the optimal path, or completely transform his strategy and drastically rethink his priorities. These extra scenes are mainly meant for raising awareness, and the answers given here, coupled with the answers given after the scene occurs, can be used to determine how well this method works in improving user efficiency and reducing user error, bringing them closer to the optimal path and, as such, reducing lostness. This method is also applied in some games, spawning monsters based on the challenge provided to the player (Lopes and Bidarra, 2011).

5.3.3 Adaptivity

Finally, once the extra scenes are in place, the adaptivity should be made dynamic. To respond to user lostness, the scores are monitored after every scene. As soon as a user's score in a category drops below a certain threshold, the next scene the user will see is the extra scene for that category. For this study, the threshold has been set to 50%; most other researches either do not use dynamic adaptivity, use another code base for their adaptivity thresholds (Kickmeier-Rust et al., 2007), or have used a pilot to determine a good threshold (van Oostendorp et al., 2014). For this research, a threshold of 50% of the maximum score was decided upon; this is to prevent users from seeing a crisis scene very early on when they make a few bad choices at the start (which would happen if the threshold were to scale off maximum current score).

Finally, the complete monitoring and dynamic adaptation aspects are clear; the user plays through scenes which are grouped into categories, choices made reward points toward the appropriate categories, and when a user's scores drop below 50% of the maximum possible score for that category, the corresponding extra scene is shown, aiming at improving the user's awareness.

5.4 Test subjects

In order to keep the results representative, the test subjects are all employees in the construction industry. Initially, the idea was to place a link to the web page online and share it through Pro4all's network; however, on second thought, stratified sampling was applied: the link was shared with specific individuals, decided upon by the Sales division, to ensure a more even spread among different specific occupations.

The test subjects are required to traverse the simulation, responding to situations as if this were a real situation and under ideal circumstances. Since all scenes concern construction topics, users are involved in the planning and designing, along with making process decisions and selecting work methods. Afterwards, they are to fill in a short questionnaire, which asks for their age, gender, job, years of experience, and education level and -field.

The optimal path through the simulation is defined in advance through the experience of experts in the construction field. When test subjects get the Dynamic Adaptivity flag and deviate too much from the optimal path, extra scenes will hopefully encourage them to return to it.

The goal for the thesis is to study the adherence to the optimal path participants display; how well do they follow the path, and where do they leave it? Will dynamic adaptivity in the form of additional scenes aid the users in following the optimal path? The dynamic adaptivity materialises itself through crisis situations provided to the participants; do they change their responses, considering they have been asked to respond to situations as they would under ideal circumstances?

5.5 Procedure

As can be seen in 5.3, users enter the website on a very empty main screen, only containing the "Start"-button. When clicked, this button leads them to the introduction. This shortly introduces the tasks awaiting, simply stating this is a building construction project and the user is asked to answer as they would prefer to respond if this were a real situation, and under ideal circumstances. Users are also asked to move to a quiet location and to avoid distractions. The full text can be found in Appendix C. As can be seen there, users are taught how the buttons work and how to proceed to next scenes.

After the introduction, where users can indicate they are ready to proceed (or not, which displays some texts with varying degrees of light humour), the Scene screen will accompany them through most of the simulation. Each scene will be loaded in, and answering options will be displayed. This all can be seen in 5.1, and is further explained in 5.1 as well. To keep track of the progress users have made, a progress bar with a moving truck is included, driving along the bar according to the user's scenes answered.

When the final scene has been passed, the user is shown to the Survey screen. This is quite similar to the Scene screen, but instead of a situation, question and a number of options, now there are multiple questions with text fields and radio buttons for responses, depending on the response required. When the user has filled in all the fields, they can continue and finish the simulation. The survey questions and possible answers can be seen in 9.

5.6 Materials

The front-end of the simulation is designed with HTML and CSS, alongside the functional JavaScript and a framework, Vue. The back-end is programmed in C# with ASP.NET Framework. Calls from front-end to back-end are done with Axios. The database runs on a PostgreSQL server, being approached with the Dapper library from the back-end.

Chapter 6

Results

Having contacted over 100 potential test subjects, a total of 27 results have been obtained. Each of the participants completed the simulation and filled in the survey at the end. Of the test subjects, 24 were male, 3 were female. 15 participants (13 male, 2 female) were assigned to the Dynamic Adaptivity group, while 12 participants (11 male, 1 female) were assigned to the control group. 7 of the participants had finished MBO, a 2-4 year practical job preparation; 17 of the participants had finished HBO, a 3-4 year higher-level job preparation, also known as University of Professional Education or University of Applied Sciences; and finally, the last 3 participants had finished WO, or University of Research. All of the participants were employed in the construction industry, with job descriptions ranging from overseer or engineer to head of IT or CEO. The ages range from 24 to 60, and work experience from 3 months to 30 years.

The first point of order is comparing the scores between participants. Since a performance measure was used to determine lostness, according to (Elm and Woods, 1985), the participants' scores should correlate directly with their lostness. The maximum scores available are 100 for Digitalisation, 50 for Innovation and 40 for Quality and Risk; the obtained scores can be seen in (Table 6.1). Scoring low in this simulation shows a distinct lack of knowledge and awareness for the optimal decisions to be made in the corresponding scenes; in short, participants with low scores displayed significant lostness during the scenario. Since the scenes test a user's recognition of situations and their knowledge of the optimal practices and decisions to be made at the situations presented, this lostness can also be seen as a lack of skill in the fields tested, being Digitalisation, Innovation and Quality and Risk. Low scores, or high lostness, imply too much adherence to outdated work methods, inefficient methods, low quality and disregard for safety regulations.

Scores	Digitalisation	Innovation	Quality and Risk	Total
	100	40	20	160
	51	35	20	106
	73	35	25	133
	83	30	20	133
	85	45	20	150
	90	45	10	145
	75	45	35	155
	73	30	20	123
	96	45	35	176
	61	45	20	126
	80	30	30	140
	83	35	20	138
	100	50	25	175
	75	35	10	120
	66	40	13	119
	83	45	20	148
	86	35	10	131
	84	35	0	119
	85	30	10	125
	100	30	20	150
	95	45	35	175
	81	45	30	156
	100	50	30	180
	90	35	20	145
	88	30	30	148
	85	50	30	165
	73	35	20	128
Mean	83	39	21	143
Max	100	50	40	190

Table 6.1: Scores per participant, on average, and maximum

Correlations

	Gndr	Educ	Age	WorkYrs	Digi	Inno	QnR	AvgDur	TotScr
Gndr	1	-.125	-.401	-.318	-.020	-.304	.011	-.074	-.110
Educ	-.125	1	-.272	-.124	.035	.196	.540	-.005	.324
Age	-.401	-.272	1	.749	.066	-.133	-.330	-.201	-.150
WorkYrs	-.318	-.124	.749	1	-.132	-.268	-.367	-.155	-.332
Digi	-.020	.035	.066	-.132	1	.274	.193	-.257	.783
Inno	-.304	.196	-.133	-.268	.274	1	.376	.010	.669
QnR	.011	.540	-.330	-.367	.193	.376	1	-.209	.683
AvgDur	-.074	-.005	-.201	-.155	-.257	.010	-.209	1	-.244
TotScr	-.110	.324	-.150	-.332	.783	.669	.683	-.244	1

Table 6.2: The results of the correlation test in SPSS

6.1 Correlations

A bivariate Pearson Correlation has been calculated using SPSS, where the education levels were labeled 1, 2 and 3, while the genders were simply labeled 1 and 2. The results are displayed in Table 6.2.

Obviously, the strongest correlations are between Age and WorkYears, a correlation of 0.749, and the correlations of 0.783, 0.669 and 0.683 between TotalScore and the three separate scores; although the question asked for the years of work in the current function, longer years require older participants, and of course separate scores count towards the total score.

The second strongest correlation is a lot more interesting, as it is a correlation of 0.54 between Education and the Quality / Durability score. This does provide new information; apparently, the more educated participants are more concerned with the environment and realise investing into quality improvements and risk reduction pays off in the long run. This is also seen in the correlation of 0.324 between Education and TotalScore, once again showing the higher total score due to higher separate scores.

The third strongest correlation is one that could very well be specific to our user pool; a correlation of -0.401 between Gender and Age, implying women in this group are younger than men, which is correct. However, whether this is a general trend or not is hard to determine, especially since there were only 3 female participants.

The other correlations are not very significant, but the following correlations can be interesting nonetheless:

- A correlation of -0.304 between gender and the innovation score, making our female participants less innovative than the male ones;
- A correlation of -0.272 between age and education, possibly hinting at highly educated people leaving the branch when gaining experience more so than lower educated people;
- Correlations of -0.330 between age and quality, -0.367 between workYears and quality, and -0.268 between workYears and innovation; apparently, older or more experienced participants were less likely to innovate or go out of their way for higher quality or risk reduction, highlighting a potential structual problem in the construction branch. This is also reflected in the -0.332 correlation between WorkYears and TotalScore;
- Correlations of 0.274, 0.193 and 0.376 between the digitalisation, innovation and quality scores, implying participants were usually either scoring well overall, or not well at all.

6.2 Dynamic Adaptivity

The dynamic adaptivity results, responding to a high displayed lostness, were not especially informative - four participants who were assigned to this group, actually scored low enough to be given any extra scenes, all on the topic of quality / risk. This, however, was always after the third-to-last or second-to-last scene, giving little room to show any potential improvement. On the extra scene, two of the participants responded with the moderate answer, one chose the most optimal course, while the last one gave the sub-optimal answer. The last one also scored the lowest on the next question, earning him a score of 0 on the quality questions. The two moderate responses both earned 10 points, and the most optimal response finished with 13 points. This seems to indicate the participants did not change their ways much, but, once again, due to the late occurrence of these scenes, it is hard to be certain in such statements, and the difference in final scoring does not seem significantly impacted either.

Group	Digitalisation	Innovation	Quality and Risk	Total
Control Before	73.8	39.6	15	128.4
Control After	7.1	0	8.7	15.8
Group 1 Before	73	35	0	108
Group 1 After	6.5	0	5	11.5

Table 6.3: Scoring before and after scene 18

Group	Digitalisation	Innovation	Quality and Risk	Total
Control Before	73.8	39.6	23.7	137.1
Control After	7.1	0	0	7.1
Group 2 Before	71.5	37.5	11.5	120.5
Group 2 After	4.5	0	0	11.5

Table 6.4: Scoring before and after scene 19

At this point, the participants can be divided into three groups: the group who did not encounter any extra scenes (which will include both the control group and the dynamic adaptivity group who did not receive extra scenes, since these users encountered exactly the same questions and answers as the control group, called control group), the group who encountered an extra scene as their third-to-last scene (as scene 18, noted as group 1 after this), and the group who encountered an extra scene as their second-to-last scene (as scene 19, mentioned as group 2 from now on). The following scores will be noted as digitalisation/innovation/quality and risk, for shortness's sake, and can also be seen in (Table 6.3). Before scene 18, the first group scored an average of 73.8/39.6/15, and an average of 7.1/0/8.7 afterwards. The two members in the second group scored an average of 73/35/0 before scene 18, and an average of 6.5/0/5 afterwards. The average score of 0 on the quality and risk category is the reason these users encountered this extra scene, and their answers to it clearly indicate their disposition afterwards; the participant who answered sub-optimally scored 0 points again, while the participant who begrudgingly changed his ways scored the full 10 points afterwards. However, there are enough participants who chose the same options without having encountered extra scenes, and as such, it is difficult to obtain any useful information from a statistical analysis.

The other distinction that can be made is the difference between the control group and group 2. Their scores can be seen in (Table 6.4). The control group, who did not encounter any extra scenes, scored an average of 73.8/39.6/23.7 before scene 19, and an average of 7.1/0/0 afterwards. The two members in group 2, who encountered the extra scene before scene 19, scored an average of 71.5/37.5/11.5 before this extra scene, and an average of 4.5/0/0 afterwards. The difficulty here is twofold: firstly, the 4.5 is not too different from many scores scored by users from the control group, but secondly, and more importantly, the categories for the extra scene and the last scene are completely different; while the participants scored low in quality and risk, the final scene covered the category of digitalisation, and as such, the extra scene should logically not influence the scoring, and did not seem to do this either. The users already scored slightly lower than average on digitalisation, and did so after the extra scene as well.

In the end, the average score on Digitalisation was 83 out of the maximum score of 100, on innovation, this was 39 with the maximum being 50, and on quality and risk it was 21 out of the 40 total. This could very well be due to the pool of participants - customers of Pro4all are already moving in a digitalised direction, and would be familiar with taking these decisions already. The highest scores were 100/100 for digitalisation, 50/50 for innovation and 35/40 for quality and risk. The lack of a 40/40 quality and risk score can be attributed to a question concerning safety regulations on the construction site, where nearly all users decided a general awareness meeting would be sufficient to prevent accidents from occurring; this is definitely not the case, awareness alone is not enough to prevent accidents at all. Finally, the lowest scores were 51/100 for digitalisation, 30/50 for innovation, and 0/40 for quality and risk. The high minimum scores for digitalisation and innovation imply some questions were too straightforward or not diverse / challenging enough.

6.3 Timing

So far, the focus has been on scores and the lostness and presence of extra scenes. However, the time participants have taken for each scene has also been taken into account, and the AvgDuration (Table 6.2) represents the average time a participant has waited before making his next decision. Taking this in consideration can be seen as the distinction between the plain performance measure and a secondary lostness measure, instead of simply looking at the score users manage to achieve, and should more closely relate to lostness; after all, comparing to physical lostness, a user can simply arrive two hours late after having wandered through town, which means he was lost for two hours, but would still score full points. As such, the score alone can be a good basis to provide extra scenes and adaptivity, but the score coupled with the time users took should fit users' lostness more closely. With the data collected, other factors seem not to point towards lostness in the same manner, leaving these two factors as the main contributing factors to the lostness measure.

The correlations do not seem to yield significant results, but there are still some that might be interesting. AvgDuration correlates negatively with every other value, after all, except for Innovation (with a score of 0.10, which is of no significance at all). Scoring -0.257 with digitalisation, -0.209 with Quality and -0.244 with TotalScore, it is clear that quick decisions impact the result, however small these influences may be. This also seems to indicate lostness is present in our simulation; users who are not lost will answer a question when having read the text and these users will be generally correct, while users who are lost might consider their options and end up still picking a sub-optimal option.

Chapter 7

Conclusion

In this study, the lostness of professionals in the construction industry was measured through a tailored simulation. The simulation tested the participants' knowledge, the application thereof, and skills in specific aspects of the field, and when a participant scored badly on a category out of digitalisation, innovation or quality and risk, the participant was provided with a crisis situation, to draw their attention to the possible outcomes of their current path and lead them back to the optimal path.

Three research questions were posed at the start of this study, and with the results being as they are, these can be answered to a certain extent now.

- How can lostness be measured accurately?
  Lostness, as already noted by earlier work such as (Elm and Woods, 1985), should be obtained using objective measures taken from users' performance. The approach used in this study, using the score during the simulation and also analysing the time required afterwards, seems to follow closely along with displayed lostness and the lack of knowledge and awareness that follow along with it.

- Is this measure accurate, and could other measures improve the method?
  As seen in 6.3, the time spent on tasks did correlate negatively with the score values, indicating it could play a vital role in the measuring of lostness. The achieved score, however, definitely played its part, and turned out to be a good choice for this study.

- Can the user be assisted dynamically?
  It turned out to be possible to include dynamic adaptivity within the scope of this project without too much trouble, aiding the user when awareness or knowledge was lacking. Although the dynamic adaptivity did not come into play as much as was expected, the functionality worked as planned and could easily be expanded.

The hypothesis was that a performance measure should provide an accurate lostness value; the answers to the extra scenes, coupled with other answers the participants gave, seem to justify this hypothesis. Answer score is definitely a concrete performance measure, making this a valid strategy. Finally, dynamic adaptivity successfully got implemented, but the effect it had on the users was problematic to determine due to the moment of adaptivity being near the end of the simulation, leaving little space to improve for the so inclined.

In addition to the answering of the research questions, interesting correlations have been found. Users tend to score similarly on the different categories, either scoring well overall, or missing points in all categories. More educated users are more concerned with quality, durability and worker safety, while older or more experienced participants tend to value innovation and quality and risk less, possibly implying a routine or difficulty to adapt to newer methods, even when these could be beneficial. These results can definitely be useful in the construction industry, pointing out flaws which, when fixed, could integrally improve the industry as a whole.

Chapter 8

Discussion

Along with researches such as (Elm and Woods, 1985), (Smith, 1996), (Gwizdka and Spence, 2007) and (van Oostendorp et al., 2014), this study has found that using objective performance measures as a measure of lostness leads to accurate and realistic results. Using the scores achieved by participants, the knowledge or awareness could be measured, which in turn could lead to the introduction of extra scenes which could help in this regard.

The dynamic adaptivity did get executed when expected; however, users scored better than expected, which led to a low number of participants who actually activated an extra scene. The decision to use a threshold of 50%, therefore, might have been erroneous or too simplistic. A better approach could be to implement multiple extra scenes for each category with thresholds at different points, such as 85%, 65% and 50%; these values are more educated than the 50% used in this study, since many users scored between 65% and 85%, which could then prompt one extra scene for that category. The extra scenes' severity could then be scaled to the threshold value, starting with minor issues which grow to crisis level as more thresholds are reached; when the user reaches 85% of the maximum score, a small problem arises, and if the user continues to miss points, at 65%, a more serious issue is encountered, while at 50%, a critical situation is presented. Another option would be to use a pilot to find a good threshold value; with the information from this study, if a single threshold was desired, using 75% would be a good start, although separate values for the different categories could be beneficial. The most important part would be to place the extra scenes more towards the front, instead of at the end of the simulation; with the method used in this thesis, that translates to a higher number as a threshold, which also makes more users encounter the extra scene.

The challenge or diversity of the questions probably also contributed to the scoring and threshold issues; although the scenario was designed to be challenging for anyone, perhaps the test subjects, being professionals from the construction industry who are connected to (Pro4all, 2018), were too familiar with these specific issues. A more general sample from the construction industry might yield more interesting results, while the question difficulty could also introduce more thought-provoking scenes.

A major self-imposed difficulty is the linearity of the simulation. The goal was to research the measure of lostness in the knowledge, knowledge application and skills in the areas of digitalisation, innovation, quality and risk management of employees in the construction branch; however, due to the path through the simulation always being the same one, and the extra scenes being the only deviation, many lostness measures can not be used. This turned out to hamper the analysis as well, since the performance measure was almost the same as the lostness measure. Instead of this linearity, a more diverse simulation would have been preferable; various paths, differing in length according to choices made, would assist in determining and proving lostness. Removing the image opening by clicking them also removed the last measurable increase in nodes visited, restricting the retrievable information.

In addition, the decision to not perform a knowledge test of any kind limits the conclusions that could be drawn with this information. These tests are performed in most lostness researches, such as (van Oostendorp et al., 2014), to have an objective, research independent measure to compare the lostness results with. Since this type of test did not seem to be recognised in the construction branch, this thesis did not perform one, which led to a missing comparison measure. Even better than a knowledge test of this type would be to gather objective performance data from the users in their own work environment, to use as an independent measure; however, to preserve anonymity and remain within the scope, this was impossible to include. As such, the timing is the best independent measure collected.

The analysis of information obtained from the simulation led to interesting insights concerning the construction branch, including the following highlights:

- More educated users score higher on Quality and Risk, being more concerned with the environment and worker safety than their lower educated colleagues. This could be seen as a major point, since the environment's impact in every industry is becoming increasingly tangible and worker safety is an ongoing issue.

- Age and the years in current position have negative correlations with the Innovation and Quality and Risk scores. This could be due to habits and mannerisms that have cemented themselves into position with more experienced professionals, but, in the case of Innovation, this could impede progress and advancement into a more optimised industry; while in the case of Quality and Risk, this could harm the environment or severely endanger workers in projects.

- The separate scores have positive correlations between each other, meaning participants were generally not specialised in any of the categories. This could also indicate an increase in knowledge in one field of expertise is directly applicable in other fields.

Chapter 9

Future Work

This thesis has found interesting results, but much more could be done in this field. Follow-up research with other thresholds for dynamic adaptation, more varying extra scenes, perhaps other categories or a different scenario, and a less linear path could definitely provide more insight into lostness and whether providing extra scenes to raise awareness can help users and reduce lostness.

On the other hand, adapting this work and using the framework and simulation while not supplying extra scenes, but using other forms of dynamic adaptivity (such as removing the simpler scenes, like (van Oostendorp et al., 2014) describes, or changing the scenes provided according to a user profile that is being kept up-to-date, like (Pujol et al., 2013) has done) could lead into completely new studies, using the underlying basis to study lostness, learning effiency or the effect of personalising content.

(Lopes and Bidarra, 2011) contains many great examples of adaptivity used in both serious and non-serious games, and many of these could also find use in future adaptations of this work. Providing additional scenes based on the challenge provided to the player is already supported, but other changes to be considered could include the following:

- a scenario more tailored to the user's skills, changing the scene structure to skip less challenging ones while providing additional scenes that the user struggles with.

- a more robust player model, using more information to base adaptations on than only the performance measure, the game score.

- keeping track of earlier decisions and making specific characters return with their own goals and agendas, to include a broader sense of "optimal" decisions.

Bibliography

Assaf, S. A. and Al-Hejji, S. (2006). Causes of delay in large construction projects. *International journal of project management*, 24(4):349–357.

Elm, W. C. and Woods, D. D. (1985). Getting lost: A case study in interface design. In *Proceedings of the Human Factors Society Annual Meeting*, volume 29, pages 927–929. SAGE Publications Sage CA: Los Angeles, CA.

Flemming, U. and Woodbury, R. (1995). Software environment to support early phases in building design (seed): Overview. *Journal of architectural engineering*, 1(4):147–152.

Frias-Martinez, E., Chen, S. Y., and Liu, X. (2009). Evaluation of a personalized digital library based on cognitive styles: Adaptivity vs. adaptability. *International Journal of Information Management*, 29(1):48–56.

Gwizdka, J. and Spence, I. (2007). Implicit measures of lostness and success in web navigation. *Interacting with Computers*, 19(3):357–369.

Jameson, A. (2003). Systems that adapt to their users: An integrative overview. In *Tutorial presented at 9th International Conference on User Modelling*.

Jeng, J. (2008). Evaluation of the new jersey digital highway. *Information Technology and Libraries*, 27(4):17–24.

Josephson, P.-E. and Hammarlund, Y. (1999). The causes and costs of defects in construction: A study of seven building projects. *Automation in construction*, 8(6):681–687.

Kickmeier-Rust, M. D., Peirce, N., Conlan, O., Schwarz, D., Verpoorten, D., and Albert, D. (2007). Immersive digital games: the interfaces for next-generation e-learning? In *International Conference on Universal Access in Human-Computer Interaction*, pages 647–656. Springer.

Lieberman, H. and Selker, T. (2000). Out of context: Computer systems that adapt to, and learn from, context. *IBM systems journal*, 39(3.4):617–632.

Lopes, R. and Bidarra, R. (2011). Adaptivity challenges in games and simulations: a survey. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(2):85–99.

Pro4all (2018). https://www.pro4all.nl/.

Pujol, L., Katifori, A., Vayanou, M., Roussou, M., Karvounis, M., Kyriakidi, M., Eleftheratou, S., and Ioannidis, Y. (2013). From personalization to adaptivity: Creating immersive visits through interactive digital storytelling at the acropolis museum. *International Conference on Intelligent Environments*, pages 541–554.

Reichard, G. and Papamichael, K. (2005). Decision-making through performance simulation and code compliance from the early schematic phases of building design. *Automation in Construction*, 14(2):173–180.

Seaborn, K. and Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of human-computer studies*, 74:14–31.

Shute, V., Ke, F., and Wang, L. (2017). Assessment and adaptation in games. In Wouters, P. and van Oostendorp, H., editors, *Instructional techniques to facilitate learning and motivation of serious games*, pages 59–78. Springer.

Simeoni, F., Candela, L., Lievens, D., Pagano, P., and Simi, M. (2009). Functional adaptivity for digital library services in e-infrastructures: the gcube approach. In *International Conference on Theory and Practice of Digital Libraries*, pages 51–62. Springer.

Smith, P. A. (1996). Towards a practical measure of hypertext usability. *Interacting with computers*, 8(4):365–381.

Utrecht, U. (2018). https://www.uu.nl.

van Oostendorp, H. and Karanam, S. (2013). Navigating in a virtual environment with model-generated support. In *Proceedings of the 31st European Conference on Cognitive Ergonomics*, page 17. ACM.

van Oostendorp, H., Van der Spek, E. D., and Linssen, J. (2014). Adapting the complexity level of a serious game to the proficiency of players. *EAI Endorsed Trans. Serious Games*, 1(2):e5.

Appendix

Appendix A: Script

Welkom bij het projectbureau!

De opdracht die je aangenomen hebt, bestaat uit twee delen. Ten eerste is er het renoveren van een bestaand kantoorgebouw, ten tweede het nieuw bouwen van een tweede, kleiner gebouw. Voor de renovatie is op dit moment alle informatie natuurlijk al beschikbaar, maar deze is allemaal in analoge vorm. Fysieke blauwdrukken en plattegronden zijn opgeslagen, naast oude briefcommunicatie en afspraken. In hoeverre werk je analoog, en hoeveel ga je digitaal doen?

Ik wil A) alle oude informatie analoog houden en analoog verder werken.

B) alle oude informatie analoog houden, maar digitaal verder werken.

C) alle oude informatie digitaliseren en digitaal verder werken.

Denken in dimensies

01:01De oude informatie is, met plattegronden en blauwdrukken, natuurlijk tweedimensionaal opgeslagen. De mappen met alle versies kosten wel wat tijd om door te spitten, maar uiteindelijk heb je een goed overzicht van wat er beschikbaar is. Al deze informatie is uiteraard 2D, wat makkelijk is met analoge informatie, maar nu je alles bij elkaar hebt, kun je ook 3D-tekeningen laten maken.—01:02De oude informatie is, met plattegronden en blauwdrukken, natuurlijk tweedimensionaal opgeslagen. De mappen met alle versies kosten wel wat tijd om door te spitten, maar uiteindelijk heb je een goed overzicht van wat er beschikbaar is. Aangezien je digitaal verder gaat werken, is dit een goed moment om alles samen te nemen en ook meteen een 3D-model te maken.—01:03Het digitaliseren van de oude informatie kost enige tijd, maar daarna zijn alle plattegronden en blauwdrukken makkelijker te navigeren. Nu alles digitaal staat, en aangezien je ook digitaal verder gaat werken, kun je nu relatief eenvoudig 3D-modellen laten maken.—

Ik wil

A) blijven werken in 2D.

B) 3D-modellen laten maken.

Nieuwbouw

Het nieuwbouwgedeelte van de opdracht zou gezien kunnen worden als een heel ander, nieuw project. Bij het renoveren van het huidige gebouw, koos je ervoor om 01:01de gegevens analoog te houden, analoog verder te werken—01:02de gegevens analoog te houden, digitaal verder te werken—01:03de gegevens te digitaliseren en digitaal verder te werken— en 02:04te blijven werken in 2D—02:053D-01:01tekeningen—01:02modellen—01:03modellen— te laten maken—. Een nieuw project is natuurlijk 01:01—01:02—01:03nog —makkelijker om volgens jouw voorkeuren op te zetten; zou je hierbij je werkwijze veranderen?

Ik zou hier A) op dezelfde manier als bij de renovatie werken.

B) meer analoog werken dan bij de renovatie.

C) meer digitaal werken dan bij de renovatie.

Contact

Nu de werkvorm vastgelegd is, kan het ontwerpen echt beginnen. Natuurlijk is er nog wel de vraag welk communicatieplatform je wil gebruiken en hoe verschillende partijen documenten zullen delen. Wat heeft jouw voorkeur?

Ik gebruik graag

A) Mail.

B) WeTransfer.

C) een online platform.

D) Post.

Risicoanalyse

De kranten staan vandaag vol met artikelen over een ongeval op een bouwplaats, niet ver hier vandaan. Jouw directie komt naar je toe met de vraag hoe jij ervoor gaat zorgen dat de risico"s te overzien blijven en hoe je bepaalt welke risico"s voorrang krijgen.

Hier

A) stel ik uitvoerende partijen voor aan.

B) schakel ik een bouwdirectie voor in.

C) voer ik zelf een analyse voor uit.

BREEAM

BREEAM is een instrument om de duurzaamheid van gebouwen te meten. Een hogere score kan voordelig zijn voor de toekomst, maar de eisen zullen wel een hogere investering vereisen. Probeer je een betere BREEAM-score te halen voor het renovatieproject, of houd je het project liever standaard?

A) Ik ga voor een zo hoog mogelijke BREEAM-score.

B) Ik probeer een balans te vinden tussen investering en score.

C) Ik houd de kosten laag door niet specifiek naar een score te streven.

Zero-emission

Na een kort onderzoek, blijkt dat het nieuwe gebouw ook als zero-emission gebouw gebouwd kan worden. Dit vergt natuurlijk een hogere investering dan een hoge BREEAM-score, maar het kan op termijn ook veel geld schelen tijdens de exploitatie. Ga je voor zero-emission, 06:16aangezien je ook veel moeite doet voor een hoge BREEAM-score?—06:17wat toch een stapje zwaarder is dan een redelijke BREEAM-score?—06:18hoewel dat nog meer investering kost dan een BREEAM-score?—

A) Ik ga voor zero-emission.

B) Ik ga niet voor zero-emission.

Checklists

Tijdens het gehele uitvoeringstraject vinden er regelmatig controles plaats, vele hiervan op basis van checklists: arbo, V&G, uitvoering, enz. Hoe voert u deze controles uit?

A) Ik doe dit op basis van papier en mail de resultaten door aan de betrokkenen.

B) Ik voer de controles en opnames digitaal uit, de betrokkenen worden direct geïnformeerd.

Ontwerpteam

Nu het conceptuele ontwerp grotendeels afgerond is en de technische specificaties relevant worden, is het de vraag welke partijen je hierbij inschakelt. In hoeverre maak je nu al keuzes voor later in het project?

A) Ik regel de technische specificaties vooral zelf.

B) Ik betrek uitvoerende partijen, zodat hun inspraak nu al gehoord wordt.

C) Ik werk verder met het team waarmee ik ook het concept heb opgesteld.

Kwaliteit

Kwaliteitsborging gaat binnenkort gemeengoed worden. Op welke wijze denkt u dit te gaan uitvoeren?

A) Wij gaan externe partijen inhuren die dit voor ons moeten gaan verzorgen.

B) Wij gaan uitgebreide checklists aanschaffen afgestemd op de gebouwen/woningen die we bouwen. De opnames doen we analoog.

C) Wij gaan uitgebreide checklists digitaal aanschaffen, afgestemd op de gebouwen / woningen die we bouwen. We nemen digitaal op, inclusief de opbouw van het gebouwdossier.

Industrialisatie

Tegenwoordig is het mogelijk om onderdelen van een huis in een fabriek te laten bouwen, zodat het op locatie alleen nog maar in elkaar gezet hoeft te worden. Een artikel in de krant heeft een aantal werknemers overtuigd van het nut hiervan. Ziet u dit als optie voor het nieuwe gebouw?

Ik wil

A) het gebouw zelf bouwen, zoals het hoort.

B) zoveel mogelijk prefab-onderdelen gebruiken.

C) de eenvoudigste onderdelen prefab gebruiken, maar de rest op locatie bouwen.

Communicatie uitvoeringsontwerp

Nu meerdere partijen hun aandeel geleverd hebben, kan het bouwproject beginnen. Aanpassingen aan plattegronden en blauwdrukken zijn nu veelvuldig en komen van alle kanten; hoe zorg je dat deze veranderingen alle partijen goed bereiken?

Wij gaan werken met

A) post.

B) e-mail.

C) een online platform.

Connectiviteit op de bouwplaats

De uitvoering kan bijna beginnen! Op het moment dat je gebeld wordt door één van de opzichters, merk je dat de verbinding slecht is, en dat het gesprek tenslotte helemaal wegvalt. Er blijkt zeer beperkt telefoonbereik te zijn op en rond de bouwplaats.

Hier doe ik

A) wel een keer wat aan, iedereen weet wat hij moet doen.

B) binnenkort wat aan, het is nuttig om te kunnen bellen als er iets is.

C) onmiddellijk iets aan, een goede verbinding is essentiëel voor het soepele verloop van de werkzaamheden.

Digitaal werken tijdens de uitvoering

Nu de uitvoering echt gestart is, is het uiteraard belangrijk dat er zo efficiënt mogelijk gewerkt wordt. In welke mate kan technologie daarbij een rol spelen?

A) Bouwtekeningen op papier zijn meer dan genoeg.

B) Telefoonoverleg is onmisbaar.

C) Alles online zetten levert tijdwinst op.

Augmented Reality / Clash Control

De projectleider wordt benaderd door een bedrijf dat zich specialiseert in technologie voor de bouw. Zij zijn in te huren om op de bouwplaats met augmented reality problemen en blokkeringen te vinden en die te verhelpen, en het gebruik van hun software kan ook voor onbepaalde tijd gekocht worden.

Dit zou ik

A) graag voor een paar dagen doen, maar niet langer.

B) kopen en langere tijd gebruiken.

C) niet doen, wij voorkomen zelf de punten van overlap.

Plaatsbepaling op de bouw

Nieuwe ontwikkelingen maken het mogelijk om je plaats via GPRS accuraat te bepalen. Hierdoor hoeft men niet meer zelf te meten, maar wordt dat voor hen gedaan. Zou je deze nieuwe techniek willen gebruiken?

A) Nee, wij kunnen zelf prima meten.

B) Ja, hier zou ik zeker gebruik van maken.

Veiligheid op de bouwplaats

De meeste ongelukken gebeuren op de bouwplaats; dit heb je regelmatig gehoord, maar ook nu weer is er een onderzoek gedaan waar deze conclusie uit komt. Een probleem is dat veel gevaren heel tijdelijk zijn en dat het veel tijd kost om deze tijdelijke gevaren aan te geven, terwijl ze binnenkort geen risico"s meer met zich meedragen. In hoeverre handhaaf je striktere veiligheidsregels?

Ik zorg ervoor dat

A) mensen bij overtreding gele en rode kaarten krijgen.

B) men zich bewust is van de gevaren en reglementen.

C) mensen een boete krijgen wanneer ze zich niet aan de regels houden.

Kwaliteitscontrole

Nu de uitvoering ten einde loopt, willen alle partijen natuurlijk zeker zijn van de geleverde kwaliteit. Eerder heb je gekozen om 10:26de kwaliteitscontroles uit te besteden—10:27kwaliteitscontroles met papieren checklists uit te voeren—10:28kwaliteitscontroles met digitale checklists uit te voeren—, hoe zou je de uiteindelijke kwaliteit nog controleren?

10:26A) Aangezien dit uitbesteed is, is dit de verantwoordelijkheid van de externe partij.¡br¿B) Wij controleren telkens de rapporten van de externe partij op correctheid.¡br¿C) Wij lopen zelf nog een laatste controleronde en leggen alles vast op een uitgebreid formulier.¡br¿D) Wij lopen zelf nog een laatste controleronde en leggen alles vast met foto"s en een formulier.— 10:27A) Wij lopen geen extra controleronde; dit hebben wij tijdens de uitvoering regelmatig gedaan.¡br¿B) Wij gaan alle formulieren nog eens na op fouten of discrepanties.¡br¿C) Wij lopen nog een laatste controleronde met een uitgebreid formulier.¡br¿D) Wij lopen nog een laatste controleronde waar wij foto"s nemen, naast een formulier in te vullen.— 10:28A) Alles ligt al digitaal vast, dus wij maken geen extra rondes.¡br¿B) Wij controleren de eerdere checklists op fouten of discrepanties.¡br¿C) Wij lopen nog een laatste controleronde met een uitgebreid formulier.¡br¿D) Wij lopen nog een laatste controleronde waar wij foto"s nemen, naast een formulier in te vullen.—

Archivering

Nu het project afgerond wordt, is de vraag natuurlijk hoe je al je gegevens wilt archiveren. Aangezien je 01:01compleet analoog—01:02vooral digitaal—01:03compleet digitaal— werkt, heb je je gegevens nu 01:01allemaal op papier—01:02voornamelijk digitaal—01:03allemaal digitaal—. Hoe zou je je gegevens willen archiveren?

Ik zou

A) 01:01alle papieren in mappen in onze archieven plaatsen—01:02alles uitprinten en met de oude documenten in mappen in onze archieven plaatsen—01:03alles uitprinten en in mappen in onze archieven plaatsen—.

B) 01:01alle documenten inscannen en op onze server opslaan—01:02de oude documenten inscannen en alle gegevens op onze servers opslaan—01:03alle gegevens op onze servers opslaan—.

C) 01:01alle documenten inscannen en online opslaan—01:02de oude documenten inscannen en alle gegevens online opslaan—01:03alle gegevens online opslaan—.

D) 01:01alle documenten inscannen en op onze server opslaan, maar ook de papieren in ons archief zetten—01:02de oude documenten inscannen en alles op onze server opslaan, maar ook de digitale bestanden uitprinten en alle papieren in ons archief zetten—01:03alles uitprinten en in het archief zetten, maar ook op onze server opslaan—

Einde

Heel erg bedankt! De simulatie is nu helemaal doorlopen! Op de volgende pagina staan nog 5 vragen; deze zijn van essentiëel belang voor het onderzoek dat ik hierna ga doen. Deze informatie blijft anoniem en kan onmogelijk aan u verbonden worden.

Gaat u alstublieft naar de korte vragenlijst.

Digitalisering

Een bezorgde werknemer komt naar je toe. Blijkbaar is er door een mismatch van plattegronden een probleem ontstaan, wat de plannen behoorlijk in de war brengt. Een misplaatste ballustrade, een onjuist geplaatste trap - de uitleg blijft vaag, maar de planning loopt hierdoor een klap op van een paar dagen, zo niet meer. Hoe zorg je ervoor dat dit niet weer gebeurt?

A) Niets, dit is éénmaal gebeurd en iedereen weet wel beter dan dit nog een keer voor te laten komen.

B) Ik breng persoonlijk de plattegronden bij iedereen langs; zo heeft iedereen de juiste versie.

C) Ik zorg voor een organisatiesysteem, waardoor men kan controleren of de huidige plattegrond de juiste is.

D) Ik zorg dat de plattegronden digitaal komen te staan, zodat de werknemers automatisch de juiste versie krijgen.

Innovatie

Investeerders hebben een spoedvergadering bijeen geroepen. Zij hebben gehoord dat je nieuwe ontwikkelingen liever niet meeneemt, maar zij zouden graag zien dat dit wel gebeurt. Deze ontwikkelingen komen ten goede van zowel de kwaliteit als, wellicht belangrijker, de prijs. Hoe reageer je hierop?

A) Dit is mijn project en ik werk op een manier waardoor ik het overzicht houd en weet dat ik goed werk lever.

B) Als zij dat graag willen, dan werken we wel met meer innovatie, maar dit gaat ten koste van mijn tevredenheid.

C) Dit is juist wat ik al hoopte, met de steun van de investeerders kan ik dit project nu moderniseren.

Kwaliteit

De gemeente heeft meegekregen dat dit project niet genoeg rekening houdt met kwaliteit, duurzaamheid en veiligheid. Een boze woordvoerder beschuldigt je ervan dat je teveel aan je portemonnee denkt en te weinig aan het welzijn van je medewerkers en toekomstige gebruikers van dit pand. Wat doe je met deze situatie?

A) Ik werk zoals ik dat gewend ben en ik weet prima wat ik doe.

B) Ik zal meer rekening houden met deze punten, maar de investeerders zullen hier niet blij mee zijn

C) De investeerders zullen inzien dat extra investering ook meer exploitatie en minder risico"s met zich meebrengt.

Appendix B: Survey

- Hoe oud bent u?
- Welk geslacht heeft u?
  - Man
  - Vrouw
- Welke functie heeft u op dit moment?
- Hoe lang werkt u al in deze functie?
- Wat is de hoogste opleiding die u voltooid heeft? (Naam studie en niveau)
  - Middelbare school of lager
  - VMBO
  - MBO
  - HBO
  - WO of hoger

Appendix C: Introduction

Welkom in deze simulatie. Dit project doorloopt een bouwproces, en ik zou van u willen vragen om keuzes te maken waar u onder ideale omstandigheden de voorkeur aan zou geven. Dit kan ik vervolgens onderzoeken voor mijn masterscriptie aan de Universiteit Utrecht, in samenwerking met Pro4all. Het doorlopen van dit programma duurt ongeveer 10 minuten. Internet Explorer wordt niet ondersteund; u kunt hiermee het programma proberen te gebruiken, maar dit raad ik niet aan.

Deze applicatie gebruikt cookies. Het enige wat deze bijhouden, is een ID waardoor ik achteraf al uw antwoorden aan elkaar kan koppelen. Aan het eind zal ik nog een aantal gegevens over uw werkervaring en situatie vragen; de gegevens blijven anoniem, en uw naam of bedrijf blijven mij onbekend.

Ik zou u willen verzoeken deze simulatie in één keer te doorlopen. Doet u dit alstublieft in een rustige ruimte waar u niet wordt afgeleid, zonder (telefoon)gesprekken tussendoor, enzovoorts. Op deze manier ben ik er zekerder van dat de antwoorden goed doordacht zijn.

De knoppen hieronder blijven in het scherm; als een knop rood is, dan kunt u deze gebruiken om uw antwoord te geven. De gele knop met een naam is een afbeelding, die extra informatie of visualisatie kan geven. Probeert u er eens op te klikken.

Bent u klaar om naar de vragen te gaan?

- Ja
- Nee

Appendix D: Scoring

Scene Number	Scene Name	Digitalisation	Innovation	Quality
1	Welkom bij het projectbureau!	x		
2	Denken in dimensies	x		
3	Nieuwbouw	x		
4	Contact	x		
5	Risicoanalyse		x	
6	BREEAM			x
7	Zero-emission			x
8	Checklists	x		
9	Ontwerpteam		x	
10	Kwaliteit	x		
11	Industrialisatie		x	
12	Communicatie uitvoeringsontwerp	x		
13	Connectiviteit op de bouwplaats		x	
14	Digitaal werken tijdens de uitvoering	x		
15	Augmented Reality / Clash Control		x	
16	Plaatsbepaling op de bouw	x		
17	Veiligheid op de bouwplaats			x
18	Kwaliteitscontrole			x
19	Archivering	x		
20	Einde			
21	Digitalisering			
22	Innovatie			
23	Kwaliteit			
0	Total	11	6	5